Labor Market Frictions, Human Capital Accumulation, and Consumption Inequality

Michael Graber[*] Jeremy Lise[†]

December 2015

Preliminary and Incomplete[‡]

Abstract

We develop a frictional model of the labor market with stochastic human capital accumulation and incomplete markets. The stochastic process for human capital may be heterogeneous across workers as well as depend on the firm type where the worker is employed. We establish nonparametric identification of the model, and estimate it using matched employer-employee data from Germany. We provide a decomposition of lifecycle inequality in earnings and consumption resulting from heterogeneity in ability, heterogeneity in the rate of human capital accumulation, search frictions (the random allocation of similar workers to different jobs) and the interaction of human capital and search friction (a worker's history of job opportunities may differentially affect her human capital accumulation opportunities).

[*]University College London; Institute for Fiscal Studies. Email uctpmrg@ucl.ac.uk

[†]University College London. Email j.lise@ucl.ac.uk

[‡]This draft was prepared as a discussion paper for the Human Capital and Economic Opportunity conference on Human Capital and Inequality, December 15–16, 2015 at the University of Chicago. Please do not cite without permission from the authors. Michael Graber gratefully acknowledges financial support from the European Union's 7th Framework Programme (FP7/2007-2013) under grant agreement no. 262608.

1 Introduction

Two well established facts in the literature which analyses the lifecycle patterns of income and consumption is that both the variance of log wages and the variance of log consumption increase approximately linearly with age, with the slope for consumption rising less than for wages or earnings. In their review, Meghir and Pistaferri (2011) argue this is a robust feature of the data that any model must confront. The linear rise in the variance of log wages is consistent with both the accumulation of substantial permanent shocks to wages (MaCurdy, 1982) or heterogeneity across workers in the rate of human capital accumulation (Lillard and Weiss, 1979). Within the lifecycle permanent income hypothesis model, the linear rise in the variance of log consumption requires substantial permanent uncertainty on the part of the workers. This is consistent with substantial permanent shocks to wages but not with heterogeneous income growth across workers, unless workers do not know their growth rates and must learn about them, leading them to behave as if they faces substantial permanent uncertainty (Guvenen, 2007).

Recently, Arellano, Blundell, and Bonhomme (2015) and Guvenen, Karahan, Ozkan, and Song (2015) have highlighted that year over year changes in earnings are far from symmetric and are certainly far from Normally distributed. Of particular interest for our work is the fact that wage or earnings changes display substantial negative skewness and substantial kurtosis. In other words, wage changes tend to be very small on average and the large changes tend to be negative. In addition, Guvenen et al. (2015) find that the skewness of year-over-year changes becomes more negative as they condition on higher quantiles of the earnings distribution. This increasing negative skewness, as the authors of both papers highlight, is consistent with the earnings dynamics that come from a standard job ladder model.

The asymmetry of earnings shocks implied by the job ladder model, and the implications for consumption and savings decisions has previously been explored by Lise (2013). He finds that this increasing negative skewness of wage shocks implies that workers at the top of the job ladder increase their savings rate substantially to insure against the increasing down-side risk of falling off the job ladder. The model in Lise (2013) does well in replicating the cross sectional dispersion in wages, assets and consumption, however, it is inconsistent with the well established linear increase in log wages and consumption. The job ladder model, combined with the empirical rates at which workers change jobs, implies that a the distribution of wages for a cohort of workers will look stationary after roughly 10 years, implying that the variance of wage becomes flat after this period as does the variance of consumption. The basic job ladder model implies far too much stationarity to be consistent with the facts about the age profile of the variance of wages or consumption.

In this paper we develop a model in which inequality in earnings and consumption results from the interaction between heterogeneity of workers, heterogeneity of jobs, shocks to human capital, and labor market frictions. Specifically, we allow for permanent differences across individuals in productivity, the ability to acquire human capital, as well as the ability to find and keep jobs. We consider jobs which differ in both in productivity and the extent to which they facilitate further human capital accumulation for workers. Shocks to human capital are permanent, they are carried by the worker even when she changes jobs. Labor market frictions induce transitory shocks as workers stochastically move up the job ladder. They are transitory in the sense that the direct productivity effect of the job on the worker's wage disappears when the worker leaves the job, either through a spell of unemployment or by a move directly to another job. However, depending on the extent to which jobs differ in the degree to which they facilitate human capital accumulation, and the severity of frictions, there may be permanent effects on the workers human capital and hence earnings arising from differential labor market histories.

Exiting work combining job search and human capital accumulation includes Bunzel, Christensen, Kiefer, and Korsholm (1999); Rubinstein and Weiss (2006); Barlevy (2008); Yamaguchi (2010); Burdett, Carrillo-Tudela, and Coles (2011); Veramendi (2011); Bowlus and Liu (2013) and Bagger, Fontaine, Postel-Vinay, and Robin (2014). While some of these models include features not included here, they all focus on the contribution of human capital accumulation and job search (and in some cases learning about own ability) for explaining the age profile of mean wages. None of these papers has a model of consumption and savings, and none focus on the facts about the linear increase in the variance of log wages and consumption with age or the asymmetry of wage changes discussed above.

Our model is closest in spirit to Bagger, Fontaine, Postel-Vinay, and Robin (2014), and as such the model touches on many of the same questions in the literature such as the effect of experience verses job tenure for wage growth (see, for example, Abraham and Farber, 1987; Altonji and Shakotko, 1987; Topel, 1991; Buchinsky, Fougère, Kramarz, and Tchernis, 2010); on earnings dynamics, a topic for which there is a very large existing literature (see, for example, MaCurdy, 1982; Abowd and Card, 1989; Topel and Ward, 1992; Meghir and Pistaferri, 2004; Browning, Ejrnaes, and Alvarez, 2010; Altonji, Smith, and Vidangos, 2013); as well as the recent literature on the asymmetry of wage changes (Arellano, Blundell, and Bonhomme, 2015; Guvenen, Karahan, Ozkan, and Song, 2015).

The model is able to reproduce the linear age profile for the variance of log wages and consumption, as well as the negative skewness and excessive kurtosis of wage changes, including the increasing negative skewness of wage changes conditional on the previous wage level. We decompose the age profile for the variance of log wage and consumption, as well as conditional skewness and kurtosis of wage changes into that due to fixed worker heterogeneity, in both levels and growth rates, of human capital; the effect of firms via the job ladder; and the stochastic shocks to human capital. We find that 77 to 80 percent of the rise in the variances can be accounted for by permanent shocks to human capital, with the uncertainty created by the job ladder contributing the remaining 20 to 23 percent. Heterogeneity across workers in the productivity levels contributes to the level of dispersion, but not to the increase. Heterogeneity across workers in the growth rate of human capital accounts for a small fraction of the increase in the variance for wages, but none of the increase in the variance for consumption. On the other hand, firm heterogeneity manifested through the job ladder accounts for the entirety of the negative skewness and excess kurtosis in the conditional wage changes; the contribution of fixed worker heterogeneity and shocks to human capital is negligible.

The paper proceeds as follows: we present the model in Section 2; we describe the matched employer-employee data in Section 3; we establish non-parametric identification of the model in Section 4; we present estimates based on the non-parametric estimator in Section 5; In Section 6 we decompose the increase in the variance of earnings and consumption as a cohort ages into the components due to heterogeneity, job search frictions and their interaction. Section 7 concludes.

2 The Model

2.1 The Environment

Agents. Time is continuous and indexed by t. The economy is populated by a unit mass of workers characterized by their human capital $\mathbf{h}_t = (h_0, h_{1t})$. The first component $h_0 \in \mathbb{R}^+$ is a constant worker heterogeneity parameter that reflects permanent differences in ability between workers. The second component $h_{1t} \in \mathbb{R}^+$ is a worker's accumulated human capital at time t. We denote by $L(h_0)$ the exogenous measure of workers with a type weakly below h_0. Workers are risk averse and maximize expected lifetime utility, ordering consumption paths according to

$$\mathbb{E}_0 \int_0^\infty e^{\rho t} u(c_t) \mathrm{d}t,$$

where ρ is the subjective rate of time preference and c_t is instantaneous consumption at time t. There is a single riskless asset a with interest rate r which allows workers to transfer resources over time. The change of wealth over time is then given by the return on assets,

ra_t plus income, i_t less consumption expenditure,

$$\mathrm{d}a_t = (ra_t + i_t - c_t)\mathrm{d}t \quad \text{subject to } a_t \geq \underline{a},$$

where $\underline{a}$ is the lower bound on assets. Workers can either be employed and earning an endogenous wage, unemployed and receiving benefits, or out of the the labour force and receiving no income. Workers leave the labour force at exogenous rate ξ and are replaced by an equal mass of unemployed labour market entrants.

On the other side of the labor market there is a continuum of firms indexed by their productivity type $y \in [\underline{y}, \overline{y}] \subset \mathbb{R}^+$. The measure of firms with productivity weakly below y is exogenous and denoted by $\Gamma(y)$. Firms are risk-neutral and maximize the present value of profits.

Matching and Production. Workers search for jobs when unemployed and for better jobs when they are employed. Search is random and all workers sequentially sample from the same exogenous firm type distribution $\Gamma(y)$ at type-dependent Poisson rate $\lambda(h_0)$. A match between a firm of type y and a worker with skill $\mathbf{h}_t$ produces output $f(\mathbf{h}_t, y)$ given by,

$$f(\mathbf{h}_t, y) = h_0 h_{1t} y.$$

A match becomes unprofitable when it is hit by an adverse idiosyncratic productivity shock. The Poisson rate at which these shocks occur depends on the fixed worker type h_0 and is denoted by $\delta(h_0)$.

When unmatched, workers produce home production $b(\mathbf{h}) = h_0 h_{1t} b$. Given the assumed proportionality of market and home production to a worker's human capital, no unemployed worker would ever accept a job from a firm with productivity below $y < b$, and thus the support of the effective sampling distribution is $\underline{y} = b$.

