

# The Cost of Job Loss

KENNETH BURDETT

*University of Pennsylvania*

CARLOS CARRILLO-TUDELA AND MELVYN COLES

*University of Essex*

*First version received February 2016; Editorial decision January 2020; Accepted March 2020 (Eds.)*

This article identifies an equilibrium theory of wage formation and endogenous quit turnover in a labour market with on-the-job search, where risk averse workers accumulate human capital through learning-by-doing and lose skills while unemployed. Optimal contracting implies the wage paid increases with experience and tenure. Indirect inference using German data determines the deep parameters of the model. The estimated model not only reproduces the large and persistent fall in wages and earnings following job loss, a new structural decomposition finds foregone human capital accumulation (while unemployed) is the worker's major cost of job loss.

*Key words:* Job search, Human capital accumulation, Job loss

*JEL Codes:* J63, J64, J41, J42.

## 1. INTRODUCTION

Following [Topel \(1990\)](#) and [Ruhm \(1991\)](#) an important reduced form literature quantifies the surprisingly large and persistent earnings losses that follow layoff—the cost of job loss. For example, the estimates of [Jacobson \*et al.\* \(1993\)](#) and [Couch and Placzek \(2010\)](#) suggest that, six years after displacement, worker losses are between 13% and 25% of their pre-displacement earnings. The aim of our article is to provide a new structural decomposition of such earnings losses. We do so by extending [Burdett and Coles \(2003\)](#), which considers equilibrium wage contracting with on-the-job search, to the case of learning-by-doing while employed and skill loss while unemployed. The equilibrium approach is powerful not only because structural estimation of the model yields wage outcomes that are consistent with observed tenure and experience effects, it also consistently explains why wages paid are disperse across firms, why higher wage paying firms raise wages more slowly with tenure (see [Abowd \*et al.\* 1999](#); [Bagger \*et al.\* 2014](#)), while taking into account endogenous quit turnover—that workers typically quit for better paid employment.

Because the equilibrium market structure is consistent with the [Jacobson \*et al.\* \(1993\)](#) statistical approach, we use it to decompose the estimated cost of job loss into three constituent parts: (i) job ladder losses, (ii) human capital losses, and (iii) employment gap effects (the laid-off worker is more likely to be unemployed in the future with zero earnings). Using German data and consistent with the large earnings losses described above, the estimated cost of being laid-off is

large, being around 8–9% of expected lifetime discounted earnings. Typically unemployment policy focusses on compensating workers for foregone earnings while unemployed. But an important policy insight here is human capital loss, which is mainly due to foregone learning-by-doing (stalled career progression), is by far the biggest component of the cost of job loss.<sup>1</sup> Indeed the insight applies more widely for [Adda \*et al.\* \(2017\)](#) make a related point in the context of the career costs of raising children and gender wage gaps.

This article builds on the [Burdett and Mortensen \(1998\)](#) framework on equilibrium wage formation and labour turnover in frictional labour markets. This literature not only provides a structural interpretation for how wages evolve over individual worker careers, it also explains the surprisingly large variation in wage outcomes across firms and across workers (*e.g.* [Mortensen 2003](#)). Much of the recent literature adopts the sequential auction approach of [Postel-Vinay and Robin \(2002\)](#) for this yields a highly tractable econometric framework (see [Robin 2011](#); [Lise and Robin, 2015](#); [Bagger and Lentz 2019](#), among many others). The most closely related work which also considers learning-by-doing includes [Bagger \*et al.\* \(2014\)](#), [Jarosh \(2015\)](#), [Krolikowski \(2017\)](#), and [Jung and Kuhn \(2019\)](#). The contracting approach considered here instead assumes outside offers are not observed by the employer, and so there is no offer matching. Firms thus compete in optimal contracts and employees quit on finding preferred employment elsewhere. Tenure effects naturally arise in this framework because firms backload wages to reduce worker quit incentives (*e.g.* [Burdett and Coles 2003](#); [Stevens 2004](#); [Carrillo-Tudela 2009](#)). By allowing learning-by-doing while employed, wages exhibit both experience and tenure effects (*e.g.* [Altonji and Shakotko 1987](#); [Topel 1991](#); [Dustmann and Meghir 2005](#)). By allowing that human capital might decay while unemployed (*e.g.* [Pissarides 1992](#); [Ljungqvist and Sargent 1998](#)), we develop an equilibrium framework which is ideal for identifying the cost of job loss. Unfortunately although the decision theory is relatively straightforward, the equilibrium fixed point problem is formidable. It not only requires identifying a non-degenerate set of equilibrium (recursive, dynamic) wage contracts posted by firms, there are *ex ante* heterogeneous workers who all use optimal job search strategies (which are best responses to the equilibrium set of posted wage contracts) along with an endogenous joint distribution of employment, worker productivities, and tenures which is an evolving (infinitely dimensional) aggregate state variable. By adapting the notion of timeless equilibrium in [Woodford \(2003\)](#), we not only describe a tractable equilibrium framework, we provide a closed form characterization of equilibrium. Most importantly the market equilibrium yields an econometric wage structure which is consistent with the statistical frameworks of [Abowd \*et al.\* \(1999\)](#) and [Jacobson \*et al.\* \(1993\)](#).

A surprising takeaway is the central role played by learning-by-doing in explaining observed wage dynamics. For example, [Davis and Von Wachter \(2011\)](#) was the first to evaluate the cost of job loss using an equilibrium search model. But because that paper did not allow learning-by-doing, it could not explain the large measured earnings losses that follow layoff. Similarly, [Hornstein \*et al.\* \(2011\)](#) argue the equilibrium search framework does not seem consistent with the empirical *Mm* ratio, but that paper also does not consider learning-by-doing in conjunction with optimal contracts. Because unemployed workers wish to purchase “learning-by-doing” as an investment into higher future wages, equilibrium here finds the unemployed will indeed accept low starting wages consistent with the *Mm* ratio. There is also a very large literature on optimal unemployment insurance, though very little considers learning-by-doing. For example, recent work in the [Shavell and Weiss \(1979\)](#) unemployment insurance literature proposes an income tax increment when the laid-off worker finds work (*e.g.* [Hopenhayn and Nicolini 1997](#);

1. See [Schmieder \*et al.\* \(2016\)](#) who make a related point but instead using IV estimators based on government policy changes in the duration of the UI scheme.

Shimer and Werning 2008).<sup>2</sup> This proposal, however, does not take into account the laid-off worker's main loss is already reduced future wages. An integrated policy analysis is clearly an important direction for future research.

Our approach also identifies an important statistical issue for the Jacobson *et al.* (1993) framework. Although in this framework it is possible to difference out the worker-fixed effect, we show the estimated cost of job loss depends on job turnover parameters. Furthermore, the German data used here find that school leavers with few qualifications face much higher layoff rates over their careers than do workers with higher level qualifications (*e.g.* Adda *et al.* 2013; Burdett *et al.* 2016). Conversely, well-qualified workers have higher outside job offer rates and so enjoy greater job ladder gains while employed (through promotion and job shopping). Because different types have different costs of being laid-off, it is thus necessary to disaggregate the data when estimating the cost of job loss. Here, we find that just disaggregating into three different educational attainment groups already yields very good results.

Section 2 describes the model, and Sections 3 and 4 characterize and establish the existence of timeless equilibria. Section 5 describes the data and estimates the model using indirect inference. Section 6 then uses the Jacobson *et al.* (1993) methodology to estimate the cost of job loss both on the data and on the model simulated data. The results are remarkably well aligned, and we use the model to provide a structural decomposition of those costs. Appendix A contains the longer proofs. Appendix B provides a full description of the data, simulations, and the estimation procedures.

## 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 are indexed by  $j \in [0, 1]$ , are equally productive with a constant returns to scale technology. Workers are ex ante heterogeneous with general human capital  $k \in (0, \infty)$ . A worker  $k$  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  $\phi \geq 0$ . 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 generalize Burdett and Coles (2003) by allowing each firm  $j \in [0, 1]$  to precommit 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)  $\tau$ , human capital  $k$  and aggregate productivity  $A$  at that date. 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  $k_0 e^{\rho(t-t')}$  with aggregate productivity  $A(t)$ . Should an employee  $(\tau, k, A)$  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

2. The underlying idea is the UI system additionally operates a loans program—the worker is given more generous UI while unemployed but only as a loan. The loan is repaid through a tax increment when the worker is re-employed.

contract  $w_{jt}(\cdot)$  with current tenure  $\tau$ , and compares it to the value of being employed at the outside firm  $j'$  on contract  $w_{j't}(\cdot)$  but with zero tenure. No recall implies the worker quits if the latter contract yields greater value. Note this contracting approach rules out offer matching; *e.g.* Postel-Vinay and Robin (2002). The simplest justification is that outside job offers are not observed by the employer. We further suppose an equal treatment rule—that anti-discrimination legislation requires the firm’s wage policy to offer the same wage to equally productive workers with the same seniority. Thus should an employee receive a preferred outside offer, the worker is let go and the firm hires replacement employees on the company wage contract.

There are exogenous job destruction shocks which imply employed workers are laid-off into unemployment at rate  $\delta$ . There are also exogenous “godfather” shocks which occur at rate  $\lambda_q$ . Should a godfather shock occur, the worker quits exogenously to a randomly generated outside offer. Bagger and Lentz (2019) motivate the godfather shock process by assuming laid-off employees must be given notice. Thus  $\lambda_q/(\delta + \lambda_q)$  might be considered the fraction of laid-off employees given notice who obtain an outside job offer before the notice expires.<sup>3</sup> 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  $\sigma > 0$ .

Equilibrium requires each company  $j$ ’s wage policy  $w_{jt}(\cdot)$  maximizes expected discounted profit given the set of contracts posted by all other firms and the optimal quit strategies of workers. Different to Burdett and Coles (2003), the equilibrium set of optimal contracts depend on the joint distribution of employment across firms  $j$ , the tenures of employees within each firm  $j$  and their corresponding human capital which evolves endogenously over time. An added difficulty with all dynamic precommitment games is that when firms precommit to their optimal contracts at date zero, those choices depend on the initial aggregate state  $\chi_0$  which, in turn, generates complex and uninteresting non-stationary wage dynamics. For tractability, we will follow Woodford (2003) and define “timeless” equilibria in which each firm  $j \in [0, 1]$  precommits to an optimal 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, experience, and aggregate productivity). The timeless equilibrium essentially describes the stationary (ergodic) growth path of the economy.

Although the framework allows firms to offer general contracts of the form  $w_j(\tau, k, A)$ , the following establishes the existence of a particular class of equilibria: those where contracts  $w_j(\cdot) = Ak\tilde{\theta}_j(\tau)$  are fully optimal, where  $\tilde{\theta}_j(\tau)$  describes the wage rate paid by firm  $j$  to an employee with tenure  $\tau$ . That is not to say other contracting equilibria  $w_j(\cdot)$  do not exist. But this class of equilibria is particularly interesting for it yields a structural log-linear wage equation of the form:

$$\log w_{ijt} = \log k_i + \log \tilde{\theta}_j(0) + \rho x_{it} - \phi Z_{it} + \log \frac{\tilde{\theta}_j(\tau_{it})}{\tilde{\theta}_j(0)} + \log A_t.$$

Specifically there are worker  $i$  and firm  $j$  fixed effects (worker  $i$ ’s initial human capital, firm  $j$ ’s starting [log] wage rate  $\log \tilde{\theta}_j(0)$ ), 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. Furthermore, the equilibrium is consistent with the AKM definition of exogenous mobility: that worker  $i$ ’s tenure  $\tau_{it}$  and place of employment  $j = J(i, t)$  at date  $t$  are sufficient information to predict worker  $(i, t)$ ’s quit rate. The market equilibrium is thus consistent with the AKM approach. Because, it is also consistent with the Jacobson *et al.* (1993), Sianesi (2004) literature, we can then use the

3. Bagger and Lentz (2019) additionally allow that a worker might receive more than one random outside offer during the notice period. We abstract from this possibility.

structurally estimated model to decompose the cost of job loss into its constituent effects. As discussed in Section 6, we will find the distribution of firm starting wages  $\{\log \tilde{\theta}_j(0)\}_{j \in [0, 1]}$  and firm-specific tenure effects  $\{\log \frac{\tilde{\theta}_j(\tau_{ii})}{\tilde{\theta}_j(0)}\}_{j \in [0, 1]}$  determine the *temporary* wage losses employees face when laid-off into unemployment. The human capital dynamics instead determine the *permanent* losses.

**Definition of equilibrium:** Equilibrium is a set of (timeless) contracts  $\{\tilde{\theta}_j(\tau)\}_{\tau \in [0, \infty)}$  for each firm  $j \in [0, 1]$  such that:

- (i) contract  $w_j(\tau, k, A) = Ak\tilde{\theta}_j(\tau)$  maximizes expected discounted lifetime profit for each firm  $j \in [0, 1]$ , (*i.e.* no more general contract  $w_j(\tau, k, A)$  exists which increases firm profit), where
- (ii) all workers use optimal job search strategies given the market set of posted wage contracts  $w_j(\cdot)$  for all  $j \in [0, 1]$  and
- (iii) the joint distribution of employment, tenures, wages, and human capital is consistent with optimal job search, the set of contracts posted and the ergodic limit of the economy.

Burdett and Mortensen (1998) is an early example of this equilibrium concept. In that equilibrium each firm  $j$  posts a [timeless] fixed wage  $\bar{w}_j$ , workers use optimal job search strategies (given the set of posted wages) and the distribution of employment and wages is consistent with the ergodic limit of the economy (*i.e.* with steady state turnover). Here, instead we allow competition in general (timeless) contracts  $w_j(\cdot)$  and extend the definition of equilibrium to allow growth which, though exogenous here, might be endogenous in future applications.<sup>4</sup>

We identify such equilibria using the following approach. The next section considers optimal worker behaviour given all firms post contracts consistent with equilibrium; *i.e.* each firm  $j \in [0, 1]$  posts a contract of the form  $w_j = Ak\tilde{\theta}_j(\tau)$ . Given the resulting worker turnover, we then identify the equilibrium set of contracts  $\{\tilde{\theta}_j(\cdot)\}$  such that there is no deviating general contract  $w_j(\tau, k, A)$  which is profit increasing.

### 3. WORKER OPTIMALITY

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

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

$$rV(\tau, k, A|\theta) = \frac{\theta(\tau)^{1-\sigma} (Ak)^{1-\sigma}}{1-\sigma} + \frac{\partial V}{\partial \tau} + \rho k \frac{\partial V}{\partial k} + \gamma_{AA} \frac{\partial V}{\partial A}$$

4. Moscarini and Postel-Vinay (2013) more generally considers aggregate stochastic shocks and instead defines “stationary” equilibria with arbitrary initial conditions. An important difference, however, is their firms are restricted to suboptimal contracts (firms cannot offer wage tenure contracts). Given the restricted contract domain, a special property of the stationary equilibrium is the existence of a set of initial sidepayments between firms and employees (at date zero) which allows the market to jump straight to the timeless solution.

$$\begin{aligned}
& + \delta[V^U(k, A) - V(\tau, k, A|\theta)] \\
& + \lambda_1 E_j[\max\{V(\tau, k, A|\theta), V(0, k, A|\tilde{\theta}_j)\} - V(\tau, k, A|\theta)] \\
& + \lambda_q E_j[\max\{V^U(k, A), V(0, k, A|\tilde{\theta}_j)\} - V(\tau, k, A|\theta)].
\end{aligned}$$

In words, the flow value of being employed on contract  $\theta$  equals the flow utility 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  $\rho$ ), (iii) aggregate productivity increases (at rate  $\gamma_A$ ), (iv) a layoff shock occurs at rate  $\delta$ , (v) a randomly drawn outside offer  $\tilde{\theta}_j$  is received at rate  $\lambda_1$ , and (vi) an exogenous quit occurs at rate  $\lambda_q$  (where the worker quits into unemployment if the offer  $\tilde{\theta}_j$  has too low value).

Similar arguments imply the value of being unemployed satisfies

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

The restrictions to a CRRA utility function and definition of equilibrium imply the critical simplifying property: the value functions are separable in productivity  $Ak$  where

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

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

Let  $U_0 = U(0|\theta)$  denote the starting value of representative contract  $\theta(\cdot)$ . As search is random, let  $F(U_0)$  denote the fraction of offered contracts  $\{\tilde{\theta}_j(\cdot)\}$  whose starting value  $U(0|\tilde{\theta}_j) \leq U_0$ . Substituting out  $V(\cdot|\theta) = (Ak)^{1-\sigma} U(\tau|\theta)$  and  $V^U(\cdot) = (Ak)^{1-\sigma} U^U$  in the Bellman equations above yields the following expressions for  $U(\cdot|\theta)$  and  $U^U$ :

$$\begin{aligned}
& [r + \delta + \lambda_q - (\rho + \gamma_A)(1 - \sigma)]U - \frac{dU}{d\tau} \\
&= \frac{\theta(\tau)^{1-\sigma}}{1-\sigma} + \delta U^U + \lambda_1 \int_U^{\bar{U}} [1 - F(U_0)] dU_0 + \lambda_q \int_U^{\bar{U}} U_0 dF(U_0)
\end{aligned} \tag{2}$$

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

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:** *Equilibrium and optimal job search for any worker  $k$  implies:*

(a) *while unemployed, the worker accepts a contract offer  $\tilde{\theta}_j(\cdot)$  if and only if its starting value  $U_0 \geq U^U$ ;*

(b) while employed with contract value  $U(\tau|\theta)$ , the worker accepts an outside job offer  $\tilde{\theta}_j$  if and only if it offers greater contract value  $U(0|\tilde{\theta}_j) > U(\tau|\theta)$ . The worker quits 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  $[\underline{U}, \bar{U}]$  denote its support. As Claim 1 implies equilibrium turnover is independent of  $k$ , it implies for any given type  $k$ , that the distribution of contract values across workers of type  $k$  is also  $G(\cdot)$ ; i.e. the distribution of contract values across the entire population is independent of  $k$ .

#### 4. OPTIMAL CONTRACTS IN A TIMELESS EQUILIBRIUM

Consider any contract  $w_j(\tau, k, A)$  in a timeless equilibrium. Clearly with no loss of generality any such contract can be rewritten as  $w_j(\cdot) = Ak\theta_j(\tau, k, A)$ . 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 job 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  $w_j(\cdot)$  is equivalent to a wage rate paid  $\theta_j(\tau | k_0, A_0)$  which varies with tenure but firm  $j$  potentially discriminates across types  $(k_0, A_0)$  when hired.