Human Capital Accumulation. On the worker side, a match between a firm y and a worker $\mathbf{h}$ contributes to human capital development. Human capital h_{1t} evolves stochastically over time according to a diffusion process on a bounded interval $[\underline{h}_1, \overline{h}_1] \subset \mathbb{R}^+$ with reflecting barriers:

$$\mathrm{d}h_{1t} = \mu(\mathbf{h}_t, y)\mathrm{d}t + \sigma(\mathbf{h}_t)\mathrm{d}\mathcal{B}_t, \tag{1}$$

where $d\mathcal{B}_t \equiv \lim_{\Delta t \to 0} \sqrt{\Delta t}\varepsilon_t$ with $\varepsilon_t \sim \mathcal{N}(0,1)$ is the increment of a standard Brownian motion with $\sigma(\mathbf{h}) \geq 0$ governing the diffusion of the process. The drift rate $\mu(.,.)$ is assumed to depend on the worker type and the firm type and is a continuous function assumed to be non-decreasing in the firm type, i.e. $\frac{\partial}{\partial y}\mu(\mathbf{h}_t, y) \geq 0$.[1]

2.2 Wage Contracts

We will restrict our attention to wage contracts that are of the piece rate form, implying that at any point in time t the wage paid to the worker is restricted to be a share of the total output of the worker-firm match. Let $0 \leq \theta_t \leq 1$ denote the contractual piece-rate at time t. A worker $\mathbf{h}_t$ employed at firm y then receives a wage given by

$$w_t = \theta_t f(\mathbf{h}_t, y) = \theta_t h_0 h_{1t} y.$$

Next consider the very same worker receiving an outside offer from a firm of type y'. The incumbent firm y and the firm y' Bertrand compete over the worker's services similar to the piece rate version of the sequential auction framework of Postel-Vinay and Robin (2002) implemented in Bagger, Fontaine, Postel-Vinay, and Robin (2014). Since the human capital accumulation rate is assumed to be non-decreasing in the firm type, the worker will move to firm y' whenever $y' > y$. The most the current firm can offer to the worker is a wage that equals the total output of the current match, equivalent to a piece-rate equal to one. We assume that the firm y' is willing to offer the worker a contract that matches this wage offer.[2] On the other hand a worker at firm y who is being poached by a less productive firm $y' \leq y$ stays with the current employer. The worker's piece-rate only gets updated if the outside firm is productive enough to be able to offer the worker a higher wage, all other things being equal. Specifically this implies that the piece-rate offered to the worker receiving an outside offer from firm y' at time $\tau > t$ is given by

$$\theta_\tau = \begin{cases} \frac{y}{y'} & \text{if } y' \geq y \\ \max\left\{\theta_t, \frac{y'}{y}\right\} & \text{if } y' < y \end{cases}.$$

Some discussion of our choice of wage setting is warranted. The main benefit of our chosen wage determination is that identification and interpretation is transparent. Theoretically, our wage setting process satisfies individual rationality since it guarantees that match formation and separations are efficient and that the participation constraint of both the worker and firm

[1] We treat the relationship between the speed of human capital accumulation and the productivity of a job as an exogenous technology to be estimated. Recently, in a closely related model, Lentz and Roys (2015) show that with heterogeneous firms the optimal employment contract features more productive firms providing more training.

[2] Note here we depart from Bagger, Fontaine, Postel-Vinay, and Robin (2014) since we compare flow output as opposed to present values to determine the updated piece-rate.

are always satisfied at the prevailing piece rate contract. However, our restricted contracts are unlikely to be optimal for at least two reasons. First, since workers are risk averse firms may be able to extract more of the match surplus by offering a contract that is smoother than that implied by the piece rate.[3] The extent to which the worker values this additional smoothness would depend on how close she is to the borrowing constraint, implying that any such contract would need to condition on the worker's asset position. Second, in the case where a firm with higher productivity also provides faster human capital accumulation, it could earn higher profit by offering a wage lower than than implied by the piece-rate as the worker will also value the higher rate of human capital accumulation. Again, how much the worker values wages to day verses more human capital tomorrow will depend on her asset position, which the firm would need to know to make the profit maximizing offer. Wage contracts that condition on the asset position of a worker strike us as a poor representation of reality. While we do give up on optimality, our piece-rate contracts ensure wages reflect worker productivity, firm productivity, and the competition for workers' services, and satisfy all participation constraints. We also have the additional benefit of transparent identification and interpretation of the model structure.

2.3 Worker Value

Let $W(a, \mathbf{h}, y, \theta)$ denote the value of employment for a worker of type $\mathbf{h}$ with assets a employed by a type-y firm under a contract that specifies a piece-rate θ. We assume that unemployment is equivalent to employment at the lowest productive firm when the worker receives the total output as wage, i.e. the value of unemployment is given by $W(a, \mathbf{h}, \underline{y}, 1)$. The Hamilton-Jacobi-Bellman equation can then be written as[4]

$$\begin{aligned}
[\rho + \lambda(h_0) + \delta(h_0) + \xi]W(a, \mathbf{h}, y, \theta) = \max_{a-\underline{a}\geq c\geq 0} \Bigg\{ & u(c) + \frac{\partial}{\partial a} W(a, \mathbf{h}, y, \theta)[ra + \theta f(\mathbf{h}, y) - c] \\
& + \mu(\mathbf{h}, y) \frac{\partial}{\partial h_1} W(a, \mathbf{h}, y, \theta) + \frac{\sigma(\mathbf{h})^2}{2} \frac{\partial^2}{\partial h_1^2} W(a, \mathbf{h}, y, \theta) + \delta(h_0) W(a, \mathbf{h}, \underline{y}, 1) + \xi R(a, \mathbf{h}) \\
& + \lambda(h_0) \int \max\{W(a, \mathbf{h}, y', \frac{y}{y'}), W(a, \mathbf{h}, y, \max\{\theta, \frac{y'}{y}\})\} d\Gamma(y') \Bigg\}. \quad (2)
\end{aligned}$$

[3]For example, Burdett and Coles (2003) characterize the optimal wage-tenure contract in a wage posting setting and Lamadon (2014) characterizes the optimal contract when search is directed, there are shocks to both worker and firm productivity. In both cases the optimal wage contract is substantially smoother than would be the case with a share of output. However, workers have no access to borrowing or savings technology so cannot do any consumption smoothing on their own, which is an important feature of our model.

[4]The derivation can be found in Appendix A.1.

The first-order necessary condition for optimal consumption choices states that the marginal utility from consumption equals the shadow price of wealth:

$$u'(c) = \frac{\partial}{\partial a} W(a, \mathbf{h}, y, \theta), \tag{3}$$

In addition, recall that the stochastic process for human capital is bounded by reflecting barriers. One can show that this gives rise to the following boundary conditions:

$$\frac{\partial}{\partial h_1} W(a, \underline{\mathbf{h}}, y, \theta) = \frac{\partial}{\partial h_1} W(a, \overline{\mathbf{h}}, y, \theta) = 0, \tag{4}$$

where $\underline{\mathbf{h}} = (h_0, \underline{h}_1)$ and $\overline{\mathbf{h}} = (h_0, \overline{h}_1)$.

When the worker makes a transition into retirement, he receives a constant flow of payments $b(\mathbf{h})$. The flow value of retirement is given by

$$\rho R(a, \mathbf{h}) = \max_{a-\underline{a} \geq c \geq 0} \{u(c) + \frac{\partial}{\partial a} R(a, \mathbf{h})[ra + b(\mathbf{h}) - c]\},$$

with first-order condition

$$u'(c) = \frac{\partial}{\partial a} R(a, \mathbf{h}).$$

2.4 Cross sectional distributions

The following flow equations define the stationary equilibrium distributions:

Unemployment Let the distribution of unemployed of type h_0 be $\mathrm{u}(h_0)$. In the stationary equilibrium,

$$[\lambda(h_0) + \xi]\mathrm{u}(h_0) = \delta(h_0)[1 - \mathrm{u}(h_0)] + \xi\ell(h_0),$$

where the flows out of unemployment come from transitions into employment and retirement, and the inflow comprises separations from employment and newly born workers. These flows produce the stationary unemployment rate by worker type:

$$\mathrm{u}(h_0) = \frac{\delta(h_0) + \xi}{\delta(h_0) + \xi + \lambda(h_0)}.$$

Employment The stationary flows in and out of employment are given by

$$[\delta(h_0) + \lambda(h_0)\overline{\Gamma}(y) + \xi]\,[1 - \mathrm{u}(h_0)]\,G(y|h_0) = \lambda(h_0)\Gamma(y)\mathrm{u}(h_0),$$

which, after substituting for u(h_0), defines the stationary relationship between the sampling distribution and the cross sectional distribution:

$$G(y|h_0) = \frac{[\delta(h_0) + \xi]\Gamma(y)}{\delta(h_0) + \xi + \lambda(h_0)\overline{\Gamma}(y)} \iff \Gamma(y) = \frac{[\delta(h_0) + \xi + \lambda(h_0)]G(y|h_0)}{\delta(h_0) + \xi + \lambda(h_0)G(y|h_0)}. \tag{5}$$

3 Data

We use data from two different sources. To estimate the employment and earnings dynamics, we use German administrative data provided by the Research Data Centre of the German Federal Employment Agency.[5] The data is a 2% random sample of all individuals in Germany which have been employed subject to social security (around 80% of the total workforce) between 1974 and 2010. It provides us with daily histories of earnings and benefits, and background information such as birth year, gender, nationality and education. The data also contains an establishment identifier which allows us to link characteristics on the establishment such as worker flows at an annual frequency using the Establishment History Panel. In the following we exclusively focus on Western Germany due to the additional regulations and restructuring of Eastern Germany after the reunification. We select a sample of male labour market entrants allowing us to construct precise measures of actual experience and firm tenure from labor market entry onwards. Our sample consists of 212,380 workers, which is split into three mutually exclusive skill groups. We define low-skilled workers (18.85%) as workers with less than 15 months of apprenticeship training and no higher education. Medium-skilled workers (55.84%) are workers who either have more than 15 months of apprenticeship training or a high-school degree, but no higher education. High-skilled workers (25.30%) are workers who received a degree from a technical college or university.[6]

The second data source we use is the Income and Expenditure Survey (EVS) collected by the Federal Statistical Office of Germany. The EVS is a repeated cross section data set that is carried out every 5 years available from 1978. [TBC]

4 Identification

This subsection lays out the formal identification arguments for the key distributions and parameters of the model.

[5] We use the weakly anonymous version of the Sample of Integrated Labour Market Biographies (SIAB7510).

[6] Further details on the construction of the sample can be found in Appendix B.1.

Worker type distribution $L(h_0)$. We normalize the location of the firm type distribution, $\underline{y} = 1$ and we normalize a worker's human capital $h_{1t_0} = 1$ at the time when she takes up her first job in t_0. This implies that the worker's skills upon taking up the first job is $\mathbf{h}_{t_0} = (h_0, 1)$. The worker's initial wage contract as specified by the model implies a piece-rate $\theta_{t_0} = y^{-1}$ when the worker is hired by firm y. Since unemployment is equivalent to working with the lowest productivity firm ($h_0 h_{1t} b = h_0 h_{1t} \underline{y}$), her initial wage is

$$w_{t_0} = h_0,$$

and our estimate of a worker type is given by $\hat{h}_{0i} = w_{it_0}$. Thus, the distribution of fixed-worker types is identified by the distribution of initial wages. A nonparametric estimate for the distribution of fixed worker types is then the empirical cdf

$$\hat{L}(h_0) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}(w_{it_0} \leq h_0).$$

Sampling distribution of firm types $\Gamma(y) = \Gamma(y(p))$. With matched employer-employee data we have a way to directly estimate the rank of firm types based on their ability to hire from other firms. Since all workers prefer higher y firms, the ability of a firm to poach workers from other firms is increasing in y.[7]

Let $p(y)$ be the rank, normalized to lie in [0,1], of the share of all hires that a firm of type y poaches from other firms:

$$p(y) = \text{rank } \left[\frac{\int \lambda(h_0)(1-u(h_0))G(h_0, y)\mathrm{d}h_0}{\int \lambda(h_0)u(h_0) + \lambda(h_0)(1-u(h_0))G(h_0, y)\mathrm{d}h_0} \right] \in [0, 1],$$

which is monotonically increasing in y. This relationship can be inverted allowing us to write $y(p)$, and thus $G(h_0, y(p)) = G_p(h_0, p)$.

We can estimate the type of each firm j by

$$\hat{p}_j = \text{rank} \left[\frac{\text{number of workers hired from other firms}}{\text{total number of workers hired}} \right] \to [0, 1].$$

A direct nonparametric estimate of the sampling distribution $\Gamma(y(p))$ is available from the distribution of jobs accepted out of unemployment.[8]

[7] This is a robust implication of the job ladder model that holds both in steady state (Burdett and Mortensen, 1998) and out of steady-state (Moscarini and Postel-Vinay, 2013). Bagger and Lentz (2014) use this monotonicity to help identify firm types in a related model without human capital accumulation where wages are not fully informative about how workers rank firms.

[8] An alternative estimator is available by first estimating the cross-sectional distribution of worker-types

Poisson arrival rates. Given estimates for the ranking of firms, worker fixed effects and a non-parametric estimate of the sampling distribution Γ, the Poisson arrival rates are identified from observed job durations. Consider the following sampling frame in which we include the first job spell upon entering the labor market for each worker in our sample to construct the following observation for each worker i:

$$\mathbf{x}_i = (h_{0i}, p_{it_0}, d_i, \tau_{EEi}, \tau_{EUi}, \tau_{ERi}, c_i),$$

where h_{0i} is the worker type, p_{it_0} is the firm type of this job, $d_i \leq T$ is the duration in the initial job spell before either the first transition or the end of the sample window, $c_i = 1$ if $d_i = T$ and zero otherwise, $\tau_{EEi} = 1$ an worker is observed making a job to job change, $\tau_{EUi} = 1$ if an worker makes a transition into unemployment, and $\tau_{ERi} = 1$ if a worker is observed transiting into retirement.

Conditional on worker and firm type (h_{0i}, p_{it_0}) the likelihood contributions for initial job spells are:

$$\begin{aligned}\ell(\mathbf{x}_i) = \gamma(p_{it_0}) \times [\delta(h_{0i}) + \xi + \lambda(h_{0i})\overline{\Gamma}(p_{it_0})]^{1-c_i} \\ \times\, \mathrm{e}^{-[\delta(h_{0i})+\xi+\lambda(h_{0i})\overline{\Gamma}(p_{it_0})]d_i} \times \left(\frac{\delta(h_{0i})}{\delta(h_{0i}) + \xi + \lambda(h_{0i})\overline{\Gamma}(p_{it_0})}\right)^{\tau_{EUi}} \\ \times \left(\frac{\lambda(h_{0i})\overline{\Gamma}(p_{it_0})}{\delta(h_{0i}) + \xi + \lambda(h_{0i})\overline{\Gamma}(p_{it_0})}\right)^{\tau_{EEi}} \times \left(\frac{\xi}{\delta(h_{0i}) + \xi + \lambda(h_{0i})\overline{\Gamma}(p_{it_0})}\right)^{\tau_{ERi}}.\end{aligned}$$

The likelihood is a function of the transition rate functions $\lambda(h_0)$ and $\delta(h_0)$ and the retirement rate ξ, where we plug in our first stage non-parametric estimate of the sampling distribution for Γ.[9]

across job types:

$$\hat{G}(h_0, y(p)) = \hat{G}(h_0, p) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}(h_{0i} \leq h_0) \times \mathbf{1}(p_i \leq p),$$

where p_i is the firm type a worker is employed at and N is the number of workers in a cross sectional sample, and the obvious estimator for $\hat{G}(y(p)|h_0)$. Using the equilibrium steady-state relationship between the sampling distribution and the cross sectional distribution (5) provides an estimate of the sampling distribution $\Gamma(y;\cdot)$ conditional on $\delta(h_0)$, $\lambda(h_0)$ and ξ:

$$\hat{\Gamma}(y; \delta(h_0), \lambda(h_0), \xi) = \frac{[\delta(h_0) + \xi + \lambda(h_0)]\hat{G}(y|h_0)}{\delta(h_0) + \xi + \lambda(h_0)\hat{G}(y|h_0)}.$$

We prefer to use the direct estimate of $\Gamma(y(p))$ based on jobs accepted out of unemployment as it does not impose the steady-state relationship (5), although it does, of course, assume Γ is constant over our sample period.

[9] Alternatively, one can directly estimate $\delta(h_0)$ as the employment to unemployment transition rate by worker type rate, and the type dependent unemployment rate $u(h_0)$ in the data. Then estimate $\lambda(h_0)$ using

Human capital accumulation and the contribution of firm type to wages In order to identify the deterministic and stochastic components of the human capital accumulation function, as well as the contribution to wages of firm type, we require wage observations on two consecutive job spells, plus information on where the worker transited from. An individual observation is given by

$$\mathbf{x}_{i,j,j',j''} = (h_{0i}, \log w^s_{ij_1}, \log w^s_{ij_2}, d_{ij_1}, p_{ij_1}, p_{ij_2}, p_{ij_0}),$$

where h_{0i} is the estimated fixed worker type; $w^s_{ij_1}$ and $w^s_{ij_2}$ are the starting wages for the worker at jobs j_1 and j_2; d_{ij_1} is the duration of the job spell at firm j_1; p_{ij_1} and p_{ij_2} are the estimated firm types for the jobs corresponding to starting wages $w^s_{ij_1}$ and $w^s_{ij_2}$; and p_{ij_0} is the firm type corresponding to the job held prior to p_{ij_1}.

The protocol for setting the piece-rate implies that the starting wages at jobs j_1 and j_2 are given by:

$$\begin{aligned}
\log w^s_{ij_1t} &= \log y(p_{ij_0}) + \log h_{0i} + \log h_{1it} \\
\log w^s_{ij_2t+d} &= \log y(p_{ij_1}) + \log h_{0i} + \log h_{1it+d},
\end{aligned}$$

where we note that the starting wage in job j_1 depends on the firm productivity of job j_0 and, similarly, the starting wage of job j_2 depends on the firm productivity of job j_1. Now, if we condition on $p_{ij_0} = p_{ij_1}$, the change in starting wages between job j_1 and j_2 reflects only human capital accumulated at job j_1:

$$\log w^s_{ij_2t+d} - \log w^s_{ij_1t} = \log h_{1it+d} - \log h_{1it}.$$

Suppose that human capital follows

$$\mathrm{d}h_{1t} = \mu(h_0, y)h_{1t}\mathrm{d}t + \sigma h_{1t}\mathrm{d}\mathcal{B}_t.$$

From Ito's lemma it follows that

$$\mathrm{d}\log h_{1t} = [\mu(h_0, y) - \frac{\sigma^2}{2}]\mathrm{d}t + \sigma\mathrm{d}\mathcal{B}_t.$$

Given this stochastic process for the evolution of $\log h_{1t}$, the difference in starting wages

the steady-state unemployment relationship $u(h_0) = \delta(h_0)/(\delta(h_0) + \lambda(h_0))$.

is Normally distributed with mean given by[10]

$$\mathbb{E}_t \left[\log w^s_{ij_2t+d} - \log w^s_{ij_1t} | p_{ij_0} = p_{ij_1}\right] = [\mu(h_{0i}, y(p_{ij_1})) - \frac{\sigma^2}{2}] d_{ij_1},$$

and variance

$$\mathrm{var}_t \left[\log w^s_{ij_2t+d} - \log w^s_{ij_1t} | p_{ij_0} = p_{ij_1}\right] = \sigma^2 d_{ij_1}.$$

For estimation we parametrize the drift allowing it to depend on the worker type and the firm type:

$$\mu(h_0, y) = \mu_0 + \mu_1 \log h_0 + \mu_2 \log y.$$

We parameterize the distribution of firm productivities to be bounded Pareto with lower support equal to one:

$$y(p) = \left(-\frac{p\overline{y}^\alpha - p - \overline{y}^\alpha}{\overline{y}^\alpha}\right)^{-\frac{1}{\alpha}}.$$

The parameters $(\mu_0, \mu_1, \mu_2, \alpha, \overline{y}, \sigma)$ can be jointly estimated by Maximum Likelihood.

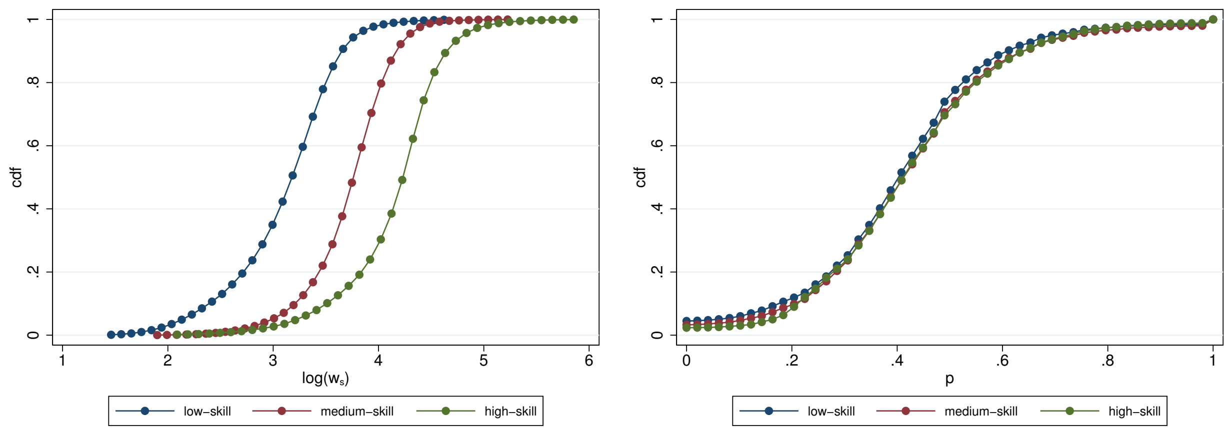
Preference parameters Given first stage estimates of the distribution of worker heterogeneity, the sampling distribution, contact rates, separation rates, and the technology of human capital accumulation, we can estimate the rate of time preference and the coefficient of relative risk by indirect inference. Specifically, we solve and simulate from the workers' problem (2) to match the mean and variance of the age profile for consumption.

5 Results

As of the current draft, the results presented in this and the following sections are preliminary as not all of the model parameters have been estimated. Specifically, the parameters of the firm productivity distribution, the human capital accumulation function and the preference parameters are based on a rough calibration and do not fully incorporate all the information from the matched employer employee data and the cross sectional data on consumption. For the simulations below, the subjective rate of time preference is set to 5 percent annual and the coefficient of relative risk aversion is set to 2.