Consider then any such contract  $\theta(\tau) = \theta_j(\tau | k_0, A_0)$ . 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  $G(\cdot)$  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  $\tau$ . Because  $F(\cdot)$  describes the distribution of starting contract values offered by all other firms, the probability this new hire remains employed by tenure  $\tau$  is

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

To determine the set of equilibrium optimal contracts, we first consider that contract which maximizes expected discounted profit conditional on hiring a new employee  $(k_0, A_0)$  with starting value  $U_0 \geq U^U$ ; i.e. we solve

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

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



[maximised] contract profit

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

Suppose now the firm is contacted by a potential employee  $(k_0, A_0)$  and the firm offers the above optimal contract  $\theta^*(\cdot)$  with starting value  $U_0 \geq U^U$ . Because  $\alpha + (1-\alpha)G(U_0)$  is the probability this contract offer is accepted, the firm's expected profit by offering  $U_0$  is then

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

The firm thus chooses  $U_0$  to maximize  $\Omega(U_0|A_0, k_0)$ . As the profit maximization problem is again simply multiplicative in  $(k_0, A_0)$ , we have established Claim 2.

**Claim 2:** *Equilibrium implies it is always optimal to offer contracts  $\theta^*(\tau|\cdot)$  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 equilibrium in which each firm  $j$  offers the same contract  $\theta_j(\cdot)$  to all potential hires  $(k_0, A_0)$ . We only consider such contracts from now on.

Theorem 1 now describes the optimal contract  $\theta^*(\tau|U_0)$  for any  $U_0 \geq U^U$ ; *i.e.* it solves the dynamic optimization problem:

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

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

**Theorem 1:** *For any  $U_0 \geq U^U$ , equilibrium implies an optimal contract  $\theta^*(\tau|U_0)$  and corresponding worker and firm [path] values  $U^*(\tau|U_0)$  and  $\Pi^*(\tau|U_0)$  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

$$\begin{aligned} & \frac{\theta^{1-\sigma}}{1-\sigma} + \theta^{-\sigma} [(1-\theta) - [r+\delta+\lambda_q - \rho - \gamma_A + \lambda_1[1-F(U)]] \Pi \\ & = [r+\delta+\lambda_q - (\rho+\gamma_A)(1-\sigma)] U - \delta U^U - \lambda_1 \int_U^{\bar{U}} [1-F(U')] dU' - \lambda_q \int_U^{\bar{U}} U_0 dF(U_0) \end{aligned} \quad (6)$$

(b)  $\Pi$  is given by

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

(c) and  $U$  evolves according to the differential equation

$$\frac{dU}{d\tau} = -\theta^{-\sigma} \frac{d\Pi}{d\tau} \quad (8)$$

with initial value  $U(0|\cdot) = U_0$ .



**Proof: See the Appendix.**

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

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

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

with  $U$  given by (8).

Equation (9) describes how the wage rate paid changes optimally with tenure, where the corresponding wage path is  $w(\tau) = A_0 k_0 e^{(\rho + \gamma_A)\tau} \theta(\tau)$ . Equation (9) and some algebra now establishes:

$$\left[ \frac{-u''}{u'^2} \right] \frac{dw}{d\tau} = \left[ A_0 k_0 e^{(\rho + \gamma_A)\tau} \right]^\sigma \lambda_1 F'(U) \Pi. \quad (11)$$

A quit at tenure  $\tau$  implies the firm loses continuation profit  $[A_0 k_0 e^{(\rho + \gamma_A)\tau}] \Pi(\tau)$ . When  $\Pi(\tau) > 0$ , (11) implies the wage paid increases within the job spell, where  $F'(U)$  is the measure of firms whose outside offer will marginally attract this worker. If  $F'(U) = 0$ , then marginally raising the wage paid at tenure  $\tau$  has no impact on the worker's quit rate and optimal consumption smoothing 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 with tenure. The scaling term  $[A_0 k_0 e^{(\rho + \gamma_A)\tau}]$  arises as the worker's value of employment at tenure  $\tau$  is  $V(\tau, \cdot) = [A_0 k_0 e^{(\rho + \gamma_A)\tau}]^{1-\sigma} U(\tau|\theta)$  while the firm's continuation profit is  $[A_0 k_0 e^{(\rho + \gamma_A)\tau}] \Pi(\tau)$ . As workers compare contracts by value  $U(\cdot)$ , 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  $(A_0, k_0)$ , while (6) describes the solution to the differential equation (9) for  $\theta(\cdot)$ .

Because  $w(\tau) = A_0 k_0 e^{(\rho + \gamma_A)\tau} \theta(\tau)$ , a constant wage paid (perfect consumption smoothing within the job spell) implies  $\theta(\tau)$  declines at rate  $\rho + \gamma_A$ . Thus although an optimal contract implies wages paid always increase within a job spell, it is not 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.

Consider first the optimal contract for the firm offering the least generous contract in the market, *i.e.* one which yields starting value  $U_0 = \underline{U}$  and suppose  $\underline{U} < U^\infty$ . As depicted in Figure 1, the wage rate paid  $\theta(\cdot)$  and contract value  $U(\cdot)$  both increase with tenure and  $U(\cdot)$  converges to  $U^\infty$  from below. Let  $\theta_1(\tau)$  denote the optimal least generous contract in the market, which we refer to as the lower baseline scale. Let  $U_1(\cdot)$  denote the corresponding path of contract values. Consider instead the optimal contract for firms which offer the most generous contract in the market,  $U_0 = \bar{U}$ , and suppose  $\bar{U} > U^\infty$ . Although the wage paid increases within the job spell,  $\theta(\cdot)$  decreases with tenure. Contract value thus falls with tenure and so  $U(\cdot)$  converges to  $U^\infty$  from above. Let  $\theta_2(\tau)$  denote the optimal most generous contract in the market, which we refer to as the upper baseline scale, and  $U_2(\cdot)$  the corresponding path of contract values.

Consider now the optimal contract  $\theta^*(\cdot|U_0)$  which yields starting contract value  $U_0 \in (\underline{U}, U^\infty)$ . As depicted in Figure 1b, 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 the critical simplification: the optimal contract  $\theta^*(\cdot|U_0)$  is simply the continuation contract starting at point  $t_0$  on the lower baseline scale; *i.e.*

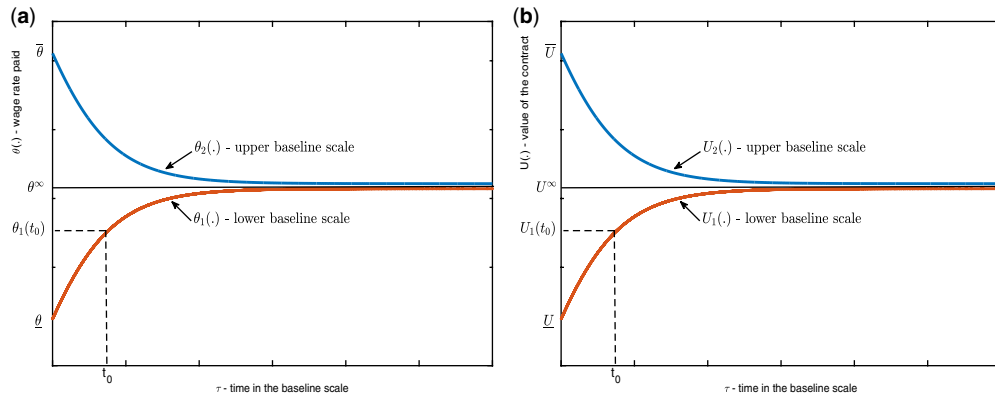


FIGURE 1

Possible sets of optimal contracts. (a)  $\theta(\cdot)$ —wage rate paid. (b)  $U(\cdot)$ —value of the contract.

$\theta^*(\tau|U_0) = \theta_1(t_0 + \tau)$ , where the wage rate paid at tenure  $\tau$  corresponds to point  $(t_0 + \tau)$  on the lower baseline scale. Let  $\Pi_1(t_0)$  denote the firm’s corresponding contract profit.

Suppose instead  $U_0 \in (U^\infty, \bar{U})$ . 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 yields contract profit  $\Pi_2(t_0)$ . It is this baseline property of the optimal contract structure which makes tractable the characterization of equilibrium.

4.1. *Characterization and existence of equilibrium*

Given a starting value  $U_0 \in [U, \bar{U}]$ , the previous section has shown that the optimal contract  $\theta^*(\cdot|U_0)$  corresponds to a baseline contract  $\theta_i(t_0 + \tau)$  with  $i = 1, 2$  and a starting point  $t_0 \geq 0$ , where  $U_i(t_0) = U_0$ . If accepted by worker  $(A_0, k_0)$ , this contract then generates profit  $A_0 k_0 \Pi_i(t_0)$ . All such contract offers then generate expected profit

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

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

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

$$\left[ \frac{\lambda_0 u + \lambda_q(1 - u) + \lambda_1(1 - u)G(U_i(t_0))}{\lambda_0 u + \lambda_q(1 - u) + \lambda_1(1 - u)} \right] \Pi_i(t_0) \leq \bar{\Omega} \text{ otherwise.}$$

To understand the approach below, note that standard recursive arguments (e.g. Spear and Srivastava 1987) suppose a contract “promises” continuation value  $U$  to an employee and then identifies  $\theta = \theta(U)$  as the optimal wage rate paid at that point. The approach here instead identifies the inverse function: let  $U = \hat{U}(\theta)$  describe the contract value enjoyed by a worker when the optimal contract pays  $\theta$ . The baseline property implies  $\hat{U}(\cdot)$  is given by  $\hat{U}(\theta) = U_i(t_0(\theta))$ , where  $t_0(\cdot)$  is the inverse function of  $\theta = \theta_i(t_0)$ , with  $i = 1, 2$ . Let  $\hat{\Pi}(\theta)$  describe the firm’s corresponding contract profit. Claim 3 reveals why this alternative approach is so useful.

**Claim 3:** Suppose  $dF(U_i(t_0)) > 0$ . Equilibrium implies

$$\widehat{\Pi}(\theta) = \frac{\sqrt{[r - \rho - \gamma_A]^2 \bar{\Omega}^2 + 4[\delta + \lambda_q] \bar{\Omega}(1 - \theta) - [r - \rho - \gamma_A] \bar{\Omega}}}{2[\delta + \lambda_q]} > 0, \tag{12}$$

where  $\theta = \theta_i(t_0)$ .

**Proof:** See the Appendix.

Claim 3 is a powerful result: it provides the closed form solution for  $\widehat{\Pi}(\theta)$ . (8) then implies  $\widehat{U}(\theta)$  is the solution to

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

and we are almost done: equilibrium simply reduces to identifying the boundary condition for (13). To do this, we transform the analysis from the time domain (how wage rates vary with tenure) to the domain of wage rates paid  $\theta \in [\underline{\theta}, \bar{\theta}]$ .

Let  $F_\theta(\theta)$  denote the distribution of starting wage rates paid by firms. Because (12) and (13) imply  $\widehat{U}(\cdot)$  is a strictly increasing function, the definition of  $F(\cdot)$  implies:

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

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

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

In the wage rates domain  $\theta \in [\underline{\theta}, \bar{\theta}]$ , the constant profit condition is now

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

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

The first step is to solve for equilibrium  $\bar{\Omega}$ . Let  $\bar{U} = \widehat{U}(\bar{\theta})$  denote the highest contract value offered by firms. A simple contradiction argument establishes  $G(\bar{\theta}) = 1$  and so substituting  $\theta = \bar{\theta}$  in (14) finds  $\widehat{\Pi}(\bar{\theta}) = \bar{\Omega}$ . Substituting  $\theta = \bar{\theta}$  and substituting out  $\widehat{\Pi}(\bar{\theta}) = \bar{\Omega}$  in (12) now determines

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

The next step, Claim 4, shows that the upper baseline scale, though consistent with optimality, does not survive equilibrium.

**Claim 4:** Equilibrium implies  $\bar{\theta} = \theta^\infty$ .

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

The final step is to tie down equilibrium  $\bar{\theta}$ . The argument used is the same as that identified in [Burdett and Coles \(2003\)](#). We first consider a candidate  $\bar{\theta}$  and Theorem 2 below constructs the corresponding equilibrium offer distribution  $F_{\theta}(\cdot|\bar{\theta})$  consistent with the equal profit condition. Claim 5 then establishes the [standard] boundary condition: equilibrium implies an unemployed worker is indifferent to accepting the lowest starting wage offer  $\underline{\theta}$ ; *i.e.*  $\widehat{U}(\underline{\theta}) = U^U$ . For  $\sigma > 1$ , Theorem 3 establishes the existence of equilibrium  $\bar{\theta}$  such that  $\widehat{U}(\bar{\theta}) = U^U$ . By finally showing that no deviating wage contract  $w = w(\tau, k, A)$  exists which can increase profit, it follows that Theorem 2 (with  $\bar{\theta}$  given by Theorem 3) describes equilibrium.

**Theorem 2:** Given  $\bar{\theta}$ , equilibrium implies  $\{\widehat{\Pi}(\cdot), \widehat{U}(\cdot), G_{\theta}(\cdot), F_{\theta}(\cdot)\}$  are given by:

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

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

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

$$1 - F_{\theta} = \theta^{\sigma} \int_{\theta}^{\bar{\theta}} \sigma \left[ \frac{1}{\theta'} \right]^{\sigma+1} \Psi(\theta') d\theta', \quad (19)$$

for all  $\theta \in [\underline{\theta}, \bar{\theta}]$ , where

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

**Proof:** See the Appendix.

The conditions of Theorem 2 depend on  $\bar{\theta}$  which is the last equilibrium variable to be determined. Claim 5 identifies the relevant boundary condition.

**Claim 5:** Equilibrium requires  $\underline{U} \equiv \widehat{U}(\underline{\theta})$  equals  $U^U$ , where  $\underline{\theta}$  is given by

$$1 - \underline{\theta} = \frac{(\delta + \lambda_q + \lambda_1)[\delta + \lambda_q + \lambda_1 + r - \rho - \gamma_A]}{(\delta + \lambda_q)[\delta + \lambda_q + r - \rho - \gamma_A]} [1 - \bar{\theta}]. \quad (20)$$

**Proof:** Standard contradiction arguments establish  $\underline{U}$  must equal  $U^U$  in a search equilibrium, and that  $G_{\theta}(\underline{\theta}) = 0$ . Putting  $\theta = \underline{\theta}$  in the constant profit condition (14), with  $\widehat{\Pi}(\underline{\theta})$  given by (16),  $\bar{\Omega}$  given by (15) and  $u = \delta/(\delta + \lambda_0)$  yields the result.

To describe the equilibrium fixed point problem, we use the following notation. First fix a candidate equilibrium value for  $\bar{\theta}$  in the range

$$\bar{\theta} \in \left( 1 - \frac{(\delta + \lambda_q)[\delta + \lambda_q + r - \rho - \gamma_A]}{(\delta + \lambda_q + \lambda_1)[\delta + \lambda_q + \lambda_1 + r - \rho - \gamma_A]}, 1 \right). \quad (21)$$

Inspection establishes that any such candidate value implies strictly positive profit ( $\bar{\Omega} > 0$ ) and  $\underline{\theta} > 0$  (strictly positive wage rates). Given this candidate choice of  $\bar{\theta}$ , let  $\widehat{F}_{\theta}(\cdot|\bar{\theta})$  denote the candidate

distribution function  $F_\theta$  implied by the conditions of Theorem 2. Given the implied distribution of contract offers, the Proof of Claim 6 now identifies the implied values of  $\underline{U}$  ( $\equiv \widehat{U}(\theta)$ ) and  $U^U$ , which we denote  $\widetilde{U}(\bar{\theta})$ ,  $\widetilde{U}^U(\bar{\theta})$ , respectively.

**Claim 6:** Given  $\bar{\theta}$  and the implied candidate distribution function  $\widetilde{F}_\theta$  then  $\widetilde{U}(\bar{\theta})$ ,  $\widetilde{U}^U(\bar{\theta})$  are given by

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

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

**Proof:** See the Appendix.

Identifying equilibrium reduces to establishing a  $\bar{\theta}$  exists where  $\widetilde{U}(\bar{\theta}) = \widetilde{U}^U(\bar{\theta})$ .

**Theorem 3:** For  $\sigma > 1$ , a  $\bar{\theta}$  satisfying (21) exists where  $\widehat{U}(\theta) = U^U$ .

**Proof:** See the Appendix.

If  $F_\theta$  is a positive increasing function (i.e. has the properties of a distribution function), then Theorems 2 and 3 fully characterize equilibria. By construction, all optimal contracts which offer starting wage rate  $\theta \in [\underline{\theta}, \bar{\theta}]$  yield the same expected profit  $\bar{\Omega} > 0$ . Consider then any deviating contract  $w_j(\cdot)$ . Because any such contract  $w_j(\cdot)$  is equivalent to a wage rate paid  $\theta_j(\tau|k_0, A_0)$  then, by construction, any such contract  $\theta_j(\cdot)$  which offers a starting value  $U_0 \in [\widehat{U}(\underline{\theta}), \widehat{U}(\bar{\theta})]$  cannot yield greater profit. Further as  $U^U = \widehat{U}(\underline{\theta})$ , any contract  $\theta_j$  which offers value  $U_0 < \widehat{U}(\underline{\theta})$  yields zero profit as all workers reject such an offer. Finally any contract  $\theta_j$  which offers  $U_0 > \widehat{U}(\bar{\theta})$  attracts no more workers than the optimal contract which offers  $U_0 = \widehat{U}(\bar{\theta})$  while the deviating contract earns strictly less profit per hire. Thus no deviating contract exists which yield greater profit and so Theorems 2 and 3 describe equilibrium.

Theorem 2 describes all equilibrium objects aside from the (lower) baseline scale  $\theta(\cdot)$ . Equations (9), (13), and (A.13) in the proof of Theorem 2 imply  $\theta(\cdot)$  is identified by the initial value problem:

$$\dot{\theta} = \frac{\lambda_1 \theta^{-\sigma} [\Psi - (1 - F_\theta)] \widehat{\Pi}}{d\widehat{U}/d\theta} - (\rho + \gamma_A)\theta$$

with  $\theta(0) = \underline{\theta}$ .