Worker type distribution $L(h_0)$. Following the identification argument above, the distribution of initial log-wages upon entering the labour market is a non-parametric estimate

[10]For a derivation see Appendix A.2



(a) The distribution of $\hat{h}_0$ by education groups.

(b) Firm type distribution

Figure 1: Estimates

for the distribution of $\hat{h}_0$ across workers. Figure 1a plots the CDF of $\hat{h}_0$ by education (skill) group. There is clear first-order stochastic dominance in the distribution of h_0 by education group.

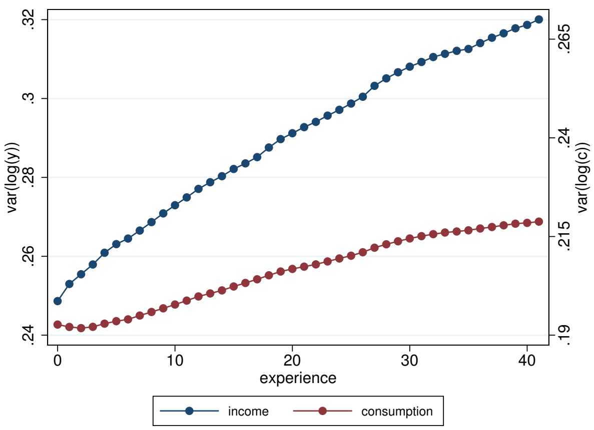
Firm type distribution $F(y(p))$. In Figure 1b we plot the estimated sampling distribution of firm types, p. This is equivalent to the sampling distribution of firm productivities. To transform to the distribution of firm productivities, note that $y(p) = \left(-\frac{p\overline{y}^{\alpha}-p-\overline{y}^{\alpha}}{\overline{y}^{\alpha}}\right)^{-\frac{1}{\alpha}}$.

Poisson arrival rates. Table 1 shows the estimated parameters for one possible specification. The ratio $\lambda(h_0)/\delta(h_0)$ measures how competitive the labor market is. A high separation rate or a low job contact rate are both indicative of strong frictions in the labor market, and the ratio $\lambda(h_0)/\delta(h_0)$ measure the expected number of contacts with firms per employment spell. Frictions are decreasing in education level (the labor market is more competitive for higher educated workers). Conditional on education,the estimated frictions are actually increasing with worker type h_0.

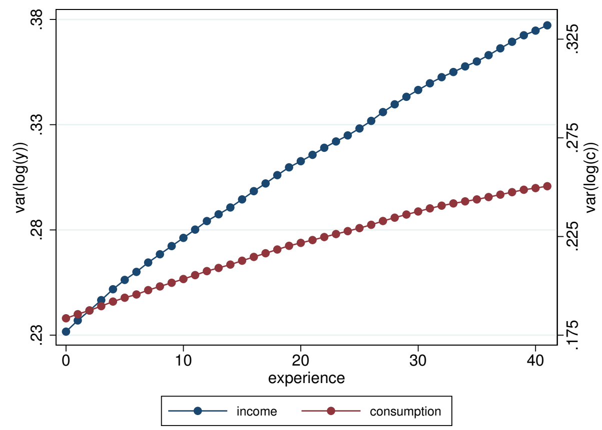
Table 1: Parameter Estimates

Parameter	low-skilled	medium-skilled	high-skilled
Contact rates: $\lambda(h_0) = \exp(\lambda_0 + \lambda_1 \log h_0)$			
λ_0	–2.632	–2.423	–2.733
	(.)	(.)	(.)
λ_1	–0.0910	–0.0243	–0.1102
	(.)	(.)	(.)
Destruction rates: $\delta(h_0) = \exp(\delta_0 + \delta_1 \log h_0)$			
δ_0	–3.585	–3.982	–4.912
	(.)	(.)	(.)
δ_1	0.1316	0.0122	–0.345
	(.)	(.)	(.)
Firm productivity distribution: $\Gamma(y) = \frac{1-y^{-\alpha}}{1-\overline{y}^{-\alpha}}$			
α	2		4
$\overline{y}$	2		3.5
Human capital, drift: $\mu(h_0, y) = \mu_0 + \mu_1 \log h_0 + \mu_2 \log y$			
μ_0	0.000		0.000
μ_1	–0.003		–0.0034
μ_2	0.001		0.0025
Human capital, variance of shocks: σ^2			
σ^2	0.0174		0.02007

Human Capital Accumulation. We estimate a parametrized drift rate following the identification arguments above, and summarized in the bottom panel of Table 1. Given a mean drift rate of zero, we obtain a reasonable fit for the calibration if the growth rate of human capital is increasing in the productivity of the firm and decreasing in the productivity of the worker. A negative correlation between the level and growth rate of worker productivity is also found by Guvenen (2009) when estimating a model of the stochastic process for



(a) Low-skilled



(b) High skilled

Figure 2: Simulated age profile for the variance of log earnings and consumption

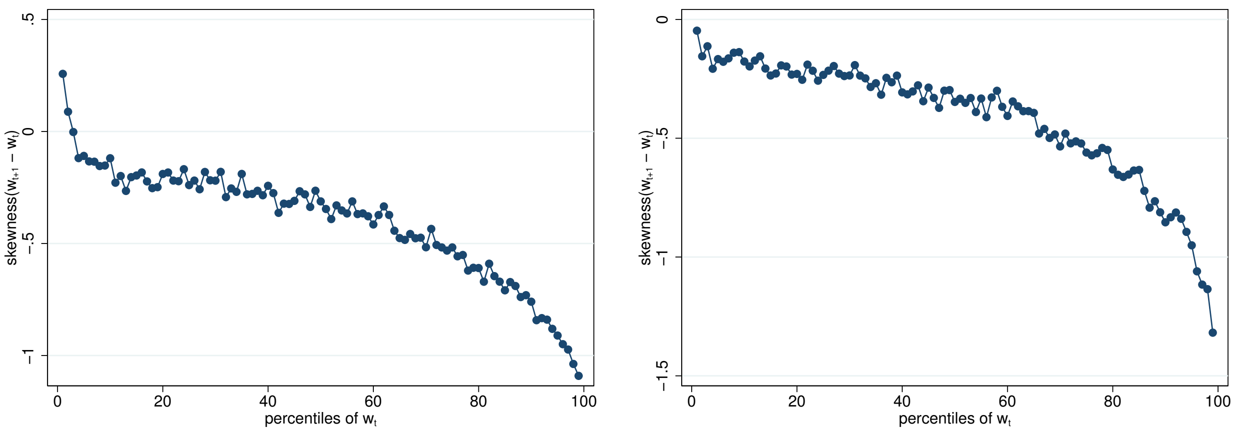
earnings allowing for heterogeneous income profiles. The variance of the Browning motion shock to human capital is similar for low and high educated workers.

5.1 Age profiles for the variance of log wages and consumption

We present model simulated age profiles for the variance of log annual wage earnings and the variance of log annual consumption in Figure 2. The model simulated profiles have the well known features: i) the level of the variance of log wages is higher than for consumption, ii) the age profiles are linear, iii) the increase for wages is approximately double the increase for consumption.

5.2 The skewness and kurtosis of conditional wage changes

In Figure 3 we plot the skewness of year over year changes in the log of annual wage income, conditional on the percentile of the wage distribution in the initial year, for data simulated from the model. For both the low and high skilled (educated) groups the skewness becomes increasingly negative as we condition on higher and higher percentiles of the initial wage distribution. The model produces a shock process in which negative shocks tend to be larger than positive shocks (except for the very bottom percentiles of the low skilled group), and they become increasingly larger as we move up the percentiles of the initial wage distribution. In the language of the model, workers tend to move gradually up the job ladder in small or moderate steps, while separating from a job leads to fall back to the bottom of the ladder. The negative change associated with falling off the job ladder is larger the higher up the



(a) Low skilled

(b) High skilled

Figure 3: Skewness of wage changes conditional on percentile of initial wage

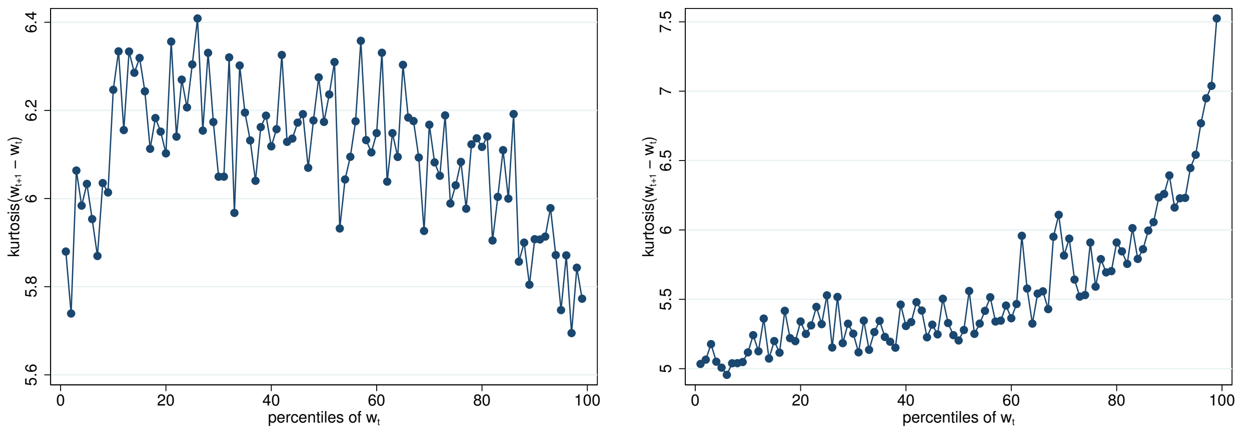
ladder the worker was employed.

We plot the kurtosis of year over year changes in the log of annual wage income, conditional on the percentile of the wage distribution in the initial year, for data simulated from the model in Figure 4. The model produces a distribution of wage changes that is substantially more peaked than what would arise from log-normal shocks (which would produce a kurtosis of 3). The pattern for the kurtosis of the change conditional on the initial wage percentile is somewhat different between the low-skilled and the high-skilled groups. For the low skilled the pattern has an inverted U shape (although the range is quite narrow), while for the high skilled group the pattern is increasing.

In summary, the model produces a stochastic process for wage changes that is consistent with the patterns highlighted by Arellano et al. (2015) and Guvenen et al. (2015). Most year over year changes in annual wage earnings tend to be small, and the large changes tend to be negative rather than positive. This pattern becomes exaggerated the higher a worker is on the job ladder.

6 Decomposition

In this section we decompose the patterns in Figures 2, 3 and 4 into the sources according to several decompositions implied by the model. A natural decomposition of the variance of log wages makes use of the model wage equation from which we have (with all variables in



(a) Low skilled

(b) High skilled

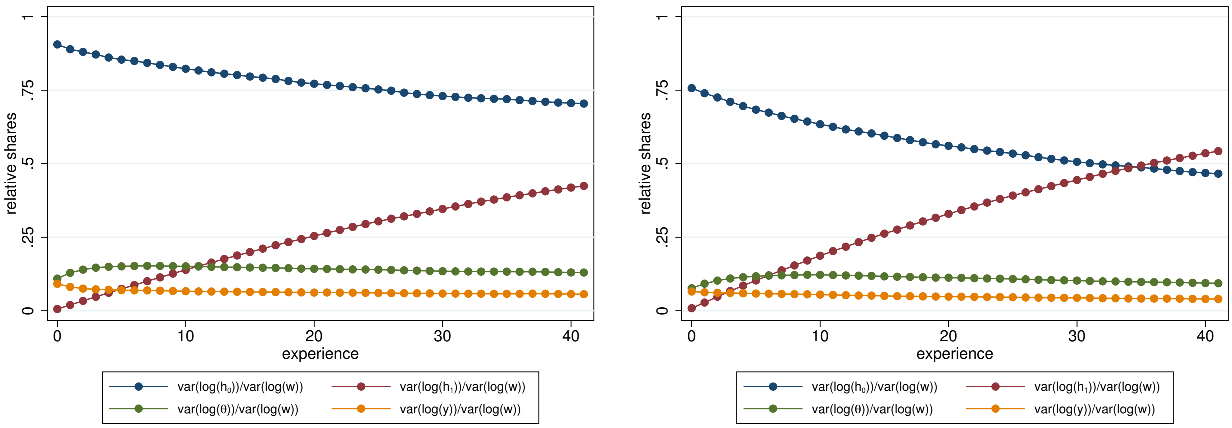
Figure 4: Kurtosis of wage changes conditional on percentile of initial wage

logs)

$$\begin{aligned}\mathrm{var}_t(w_{it}) = \mathrm{var}_t(h_{0i}) + \mathrm{var}_t(h_{1it}) + \mathrm{var}_t(y_{it}) + \mathrm{var}_t(\theta_{it}) \\ + 2\mathrm{cov}(h_{0i}, h_{1it}) + 2\mathrm{cov}(h_{0i}, y_{it}) + 2\mathrm{cov}(h_{0i}, \theta_{it}) \\ + 2\mathrm{cov}(h_{1it}, y_{it}) + 2\mathrm{cov}(h_{1it}, \theta_{it}) + 2\mathrm{cov}(h_{1it}, \theta_{it}).\end{aligned}$$

In Figure 5 plot the share of variance of log wages accounted for by each component of the wage equation. Specifically, we decompose the variance at each experience level, plotted in Figure 2, into variation in fixed worker heterogeneity h_{0i}, variation in acquired human capital h_{1it}, variation in firm heterogeneity y_{it}, and variation in piece rates. A striking feature for both the low an high skilled groups is that worker heterogeneity in initial human capital explains the bulk of the variance in wage at labor market entry. Over time, as workers differentially accumulate human capital while working, this component accounts for an increasing share of the overall wage variance. The model (at the current preliminary estimates) ascribes only a minor role for the combination of dispersion arising from the job ladder and market frictions, each contributing less than 12 percent, with the total contribution closer to 5 or 6 percent once we account for the negative correlation between firm type and the piece rate (see the covariance components in Figure 9 in the Appendix).

While this decomposition is useful, it may understate the effect of the job ladder on the rise in wage dispersion because it does not account for the fact that in the model firms also contribute differentially to the human capital accumulation of workers. The differential contribution to human capital accumulation arising from the job ladder may be more impor-



(a) low skilled

(b) high skilled

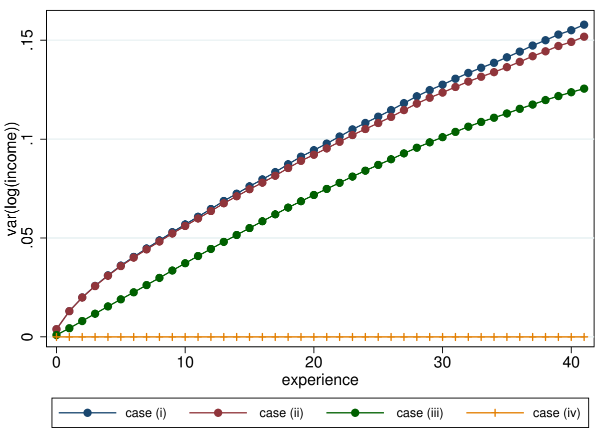
Figure 5: Decomposition of the share of the age profile of the variance of log wage
This decomposition uses the wage equation (in logs) $w_{it} = h_{0i} + h_{1it} + y_{it} + \theta_{it}$ to decompose the variance into the share due to initial worker heterogeneity, heterogeneity in accumulated human capital, the firm effect and the piece rate (which reflects market frictions).

tant than the output differences. This suggests another model decomposition in which we sequentially shut down heterogeneity in worker productivity; worker heterogeneity in human capital accumulation; heterogeneity firm productivity and human capital provision; and the shocks to human capital. This procedure also allows us to decompose the rise in consumption discretion into the same components.

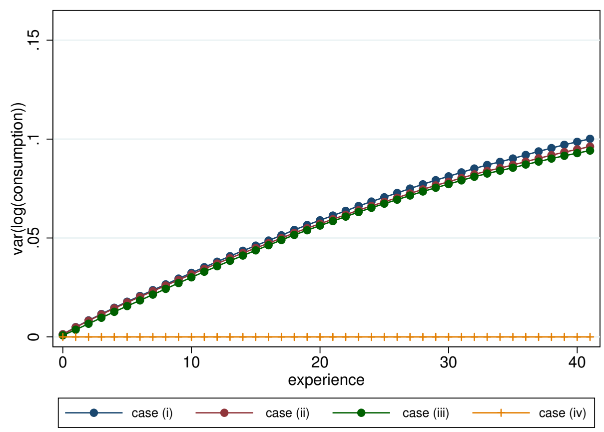
We plot this decomposition in Figure 6. Removing worker heterogeneity from production reduces the variance essentially to zero at labor market entry (the scale is the same is in Figure 2), and results in a steeper slope for the age profile. The increase in slope is a consequence of the negative correlation between the intercept and slop of the deterministic component of the human capital profiles. Removing the heterogeneity in human capital accumulation across workers reduces the slope slightly. Removing the job ladder effect on wages from both production and the human capital reduces the slope, leading to a 20 percent reduction in the rise in wage dispersion over 40 years of experience. Finally, setting the variance of shocks to human capital to zero completely flattens the age profile.

An interesting difference between the decomposition for wages and for consumption is that the job ladder accounts for about 20 percent of the rise in wage dispersion, but does not account for any of the rise in consumption dispersion. From the point of view of agents in the model, movements up the job ladder are seen as predictable or transitory while shocks to human capital are permanent. The job ladder does not contribute much to the age profile for the wage dispersion, and not at all to the age profile for consumption dispersion.

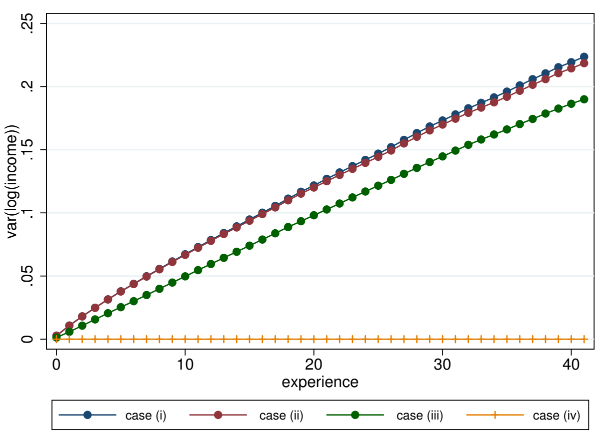
We performs the same decomposition for the skewness and kurtosis of conditional wage



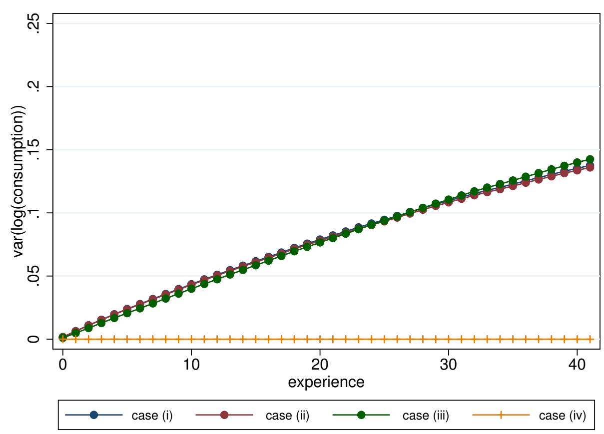
(a) Variance of log wages, Low-skilled



(b) Variance of log consumption, Low-skilled



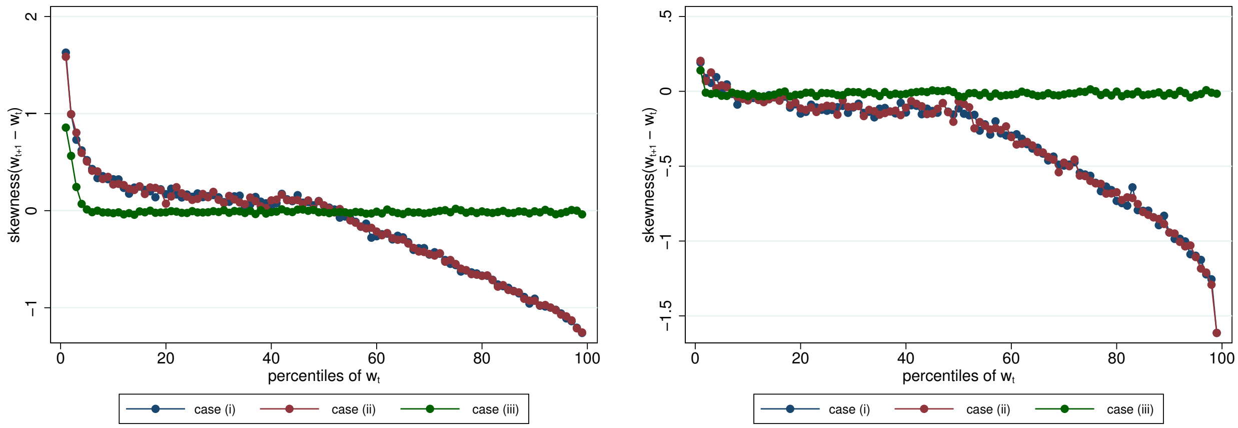
(c) Variance of log wages, High skilled



(d) Variance of log consumption, High skilled

Figure 6: Decomposing the rise in life-cycle inequality in wages and consumption

Note: case (i) removes the contribution of h_{0i} from production (and hence from the wage equation); (ii) additionally sets the coefficient on worker heterogeneity in human capital accumulation to zero; (iii) sets the firm-effect to $\underline{y}$ for all jobs; finally (iv) sets the variance of the shock to human capital to zero ($\sigma = 0$).