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

## 5. QUANTITATIVE ANALYSIS

We estimate our model using indirect inference (see [Gourieroux \*et al.\* 1993](#); [Bagger \*et al.\* 2014](#); [Bagger and Lentz 2019](#), for recent related work and Appendix B for a full description of our approach). An important feature of the data is that workers in different education groups exhibit very different turnover patterns. Hence, the equilibrium wage/tenure/experience contracts offered by firms are likely to have different properties by education group. We therefore suppose these types participate in separate submarkets and estimate the model separately for each type. We show the distinction plays an important role when trying to identify the costs of job loss.

5.1. *Data*

Our main source of information is the Sample of Integrated Labour Market Biographies (SIAB), an administrative matched employer–employee dataset developed by the German government for tax purposes. It is a 2% random sample drawn from the Integrated Employment Biographies (IEB), which comprises the universe of individuals who are either (i) in jobs that are subject to social security, (ii) in marginal employment, (iii) in benefit receipt according to the German Social Code, (iv) officially registered as a job-seeker at the German Federal Employment Agency, or (v) participating (plan to participate) in active labour market policies programmes. Individuals are followed as from 1975 or since the worker entered the labour market, whichever is the later. The data provide daily information on employment status and information on the gross daily wage/benefit, education, gender, occupation, age of the individuals, and geographical location of the place of work.<sup>5</sup> Importantly, the data also provide the unique establishment identifier employing these individuals. This establishment identifier allows us to match the worker information to that of his/her establishment. Using this information, we are able to reconstruct individuals' labour market histories as well as identify mass layoffs, which are needed for our estimates of the cost of job loss (see Section 6).

The establishment information is obtained from the Establishment History Panel (BHP), which collects information from all German establishments with at least one employee contributing to social security since 1975. The BHP provides annual information about the number of employees working in the establishment, their 3-digit industry classification, the median gross daily wage of full-time employees and the location of establishment. For convenience, we will use the terms establishment and firm interchangeably.

Using these data we restrict attention to all West German, male workers with a contributing job who entered the panel between the ages of 18 and 35. This implies that we exclude those workers who are reported as trainees, marginal part-time workers, employees in partial retirement, interns and student trainees, or in other employment status.<sup>6</sup> We divide the workers in our sample into three educational groups. (i) Low education level: workers who have up to a high school degree but no vocational training. (ii) Medium education level: workers who have up to a high school degree and hold a vocational qualification. (iii) Higher education level: workers with a university degree, either from a university of applied science (Fachhochschule), technical college (technische Hochschule) or a university. Table 1 presents the size of the data along several dimensions for these three categories.

5. The gross daily wage in the SIAB is constructed by dividing total gross earnings by the number of days employed in that job. If a worker did not leave the employer during a given year, the average gross daily wage in a job is computed annually. If the worker changed employers during the year, the gross average daily wage is computed for each employer using the time spend with the employer during that year.

6. Note that we do not consider civil servants or the self-employed as they are not covered in our data. We also exclude those individuals in the armed and police forces as well as members of parliament.

TABLE 1  
*Summary statistics by educational groups, SIAB*

	Low education	Medium education	Higher education
Number of observations	8,126,696	46,833,857	7,151,321
Number of workers	134,508	283,413	61,752
Number of establishments	176,436	488,600	69,571
Number of employment spells	339,106	920,739	108,442
Number of job spells	400,282	1,263,658	158,085
Number of unemployment spells	238,937	643,980	48,675

These data allow us to estimate workers' average unemployment, job, and employment durations, their average wage–experience profiles and measures of wage dispersion.<sup>7</sup> The SIAB, however, is not suitable to estimate firm-specific effects as many establishments have only one worker and these effects might not be properly identify. Instead, we obtain estimates of the firm-specific wage rate and its correlation with firm-specific returns to tenure from Carrillo-Tudela *et al.* (2019), who use the full IEB data containing information on the universe of German workers and their establishments.<sup>8</sup> For these exercises, we deflate the wage information using the CPI.

An important issue is that the wage data are top-coded, meaning that the data do not report the wage paid should it exceed a certain level. Although only 1.3% of the low educated and 6.0% of the medium educated group are top-coded, unfortunately 40% of the higher educated group are top-coded. We impute the missing wages using the methodology of Buetnerr and Raessler (2008) but note that 40% of worker wages in the latter group are subject to imputation error.<sup>9</sup>

We use the German Socio-Economic Panel (GSOEP) to complement the information derived from the SIAB and IEB data. The main advantage of the GSOEP for our study is that it provides information about the nature of an employer-to-employer transition. We label a “voluntary” transition as one in which the worker reported “own resignation,” “mutual agreement separation,” or “leave of absence” when leaving his job to take another job with a different employer. An “involuntary” transition is one in which the worker reported he changed employers due to “company shut down,” “dismissal,” or “temporary contract expired.” In contrast to the SIAB or IEB, the GSOEP data are household panel survey and hence is much smaller. The GSOEP started in 1984 and is updated on an annual basis.<sup>10</sup> Appendix B.4 provides detailed information on the data construction.

7. We consider a job spell as the time spent with a given employer and an employment spell as the time spent between two consecutive unemployment spells, where an unemployment spell takes into account both registered unemployment and non-participation periods. We follow this strategy as a large proportion of male workers who lost their jobs did not register as unemployed or if they did register they stopped registering soon afterwards before re-entering employment. Carrillo-Tudela *et al.* (2018) provide a statistical analysis of this feature during the 1999–2014 period. This implies that to capture the consequences of job loss we need to consider both registered and non-registered unemployment spells. See also Schmieder *et al.* (2016) for a similar practice. Throughout the analysis, we also distinguish between potential and actual labour market experience. Potential experience is defined as the sum of the overall time spent in employment and unemployment; while actual experience is the sum of the overall time spent in employment.

8. Carrillo-Tudela *et al.* (2019) were able to access the full IEB data for the 1993–2017 period and estimate wage regressions following Abowd *et al.* (1999) and Card *et al.* (2013) for different sub-periods. In Appendix B.2, we provide an account of their estimation procedure and how we use their estimates in our analysis.

9. See also Card *et al.* (2013) for alternative imputation procedures. We find that when regressing the Mincer wage equations, described below, wage imputations do not seem to make much of a difference on the average returns to experience and tenure relative to using top-coded wages.

10. Further information about the SIAB and GSOEP data can be found in Antoni *et al.* (2016) and in <http://www.diw.de/en/diw02.c.222857.en/documents.htm>, respectively.



TABLE 2  
*Targeted moments*

Moments	Low education		Medium education		Higher education	
	Data	Model	Data	Model	Data	Model
Average transitions/duration						
Unemployment spell (months)	23.89 (0.107)	22.48	17.89 (0.061)	16.84	18.89 (0.190)	17.45
Employment spell (months)	31.90 (0.139)	32.38	73.31 (0.168)	72.86	90.62 (0.372)	91.49
Job spell (months)	19.75 (0.069)	22.71	36.17 (0.050)	39.34	48.04 (0.148)	47.39
Invol./vol. EE transitions	0.630 (0.065)	0.653	0.560 (0.025)	0.540	0.400 (0.032)	0.403
AKM regression						
Corr [ $\gamma_j^0 \gamma_j^1$ ]	-0.1833 (0.022)	-0.1861	-0.1881 (0.005)	-0.1868	-0.1949 (0.019)	-0.1806
Var [ $\gamma_j^0$ ]	0.0541 (2.7e-03)	0.0629	0.0492 (6.9e-04)	0.0455	0.0423 (1.9e-03)	0.0387
Wage regressions						
Actual experience—linear term	0.00312 (3.5e-05)	0.00313	0.00304 (1.4e-05)	0.00308	0.00375 (3.0e-05)	0.00379
Actual experience—quadratic term	-5.95e-06 (9.3e-08)	-1.20e-07	-5.25e-06 (3.6e-08)	-2.14e-07	-7.37e-06 (8.3e-08)	-2.57e-07
Tenure—linear term	0.00266 (4.0e-05)	0.00200	0.00140 (3.1e-05)	0.00126	0.00075 (3.0e-05)	0.00103
Tenure—quadratic term	-5.60e-06 (1.3e-07)	-3.74e-06	-2.56e-06 (4.9e-08)	-1.93e-06	-1.86e-06 (1.1e-07)	-1.03e-06
Last unemployment dur.	-0.00106 (5.7e-05)	-0.00085	-0.00145 (2.9e-05)	-0.00123	-0.00147 (8.10e-05)	-0.00140
Wage dispersion						
Coefficient of variation	0.2320 (0.0427)	0.2301	0.1988 (0.0307)	0.1910	0.1630 (0.0085)	0.1682

Notes: Bootstrapped standard errors in parenthesis.

## 5.2. Estimation procedure

We adopt a month as the reference unit of time and set  $r=0.005$  (an annual discount rate of 6%). We set  $\gamma_A=0.0022$  to match the estimated slope of a linear trend on output per hour in Germany over the relevant time period.<sup>11</sup> This leaves a vector  $\Lambda = \{\delta, \lambda_0, \lambda_1, \lambda_q, \rho, \phi, \sigma, b\}$  of eight parameters that we recover by minimizing the sum of squared differences between a set of simulated moments from the model and their counterparts in the data, using the variance-covariance matrix of the empirical moments as a weighting matrix (see Appendix B for full details).

We target 12 statistics based on the main characteristics of the labour market to which the model is directly related. Table 2 describes those statistics for each education group. The average duration of spells and the ratio of involuntary to voluntary employer-to-employer transitions provide direct information for  $\{\lambda_0, \lambda_1, \delta, \lambda_q\}$ . The parameters  $\{\rho, \phi, \sigma, b\}$  are identified using wage information. All parameters, of course, are jointly estimated.

11. We estimate the slope of the linear trend through OLS, by regressing the log of yearly output per hour on a linear trend for the period 1991–2014. We start in 1991 to avoid the discrete change in the series introduced by the German re-unification. Similar results are obtained when using output per head of household as an alternative measure of labour productivity. Output per hour and output per head of household are directly obtained from the OECD website.

Equation (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 job spell for the worker is becoming more productive through learning-by-doing ( $k = k_0 e^{\rho\tau}$ ), outside wage 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 contract effects depends on the degree of risk aversion  $\sigma$ . For example,  $\sigma \rightarrow \infty$  implies each firm  $j$  optimally commits to a constant wage rate  $\bar{\theta}_j$ , analogous to the case considered in [Burdett and Mortensen \(1998\)](#) and there are no tenure effects. At the other end of the spectrum, risk neutral workers with  $\sigma = 0$  instead imply tenure effects can be infinitely large; e.g. step wage contracts as in [Stevens \(2004\)](#) and [Carrillo-Tudela \(2009\)](#). Given the optimal contract trades-off consumption smoothing against improved quit incentives, greater risk aversion implies flatter wage profiles, and smaller wage tenure effects. The inference process below identifies the parameterization which best explains the observed experience and tenure effects in the data, noting that firm wages are disperse, that tenure effects are firm specific, quit turnover is endogenous, and experience effects arise from two sources: (i) general human capital accumulation through learning-by-doing and (ii) workers' job-to-job transitions.<sup>12</sup>

Because the framework is consistent with the AKM approach, [Abowd et al. \(1999\)](#), we use the AKM regression equation as an auxiliary equation to measure the degree of wage dispersion across firms. Specifically the estimated AKM equation is

$$\log w_{it} = \alpha_i + \gamma_{J(i,t)}^0 + \gamma_{J(i,t)}^1 \tau_{it} + \beta_1 X_{it} + u_{it}, \quad (24)$$

where  $\alpha_i$  is the worker  $i$  fixed effect and there are two firm fixed effects—firm  $j$ 's starting wage rate  $\gamma_j^0$  and its firm specific return to tenure  $\gamma_j^1$ —where  $j = J(i, t)$  describes worker  $i$ 's place of employment at date  $t$ .  $X_{it}$  denotes a vector of covariates composed of a polynomial in potential experience and a time trend and  $u_{it}$  is the wage residual which is assumed white noise. From this regression, we take two target moments: (i) the estimated variance of the firm-fixed effects  $\gamma_j^0$  and (ii) the correlation of the estimated  $\gamma_j^0$  with the firm-specific tenure effects  $\gamma_j^1$ . These targets ensure the estimated model not only generates wage dispersion across firms consistent with the measured AKM variance in firm starting wages but also with firm-specific tenure effects which vary systematically across firms. The estimated worker-fixed effects, however, do not provide useful targeting information. The model implies the distribution of worker productivities is (asymptotically) log normal and uncorrelated with the firm-fixed effects. Although, we might fit an underlying log normal human capital distribution to the mean and variance of the AKM estimated worker-fixed effects, this yields no further useful information.

For our second auxiliary equation, we follow [Bagger et al. \(2014\)](#) by using a Mincer wage regression to describe the average market returns to experience and tenure (using actual experience, rather than potential experience, as the conditioning variable). This regression provides direct information on worker's average tenure effects and learning-by-doing,  $\rho$ , and

12. [Burdett et al. \(2016\)](#) demonstrate that workers' job shopping behaviour on its own is able to generate a positive and concave wage–experience profile as workers move to better paying jobs over time (see also [Burdett 1978](#)). Further, the same arguments that motivate the literature that tries to estimate unbiased returns to tenure ([Altonji and Shakotko 1987](#); [Topel 1991](#), among others) also imply that the returns to labour market experience could be biased if workers' experience in the labour market is correlated with an unobserved match-specific component, which is typically the case in standard job ladder models.

further informs on the job offer arrival rate of employed workers  $\lambda_1$ .<sup>13</sup> From the Mincer wage regression, we take as target moments the estimated coefficients for the returns to actual experience and tenure (linear and quadratic terms).

To infer the rate of skill loss while unemployed,  $\phi$ , we follow [Ortego-Marti \(2016\)](#) and [Jarosh \(2015\)](#) who use the auxiliary equation

$$\log w_{it}^r = \alpha_i + \beta_0 \text{Udur}_{it}^{\text{last}} + \beta_1 d_t + u_{it}, \quad (25)$$

which relates a worker's re-employment (log) wage,  $w_{it}^r$ , on last unemployment duration  $\text{Udur}_{it}^{\text{last}}$ , a worker-fixed effect,  $\alpha_i$ , and year dummies,  $d_t$ . The estimated coefficient  $\hat{\beta}_0$  thus provides a measure of skill loss  $\phi$  while unemployed and so is used as a target moment.

Finally, the parameter  $b$  is determined by identifying the extent of frictional wage dispersion. In a previous version of this article,  $b \in [0, 1]$  was identified using the *Mm* ratio [mean to minimum], defined in [Hornstein et al. \(2011\)](#), as a target moment. This was useful targeting information because [Burdett and Coles \(2003\)](#) with no learning-by-doing implies the *Mm* ratio equals one when  $b=1$  (all earn  $\theta=1$ ) and becomes arbitrarily large as  $b \rightarrow 0$  (because  $\underline{\theta} \rightarrow 0$ ). In this version, however, we instead target the coefficient of variation (on residual wages) which is a more robust statistic. The model is over-identified with 12 targets and 8 parameter values.

### 5.3. Model fit

Table 2 shows the fit of the model is very good for each education group. Because the fit is qualitatively identical across these groups, we first focus the discussion on the low educated. The estimated turnover parameters  $\{\delta, \lambda_0, \lambda_1, \lambda_q\}$ , described in Table 3 below, imply the model fits very well the average transitions/duration statistics. For these workers the data find (i) long average spells of unemployment (24 months), (ii) comparatively short employment spells (32 months), (iii) job spells average around 20 months (which implies job tenures are typically short).

A major success of this exercise is that the model captures the AKM correlation between the firm-fixed effect and the firm-specific tenure effect, which not only establishes that high wage firms do indeed raise wages more slowly with tenure, but also captures the magnitude of this effect. The model also reproduces the variance of the firm fixed effect very well. Note that in the model firm heterogeneity arises from differences in the starting values of their offered contracts and not from differences in ex ante firm-specific productivity. Nevertheless, the estimation shows that this source of heterogeneity is sufficient to capture the variation in firm-fixed effects estimated through the AKM approach.

The model generates appropriate linear experience effects, but it does not generate sufficient curvature as measured by the quadratic experience term in the Mincer wage regression. Given the rate of learning-by-doing is assumed constant, the curvature generated by the model reflects the original [Burdett \(1978\)](#) job-ladder insight: that as employed workers accumulate experience they also climb to higher wage points on the job ladder which then causes a positive and decreasing correlation between wages earned and experience (see footnote 12 for further discussion). To mitigate for possibly declining rates of learning-by-doing, we have restricted the data sample to the relatively young (entrants aged between 18 and 35) though also see Section 6.3 for further discussion.

13. Formally, [Bagger et al. \(2014\)](#) use an AKM wage equation similar to (24) but with only one firm-fixed effect. They estimate the returns to experience and tenure before estimating the firm- and worker-fixed effects on the residual wages. Specifying this regression on potential experience rather than actual experience gives very similar model parameter estimates.

TABLE 3  
*Estimated parameters*

		Low education	Medium education	Higher education
Parameters				
Job destruction	$\delta$	0.0301 (0.0003)	0.0137 (0.0001)	0.0109 (0.0001)
Unemployment offer arrival rate	$\lambda_0$	0.0412 (0.0005)	0.0553 (0.0005)	0.0538 (0.0015)
Emp. offer arrival rate—voluntary	$\lambda_1$	0.0222 (0.0009)	0.0223 (0.0004)	0.0201 (0.0007)
Emp. offer arrival rate—involuntary	$\lambda_q$	0.0060 (0.0010)	0.0040 (0.0001)	0.0029 (0.0003)
Human capital accumulation	$\rho$	0.0040 (0.0003)	0.0033 (0.0005)	0.0041 (0.0001)
Human capital depreciation	$\phi$	0.0010 (0.0002)	0.0012 (0.0002)	0.0015 (0.0001)
Risk aversion	$\sigma$	8.8279 (0.3242)	5.8672 (0.0426)	4.5384 (0.2049)
Unemployment income	$b$	0.3246 (0.0204)	0.4444 (0.0105)	0.5218 (0.0266)
Endog. variables				
Minimum wage rate	$\theta$	0.2949	0.4219	0.4694
Mass of firms at $\theta$	$F(\theta)$	0.0068	0.0060	0.0049
Average human capital growth	$\gamma_k$	0.00187	0.00241	0.00315

Notes: Standard errors in parenthesis.