(a) Low-skilled

(b) High skilled

Figure 7: Decomposing the skewness of conditional wage changes

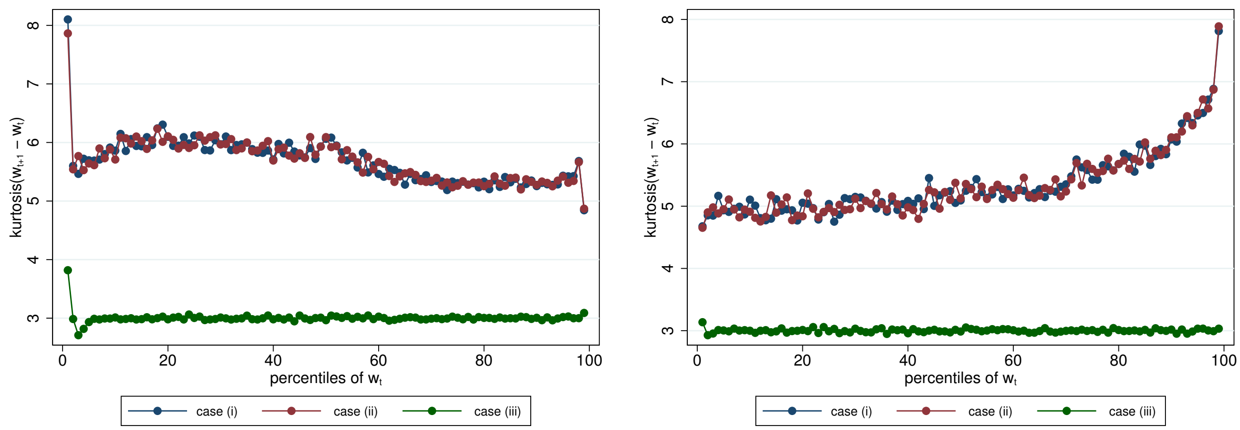
Note: case (i) removes the contribution of h_{0i} from production (and hence from the wage equation); (ii) additionally sets the coefficient on worker heterogeneity in human capital accumulation to zero; (iii) sets the firm-effect to $\underline{y}$ for all jobs.

changes. In Figures 7 and 8 we plot these decompositions. Removing fixed worker heterogeneity eliminates much of the negative skewness in wage changes below the 50th percentile of the initial wage distribution, resulting in positive skewness at the low end of the type distribution, however, the large negative skewness at the top of the initial wage distribution remains. Once remove the effect of the job ladder, the negative skewness completely disappears and wage changes appear symmetric, with the exception of positive skewness at the very bottom of the initial wage distribution.

The same pattern emerges when we look at the decomposition of kurtosis in Figure 8. There is some action at the very bottom of distribution of initial wages (for the low skilled) when we eliminate worker heterogeneity, however, once we remove the effect of the job ladder the excess skewness (above 3) in year-over-year wage changes disappears. In the absence of the job ladder wage changes appear to be essentially normally distributed, even conditional on the level of the initial wage (with the exception of the bottom of the distribution among low skilled workers).

7 Conclusion

The model is able to reproduce the linear age profile for the variance of log wages and consumption, as well as the negative skewness and excessive kurtosis of wage changes, including the increasing negative skewness of wage changes conditional on the previous wage level. We



(a) Low-skilled

(b) High skilled

Figure 8: Decomposing the kurtosis of conditional wage changes
Note: case (i) removes the contribution of h_{0i} from production (and hence from the wage equation); (ii) additionally sets the coefficient on worker heterogeneity in human capital accumulation to zero; (iii) sets the firm-effect to $\underline{y}$ for all jobs.

decompose the age profile for the variance of log wage and consumption, as well as conditional skewness and kurtosis of wage changes into that due to fixed worker heterogeneity, in both levels and growth rates, of human capital; the effect of firms via the job ladder; and the stochastic shocks to human capital. We find that 77 to 80 percent of the rise in the variances can be accounted for by permanent shocks to human capital, with the uncertainty created by the job ladder contributing the remaining 20 to 23 percent. Heterogeneity across workers in the productivity levels contributes to the level of dispersion, but not to the increase. Heterogeneity across workers in the growth rate of human capital accounts for a small fraction of the increase in the variance for wages, but none of the increase in the variance for consumption. On the other hand, firm heterogeneity manifested through the job ladder accounts for the entirety of the negative skewness and excess kurtosis in the conditional wage changes; the contribution of fixed worker heterogeneity and shocks to human capital is negligible.

Appendix

A Derivations

A.1 Hamilton-Jacobi-Bellman equation

We first derive the Bellman equation in discrete time for a small time interval Δt and then take the limit $\lim_{\Delta t \to 0}$ to derive the continuous time HJB equation.

Discrete time representation of (1). First consider the random walk representation of the human capital accumulation equation (1):

$$\Delta h_{1t} = \begin{cases} \Delta x & \text{with probability } p \\ -\Delta x & \text{with probability } 1-p \end{cases}.$$

with

$$\mathbb{E}(\Delta h_{1t}) = \mu(\mathbf{h}, y)\Delta t; \quad \text{var}(\Delta h_{1t}) = \sigma(\mathbf{h})^2 \Delta t.$$

It follows that

$$\begin{aligned} (2p-1)\Delta x &= \mu(\mathbf{h}, y)\Delta t \\ 2p(1-p)\Delta x &= \sigma(\mathbf{h})^2 \Delta t. \end{aligned}$$

Taking into account that $\Delta x \Delta t = \Delta t^2 = 0$ for small Δt, we therefore get

$$\Delta h_{1t} = \begin{cases} \sigma(\mathbf{h})\sqrt{\Delta t} & \text{with probability } \frac{1}{2}\left[1 + \frac{\mu(\mathbf{h},y)}{\sigma(\mathbf{h})}\sqrt{\Delta t}\right] \\ -\sigma(\mathbf{h})\sqrt{\Delta t} & \text{with probability } \frac{1}{2}\left[1 - \frac{\mu(\mathbf{h},y)}{\sigma(\mathbf{h})}\sqrt{\Delta t}\right] \end{cases}. \tag{6}$$

HJB equation. Consider the Bellman equation in discrete time,

$$W(a, \mathbf{h}, y, \theta) = \max_{a - \underline{a} \geq c \geq 0} \left\{ u(c)\Delta t + \frac{1}{1+\rho \Delta t} \mathbb{E} W(a', \mathbf{h}', y', \theta') \right\}, \tag{7}$$

Using a Taylor expansion, the expected continuation value can be written as:

$$\begin{aligned} \mathbb{E}W(a', \mathbf{h}', y', \theta') = \frac{\partial}{\partial a} W(a, \mathbf{h}, y, \theta)[ra + \theta f(\mathbf{h}, y) - c]\Delta t + \mu(\mathbf{h}, y)\Delta t \frac{\partial}{\partial h_1} W(a, \mathbf{h}, y, \theta) \\ + \frac{\sigma(\mathbf{h})^2}{2}\Delta t \frac{\partial^2}{\partial h_1^2} W(a, \mathbf{h}, y, \theta) + [1 - \lambda(h_0)\Delta t - \delta(h_0)\Delta t - \xi \Delta t] W(a, \mathbf{h}, y, \theta) + \delta(h_0)\Delta t W(a, \mathbf{h}, \underline{y}, 1) \\ + \xi \Delta t R(a, \mathbf{h}) + \lambda(h_0)\Delta t \int \max\{W(a, \mathbf{h}, y', \frac{y}{y'}), W(a, \mathbf{h}, y, \max\{\theta, \frac{y'}{y}\})\} dF(y') + o(\Delta t). \end{aligned}$$

Substituting into the Bellman equation and multiplying with $[1+\rho\Delta t]$ gives after rearranging,

$$[\rho+\lambda(h_0)+\delta(h_0)+\xi]W(a,\mathbf{h},y,\theta)\Delta t = \max_{a-\underline{a}\geq c\geq 0}\Bigg\{[1+\rho\Delta t]u(c)\Delta t+\frac{\partial}{\partial a}W(a,\mathbf{h},y,\theta)[ra+\theta f(\mathbf{h},y)-c]\Delta t$$
$$+\mu(\mathbf{h},y)h_1\Delta t\frac{\partial}{\partial h_1}W(a,\mathbf{h},y,\theta)+\frac{\sigma(\mathbf{h})^2}{2}\Delta t\frac{\partial^2}{\partial h_1^2}W(a,\mathbf{h},y,\theta)+\delta(h_0)\Delta t W(a,\mathbf{h},\underline{y},1)+\xi\Delta t R(a)$$
$$+\lambda(h_0)\Delta t\int \max\{W(a,\mathbf{h},y',\frac{y}{y'}),W(a,\mathbf{h},y,\max\{\theta,\frac{y'}{y}\})\}dF(y')+o(\Delta t)\Bigg\}.$$

Dividing by Δt and taking the limit $\lim_{\Delta t\to 0}$ yields the continuous time HJB (2).

Reflecting barriers. When the process for human capital h_1 reaches the upper (lower) barrier, then the process moves down (up) with probability 1 by assumption. Thus, when the process is at the upper barrier $\overline{h}_1$, the expected continuation value becomes

$$\mathbb{E}W(a',\mathbf{h}',y',\theta') = \frac{\partial}{\partial a}W(a,\overline{\mathbf{h}},y,\theta)[ra+\theta f(\overline{\mathbf{h}},y)-c]\Delta t - \sigma(\overline{\mathbf{h}})\sqrt{\Delta t}\frac{\partial}{\partial h_1}W(a,\overline{\mathbf{h}},y,\theta)$$
$$+\frac{\sigma(\overline{\mathbf{h}})^2}{2}\Delta t\frac{\partial^2}{\partial h_1^2}W(a,\overline{\mathbf{h}},y,\theta)+[1-\lambda(h_0)\Delta t-\delta(h_0)\Delta t-\xi\Delta t]W(a,\overline{\mathbf{h}},y,\theta)+\delta(h_0)\Delta t W(a,\overline{\mathbf{h}},\underline{y},1)$$
$$+\xi\Delta t R(a,\overline{\mathbf{h}})+\lambda(h_0)\Delta t\int \max\{W(a,\overline{\mathbf{h}},y',\frac{y}{y'}),W(a,\overline{\mathbf{h}},y,\max\{\theta,\frac{y'}{y}\})\}dF(y')+o(\Delta t).$$

Substitute this expression into the Bellman equation (7) and multiply with $\frac{[1+\rho\Delta t]}{\sqrt{\Delta t}}$. Taking the limit $\lim_{\Delta t\to 0}$ yields the boundary condition

$$\frac{\partial}{\partial h_1}W(a,\overline{\mathbf{h}},y,\theta)=0.$$

When we consider the lower barrier $\underline{h}_1$,we get the boundary condition

$$\frac{\partial}{\partial h_1}W(a,\underline{\mathbf{h}},y,\theta)=0.$$

A.2 Derivation of $\mathrm{var}_t \hat{h}_{1t}$

To compute the $\mathrm{var}_t \hat{h}_{1t}$, first note that

$$\begin{aligned}
\mathbb{E}_t \hat{h}_{1\tau}^2 &= \mathbb{E}_t \left[\hat{h}_{1t} + \tilde{\mu}(h_0, y')[\tau - t] + \tilde{\sigma}[\mathcal{B}_\tau - \mathcal{B}_t]\right]^2 \\
&= \mathbb{E}_t \left[\left[\hat{h}_{1t} + \tilde{\mu}(h_0, y')[\tau - t]\right]^2 + 2\left[\hat{h}_{1t} + \tilde{\mu}(h_0, y')[\tau - t]\right] \tilde{\sigma}[\mathcal{B}_\tau - \mathcal{B}_t] + \tilde{\sigma}^2[\mathcal{B}_\tau - \mathcal{B}_t]^2\right] \\
&= \left[\hat{h}_{1t} + \tilde{\mu}(h_0, y')[\tau - t]\right]^2 + \tilde{\sigma}^2 \mathbb{E}_t[\mathcal{B}_\tau - \mathcal{B}_t]^2
\end{aligned}$$

where we have again used the fact that the expected increment of the Brownian motion is zero. Furthermore note that

$$\mathbb{E}_t[\mathcal{B}_\tau - \mathcal{B}_t]^2 = \mathbb{E}_t\mathcal{B}_\tau^2 - 2\mathcal{B}_t\mathbb{E}_t\mathcal{B}_\tau + \mathcal{B}_t^2 = \mathbb{E}_t\mathcal{B}_\tau^2 - [\mathbb{E}_t\mathcal{B}_\tau]^2 = \mathrm{var}_t\mathcal{B}_\tau$$

where we have used that $\mathbb{E}_t\mathcal{B}_\tau = \mathcal{B}_t$. Hence we get

$$\mathbb{E}_t \hat{h}_{1\tau}^2 = \left[\hat{h}_{1t} + \tilde{\mu}(h_0, y')[\tau - t]\right]^2 + \tilde{\sigma}^2 \mathrm{var}_t \mathcal{B}_\tau.$$

Since the variance of the standard Brownian motion, $\mathrm{var}_t\mathcal{B}_\tau = \tau - t$, we find

$$\mathrm{var}_t \hat{h}_{1\tau} = \tilde{\sigma}^2[\tau - t].$$

We therefore conclude that the change in the time-variant component of human capital, $\hat{h}_{1\tau} - \hat{h}_{1t}$, is normally distributed with mean $\tilde{\mu}(h_0, y')[\tau - t]$ and variance $\tilde{\sigma}^2[\tau - t]$.

B Data

B.1 SIAB 7510

Unemployment. The data allows us to distinguish between three labour market states: employed, recipient of transfer payments and out of sample. Unfortunately, none of the two last categories corresponds exactly to the economic concept of unemployment. For example, participants in a training program are transfer payment recipients despite being in employment. From an administrative point of view they are considered unemployed. At the same time individuals that are registered unemployed but are no longer entitled to receive benefits appear to be out of the labour force. We therefore proxy unemployment by non-employment, i.e. the time between two employment spells.

Education. We distinguish between three skill groups which we label low-, medium- and high-skilled. We group workers according to their highest educational degree reported in the data and the time they have spent in apprenticeship training. In the raw data we do not have information on the education variable for 12.30% of all employment spells. However, since our data is longitudinal, we can impute the value by looking at past and future values of the educational variable. In particular we follow the imputation procedure IP1 suggested by Fitzenberger, Osikominu, and Völter (2006). After having imputed the education variable, only 1.56% of all employment spells have missing information on education.

Wages and Hours. Wages are reported by employers, but they are likely to be reported very precisely due to severe penalties in the case of misreporting. In the raw data we observe the average daily wage reported by the employer over the notification period. Thus, if a worker worked for more than one employer in a given year, we observe two employment spells for that worker associated with two different wages. If a worker gets promoted within a year, then this will be reflected in an increase in the average daily wage over the notification period.

Wages are right-censored at the highest level of earnings that are subject to social security contributions. We follow Dustmann, Ludsteck, and Schönberg (2009) and impute wages above the contribution limit. In particular we impute wages under the assumption that the error term in a standard wage regression is normally distributed. We estimate censored regressions for each year separately and account for heteroskedasticity by allowing variances to differ with skill and age groups.[11] Unfortunately we do not observe the actual hours worked, but we do observe whether a worker is working full-time or part-time defined as working less than 30 hours per week.

Sample Selection and Panel Structure. From the original sample, we select only male workers with information on the highest educational degree attained. For females the participation decision should be taken into account and this would require a different model. Further we restrict our attention to workers with employment spells in West-Germany, since there were additional regulations and restructuring after reunification in Eastern Germany. We further drop marginal part-time employment spells since they are recorded only since 1999 and consider only the main job defined to be the job with the highest daily wage. We define the date of entering the labour market as the date after finishing education and after all training periods and apprenticeships. We drop all persons who spent more than 7 years in apprenticeships or training and drop all registered spells before finishing education and

[11]We consider 6 age groups: < 19, 20-29, 30-39, 40-49, 50-59, > 60.

apprenticeships.[12] We also require workers to be within a certain age range when we first observe them. This ensures that we can follow workers from day one of their entry into the labour market. We require that workers without a high school degree have to be at most 19 years old when we first observe them. Those with a high school degree have to be at most 22 years old, those with a degree from a technical college 28, and those who graduated from university at most 30. In addition we also drop persons where we have missing information on either wages or the job-mover inflow ratio at the establishment level during periods of employment. These sample restrictions leave us with a sample of $212,380$ workers. We then process the spell data by taking monthly cross-sections of the data (the 15th day of each month) from which we construct a panel at monthly frequency, where gaps between two employment spells are labeled as episodes of non-employment.

C Numerical Solution

We follow Achdou, Han, Lasry, and Moll (2015) and use a finite difference method to solve the dynamic programming problem numerically. We solve for the value functions at N discrete points using linearly spaced grids for $(a, \mathbf{h}, y, \theta)$. Derivatives are approximated by finite differences, where Δ_x denotes the forward difference operator and ∇_x the backward difference operator in dimension x.

Retirement value. In the first step, we iteratively solve for the value of retirement,

$$\rho R(a, \mathbf{h}) = \max_{a-\underline{a} \geq c \geq 0} \{u(c) + \frac{\partial}{\partial a} R(a, \mathbf{h})[ra + b(\mathbf{h}) - c]\}.$$

Given a guess $R^i(a, \mathbf{h})$ the value function $R^{i+1}(a, \mathbf{h})$ is implicitly defined by

$$\frac{1}{\Delta}[R^{i+1}(a, \mathbf{h}) - R^i(a, \mathbf{h})] + \rho R^{i+1}(a, \mathbf{h}) = u(c^i(a, \mathbf{h})) + \frac{\partial}{\partial a} R^{i+1}(a, \mathbf{h})[ra + b(\mathbf{h}) - c^i(a, \mathbf{h})], \quad (8)$$

where Δ denotes the step size. We approximate the partial derivative with respect to wealth using an upwind scheme. The basic idea is to use a forward (backward) difference approximation whenever the drift of the state variable (i.e. savings) is positive (negative). We

[12] We also drop spells that occur before the age of 16 for the low-skilled, age 17 for the medium-skilled and age 24 for high-skilled.

therefore calculate savings using both forward and backward difference approximations,

$$
\begin{aligned}
s^i_\Delta &= ra + b(\mathbf{h}) - c^i_\Delta(a, \mathbf{h}) \\
s^i_\nabla &= ra + b(\mathbf{h}) - c^i_\nabla(a, \mathbf{h}),
\end{aligned}
$$

where consumption is implicitly defined by the first order condition given the current guess for the value function,

$$
\begin{aligned}
c^i_\Delta(a, \mathbf{h}) &= (u')^{-1}(\Delta_a R^i(a, \mathbf{h})) \\
c^i_\nabla(a, \mathbf{h}) &= (u')^{-1}(\nabla_a R^i(a, \mathbf{h})).
\end{aligned}
$$

The approximation of the derivative of the value function with respect to wealth is then given by

$$
\frac{\partial}{\partial a} R^{i+1}(a, \mathbf{h}) \approx \Delta_a R^{i+1}(a, \mathbf{h}) \mathbf{1}\{s^i_\Delta > 0\} + \nabla_a R^{i+1}(a, \mathbf{h}) \mathbf{1}\{s^i_\nabla < 0\} + u'(ra + b(\mathbf{h})) \mathbf{1}\{s^i_\Delta \leq 0 \leq s^i_\nabla\}, \tag{9}
$$

where the last term captures situations where the sign of the savings is undetermined for some grid points.[13] At these grid points we set savings to zero, so that the derivative of the value function is equal to the marginal utility from consuming the flow income. Similarly, consumption $c^i(a, \mathbf{h})$ is defined by

$$
c^i(a, \mathbf{h}) = c^i_\Delta(a, \mathbf{h}) \mathbf{1}\{s^i_\Delta > 0\} + c^i_\nabla(a, \mathbf{h}) \{s^i_\nabla < 0\} + [ra + b(\mathbf{h})] \mathbf{1}\{s^i_\Delta \leq 0 \leq s^i_\nabla\}. \tag{10}
$$

Substituting (9) and (10) into (8) yields a linear system of equations that can be efficiently solved by exploiting the sparse structure of the numerical problem. We then iterate on the value function until convergence.

[13] Note that the concavity of the value function implies that $\Delta_a R(a, \mathbf{h}) < \nabla_a R(a, \mathbf{h})$ and therefore $s_\Delta < s_\nabla$. Thus it can arise in some situations that $s_\Delta \leq 0 \leq s_\nabla$.