The model reproduces very well the estimated linear and quadratic tenure effects in the Mincer wage regression. It is interesting that the low-educated group have high (estimated) marginal tenure effects which are only slightly smaller than the marginal return to experience. But because the average job spell for these workers is just 20 months, such large marginal tenure effects do not yield large overall tenure effects and so, not surprisingly, we will find that the estimated job ladder effects are small for this group. Nevertheless small overall tenure effects on wages is not evidence that marginal tenure effects are unimportant. Indeed, the key feature of the frictional labour market is the existence of a job ladder, which workers climb either through internal promotion (tenure effects) or through on-the-job search.

The model also captures very well the negative relation between unemployment duration and re-employment wages and the coefficient of variation of frictional wage dispersion. Though not fitted to the  $Mm$  ratio, we find that for low educated workers the model and the data generate  $Mm$  ratios equal to 1.62 and 1.78, respectively. For medium-educated workers the  $Mm$  ratios are instead 1.56 and 1.77, while for high educated workers they are 1.49 and 1.49.

We now broaden the discussion to compare economic outcomes across the education groups. Table 3 describes the estimated parameters for each of these groups. Reflecting that those with low education have much lower average employment spells, it is important for what follows to note the inferred layoff rate  $\delta$  in the less educated sector is around three times higher than those in the more educated sectors. The inferred job offer arrival rate while unemployed is also low for the less educated, though Table 3 suggests job offer arrival rates while employed are broadly the same.

The structural estimates identify high learning-by-doing rates:  $\rho$  is equal to 4.8% per annum for those with low education, 4.0% for those with medium education, and 4.9% for those with higher education. These estimated returns are appreciably higher than those suggested by the Mincer wage regressions. This reveals an important source of bias in those Mincer wage regressions: they omit skill loss while unemployed, where estimated human capital loss rates  $\phi$  are 1.2% per annum

for the less educated, 1.4% for those with medium education, and 1.8% for those with higher education. Because actual experience is correlated with age, and so positively correlated with time unemployed, omitting skill loss while unemployed biases downwards the Mincer estimates on the return to experience. This bias is clearly highest for the low educated who are particularly liable to long spells of unemployment, and smallest for university graduates.

The estimated relative risk aversion parameters are plausible, though their values are higher than the standard ones used in the macro literature.<sup>14</sup> An important role of  $\sigma$  is to make within-firm wage variation consistent with the data. For example with risk neutral workers, the sequential auctions framework (e.g. [Postel-Vinay and Robin 2002](#)) would imply unemployed workers in the low education group accept negative starting wage rates  $\underline{\theta} = -1.03$ . This low starting wage reflects that at rate  $\lambda_1$  the worker receives an outside offer and subsequently earns  $\theta = 1$  (until layoff), and that unemployed workers are willing to “buy” valuable experience.<sup>15</sup> The sequential auction literature reduces such large within-firm wage variation by assuming workers have bargaining power; e.g. [Dey and Flinn \(2005\)](#), [Cahuc et al. \(2006\)](#), and [Bagger et al. \(2014\)](#). Here, instead firms post optimal contracts and risk averse workers prefer less wage variation over time. Given the literature typically argues that tenure effects are small, it is interesting that the degree of risk aversion required to keep tenure effects consistent with the data is not so high.<sup>16</sup>

Taking into account the average wage rate paid in the market, the values of  $b$  imply average replacement ratios of 0.67, 0.69, and 0.74 across the low, medium, and high education categories. These replacement ratios are consistent with [Krebs and Scheffel \(2013\)](#) and [Launov and Wealde \(2013\)](#) who identify replacement ratios of around 0.64.

The ratio  $\lambda_1/\delta$  measures the rate at which workers receive outside job offers relative to the rate at which they become unemployed. This is estimated at 1.63 and 1.89 for those workers in the medium and high education groups, but it is just 0.74 for those workers in the low education

14. [Gandelman and Hernandez-Murillo \(2015\)](#) survey some of the literature and find that estimates of relative risk aversion vary widely, going from 0.2 to 10 and above.

15. Specifically

$$V^U(k_0, A_0) = \frac{bA_0k_0}{r + \phi - \gamma_A},$$

because unemployed workers obtain no surplus from a job offer and skills decline at rate  $\phi$  while unemployed. Conversely, the matching offers game with equally productive firms implies an employed worker with an outside offer thereafter enjoys  $\theta = 1$  and hence value

$$\begin{aligned} V^1(k_0, A_0) &= \int_0^\infty e^{-(r+\delta+\lambda_q)t} \left[ A_0k_0 e^{(\rho+\gamma_A)t} + (\delta+\lambda_q) \frac{bA_0k_0 e^{(\rho+\gamma_A)t}}{r + \phi - \gamma_A} \right] dt \\ &= A_0k_0 \frac{1 + \frac{(\delta+\lambda_q)b}{r+\phi-\gamma_A}}{r + \delta + \lambda_q - \rho - \gamma_A}. \end{aligned}$$

An unemployed worker is hired on initial rate  $\theta_0$  until an outside offer or a job destruction shock occurs, and so:

$$V^0(k_0, A_0) = \frac{A_0k_0}{r + \delta + \lambda_q + \lambda_1 - \rho - \gamma_A} \left[ \theta_0 + \frac{(\delta + \lambda_q)b}{r + \phi - \gamma_A} + \frac{\lambda_1 \left[ 1 + \frac{(\delta + \lambda_q)b}{r + \phi - \gamma_A} \right]}{r + \delta + \lambda_q - \rho - \gamma_A} \right].$$

As no worker surplus implies  $V^0(k_0, A_0) = V^U(k_0, A_0)$ , some algebra establishes:

$$\theta_0 = \frac{r - \rho - \gamma_A}{r + \phi - \gamma_A} b - \lambda_1 \frac{[r + \phi - \gamma_A] - b[r - \rho - \gamma_A]}{(r + \delta + \lambda_q - \rho - k_A)(r + \phi - \gamma_A)}.$$

16. Although adding savings to the model is beyond the scope of this article, it is interesting to note that less educated workers, who are more likely to have low savings and so be liquidity constrained while unemployed, exhibit higher degrees of risk aversion. They also have lower values of  $b$ , perhaps reflecting that being unemployed and liquidity constrained with small mouths to feed may not be a very leisurely environment.

group. Because the latter are more likely to be laid-off into unemployment than receive an outside offer, the efficiency gain to backloading wages is much reduced. At first sight it is then surprising that the Mincer wage regression, both on the actual and simulated data, suggest that low educated workers enjoy the highest returns to tenure (see Table 2). But there is a second effect: high layoff rates in a stationary equilibrium imply a larger fraction of workers are employed at lower points on the baseline scale. In other words, workers in the low-educated group are distributed more on the steeper part of their baseline scale, while workers in the high educated group are distributed more on the flatter part. Or simply put: Mincer estimates of the average return to tenure across employed workers does not reflect the underlying (non-linear) job ladder and the (endogenous) distribution of employment across that ladder.

## 6. ESTIMATING THE COST OF JOB LOSS

Following Jacobson *et al.* (1993), a large statistical literature measures the cost of being laid-off, both in terms of lower future wages and lost earnings. This allows a direct validation test of the model: using the same reduced form techniques, are estimates of the cost of job loss using model-generated data consistent with those obtained on the actual data? We not only find the results are remarkably well-aligned, we use the model to decompose the cost of job loss into its core components:

- (i) job ladder losses: the laid-off must seek re-employment at a new firm;
- (ii) skill losses: there is foregone human capital accumulation as well as skill loss while unemployed;
- (iii) the employment gap effect: it takes time for the laid-off worker to find suitable employment.

The Jacobson *et al.* (1993) approach selects a given year  $y$  and considers those who were displaced into unemployment in that year. Let  $\Psi_{it'}$  denote the variable of interest: either measured earnings or measured log wages of agent  $i$  in calendar year  $t'$ . Losses due to job separation in year  $y$  are estimated using the diff-in-diff specification:

$$\Psi_{it'} = \alpha_i + d_{t'} + \beta X_{it'} + \sum_{t=-T_s}^{T_e} \varepsilon_t D_{it} + u_{it'}, \quad (26)$$

where  $\alpha_i$  is the worker-fixed effect,  $d_{t'}$  are year dummies,  $X_{it'}$  is a cubic on worker  $i$ 's potential experience and the  $D_{it}$  are a set of dummy variables which take value 1 if worker  $i$  was displaced in year  $y$  and  $t' - y = t$ , and is zero otherwise. The estimated parameters  $\varepsilon_t$  thus describe the displaced worker's average loss of earnings (or log wages depending on case)  $t$  years following displacement relative to a control group of those who were not displaced in year  $y$ . The error term  $u_{it'}$  is assumed white noise.

To minimize selection effects—that employers might choose which workers to layoff—the standard approach is to focus on mass layoff events. For consistency, we also adopt this approach though robustness checks find that instead considering all separations does not much affect the results. Estimating the  $\varepsilon_t$  for  $t < 0$  provides a simple check for selection effects and trending differences in the  $\Psi_{it'}$  between those who are laid-off (at future date  $y > t'$ ) and those not laid-off.

Jacobson *et al.* (1993) chooses those who are never laid-off as the control group. This is not appropriate in our framework for it conditions on workers who are ex post lucky. Instead, we adopt the Sianesi (2004) approach: for our control group we use all workers who were employed

but not laid-off in year  $y$  and so remain at risk of future layoff (see also [Davis and Von Wachter 2011](#)). We will show this choice of control group is important for identification purposes.

A crucial property of the theoretical model is that the change in worker (log) wages is independent of the worker-fixed effect. This is precisely that required for identification in the reduced form approach: given two identical workers where one is initially employed the other unemployed, the (percentage) difference in expected future earnings (or wages) is independent of the (unobserved) fixed effect. Our approach, however, also identifies an important caveat. Although it is possible to difference out the worker (productivity) fixed effect, the estimated cost of job loss still depends on underlying turnover parameters. Because [Table 3](#) shows workers in different education groups face very different job turnover parameters, it is important to disaggregate the data by education choice.

The following first estimates the cost of job loss in terms of lost earnings, and then in reduced log wages. There is, however, one more caveat. For this dataset, earnings information on the highly educated group is missing for 40% of the sample due to top-coding. Because the imputation method generates additional measurement error, the estimates of (26) have large standard errors for the high education group.<sup>17</sup> The following focusses on the results for the low and medium education groups, and we refer the reader to the Appendix for the full set of results.

### 6.1. *Reduced form cost of job loss I: earnings losses*

Because earnings in some periods are zero, the statistical literature considers (26) using gross yearly earnings as the dependent variable. We estimate (26) on the SIAB/BHP sample described in [Section 5.1](#) for the period 1981–2005 with  $T_s = 3$  and  $T_e = 15$ . As mentioned earlier, following [Jacobson et al. \(1993\)](#), [Couch and Placzek \(2010\)](#) and many others, we only consider displaced workers who are part of a mass layoff event.<sup>18</sup> We follow [Davis and Von Wachter \(2011\)](#) by defining a mass layoff event as one in which the number of workers in an establishment significantly falls during a 2-year span.<sup>19</sup> To further deal with worker selection issues, we also follow the statistical literature by only considering workers with at least 3 years of tenure in the establishment prior to the mass layoff event.<sup>20</sup> In [Appendix B.3](#), we provide further details of the identification of mass layoffs, estimation results, and robustness exercises. The results based on model-simulated data use the same estimation procedure but instead with simulated data generated by the parameters values described in [Table 3](#).

17. The sample is also very small: there are only 138 instances in which higher education workers were actually laid-off as part of a mass layoff event.

18. Note that the estimation of  $\delta$  in [Section 5.2](#) relied on all separations into unemployment, while the reduced form approach relies on mass layoffs which occur at a lower frequency. This generates a potential tension between the way the model is estimated ([Section 5.2](#)) and the analysis in this section. In [Appendix B.3](#), however, we show that this tension is small as estimating equation (26) on all separations gives similar results as the ones obtained when using mass layoffs. See [Jarosh \(2015\)](#) for a similar finding. See also [Flaen et al. \(2019\)](#) for a discussion about the validity of using mass layoff as a way to identify a random displacement event due to firm financial distress to minimize worker selection effects when estimating equation (26).

19. More precisely to be considered a mass layoff event in year  $y$ , the employer must meet the following criteria: (i) 50 or more employees in  $y-2$ ; (ii) employment reduces by 30–99% from  $y-2$  to  $y$ ; (iii) employment in  $y-2$  is no more than 130% of employment in  $y-3$ ; (iv) employment in  $y+1$  is less than 90% of employment in  $y-2$ .

20. These tenure restrictions do not play an important role in the results presented below. In particular, restricting to at least 12, 24, or 36 months of tenure in the establishment prior to a mass layoff event leads to similar post displacement patterns for each of the education groups. Further, considering only workers with full-time employment spells or pooling full-time and part-time employment spells also has a small effect in our results. The latter probably occurs because part-time spells represent a very small proportion of all spells for male workers in our sample. Following the literature, here we present the results based on workers with full-time employment spells.



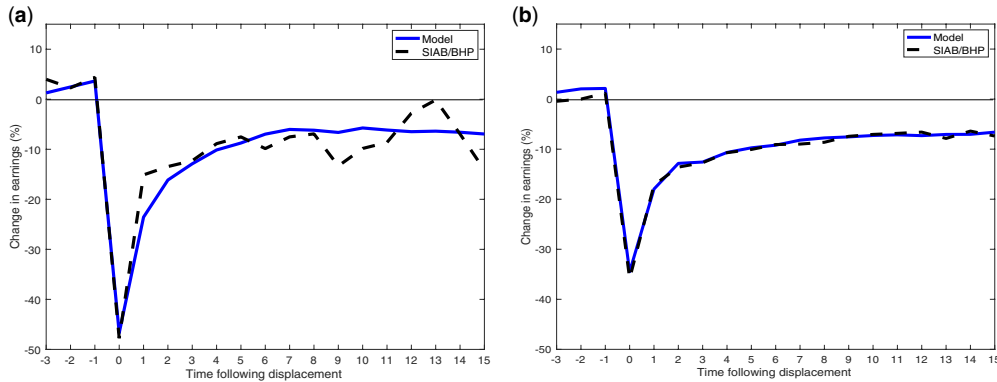


FIGURE 2

Post displacement earnings losses. (a) Low educated: data and model. (b) Medium educated: data and model.

Figure 2 describes the estimated (proportional) earnings losses  $\widehat{\varepsilon}_t^E$  following displacement, denoted  $\widehat{\varepsilon}_t^E$ , both for the SIAB/BHP data and for the simulated data.

Figure 2 shows, for both the model and the data, there are large drops in earnings immediately following layoff: initial expected losses are of the order of 40% for both education groups. This estimate would seem a little on the low side, however, for a newly laid-off worker necessarily suffers a 100% earnings reduction. This statistic reflects the presence of a temporal aggregation bias—that earnings are measured as total earnings over the accounting year. Notice then that the worker laid-off on 1 January loses a maximum of 100% earnings over the year, while the worker laid-off on 31 December loses none. Hence the average measured earnings loss  $\widehat{\varepsilon}_0^E$  through layoff is never more than 50%. Indeed an estimate of  $\widehat{\varepsilon}_0^E$  close to 0.5 for the low education group reflects these workers have very low re-employment rates. In contrast for medium-educated workers, the measured first year earnings loss is around 35%. Of course for consistency the model-simulated results are subject to the same aggregation methodology.

Figure 2 demonstrates the very steep drop in earnings in the first year following layoff. Over the next 2 years, earnings recover at a relatively fast rate but then recovery slows and measured earnings losses eventually plateau to a long run cost which is strictly below zero; *i.e.* earnings never fully recover. The results using model-simulated data are remarkably well-aligned. The structural decomposition offered by the model (see below) provides a clear explanation of the underlying economic principles.

## 6.2. Reduced form cost of job loss II: [log] wage losses

We use the same methodology but now specify (26) in log wages. Table B.1 in Appendix B.3 describes the parameter estimates  $\widehat{\varepsilon}_t^w$  (reported in percentage terms) and standard errors. Figure 3 graphs these estimates, along with a 95% confidence interval, for the low and medium educated groups.

The results using the actual data exhibit two main features: (i) workers have large and persistent displacement wage losses and (ii) those losses are bounded away from zero as  $t$  becomes large. Medium-educated workers suffer a larger wage loss immediately following displacement (point estimate  $\widehat{\varepsilon}_1^w = -11.7\%$ ) compared to low educated workers ( $\widehat{\varepsilon}_1^w = -8.1\%$ ).

Though not quite such a good fit as obtained in Figure 2, the model still captures very well the extent of the wage losses at  $t=0$  and their persistence following job displacement. Figure 3, however, seems to suggest that log wages losses are overestimated at intermediate  $t$ ,

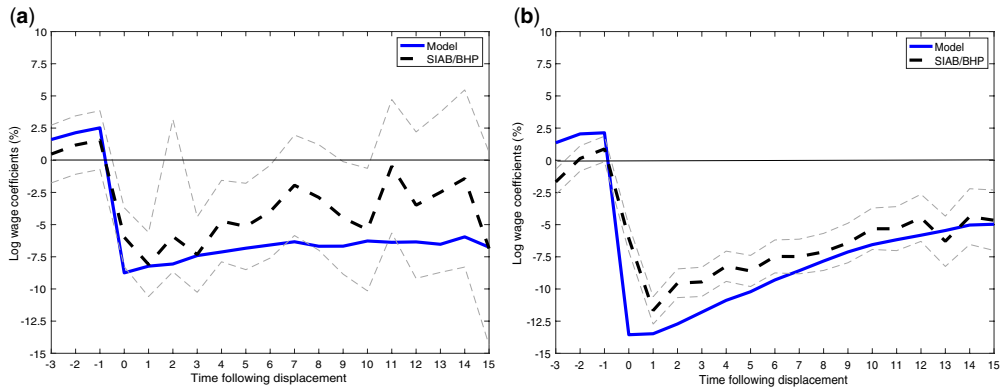


FIGURE 3

Post displacement wage losses. (a) Low educated: data and model. (b) Medium educated: data and model.

particularly for the low education group (though estimated standard errors are also large). A plausible explanation is learning-by-doing rates decline with experience in which case estimated foregone learning while unemployed is overstated.

Prior to displacement  $t < 0$  and for the actual data, Figure 3 demonstrates a clear positive spike in the estimated  $\hat{\varepsilon}_t^w$  (see also Davis and Von Wachter 2011; Jarosh 2015, for a similar pattern). Perhaps surprisingly, the estimates based on model-simulated data exhibit a similar spike (also see Figure 2 though there this effect is less pronounced). Because there are no selection effects in the model-simulated data (layoffs are random by construction), the pre-displacement trend arises because of the assumed “treated group”: they are those laid-off as part of a mass layoff event *with at least 3 years tenure*. This group is thus ex ante lucky: they previously enjoyed 3 solid years of learning-by-doing. In contrast, some workers in the control group were previously unemployed with skills loss. The difference in average skills accumulation between these two groups (prior to the date of displacement) then generates a positive trend in the pre-displacement estimates. Because the results using model-simulated data are so closely aligned to those obtained on the actual data, we now use the structurally estimated model to decompose the cost of job loss.

### 6.3. The cost of job loss: a structural decomposition

Consider two representative workers where at date  $t = 0$ , each initially has the same human capital  $k = 1$  but “control” is employed while “treated” is unemployed. We suppose control has wage rate  $\theta^C \sim G_\theta$  consistent with the ergodic distribution of wage rates paid.

The model describes a Markov process for how employment, human capital, and wage rates subsequently evolve over time. Let  $p_t^C$  and  $p_t$  denote the workers’ respective probabilities of being employed at date  $t \geq 0$ . Conditional on being employed at date  $t$ , let  $\theta_t^c$  describe the wage rate earned by control and  $k_t^c$  denote control’s human capital. Both, of course, are random variables. Conditional on being employed at date  $t$ , let  $\theta_t$  and  $k_t$  describe the earned wage rate and human capital of the laid-off worker.

We first decompose  $\varepsilon_t^w$ . Conditional on being employed, the expected *log wage gap* is

$$\begin{aligned} \varepsilon_t^w = & E[\log \theta_t^c + \log k_t^c \mid \text{control is employed at } t] \\ & - E[\log \theta_t + \log k_t \mid \text{laid-off is employed at } t]. \end{aligned}$$

Because the turnover processes are independent,  $\varepsilon_t^w$  can be re-expressed as

$$\varepsilon_t^w = E[\log \theta_t^c - \log \theta_t | \text{both employed at } t] + E[\log k_t^c - \log k_t | \text{both employed at } t].$$

We denote the first term

$$\varepsilon_t^{JL} = E[\log \theta_t^c - \log \theta_t | \text{both employed at } t],$$

which describes the laid-off worker's expected wage loss due to job ladder effects. This term not only reflects that the laid-off worker must re-climb the job ladder, but also that the control worker remains employed at  $t = 0$ .

We denote the second term

$$\varepsilon_t^{SL} = E[\log k_t^c - \log k_t | \text{both employed at } t],$$

which describes the expected wage loss due to differential skill accumulation rates while employed and unemployed. Hence, we have the decomposition

$$\varepsilon_t^w = \varepsilon_t^{JL} + \varepsilon_t^{SL}. \tag{27}$$

By providing an analytic solution for  $\varepsilon_t^{SL}$ , Proposition 2 will allow us to decomposed estimated losses  $\varepsilon_t^w$  into skill loss and job ladder losses.<sup>21</sup>

**Proposition 2.** *Equilibrium implies  $\varepsilon_t^{SL} = \mu_t^c - \mu_t$  with:*

$$\mu_t^c = \frac{2\delta \frac{\phi + \rho}{[\lambda_0 + \delta]^2} [1 - e^{-(\lambda_0 + \delta)t}] + \left[ \frac{\lambda_0}{\lambda_0 + \delta} \rho - \frac{\delta}{\lambda_0 + \delta} \phi \right] t + \frac{\delta[\delta\rho - \lambda_0\phi]}{\lambda_0[\lambda_0 + \delta]} t e^{-(\lambda_0 + \delta)t}}{1 + \frac{\delta}{\lambda_0} e^{-(\lambda_0 + \delta)t}}$$

$$\mu_t = \frac{(\delta - \lambda_0) \frac{\phi + \rho}{[\delta + \lambda_0]^2} [1 - e^{-(\lambda_0 + \delta)t}] + \left[ \frac{\lambda_0}{\lambda_0 + \delta} \rho - \frac{\phi}{\lambda_0 + \delta} \delta \right] t - \left[ \frac{\delta\rho - \lambda_0\phi}{[\lambda_0 + \delta]} \right] t e^{-(\lambda_0 + \delta)t}}{1 - e^{-(\lambda_0 + \delta)t}}.$$

**Proof:** See the Appendix.

We can then decompose  $\widehat{\varepsilon}_t^E$ , the estimated cost of job loss in earnings, by defining

$$\widehat{\varepsilon}_t^U = \widehat{\varepsilon}_t^E - \widehat{\varepsilon}_t^w,$$

which is the gap between the two different estimates of the cost of job loss. For reasons that will become clear, we refer to  $\widehat{\varepsilon}_t^U$  as the employment gap effect [that the laid-off worker is more likely to be unemployed (with zero earnings)]. Hence, we obtain the decomposition

$$\widehat{\varepsilon}_t^E \equiv \widehat{\varepsilon}_t^U + \widehat{\varepsilon}_t^{SL} + \widehat{\varepsilon}_t^{JL},$$

where:

21. Given the temporal aggregation issue described above, the skill loss estimate described in Figure 4 below is the average value of over the accounting year. For example  $\widehat{\varepsilon}_0^{SL}$  is not zero, rather it is the average value of  $\varepsilon_t^{SL}$  over  $t \in [0, 1]$ .

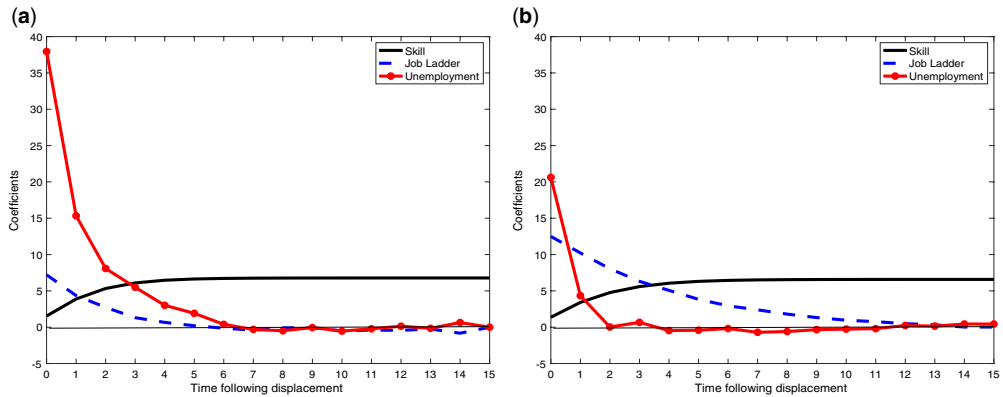


FIGURE 4

Model decomposition of earnings losses. (a) Low education. (b) Medium education.

1. Employment gap:  $\widehat{\varepsilon}_t^U \equiv \widehat{\varepsilon}_t^E - \widehat{\varepsilon}_t^w$ ;
2. Skill loss  $\widehat{\varepsilon}_t^{SL}$  given by Proposition 2;
3. Job ladder:  $\widehat{\varepsilon}_t^{JL} \equiv \widehat{\varepsilon}_t^w - \widehat{\varepsilon}_t^{SL}$ .

To be theory consistent, we decompose the estimates  $\widehat{\varepsilon}_t^E$  obtained using model-simulated data. Figure 4 graphs the resulting decomposition by year following layoff.

We stress the three key insights that follow:

**Insight 1:** as time since displacement becomes large, the cost of job loss  $\varepsilon_t^E$  converges to  $\varepsilon_t^{SL}$  (*i.e.*  $\varepsilon_t^U, \varepsilon_t^{JL} \rightarrow 0$  as  $t \rightarrow \infty$ ). Note this property does not occur if instead the control group is instead those who are never laid-off.<sup>22</sup>  $\Phi^P$  in Table 4 below describes the model predicted permanent human capital loss  $\lim_{t \rightarrow \infty} \varepsilon_t^{SL}$ . Figure 2 shows  $\Phi^P$  is fully consistent with the permanent earnings losses  $\lim_{t \rightarrow \infty} \varepsilon_t^E$  as estimated on the actual data.

**Insight 2:** the estimated cost of job loss in log wages at  $t=0$  describes the immediate drop in log wages due to falling off the job ladder. This estimated loss is consistent with  $\Phi^T$  in Table 4 below which describes the model predicted loss.

**Insight 3:** the unusual recovery path of estimated earnings losses reflects that the employment gap effect and job ladder losses decay at different rates.

**Job ladder effects  $\varepsilon_t^{JL}$**

Recall that

$$\varepsilon_t^{JL} = E[\log \theta_t^c - \log \theta_t] \text{ both employed at } t].$$

If at  $t=0$  the laid-off worker is immediately re-employed, then the wage rate earned  $\theta \sim F_\theta$ . Hence for  $t=0$ , the model implies  $\varepsilon_0^{JL} = \Phi^T$  where

$$\Phi^T = E_{G_\theta}[\log \theta] - E_{F_\theta}[\log \theta].$$

Table 4 below computes the model implied (temporary wage loss)  $\Phi^T$  for each educational group. For the medium-educated workers, the model implies an initial job ladder wage loss of 10% while the low-educated group have a 6.2% loss. This differential arises because workers in the low-educated group are more likely to be laid-off than receive an outside job offer (*i.e.* their  $\delta > \lambda_1$ )

22. In that case, the employment gap effect  $\varepsilon_t^U$  is bounded away from zero.

TABLE 4  
*The cost of job loss—model*

	Low education (%)	Medium education (%)	Higher education (%)
Temporary wage loss $\Phi^T$	6.20	10.02	9.03
Permanent wage loss $\Phi^P$	6.89	6.53	8.59

and so their job ladder effects are relatively small. Instead medium-educated workers are more likely to receive outside offers than be laid-off (*i.e.*  $\lambda_1 > \delta$ ) and so their job ladder effects are larger. Using the actual data, Figure 3 demonstrates that estimated [short run] wage losses  $\widehat{\varepsilon}_t^w$  are indeed consistent with the model-implied  $\Phi^T$ ; *i.e.* early wage losses are consistent with job ladder effects.

Now consider the limit as  $t \rightarrow \infty$ . Conditional on being employed, the distribution of the laid-off worker's wage rate gradually converges to the ergodic distribution  $G_\theta$ . Hence,  $\lim_{t \rightarrow \infty} \varepsilon_t^{JL} = 0$  as confirmed by Figure 4. This figure also shows the job ladder loss decays more slowly than does the employment gap loss  $\varepsilon_t^U$ . We return to this feature below.

### Skill loss $\varepsilon_t^{SL}$

Proposition 2 provides the analytic solution for

$$\varepsilon_t^{SL} = E[\log k_t^c - \log k_t | \text{both employed at } t].$$

Note that  $\varepsilon_t^{SL} = 0$  at  $t=0$  for there is no skill loss should the laid-off worker be immediately rehired (as in the godfather shock). Conversely as  $t \rightarrow \infty$ , Proposition 2 implies  $\varepsilon_t^{SL} \rightarrow \Phi^P > 0$ , where

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

Hence as time since displacement becomes large, expected skill loss not only depends on foregone skill accumulation rates  $\rho + \phi$  but also on job turnover rates. Not surprisingly if finding work is fast ( $\lambda_0$  large), foregone skill accumulation through becoming unemployed is small. But the measured loss also depends on job loss rates, for  $\delta$  high implies the control worker is also likely to become unemployed in the near future.<sup>23</sup> Using the parameter values in Table 3, Table 4 below describes  $\Phi^P$  (the permanent expected skills loss as  $t$  becomes large) for each education group. Despite having quite different turnover processes, the model implies  $\Phi^P$ , the expected (long run) fall in skills following displacement, is 6.5% and 6.9% for the two groups.

Figure 2 demonstrates that estimated (long run) earnings losses are fully consistent with the model estimated permanent skills loss  $\Phi^P$ . At first sight this seems surprising if it is held there are decreasing rates of learning-by-doing. But an important insight here is that the skills loss term converges to  $\Phi^P$  very quickly (see Figure 4). Like the employment gap effect discussed below, Proposition 2 and some algebra implies convergence occurs at rate  $(\lambda_0 + \delta)$  which is fast: Table 3 imply workers in the low education group have slow re-employment rates  $\lambda_0$  but high layoff rates  $\delta$  which together yield a convergence rate of  $\lambda_0 + \delta = 0.071$  per month (*i.e.* a half-life of just 10 months). Surprisingly the medium educated group has almost the same convergence rate  $\lambda_0 + \delta = 0.069$ , though for the opposite reason. Although learning-by-doing rates might decline with experience, fast convergence to  $\Phi^P$  implies (a relatively slow) decline in learning-by-doing

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

has little effect on realized skills loss. Hence, the model predicted  $\Phi^P$  is remarkably close to the (long-run) estimated earnings losses.

### Employment gap $\varepsilon_t^U$

Conditional on being employed, let  $w_t^C$  denote the expected wage earned by the control at date  $t$ ,  $w_t$  the expected wage earned by the displaced worker. The definition of  $p_t^C, p_t$  above imply the % expected earnings loss is

$$\varepsilon_t^E = \frac{p_t^C w_t^C - p_t w_t}{p_t^C w_t^C},$$

which we now decompose as

$$\varepsilon_t^E \equiv \frac{(p_t^C - p_t) w_t}{p_t^C w_t^C} + \frac{w_t^C - w_t}{w_t^C}.$$

Note that conditional on both workers being employed, the second term is the expected cost of job loss in % wages. Using instead a logarithmic approximation,  $\widehat{\varepsilon}_t^w$  provides a direct estimate of this second term. Hence the measured gap  $\widehat{\varepsilon}_t^E - \widehat{\varepsilon}_t^w$  is an estimator of the employment gap effect:

$$\varepsilon_t^U = \frac{(p_t^C - p_t) w_t}{p_t^C w_t^C}.$$

Because the laid-off worker is more likely to be unemployed at any future date  $t > 0$ , the employment gap effect  $\varepsilon_t^U$  describes how this additionally affects worker earnings losses. Now standard algebra establishes  $p_t^C - p_t = e^{-(\lambda_0 + \delta)t}$  and so the employment gap effect  $\varepsilon_t^U$  decays at rate  $\lambda_0 + \delta$ .<sup>24</sup> The intuition is simply that the “treated” worker regains employment at rate  $\lambda_0$  while “control” becomes unemployed at rate  $\delta$  and so the probability gap  $p_t^C - p_t$  decays at rate  $\lambda_0 + \delta$ . And because the above establishes this decay rate is around 7% a month (see above), the employment gap effect declines very quickly. This implies the estimated profile of earnings losses converges quickly to the estimated profile of (log) wage losses.

But why are job ladder losses seemingly more persistent? An important difference is that the control worker has wage rate  $\theta^C$  consistent with the ergodic distribution. Conversely (11) implies that when the laid-off worker is re-employed, wage tenure effects are proportional to the job offer arrival rate  $\lambda_1$ . Thus via on-the-job search and tenure effects, measured job ladder losses decay at rate  $\lambda_1$  and parameter estimate  $\lambda_1 \ll \lambda_0 + \delta$  then implies job ladder losses are more persistent than the employment gap effect.

Our final issue is to measure the relative contribution of these three channels to discounted lifetime earnings losses following displacement. This requires taking into account there is positive earnings growth over time: that average total earnings grow at rate  $\gamma_A + \gamma_k$ , where  $\gamma_k = (1 - u)\rho - u\phi$  is the average growth rate of human capital. The model estimated parameters imply annual growth rates  $\gamma_k = 2.0\%$ ,  $2.4\%$  for the low- and medium-educated groups, respectively.

Taking trend growth into account, we measure the percentage loss in lifetime earnings (LLE) due to layoff as:

$$\text{LLE} = \frac{\sum_{t=0}^{\infty} \widehat{\varepsilon}_t^E e^{[\gamma_A + \gamma_k - r]t} y_0}{\sum_{t=0}^{\infty} e^{[\gamma_A + \gamma_k - r]t} y_0}.$$

24. Specifically

$$p_t^C = \frac{\lambda_0 + \delta e^{-(\lambda_0 + \delta)t}}{\lambda_0 + \delta}, p_t = \frac{\lambda_0}{\lambda_0 + \delta} [1 - e^{-(\lambda_0 + \delta)t}].$$

TABLE 5  
Discounted earnings losses—decomposition (%)

	LLE <sup>SL</sup> (%)	LLE <sup>JL</sup> (%)	LLE <sup>U</sup> (%)
Low educated	70.67	5.51	23.82
Medium educated	80.72	11.27	8.01

Here, the denominator measures discounted lifetime earnings with  $y_0$  describing representative worker earnings at date zero, and using  $\widehat{\varepsilon}_t^E$  in the numerator then describes the (proportional) loss in discounted earnings in every period  $t \geq 0$ . Because earnings losses at year 15 are close to the permanent wage losses  $\Phi^P$  described in Table 4, we set  $\widehat{\varepsilon}_t^E = \Phi^P$  for  $t \geq 16$ . Using model-based estimates of  $\widehat{\varepsilon}_t^E$ , the expected loss in lifetime earnings due to layoff is 9.3% for the low educated and 7.9% for the medium educated. These are large losses.

We decompose this loss LLE into its constituent parts  $LLE = LLE^{SL} + LLE^{JL} + LLE^U$  by defining

$$LLE^k = \frac{\sum_{t=0}^{\infty} \widehat{\varepsilon}_t^k e^{[\gamma_A + \gamma_k - r]t}}{\sum_{t=0}^{\infty} e^{[\gamma_A + \gamma_k - r]t}}$$

with  $k = SL, JL, U$  using the model estimated  $\widehat{\varepsilon}_t^k$  for  $t \leq 15$  and, for consistency, setting  $\widehat{\varepsilon}_t^{SL} = \Phi^P$  and  $\widehat{\varepsilon}_t^{JL} = \widehat{\varepsilon}_t^U = 0$  for  $t > 15$ . Doing this yields the following decomposition of lifetime earnings losses by education groups

Table 5 reveals skill loss is by far the most important effect: it contributes over 70% of those lifetime earnings losses. For the low education category, losses due to job ladder effects are very small and the employment gap effect is correspondingly large. Instead for the medium educated, the job ladder effect is roughly of the same magnitude as the employment gap effect. For these workers falling off the job ladder implies a 0.8% fall in lifetime earnings, which is the expected cost of the godfather shock.

## 7. CONCLUSION

This article has generalized the equilibrium framework of [Burdett and Coles \(2003\)](#) to the case of learning-by-doing while employed, human capital loss while unemployed and to a “timeless equilibrium” which allows for growth. Structural estimation finds the model explains well not only the variation in firm fixed wage effects and firm wage tenure effects as estimated using the AKM methodology, but also standard Mincer-estimated returns to experience and tenure. The validation exercise finds that when the data are disaggregated by education type and using the control group advocated in [Sianesi \(2004\)](#), the structurally estimated model replicates very well the estimated cost of job loss. Results find that the estimated cost of job loss is very large (around 8–9% of lifetime earnings) and human capital loss is by far the largest component of such losses. Job ladder losses are not unimportant—they are central to explaining the (short-term) large wage drops that laid-off workers face on re-employment. But the cost of the godfather shock (a pure job ladder loss) is around 0.8% of lifetime earnings for the medium educated group and half that again for the low-educated group. Although losing 0.8% of your expected lifetime earnings is not a trivial amount, being laid-off into unemployment as part of a mass layoff event is a substantially costlier event.

A strong assumption of the model is no selection effects into layoff. This is not only necessary to make the equilibrium analysis tractable, it is also necessary for the validity of the [Jacobson et al.](#)



(1993) approach. The standard selection argument is that some workers may be more likely to be laid-off than others due to firm choice. Our approach instead presumes layoff rates are type-specific, that some types of workers are more likely to be laid-off than others. For example, a manufacturing firm's workforce might be mainly composed of (low educated) assembly line workers, and shocks to the manufacturing sector then cause such low educated types to be over-represented in any mass layoff event. Instead of a selection issue, this instead describes a composition problem which can only be resolved through disaggregation of the data. Even though the data here are just disaggregated into three educational groups, we find the results are already very good in the sense that the type-specific estimated cost of job loss is closely aligned to the theoretical predictions of the structurally estimated model. Of course future research might consider an even finer disaggregation.

That is not to say selection effects are unimportant. Indeed an interesting research question is rather than assume a firm implements a random layoff rule, what instead might be an optimal selection rule? An obvious candidate is a last in/first out seniority rule; *i.e.* the most recently hired worker is the first to be laid-off. Such a rule is not only transparent (which avoids claims of unfair dismissal), it also backloads job security with tenure and so has valuable incentive properties. Recently, [Pinheiro and Visschers \(2015\)](#), [Jarosh \(2015\)](#) find laid-off workers are more likely to be laid-off again once re-employed. That approach explains such outcomes as caused by firm heterogeneity—that firms at the bottom of the job ladder offer less secure employment. But it could simply reflect seniority protocols, that new hires are the first in line to be laid-off. A different selection rule might instead find firms asking older workers, particularly those close to retirement, to take (compensated) voluntary redundancy. Such selection based on life cycle issues, however, cannot be considered in the ageless framework analysed here.

A different way to generate a richer layoff structure is to introduce match heterogeneity of the form  $F(.) = Ak\varepsilon$ , where  $\varepsilon$  is a match-specific component which follows an exogenous geometric process. If  $\varepsilon = 1$  for all new hires and subsequently grows at a constant rate  $\gamma_F$ , then  $\gamma_F$  would describe the growth rate of firm-specific human capital. The only material difference this makes to the analysis is that firm-worker profit  $\Pi(.)$  in (11) additionally grows at rate  $\gamma_F$  and so wages would then increase more quickly with tenure. However, to keep tenure effects consistent with the Mincer wage equations, estimation would then have to increase worker risk aversion  $\sigma$  to generate a flatter wage tenure profile. Without additional wage information, distinguishing between tenure effects due to growth in firm-specific human capital and due to the backloading of wages is problematic. In this article, we have assumed no firm-specific human capital.

If instead  $\varepsilon$  is an idiosyncratic match draw and assuming  $\varepsilon$  is contractible, (11) would still describe how wages evolve within the match, but a bad match would not only imply low wages today but also low, and even negative, wage growth. Quit turnover would then depend on an employee's (idiosyncratic) wage and promotion prospects. Layoff instead occurs should match value  $\varepsilon$  fall below some reservation value  $\varepsilon^R$  (where match surplus is zero). Introducing idiosyncratic match draws then generates a rich and complex relationship between layoff rates and worker employment histories.

*Acknowledgments.* We would like to thank our co-editor and four anonymous referees for their comments and suggestions. We will also thank Jesper Bagger, Gregory Jolivet, John Ham, Giovanni Mastrobuoni, Giuseppe Moscarini, Fabien Postel-Vinay, Jean-Marc Robin, Ludo Visschers, and seminar participants at the SaM annual conference (Amsterdam), the Aix-en-Provence, Essex, and Bristol SaM workshops, the St Louis FED, the Institute of Employment Agency (IAB), Germany, and the Universities of Bonn, Konstanz, Kent, and Vienna for their comments and suggestions. Patryk Bronka provided excellent research assistance with the data work. The usual disclaimer applies. Carlos Carrillo-Tudela and Melvyn Coles acknowledge financial support from the UK Economic and Social Research Council (ESRC), award reference ES/I037628/1.

## APPENDIX

## A. PROOFS

**Proof of Theorem 1:** Substituting out  $\psi$  in the objective functions gives the dynamic optimization problem:

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

subject to starting value  $U(0|\theta) = U_0$  where (2) describes how  $U(\cdot)$  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

$$\dot{\psi}_0 = -[r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F(U)]]\psi_0. \quad (\text{A.1})$$

The dynamic optimization problem is equivalently rewritten as

$$\max_{\theta(\cdot)} \int_0^{\infty} \psi_0 [1-\theta] d\tau, \quad (\text{A.2})$$

where  $\psi_0, U$  are state variables which evolve according to the autonomous, first-order differential equations (A.1) 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

$$H = \psi_0 [1-\theta] - \xi_{\psi_0} [r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F(U)]]\psi_0 \\ + \xi_U \left[ -\left[ \frac{\theta^{1-\sigma}}{1-\sigma} + \delta\phi^{1-\sigma} U^U + \lambda_1 \int_U^{\bar{U}} [1-F(U_0)] dU_0 + \lambda_q \int_U^{\bar{U}} U_0 dF(U_0) \right] \right],$$

where  $\xi_{\psi_0}, \xi_U$  are the respective costate variables. The maximum principle yields the following necessary conditions for optimality:

$$\theta^{-\sigma} = -\frac{\psi_0}{\xi_U} \quad (\text{A.3})$$

$$\dot{\xi}_{\psi_0} = -[1-\theta] + \xi_{\psi_0} [r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F(U)]] \quad (\text{A.4})$$

$$\dot{\xi}_U = -[\xi_{\psi_0} \lambda_1 F'(U) \psi_0 + \xi_U [[r+\delta+\lambda_q - [\rho + \gamma_A](1-\sigma)] + \lambda_1 [1-F(U)]]] \quad (\text{A.5})$$

along with autonomous differential equations (A.1), (2) for  $\dot{\psi}_0$  and  $\dot{U}$ . As we do not wish to assume  $F$  is differentiable, however, we drop condition (A.5) and instead note that as the objective function in (A.2) does not depend explicitly on tenure, optimality also implies

$$H = 0 \quad (\text{A.6})$$

(e.g. Leonard and Long 1992, p. 298). Now integrating (A.4) forward yields:

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

where  $B_0$  is the constant of integration and  $\Pi(\cdot)$  is the firm's continuation profit as defined in Theorem 1. Equation (A.3) implies  $\xi_U = -\psi_0 \theta^{\sigma}$ . Substituting out  $\xi_U$  and  $\xi_{\psi_0}$  in the definition of  $H$ , the restriction  $H = 0$  yields the optimality condition:

$$0 = [1-\theta] - [r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1[1-F(U)]] \left[ \Pi(t) + B_0 e^{\int_0^t \{r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1(1-F(U(x)))\} dx} \right] \\ - \theta^{\sigma} \left[ [r+\delta+\lambda_q - [\rho + \gamma_A](1-\sigma)] U - \frac{\theta^{1-\sigma}}{1-\sigma} - \delta\phi^{1-\sigma} U^U - \lambda_1 \int_U^{\bar{U}} [1-F(U_0)] dU_0 - \lambda_q \int_U^{\bar{U}} U_0 F(U_0) \right]. \quad (\text{A.7})$$

Now the restriction  $r+\delta-\rho-\gamma_A > 0$  ensures the exponential term becomes arbitrarily large as  $\tau \rightarrow \infty$ . As  $\Pi$  and  $U$  must be bounded, then (A.7) implies  $B_0 = 0$ . Equation (A.7) 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 contract value more than  $U$  while the right hand side describes the flow out through job separation. As (8) and (10) together imply

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

rearranging the previous expression yields

$$\frac{dG}{dU} = \frac{(1-u)\delta(1-G(U))-u\lambda_0[1-F(U)]-(1-u)G(U)\lambda_1[1-F(U)]}{(1-u)\theta^{-\sigma} [1-\theta-[r+\delta-\rho-\gamma_A+\lambda_1(1-F(U))]\Pi]}, \quad (\text{A.8})$$

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

$$\frac{d\widehat{U}}{d\theta} = -\theta^{-\sigma} \frac{d\widehat{\Pi}}{d\theta}, \quad (\text{A.9})$$

and using (A.8) to substitute out  $\widehat{\Pi}'(U)$  and  $G'(U)$  gives

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

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

$$\begin{aligned} & [\lambda_0 u + \lambda_q(1-u) + \lambda_1(1-u)G] \{1-\widehat{\theta} - [r+\delta+\lambda_q-\rho-\gamma_A+\lambda_1]\widehat{\Pi}\} \\ & = \lambda_1(1-u)\widehat{\Pi} \left[ \delta(1-G) - \frac{u}{1-u} \lambda_0 - G\lambda_1 - G\lambda_q \right]. \end{aligned} \quad (\text{A.10})$$

But the constant profit condition also implies

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

Using this to substitute out  $G$  in (A.10) and substituting out  $u = \delta/(\delta + \lambda_0)$  yields the quadratic equation

$$\widehat{\Pi}^2[\delta + \lambda_q] + [r - \rho - \gamma_A]\widehat{\Omega}\widehat{\Pi} - (1-\widehat{\theta})\widehat{\Omega} = 0. \quad (\text{A.11})$$

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:** Equations (13) and (16) imply

$$\frac{d\widehat{U}}{d\theta} = \frac{\theta^{-\sigma}}{\left[ [r-\rho-\gamma_A]^2 + 4[\delta+\lambda_q][[\delta+\lambda_q+r-\rho-\gamma_A]\frac{1-\theta}{1-\widehat{\theta}}] \right]^{1/2}}, \quad (\text{A.12})$$

whose solution is given by (17). Given  $\widehat{\Pi}(\theta)$ , the constant profit condition (14) implies (18).

To determine the equilibrium distribution of offers  $F_\theta$ , standard turnover arguments imply  $G_\theta$  must satisfy

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

where the left-hand side describes the flow of workers into employment with wage rate more than  $\theta$  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  $\theta$  and  $G_\theta, \widehat{\Pi}$  described in the theorem, a lot of algebra finds the turnover equation for  $G_\theta$  implies the following first-order differential equation for  $F$ :

$$\frac{\theta F'_\theta}{\sigma} + (1 - F_\theta) = \Psi(\theta), \tag{A.13}$$

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  $F_\theta$  while using (12) it is easy to show that  $\Psi(\theta) > 0$  for all  $\theta \in [\underline{\theta}, \bar{\theta}]$ . This completes the proof of Theorem 2.

**Proof of Claim 6:** Integration by parts finds

$$\begin{aligned} \int_{\underline{\theta}}^{\bar{\theta}} \widehat{U}(\theta) dF_\theta(\theta) &= \overline{U} - \int_{\underline{\theta}}^{\bar{\theta}} \frac{d\widehat{U}}{d\theta} F_\theta(\theta) d\theta \\ &= \overline{U} - \int_{\underline{\theta}}^{\bar{\theta}} \frac{[\theta]^{-\sigma} F_\theta(\theta) d\theta}{[r - \rho - \gamma_A]^2 + 4[\delta + \lambda_q][\delta + \lambda_q + r - \rho - \gamma_A] \frac{1-\theta}{1-\bar{\theta}}} d\theta. \end{aligned} \tag{A.14}$$

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

$$\overline{U} = \underline{U} + \int_{\underline{\theta}}^{\bar{\theta}} \frac{\theta^{-\sigma} d\theta}{[r - \rho - \gamma_A]^2 + 4[\delta + \lambda_q][\delta + \lambda_q + r - \rho - \gamma_A] \frac{1-\theta}{1-\bar{\theta}}} d\theta. \tag{A.15}$$

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

$$\frac{\bar{\theta}^{1-\sigma}}{1-\sigma} = [r + \delta + \lambda_q - [\rho + \gamma_A](1-\sigma)] \overline{U} - \delta \underline{U} - \lambda_q \int_{\underline{\theta}}^{\bar{\theta}} \widehat{U}(\theta) dF_\theta(\theta). \tag{A.16}$$

Using (A.14) and (A.15) to substitute out  $\overline{U}$  implies:

$$[r - [\rho + \gamma_A](1-\sigma)] \underline{U} = \frac{\bar{\theta}^{1-\sigma}}{1-\sigma} - \int_{\underline{\theta}}^{\bar{\theta}} \frac{[r + \delta - [\rho + \gamma_A](1-\sigma) + \lambda_q F_\theta(\theta)] \theta^{-\sigma} d\theta}{[r - \rho - \gamma_A]^2 + 4[\delta + \lambda_q][\delta + \lambda_q + r - \rho - \gamma_A] \frac{1-\theta}{1-\bar{\theta}}} d\theta. \tag{A.17}$$

Equation (3) with  $U^U = \underline{U}$  [Claim 5], (A.14) and substituting out  $\overline{U}$  using (A.15) implies

$$(r + \phi(1-\sigma) - \gamma_A(1-\sigma)) U^U = \frac{b^{1-\sigma}}{1-\sigma} + \int_{\underline{\theta}}^{\bar{\theta}} \frac{\theta^{-\sigma} \lambda_0 [1 - F_\theta(\theta)]}{[r - \rho - \gamma_A]^2 + 4[\delta + \lambda_q][\delta + \lambda_q + r - \rho - \gamma_A] \frac{1-\theta}{1-\bar{\theta}}} d\theta. \tag{A.18}$$

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  $\bar{\theta} \rightarrow 1$ , (20) implies  $\underline{\theta} \rightarrow 1$  and so all wage rates paid lie in a neighbourhood of 1. Frictions  $\lambda_0 < \infty$ ,  $b < 1$ , and  $\phi \geq 0$  ensure the value of being unemployed  $\widetilde{U}^U(\bar{\theta}) < \widetilde{U}(\bar{\theta})$  in this limit.

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

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

then  $\underline{\theta} \rightarrow 0^+$  and (A.17) implies  $\widetilde{U}(\bar{\theta}) \rightarrow -\infty$  in this limit. As (A.18) implies  $\widetilde{U}^U \geq \frac{b^{1-\sigma}}{1-\sigma} / (r + \phi(1-\sigma) - \gamma_A(1-\sigma))$  and so is finite, then  $\widetilde{U}^U(\bar{\theta}) > \widetilde{U}(\bar{\theta})$  in this limit. As these are continuous functions for  $\bar{\theta}$  satisfying (21), a  $\bar{\theta}$  satisfying (21) exists such that  $\widetilde{U}(\bar{\theta}) = \widetilde{U}^U(\bar{\theta})$ . This completes the proof of Theorem 3.

**Proof of Proposition 1.**

Consider two representative workers with human capital  $k_0$  at time  $t=0$ , where one is employed (control) and the other is unemployed (laid-off). Human capital evolves stochastically where

$$d[\log k] = \rho dt \text{ while employed,}$$

$$d[\log k] = -\phi dt \text{ while unemployed,}$$

and a worker switches between employment states according to an independent Poisson process with transition parameters  $\lambda_0, \delta$ . Note the  $d[\log k]$  dynamics are independent of  $\log k$ .

Let  $t'$  denote the first date when either worker transits employment state; *i.e.* either the employed worker (control) becomes unemployed, or the laid-off worker finds employment. The (independent) Poisson processes implies this occurs with probability density  $(\lambda_0 + \delta)e^{-(\lambda_0 + \delta)t'}$ . At this date  $t'$ , their difference in  $\log k$  is  $(\rho + \phi)t'$ . Because both workers are in the same employment state at date  $t'$  then, in the continuation  $t > t'$ , both workers face the same  $d[\log k]$  process and so, in expectation, there is no further change in the difference in their  $[\log]$  human capitals. Hence the expected difference in  $\log k$ , at date  $t$ , is

$$\Phi_t = \int_0^t [(\rho + \phi)t'](\lambda_0 + \delta)e^{-(\lambda_0 + \delta)t'} dt' + e^{-(\lambda_0 + \delta)t}(\rho + \phi)t,$$

where  $(\rho + \phi)t'$  describes the expected difference in  $\log k$  at date  $t$  for first transitions which occur at date  $t' < t$ , while the second term describes the difference in  $\log k$  if the first transition has not occurred by time  $t$ . Integration by parts yields

$$\Phi_t = \frac{\rho + \phi}{\lambda_0 + \delta} [1 - e^{-(\lambda_0 + \delta)t}]$$

and taking the limit  $t \rightarrow \infty$  yields Proposition 1.

**Proof of Proposition 2.**

First consider the control group and, for ease of exposition, suppose there is a unit measure of workers who are employed at date  $t=0$ , each with human capital  $k_0=1$ . Fraction  $p_t^C$  of these workers are employed at date  $t$ , where

$$p_t^C = \frac{\lambda_0 + \delta e^{-(\lambda_0 + \delta)t}}{\lambda_0 + \delta}.$$

Log human capital per worker evolves according to

$$d[\log k] = \rho dt \text{ while employed}$$

$$d[\log k] = -\phi dt \text{ while unemployed}$$

and workers have transition rates  $\lambda_0, \delta$  between employment states.

Define  $X_t^U$  as the integral of  $\log k$  across these workers who are unemployed at date  $t$ ; *i.e.*

$$X_t^U = \int_i [\log k_i] m_i^U(t) di,$$

where  $m_i^U(t)$  is the date  $t$  measure of unemployed workers with  $\log k_i$  and  $\int_i m_i^U(t) di = [1 - p_t^C]$  is the date  $t$  measure of unemployed workers. Similarly define  $X_t^E$  as the integral of  $\log k$  across workers who are employed at date  $t$ . Over arbitrarily small time period  $dt > 0$ , the turnover processes imply:

$$X_{t+dt}^U = \int_i [\log k_i - \phi dt][1 - \lambda dt] m_i^U(t) di + \int_i [\log k_i][\delta dt] m_i^E(t) di + o(dt),$$

where the first term describes the period  $t+dt$  contribution of those previously unemployed workers who remain unemployed (and their  $\log k$  falls by  $\phi dt$  over this period), the second describes entry from those workers who were previously employed, and the  $o(dt)$  term reflects the Poisson approximation. Hence

$$X_{t+dt}^U = [1 - \lambda dt] X_t^U - \phi dt [1 - p_t^C] + \delta dt X_t^E + o(dt).$$

Re-arranging appropriately and letting  $dt \rightarrow 0$  yields the differential equation:

$$\dot{X}^U = -\lambda_0 X^U + \delta X^E - [1 - p^C] \phi.$$

Repeating the argument establishes  $X^E$  evolves according to the differential equation:

$$\dot{X}^E = \lambda_0 X^U - \delta X^E + p^C \rho.$$

We thus have a pair of linear differential equations for  $X^U, X^E$ , where  $p^C$  is described above and have initial conditions  $X^U = X^E = 0$ . Standard arguments now apply which yield

$$X_t^E = \frac{\lambda_0}{\lambda_0 + \delta} \left[ \begin{aligned} &2\delta \frac{\phi + \rho}{[\lambda_0 + \delta]^2} [1 - e^{-(\lambda_0 + \delta)t}] + \left[ \frac{\lambda_0}{\lambda_0 + \delta} \rho - \frac{\delta}{\lambda_0 + \delta} \phi \right] t \\ &+ \frac{\delta [\delta \rho - \lambda_0 \phi]}{\lambda_0 [\lambda_0 + \delta]} t e^{-(\lambda_0 + \delta)t} \end{aligned} \right].$$

Because measure  $p_t^C$  of workers are employed, then the mean value of log  $k$  across employed workers at date  $t$  is

$$\mu_t^C = \frac{X_t^E}{p_t^C},$$

which yields  $\mu_t^C$  described in Proposition 2. The same argument, but with  $p_t^C$  replaced by  $p_t = \frac{\lambda_0}{\lambda_0 + \delta} [1 - e^{-(\lambda_0 + \delta)t}]$ , determines  $\mu_t$ , defined as the average value of log  $k$  across employed workers at date  $t$  who were instead unemployed at  $t=0$ . This completes the proof of Proposition 2.

## B. QUANTITATIVE ANALYSIS

### B.1. Simulation

To estimate the parameters of the model separately for each education group, we use the indirect inference formula

$$\hat{\Lambda} = \arg \min_{\Lambda} (\mathbf{M}^D - \mathbf{M}^S(\Lambda))' \mathcal{W} (\mathbf{M}^D - \mathbf{M}^S(\Lambda)), \tag{B.1}$$

where  $\mathbf{M}^D$  is an  $12 \times 1$  vector of data moments as described in Table 2,  $\mathbf{M}^S(\Lambda)$  is an  $12 \times 1$  vector of the same moments obtained from the simulations, which are a function of the  $8 \times 1$  vector of parameters to estimate  $\Lambda = \{\delta, \lambda_0, \lambda_1, \lambda_q, \rho, \phi, \sigma, b\}$ , and  $\mathcal{W}$  is an  $12 \times 12$  weighting matrix. To obtain the empirical moments in  $\mathbf{M}^D$ , we use data drawn from the SIAB/BHP and the GSEOP as described in the main text and in Sections B.2 and B.3, below. To obtain the simulated moments in  $\mathbf{M}^S(\Lambda)$ , we first compute the equilibrium of our model, then simulate workers' employment histories, and then compute each moment from this data. As a weighting matrix we use the inverse of the variance-covariance matrix of the data moments. We also used instead a diagonal matrix containing only the variance of the data moments, obtaining very similar results. The variance and covariances are obtained by bootstrapping using SIAB and GSEOP data. In the case of the firm fixed effect and its correlation with the firm-specific slope obtained from Carrillo-Tudela *et al.* (2019), we approximate their variances and covariances (and standard errors) using AKM estimates from the SIAB. This is the best we can do given that we cannot access directly the full BHP dataset. Although we acknowledge that this is not ideal given the potential lack of identification of the firm fixed effects using the SIAB, these variance and covariances show that the AKM estimates are nevertheless very precisely estimated.

For a set of parameter values, computing the equilibrium implies picking a  $\bar{\theta}$  satisfying (21), then using Theorem 2 to compute  $F_{\bar{\theta}}$  over  $[\underline{\theta}, \bar{\theta}]$  with  $\bar{\theta}$  given by (20). Then computing  $\underline{U}(\bar{\theta}), \bar{U}^U(\bar{\theta})$  as defined in Claim 6. The equilibrium value of  $\bar{\theta}$  is then determined by  $\underline{U}(\bar{\theta}) = \bar{U}^U(\bar{\theta})$ . Using the corresponding value of  $\underline{\theta}$ , we then solve the differential equation describing the evolution of  $\theta$  to obtain the baseline scale.

Given these equilibrium outcomes, we simulate the employment histories of 15,000 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  $\lambda_0$ . We obtain his unemployment duration by drawing a random number,  $r \in [0, 1]$  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  $tu = -\log(1 - r)/\lambda_0$ . After deriving  $tu$ , we sample a position in the baseline 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  $\theta$ .

When the worker is employed, he faces three shocks: a "godfather" reallocation shock, a job offer shock and a job layoff shock. All these shocks follow Poisson process with rates,  $\lambda_q, \lambda_1$ , and  $\delta$ , 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 transitioning 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 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 godfather 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  $\rho$  and human capital depreciation occurs at rate  $\phi$ .

We follow workers for over 40 years to guarantee that we converge to the ergodic distributions for each element of  $M^S$ . To compute the transition moments, we use average durations, except for the average ratio of involuntary to voluntary employer-to-employer transitions, for which we compute the average number of involuntary and voluntary transitions. To compute the average returns to experience and tenure, we construct a panel resembling the SIAB/BHP structure and regress log wages on a constant, a quadratic on experience and a quadratic on tenure. To compute the firm-specific wage rate and its correlation with the firm-specific tenure profile we estimate the AKM equation described in (24) using a 1–10 ratio between the number of firms (establishments) and workers. We have done several robustness exercises whereby we increased the number of firms to obtain a 1–5 ratio and decreased the number of firms to obtain a 1–50 ratio without observing any meaningful change in our results. We also use the simulated panel to compute the coefficient of variation and the  $Mm$  ratio as measures of frictional wage dispersion. The latter two follow the same procedure as we use to compute these moments from the data.

After the simulated moments are computed, we obtain the solution to the loss function, (B.1). 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. For our minimization algorithm, we first use simulating annealing to perform a global search and then use a constrain minimization procedure to perform a local search. Once the parameters that solve (B.1) are recovered, their standard errors can be calculated by using

$$\text{var}(\hat{\Lambda}) = [J'WJ]^{-1} J'W\mathbf{W}WJ[J'WJ]^{-1},$$

where  $J = \partial \mathbf{M}^S(\Lambda) / \partial \Lambda$  evaluated at  $\hat{\Lambda}$  and  $\mathbf{W} = \text{var}(\mathbf{M}^D - \mathbf{M}^S(\Lambda))$ , which reduces to  $(1 + 1/K)\text{var}(\mathbf{M}^D)$  at the null where  $K$  is the ratio of the number of simulations to the number of data points.

## B.2. Data moments

In this section, we describe the procedure we follow to compute the data moments obtained from the auxiliary equations. In particular, to compute the firm-specific wage rate and its correlation with the firm-specific (linear) tenure profile as described in the main text, Carrillo-Tudela *et al.* (2019) estimate equation (24) using the universe of German *full-time* workers available through the BHP. Given the lack of information on hours worked in the administrative data adding part-time workers would not be advisable (see also Card *et al.* 2013, for such a practise). Carrillo-Tudela *et al.* (2019) estimate separate AKM regression for each of the following four time windows: 1993–1999, 1998–2004, 2003–2010, and 2010–2017. They use the AKM estimates based on the last window for their main analysis. AKM estimates based on previous windows are used to understand the pattern of the worker and firm fixed effects since the German re-unification. Once estimated from the full sample, the fixed effects were then subdivided using the same education groups we use in the SIAB. Summary statistics and correlation tables were generated for each education group and time windows for the AKM estimates. Here, we use the results from the first to time windows as they overlap with the time period we use in the SIAB. The variance of the firm fixed effect and the correlation between the firm fixed effect and the firm-specific (linear) tenure profile for each 1993–1999 and 1998–2004 windows are (0.0473, 0.0652) and (–0.1833, –0.1901) among low skilled workers; (0.0489, 0.0578) and (–0.1873, –0.1768) among medium skilled workers; and (0.0593, 0.0387) and (–0.1850, –0.2237) among high skilled workers. To obtain a single estimate for the variance and the correlation for each education group, we obtain the average across time windows using the number of observations in each education group/time window as weights. These results are the ones reported in Table 2. As robustness we also used the estimates based the 1993–1999 and 1998–2004 periods on their own, without any meaningful change in the model's parameter estimates.

To compute the average returns to experience and tenure we estimate the following standard Mincer wage equation based on the SIAB data:

$$\log w_{it} = \beta X_{it} + \mu_{it}, \quad (\text{B.2})$$

where  $X$  is a vector of covariates consisting of a quadratic on actual experience, a quadratic on tenure and year dummies, and  $\mu_{it}$  denotes the error term. To compute the correlation between the last completed unemployment duration and re-employment wages, we estimate equation (25) as described in the text using a fixed effect estimator also using the SIAB data.

To compute the coefficient of variation and the  $Mm$  ratio, we follow Hornstein *et al.* (2007) and first estimate the wage equation (B.2) for each year of the sample period and education group using OLS and the same covariates in  $X$ . We then eliminate unobserved worker heterogeneity from wages by using the individual residuals  $\hat{\eta}_{it}$  and their individual specific mean  $\bar{\eta}_i = \sum_{t=1}^{N_i} \hat{\eta}_{it} / N_i$ . The vector  $\{\bar{\eta}_i\}_{i=1}^N$  then captures the wage variation due to fixed unobserved individual factors. Finally, we use the estimated distribution of transformed wages,  $\tilde{w}_{it} = \exp(\hat{\eta}_{it} - \bar{\eta}_i)$ , across individuals and time



to calculate the coefficient of variation and *Mm* ratio for each education group. For each education group, we estimate a set of three *Mm* ratios using the minimum observed wage and the wage at the first percentile. Given that the ratios using the minimum observed wages are implausibly high, we report in the main text the one based on the first percentile. Once again we use the SIAB data for this purpose.

Finally, as described in the main text, we also use the SIAB data to compute the average employment, unemployment, and job spell durations, but use the GSEOP to compute the ratio of involuntary to voluntary employer-to-employer transitions.

### B.3. Empirical wage/earnings losses

Our analysis of wage/earnings losses focus on a sample of male, West German workers. Following [Jacobson \*et al.\* \(1993\)](#), [Couch and Placzek \(2010\)](#) and many others using administrative data, to reduce selection effects in our main analysis we only consider displaced workers who were part of a mass layoff event in year  $y$ . The identification of a mass layoff follows the [Davis and Von Wachter's \(2011\)](#) criteria such that to qualify as a mass layoff event in year  $y$ , the employer must meet: (i) 50 or more employees in year  $y-2$ ; (ii) employment decreases by 30% to 99% from years  $y-2$  to  $y$ ; (iii) employment in year  $y-2$  is no more than 130% of employment in year  $y-3$ ; (iv) employment in year  $y+1$  is less than 90% of employment in year  $y-2$ . Due to data anonymization, the number of employees in a given establishment is not exact, but given in eight bands. We use the smallest value in a given band as a proxy for the number of employees, *i.e.* 1, 5, 10, 20, 50, 100, 200, or 500.

To further reduce selection effects, in our baseline results we consider workers with at least 3 years of tenure in the establishment prior to the mass layoff event in the establishment prior to the mass layoff event. However, we experimented with other tenure requirements. In particular, we considered at least 12 months or at least 24 months of tenure in the establishment prior to separation, obtaining very similar results across different specifications. We also experimented between using only full-time employment spells or pooling together full-time and part-time employment spells, but this makes little difference to our estimates due to the small proportion of part-time workers in our sample. We also impose that workers should be re-employment at most after 36 months after displacement. The window during which displacements are recorded is 1981–2005, with 3 pre-separation periods, and 15 post-separation periods. We find that there are 815 instances of mass layoffs among the low-educated groups of workers, 2,975 instances among the medium educated group, and 163 instances among the higher educated group. As standard in the literature, we use all workers who did not lose their jobs as our control group.

Following [Davis and Von Wachter \(2011\)](#), we estimate equation (26) in the main text separately for different value of  $y$ , which implies that for each coefficient  $\varepsilon_t^w$  in (26) we obtain a distribution of its estimates from which we take the average. Alternatively, we estimate equation (26) by pooling all the years. The results obtained using these two methods are very similar. The advantage of the latter is that we can report the standard errors for each coefficient. Table B.1 reports these latter results for the pooled sample (all workers) as well as for each of the three individual education groups when using (log) wages. For this exercise, annual wages are constructed by averaging the real daily wages reported during months in which the establishment identifier is present (months of employment) in a given year  $y$ . Table B.2 instead reports the estimated coefficient when estimating (26) on earnings by individual education group. In this case, earnings are created by adding the zero wages during the months in which the establishment identifier was missing (months of unemployment) to the real daily wage when the establishment identifier is present (months of employment). Annual earnings in year  $y$  are constructed as the mean of earnings across all months in year  $y$ , including the zeros.

As discussed in the main text, there is a potential tension between the way we estimate the model, where all employment-to-unemployment (EU) transitions are used to identify  $\delta$ , and the analysis referring to the long-term earnings/wage losses of displaced workers, where we use mass layoffs to perform to identify displacement. Here, we show that this tension appears to be small as estimating equation (26) using either mass layoff events or all EU transitions yields similar results. Figure B.1 depicts the earnings losses of the low and medium education groups using either all EU transitions or mass-layoffs; while Figure B.2 depicts the (log) wage loss of these education groups when using either all EU transitions or mass-layoffs. These graphs show that losses due to mass layoffs are slightly larger than when considering all EU separations, particularly for wages, but both follow very similar patterns. [Jarosh \(2015\)](#) using the same data (but a different sample) obtains a similar conclusion.

### B.4. Data construction

Since nearly all of our data work arises from the SIAB/BHP, we start by describing the main features in the construction of our SIAB/BHP sample. We merge the individual files with the establishment files on the establishment identifier and

TABLE B.1  
Log wage losses, SIAB/BHP

log real wage	All workers		Low educated		Medium educated		Higher educated	
	Coef. (%)	Std. Err.	Coef. (%)	Std. Err.	Coef. (%)	Std. Err.	Coef. (%)	Std. Err.
$\varepsilon_{-3}$	-1.75***	0.004	0.48	0.011	-1.68	0.005	0.75	0.014
$\varepsilon_{-2}$	-0.06	0.004	1.18	0.012	0.14	0.005	-0.26	0.015
$\varepsilon_{-1}$	0.70	0.004	1.56	0.012	0.88*	0.005	0.69	0.014
$\varepsilon_0$	-6.59***	0.005	-5.98***	0.012	-6.07***	0.005	-4.02**	0.016
$\varepsilon_1$	-11.34***	0.005	-8.10***	0.013	-11.67***	0.005	-3.61**	0.018
$\varepsilon_2$	-9.80***	0.005	-5.92***	0.014	-9.55***	0.006	-4.00**	0.016
$\varepsilon_3$	-9.70***	0.005	-7.33***	0.015	-9.45***	0.006	-1.49	0.016
$\varepsilon_4$	-8.56***	0.005	-4.73***	0.016	-8.23***	0.006	-2.62	0.017
$\varepsilon_5$	-8.95***	0.005	-5.14***	0.017	-8.60***	0.006	-3.19	0.018
$\varepsilon_6$	-7.82***	0.006	-3.99**	0.018	-7.47***	0.007	-2.22	0.019
$\varepsilon_7$	-7.28***	0.006	-1.94	0.020	-7.48***	0.007	-2.56	0.020
$\varepsilon_8$	-7.35***	0.006	-2.89	0.021	-7.12***	0.007	-1.95	0.021
$\varepsilon_9$	-7.07***	0.007	-4.48***	0.022	-6.42***	0.008	-0.75	0.023
$\varepsilon_{10}$	-6.31***	0.007	-5.39***	0.024	-5.32***	0.008	0.08	0.024
$\varepsilon_{11}$	-5.62***	0.008	-0.46	0.026	-5.32***	0.009	0.41	0.027
$\varepsilon_{12}$	-5.11***	0.008	-3.49	0.029	-4.47***	0.009	-0.07	0.028
$\varepsilon_{13}$	-6.49***	0.009	-2.49	0.032	-6.29***	0.010	-2.02	0.030
$\varepsilon_{14}$	-5.18***	0.009	-1.42	0.035	-4.38***	0.011	-10.02***	0.033
$\varepsilon_{15}$	-5.59***	0.010	-6.86*	0.038	-4.66***	0.012	-4.23	0.040

Notes: \*Significant at a 10%, \*\*significant at a 5%, \*\*\*significant at a 1%.

TABLE B.2  
Earnings losses, SIAB/BHP

Real earnings	All workers		Low educated		Medium educated		Higher educated	
	Coef. (%)	Std. Err.	Coef. (%)	Std. Err.	Coef. (%)	Std. Err.	Coef. (%)	Std. Err.
$\varepsilon_{-3}$	-0.646	0.444	2.034*	0.912	-0.424	0.515	3.829*	2.191
$\varepsilon_{-2}$	-0.364	0.445	1.166	0.920	0.010	0.517	3.541	2.201
$\varepsilon_{-1}$	0.812*	0.447	2.238*	0.926	1.181**	0.520	3.156	2.191
$\varepsilon_0$	-36.006***	0.464	-24.382***	0.927	-35.607***	0.537	-54.738***	2.405
$\varepsilon_1$	-16.721***	0.480	-7.714***	1.010	-17.053***	0.560	-17.265***	2.679
$\varepsilon_2$	-13.753***	0.492	-6.865***	1.109	-13.488***	0.584	-16.748***	2.449
$\varepsilon_3$	-12.410***	0.502	-6.270***	1.188	-12.612***	0.597	-2.219	2.495
$\varepsilon_4$	-10.895***	0.519	-4.531***	1.280	-10.628***	0.617	-5.956**	2.559
$\varepsilon_5$	-9.982***	0.535	-3.841***	1.359	-9.983***	0.635	-4.917*	2.680
$\varepsilon_6$	-9.353***	0.565	-5.034***	1.460	-9.008***	0.672	-4.328	2.895
$\varepsilon_7$	-8.794***	0.597	-3.838*	1.582	-8.915***	0.709	-3.541	2.994
$\varepsilon_8$	-8.577***	0.634	-3.532**	1.661	-8.558***	0.756	-3.153	3.154
$\varepsilon_9$	-8.039***	0.673	-6.807***	1.770	-7.375***	0.805	-2.297	3.458
$\varepsilon_{10}$	-8.138***	0.709	-5.020***	1.929	-6.996***	0.845	-1.535	3.577
$\varepsilon_{11}$	-7.702***	0.757	-4.402**	2.101	-6.790***	0.900	-1.691	4.052
$\varepsilon_{12}$	-6.803***	0.806	-1.423	2.306	-6.541***	0.962	-1.273	4.185
$\varepsilon_{13}$	-7.450***	0.862	-0.072	2.522	-7.797***	1.028	-3.935	4.595
$\varepsilon_{14}$	-7.560***	0.951	-3.548	2.789	-6.339***	1.137	-10.544**	5.039
$\varepsilon_{15}$	-7.873***	1.038	-7.216***	3.034	-7.320***	1.236	-8.674	6.051
Constant	99.366***	0.816	51.040***	0.801	99.256***	0.705	115.716***	2.538

Notes: \*Significant at a 10%, \*\*significant at a 5%, \*\*\*significant at a 1%.

year. Then using adjusted beginning and end of spell date we compute the months when the spell started and ended.<sup>25</sup> As there can be many spells in the same month, especially when a worker moves between two labour market states, we

25. We have to adjust the beginning and end of a spell as in the original data spells are split in such a way that they do not overlap; however, more than one spell in any given period is possible.

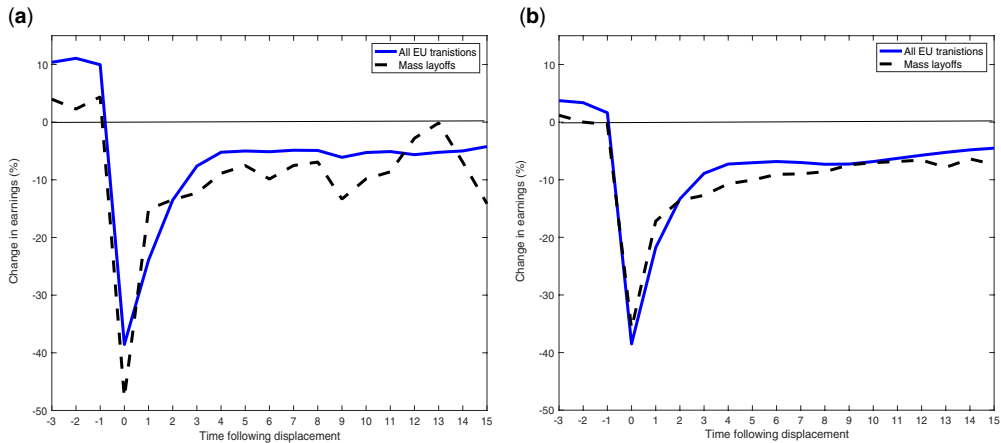


FIGURE B.1  
Post Displacement Earnings Losses - All EU transitions and Mass-layoffs

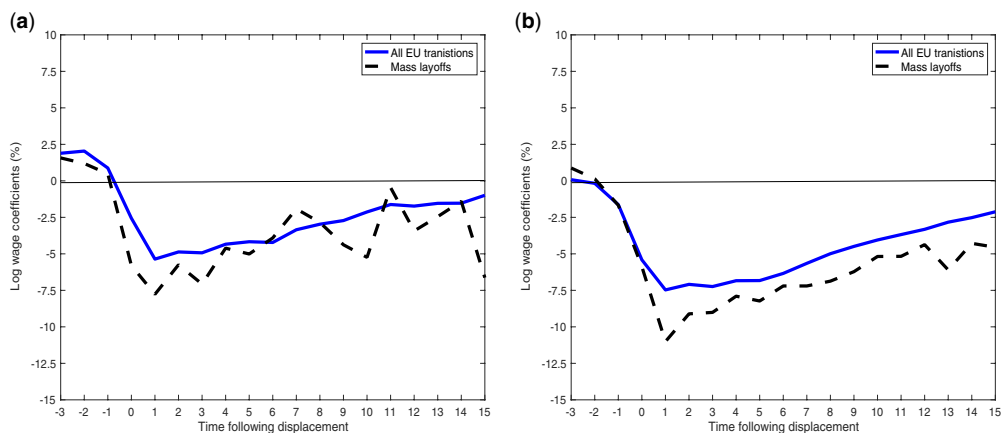


FIGURE B.2  
Post displacement wage losses—all EU transitions and mass-layoffs. (a) Low educated. (b) Medium educated.

calculate, for months at the beginning and end of a spell, how many days are part of the spell (on each “side”), assuming each month has 31 days. If there are two spells with the same individual identifier and monthly date, we assign a “repeated spell” dummy variable to the second and following spells. If the repeated spell is shorter (in days) than the previous one, we assign it a dummy “repeated spell short” that is equal to one. If the repeated spell has daily wage lower than the previous spell, we assign the dummy “repeated spell low wages” to be equal to one. If the repeated spell starts on the same daily date as the previous spell, and lasts for the same number of months, we assign a “repeated spell same time” dummy and set it equal to one. We then use the duration of spells in months to construct a monthly panel. To determine which observations to keep, first we drop repeated spells which started on the same date and have the same duration in months, and lower daily wages than the previous spell. If two spells still coexist in a given month, we drop the repeated one with shorter duration.

The above procedure allows us to uniquely identify observations by person number and monthly date combination. Place of work is used to split individuals between East and West Germany, where we only keep the latter. This classification is based on the federal state in which individual work. We classify Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt, and Thuringen as “East” and the remaining federal states are classified as “West.” Observations not assigned to either region are dropped. Further, spells from the data sources Benefit Recipient History File, Job-Search

History File, Unemployment Benefit II Recipient History File, Job-Search History File XASU, and Participants-in-Measures History File are dropped, such that only the spells coming from the Employee History File (BeH) are kept. We also drop all female and/or foreign workers from the sample. Only people between the ages of 18 and 35 at the time of the first observation are kept, where age is generated as the difference between current year and year of birth. We also keep those spells that are liable to social security, which implies that spells of trainees, marginal part-time workers, employees in partial retirement, interns and student trainees, and workers having “other” employment status are dropped. Additionally soldiers, border guards, police, and other related professions as well as members of parliament, ministers, and civil servants are also dropped.

To construct the education groups, we create eight education groups based on combinations of formal schooling and vocational qualifications. Group 1: No school leaving certificate and no vocational training. Group 2: Up to intermediate school leaving certificate and no vocational training. Group 3: No school leaving certificate but vocational training. Group 4: Up to intermediate school leaving certificate plus vocational training. Group 5: Up to upper secondary school leaving certificate but no vocational training. Group 6: Up to upper secondary school leaving certificate but no vocational training plus vocational training. Group 7: University of applied sciences. Group 8: University. We label Groups 1, 2, and 5 as “Low education”; Groups 3, 4, and 6 are labelled “Medium education”; and Groups 7 and 8 are labelled “Higher education.” This implies that the low education group have up to secondary school leaving certificate (13 years of schooling) but no vocational qualifications. The medium education group also have up to secondary school leaving certificate (13 years of schooling) but additionally hold some kind of vocational qualifications (in company vocational training/external vocational training/technical school vocational training/technical school advanced vocational training. The higher education group have a university of applied sciences or university degree.

Employment spells are obtained as the number of consecutive months of employment, defined as observations not missing establishment identifier. In case of more than one spell in a month, the spell with higher daily wages would take precedence, so gaps up to one month would not break an employment spell. If the whole month is missing, it will be considered an unemployment spell. The average employment spell is calculated as an average of the average length of employment spells for each person. Job spells are obtained as a number of consecutive months in which an individual was employed and the establishment identifier remained constant. Otherwise they are constructed in the same way as the employment spells. Unemployment spells are the consecutive months for an individual, in which the establishment identifier is missing. Employer-to-employer transitions are recorded when there is a change in establishment identifier, but worker remains employed. Employment-to-unemployment (unemployment-to-employment) transitions are recorded when worker moves from employment (unemployment) to unemployment (employment).

In the case of the GSEOP, we construct a sample that is as close as possible to the SIAB sample described above. In these data, employer-to-employer transitions are defined as a transition from full-time employment to full-time employment when the reason for job change given by the worker reported either “job with new employer” or “company taken over.” A voluntary employer-to-employer transition is one where the worker gave one of the following reasons for the termination of the previous job: “own resignation,” “mutual agreement,” or “leave of absence.” An involuntary employer-to-employer transition is one where the worker reported “company shut down,” “dismissal,” or “temporary contract expired.”

## REFERENCES

- ABOWD, J., KRAMARZ, F. and MARGOLIS, D. N. (1999), “High Wage Workers and High Wage Firms”, *Econometrica*, **67**, 251–333.
- ADDA, J., DUSTMANN, C. and STEVENS, K. (2017), “The Career Costs of Children”, *Journal of Political Economy*, **125**, 293–337.
- ADDA, J., DUSTMANN, C., MEGHIR, C. *et al.* (2013), “Career Progression, Economic Downturns, and Skills” (NBER Working Paper Series, WP 18832, NBER, Cambridge, USA).
- ALTONJI, J. G. and SHAKOTKO, R. A. (1987), “Do Wages Rise with Job Seniority?”, *Review of Economic Studies*, **54**, 437–459.
- ANTONI, M., GANZER, A. and VON BERGE, P. (2016), “Sample of Integrated Labour Market Biographies” (FDZ-Datenreport, 04/2016, Bundesagentur für Arbeit, Germany).
- BAGGER, J. and LENTZ, R. (2019), “An Empirical Model of Wage Dispersion with Sorting”, *Review of Economic Studies*, **86**, 153–190.
- BAGGER, J., FONTAINE, F., POSTEL-VINAY, F. *et al.* (2014), “A Feasible Equilibrium Search Model of Individual Wage Dynamics with Human Capital Accumulation”, *American Economic Review*, **104**, 1551–1596.
- BUETTNER, T. and RAESSLER, S. (2008), “Multiple Imputation of Right-Censored Wages in the German IAB Employment Sample Considering Heteroscedasticity” (IAB-Discussion Paper, No. 44/2008).
- BURDETT, K. (1978), “A Theory of Employee Search and Quits”, *American Economic Review*, **68**, 212–220.
- BURDETT, K., CARRILLO-TUDELA, C. and COLES, M. G. (2016), “Wage Inequality: A Structural Approach”, *Review of Economic Dynamics*, **16**, 20–37.
- BURDETT, K. and COLES, M. G. (2003), “Equilibrium Wage-Tenure Contracts”, *Econometrica*, **71**, 1377–1404.

- BURDETT, K. and MORTENSEN, D. T. (1998), "Wage Differentials, Employer Size, and Unemployment", *International Economic Review*, **39**, 257–273.
- CAHUC, P., POSTEL-VINAY, F. and ROBIN, J.-M. (2006), "Wage Bargaining with On-the-job Search: Theory and Evidence", *Econometrica*, **74**, 323–364.
- CARD, D., HEINING, J. and KLINE, P. (2013), "Workplace Heterogeneity and the Rise of West German Inequality", *Quarterly Journal of Economics*, **128**, 967–1015.
- CARRILLO-TUDELA, C., GARTNER, H., KAAS, L. *et al.* (2019), "Sorting Through Search Channels" (Mimeo, University of Essex).
- CARRILLO-TUDELA, C., LAUNOV, A. and ROBIN, J.-M. (2018), "The Fall in German Unemployment: A Flow Approach" (CEPR, Discussion Papers, DP12846).
- CARRILLO-TUDELA, C., (2009), "An Equilibrium Search Model with Optimal Wage-Experience Contracts", *Review of Economic Dynamics*, **12**, 108–128.
- COUCH, K. A. and PLACZEK, D. W. (2010), "Earnings Losses of Displaced Workers Revisited," *American Economic Review*, **100**, 572–89.
- DAVIS, S. J. and VON WACHTER, T. (2011), "Recessions and the Costs of Job Loss", *Brookings Papers on Economic Activity*, **43**, 1–72.
- DEY, M. S., and FLINN, C. J. (2005), "An Equilibrium Model of Health Insurance Provision and Wage Determination", *Econometrica*, **73**, 571–627.
- DUSTMANN, C. and MEGHIR, C. (2005), "Wages, Experience and Seniority", *Review of Economic Studies*, **72**, 77–108.
- FLAAEN, A., SHAPIRO, M. D. and SORKIN, I. (2019), "Reconsidering the Consequences of Worker Displacements: Firm versus Worker Perspective", *American Economic Journal: Macroeconomics*, **11**, 193–227.
- GANDELMAN, N. and HERNANDEZ-MURILLO, R. (2015), "Risk Aversion at the Country Level", *Federal Reserve Bank of St. Louis Review*, **97**, 53–66.
- GOURIEROUX, C., MONFORT, A. and RENAULT, E. (1993), "Indirect Inference", *Journal of Applied Econometrics*, **8**, 885–118.
- HOPENHAYN, H. A. and NICOLINI, J. P. (1997), "Optimal Unemployment Insurance," *Journal of Political Economy*, **105**, 412–438.
- HORNSTEIN, A., KRUSELL, P. and VIOLANTE, G. L. (2011), "Frictional Wage Dispersion in Search Models: A Quantitative Assessment", *American Economic Review*, **101**, 2873–2898.
- HORNSTEIN, A., KRUSELL, P. and VIOLANTE, G. L. (2007), "Frictional Wage Dispersion in Search Models: A Quantitative Assessment" (Mimeo, Department of Economics, NYU).
- JACOBSON, L., LA LONDE, R. and SULLIVAN, D. (1993), "Earnings Losses of Displaced Workers", *American Economic Review*, **83**, 685–709.
- JAROSH, G. (2015), "Searching for Job Security and the Consequences of Job Loss" (Mimeo, University of Chicago, USA).
- JUNG, P. and KUHN, M. (2019), "Earnings Losses and Labor Mobility Over The Lifecycle". *Journal of the European Economic Association*, **17**, 678–724.
- KREBS, T. and SCHEFFEL, M. (2013), "Macroeconomic Evaluation of Labor Market Reform in Germany" (IMF Working Paper, WP/13/42).
- KROLIKOWSKI, P. (2017), "Job Ladders and Earnings of Displaced Workers", *American Economic Journal: Macroeconomics*, **9**, 1–31.
- LAUNOV, A. and WEALDE, K. (2013), "Estimating Incentive and Welfare Effects of Non-stationary Unemployment Benefits", *International Economic Review*, **54**, 1159–1198.
- LEONARD, D. and LONG, N. V. (1992), *Optimal Control Theory and Static Optimization in Economics* (Cambridge, UK: Cambridge University Press).
- LISE, J. and ROBIN, J.-M. (2017), "The Macro-dynamics of Sorting between Workers and Firms", *American Economic Review*, **107**, 1104–1135.
- LJUNGQVIST, L. and SARGENT, T. (1998), "The European Unemployment Dilemma", *Journal of Political Economy*, **106**, 514–550.
- MORTENSEN, D. T. (2003), *Wage Dispersion: Why are Similar Workers Paid Differently?* (Cambridge, MA, USA: MIT Press).
- MOSCARINI, G. and POSTEL-VINAY, F. (2013), "Stochastic Search Equilibrium", *Review of Economic Studies*, **80**, 1545–1581.
- ORTEGO-MARTI, V. (2016), "Unemployment History and Frictional Wage Dispersion", *Journal of Monetary Economics*, **78**, 5–22.
- PINHEIRO, R. and VISSCHERS, L. (2015), "Unemployment Risk and Wage Differentials", *Journal of Economic Theory*, **157**, 396–424.
- PISSARIDES, C. A. (1992), "Loss of Skill During Unemployment and the Persistence of Employment Shocks", *Quarterly Journal of Economics*, **107**, 1371–1391.
- POSTEL-VINAY, F. and ROBIN, J.-M. (2002), "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity", *Econometrica*, **70**, 2295–2350.
- ROBIN, J.-M. (2011), "On the Dynamics of Unemployment and Wage Distributions", *Econometrica*, **79**, 1327–1355.
- RUHM, C. J. (1991), "Are Workers Permanently Scared by Job Displacements?", *American Economic Review*, **81**, 319–324.

- SCHMIEDER, J. F., VON WACHTER, T. and BENDER, S. (2016), "The Effect of Unemployment Benefits and Nonemployment Durations on Wages", *American Economic Review*, **106**, 739–777.
- SIANESI, B. (2004), "An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s", *Review of Economic and Statistics*, **86**, 133–155.
- SHAVELL, S. and WEISS, L. (1979), "The Optimal Payment of Unemployment Insurance Benefits over Time", *Journal of Political Economy*, **87**, 1347–1362.
- SHIMER, R. and WERNING, I. (2008), "Liquidity and Insurance for the Unemployed", *American Economic Review*, **98**, 1922–1942.
- SPEAR, S. E. and SRIVASTAVA, S. (1987), "On Repeated Moral Hazard with Discounting", *The Review of Economic Studies*, **54**, 599–617.
- STEVENS, A. H. (1997), "Persistent Effects of Job Displacement: The Importance of Multiple Job Losses", *Journal of Labor Economics*, **15**, 165–188.
- STEVENS, M. (2004), "Wage-Tenure Contracts in a Frictional Labour Market: Firms' Strategies for Recruitment and Retention", *Review of Economic Studies*, **71**, 535–551.
- TOPEL, R. (1991), "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority", *Journal of Political Economy*, **99**, 145–176.
- TOPEL, R. (1990), "Specific Capital and Unemployment: Measuring the Costs and Consequences of Job Loss", *Carnegie-Rochester Conference Series on Public Policy*, **33**, 181–214.
- WOODFORD, M. (2003), *Interest and Prices: Foundations of a Theory of Monetary Policy* (Princeton, USA: Princeton University Press).