Worker value. Given the retirement value, we can solve for the worker value in the second step. Given a guess $W^i(a, \mathbf{h}, y, \theta)$ the update $W^{i+1}(a, \mathbf{h}, y, \theta)$ is implicitly defined by

$$\frac{1}{\Delta}[W^{i+1}(a,\mathbf{h}) - W^i(a,\mathbf{h},y,\theta)] + [\rho + \lambda(h_0) + \delta(h_0) + \xi]W^{i+1}(a,\mathbf{h},y,\theta) = u(c^i(a,\mathbf{h},y,\theta))$$
$$+\frac{\partial}{\partial a}W^{i+1}(a,\mathbf{h},y,\theta)[ra+\theta f(\mathbf{h},y)-c^i(a,\mathbf{h},y,\theta)]+\mu(\mathbf{h},y)\frac{\partial}{\partial h_1}W^{i+1}(a,\mathbf{h},y,\theta)+\frac{\sigma(\mathbf{h})^2}{2}\frac{\partial^2}{\partial h_1^2}W(a,\mathbf{h},y,\theta)$$
$$+\delta(h_0)W^{i+1}(a,\mathbf{h},\underline{y},1)+\xi R(a,\mathbf{h})+\lambda(h_0)\int \max\{W^{i+1}(a,\mathbf{h},y',\frac{y}{y'}),W^{i+1}(a,\mathbf{h},y,\max\{\theta,\frac{y'}{y}\})\}dF(y'), \tag{11}$$

where Δ denotes the step size. We approximate the first order partial derivatives using an upwind scheme. Specifically, we approximate the first order partial derivative with respect to human capital h_1 as follows:

$$\frac{\partial}{\partial h_1}W^{i+1}(a,\mathbf{h},y,\theta) \approx \Delta_{h_1}W^{i+1}(a,\mathbf{h},y,\theta)\mathbf{1}\{\mu(\mathbf{h},y) \geq 0\}+\nabla_{h_1}W^{i+1}(a,\mathbf{h},y,\theta)\mathbf{1}\{\mu(\mathbf{h},y) < 0\}, \tag{12}$$

while the second order partial derivative with respect to human capital h_1 is approximated by a central difference. In addition, we have to take into account that the diffusion process is reflected at the boundaries. The reflection implies the boundary conditions

$$\frac{\partial}{\partial h_1}W^{i+1}(a,\underline{\mathbf{h}},y,\theta) = \frac{\partial}{\partial h_1}W^{i+1}(a,\overline{\mathbf{h}},y,\theta) = 0, \tag{13}$$

where $\underline{\mathbf{h}} = (h_0, \underline{h}_1)$ and $\overline{\mathbf{h}} = (h_0, \overline{h}_1)$. In terms of the partial derivative with respect to wealth, we proceed as already explained above. We first calculate savings to determine the sign of the drift. Given the current guess for the value function, savings based on the forward and backwards difference approximations are given by

$$s^i_\Delta = ra + \theta f(\mathbf{h}, y) - c^i_\Delta(a,\mathbf{h},y,\theta)$$
$$s^i_\nabla = ra + \theta f(\mathbf{h}, y) - c^i_\nabla(a,\mathbf{h},y,\theta),$$

where consumption is implicitly defined by the first order condition,

$$c^i_\Delta(a,\mathbf{h},y,\theta) = (u')^{-1}(\Delta_a W^i(a,\mathbf{h},y,\theta))$$
$$c^i_\nabla(a,\mathbf{h},y,\theta) = (u')^{-1}(\nabla_a W^i(a,\mathbf{h},y,\theta)).$$

This yields the following approximation for the shadow value of wealth:

$$\frac{\partial}{\partial a}W^{i+1}(a,\mathbf{h},y,\theta) \approx \Delta_a W^{i+1}(a,\mathbf{h},y,\theta)\mathbf{1}\{s^i_\Delta > 0\} + \nabla_a W^{i+1}(a,\mathbf{h},y,\theta)\mathbf{1}\{s^i_\nabla < 0\} \\ + u'(ra + \theta f(\mathbf{h},y))\mathbf{1}\{s^i_\Delta \leq 0 \leq s^i_\nabla\}, \quad (14)$$

where the last term again captures situations where the sign of the savings is undetermined for some grid points. At these grid points we set savings to zero, so that the derivative of the value function is equal to the marginal utility from consuming the flow income. Similarly, consumption $c^i(a,\mathbf{h},y,\theta)$ is defined by

$$c^i(a,\mathbf{h},y,\theta) = c^i_\Delta(a,\mathbf{h},y,\theta)\mathbf{1}\{s^i_\Delta > 0\} + c^i_\nabla(a,\mathbf{h},y,\theta)\{s^i_\nabla < 0\} + [ra + \theta f(\mathbf{h},y)]\mathbf{1}\{s^i_\Delta \leq 0 \leq s^i_\nabla\}. \quad (15)$$

Finally, the integral in (11) is approximated by a finite sum,

$$\int \max\{W^{i+1}(a,\mathbf{h},y',\frac{y}{y'}), W^{i+1}(a,\mathbf{h},y,\max\{\theta,\frac{y'}{y}\})\}dF(y') \\ \approx \sum_{y'} \max\{W^{i+1}(a,\mathbf{h},y',\frac{y}{y'}), W^{i+1}(a,\mathbf{h},y,\max\{\theta,\frac{y'}{y}\})\}\hat{f}(y'), \quad (16)$$

where $\hat{f}(.)$ is an appropriate probability mass function. Substituting (12), (13), (14), (15) and (16) into (11) yields a linear system of equations that can be efficiently solved by exploiting the sparse structure of the numerical problem. We then iterate on the value function until convergence.

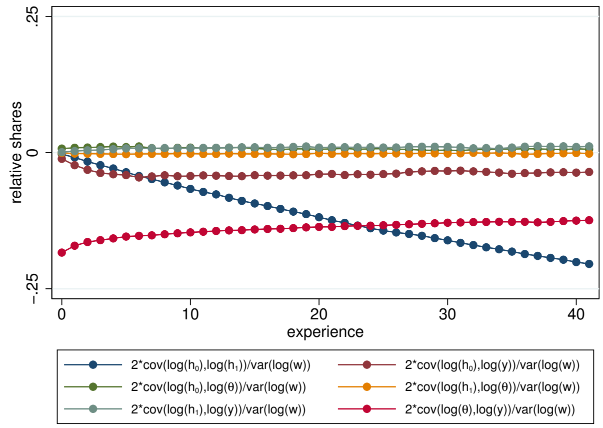
References

Abowd, J. M., and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–45.

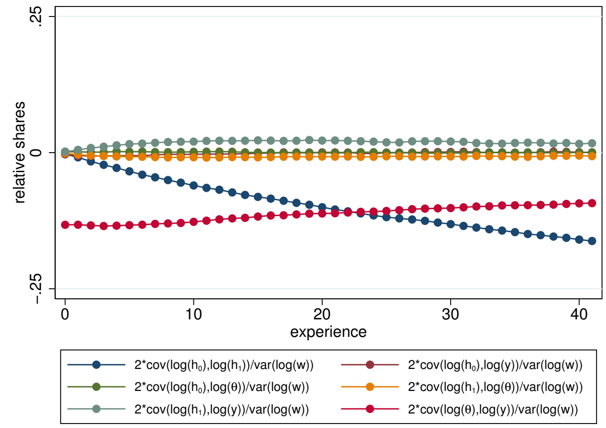
Abraham, K. G., and H. S. Farber (1987): "Job Duration, Seniority, and Earnings," *The American Economic Review*, 77(3), 278–297.

Achdou, Y., J. Han, J.-M. Lasry, and B. Moll (2015): "Heterogeneous Agent Models in Continuous Time," unpublished manuscript.

Altonji, J. G., and R. A. Shakotko (1987): "Do Wages Rise with Job Seniority?," *Review of Economic Studies*, 54(3), 437–59.



(a) low skilled



(b) high skilled

Figure 9: Covariance components for the decomposition of the share of the age profile of the variance of log wage

This decomposition uses the wage equation (in logs) $w_{it} = h_{0i} + h_{1it} + y_{it} + \theta_{it}$ to decompose the variance into the share due to initial worker heterogeneity, heterogeneity in accumulated human capital, the firm effect and the piece rate (which reflects market frictions).

Altonji, J. G., A. Smith, and I. Vidangos (2013): "Modeling Earnings Dynamics," *Econometrica*, 81, 1395–1454.

Arellano, M., R. Blundell, and S. Bonhomme (2015): "Earnings and consumption dynamics: a nonlinear panel data framework," Discussion paper, cemmap CWP53/15.

Bagger, J., F. Fontaine, F. Postel-Vinay, and J.-M. Robin (2014): "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics," *American Economic Review*, 104(6), 1551–96.

Bagger, J., and R. Lentz (2014): "An empirical model of wage dispersion with sorting," Discussion paper, National Bureau of Economic Research WP 20031.

Barlevy, G. (2008): "Identification of Search Models using Record Statistics," *Review of Economic Studies*, 75(1), 29–64.

Bowlus, A. J., and H. Liu (2013): "The contributions of search and human capital to earnings growth over the life cycle," *European Economic Review*, 64, 305 – 331.

Browning, M., M. Ejrnaes, and J. Alvarez (2010): "Modelling Income Processes with Lots of Heterogeneity," *Review of Economic Studies*, 77(4), 1353–1381.

Buchinsky, M., D. Fougère, F. Kramarz, and R. Tchernis (2010): "Interfirm Mobility, Wages and the Returns to Seniority and Experience in the United States," *The Review of Economic Studies*, 77(3), 972–1001.

Bunzel, H., B. Christensen, N. Kiefer, and L. Korsholm (1999): "Equilibrium Search with Human Capital Accumulation," unpublished manuscript.

Burdett, K., C. Carrillo-Tudela, and M. G. Coles (2011): "Human Capital Accumulation and Labor Market Equilibrium," *International Economic Review*, 52(3), 657–677.

Burdett, K., and M. Coles (2003): "Equilibrium Wage-Tenure Contracts," *Econometrica*, 71(5), 1377–1404.

Burdett, K., and D. T. Mortensen (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39(2), 257–273.

Dustmann, C., J. Ludsteck, and U. Schönberg (2009): "Revisiting the German Wage Structure," *The Quarterly Journal of Economics*, 124(2), 843–881.

Fitzenberger, B., A. Osikominu, and R. Völter (2006): "Imputation Rules to Improve the Education Variable in the IAB Employment Subsample," *Schmollers Jahrbuch: Journal of Applied Social Science Studies*, 126(3), 405–436.

Guvenen, F. (2007): "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *American Economic Review*, 97(3), 687–712.

——— (2009): "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, 12(1), 58–79.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2015): "What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk?," Discussion paper, NBER WP 20913.

Lamadon, T. (2014): "Productivity shocks, Optimal Contracts and Income Dynamics," unpublished working paper.

Lentz, R., and N. Roys (2015): "Training and Search On the Job," unpublished working paper.

Lillard, L. A., and Y. Weiss (1979): "Components of Variation in Panel Earnings Data: American Scientists 1960-70," *Econometrica*, 47(2), 437–454.

LISE, J. (2013): "On-the-Job Search and Precautionary Savings," *Review of Economic Studies*, 80(3), 1086–1113.

MACURDY, T. (1982): "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, 18(1), 82–114.

MEGHIR, C., AND L. PISTAFERRI (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72(1), 1–32.

——— (2011): "Chapter 9 - Earnings, Consumption and Life Cycle Choices*," vol. 4, Part B of *Handbook of Labor Economics*, pp. 773–854. Elsevier.

MOSCARINI, G., AND F. POSTEL-VINAY (2013): "Stochastic Search Equilibrium," *Review of Economic Studies*, 80(4), 1545–1581.

POSTEL-VINAY, F., AND J.-M. ROBIN (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70(6), pp. 2295–2350.

RUBINSTEIN, Y., AND Y. WEISS (2006): "Chapter 1 Post Schooling Wage Growth: Investment, Search and Learning," vol. 1 of *Handbook of the Economics of Education*, pp. 1 – 67. Elsevier.

TOPEL, R. H. (1991): "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority," *Journal of Political Economy*, 99(1), 145–76.

TOPEL, R. H., AND M. P. WARD (1992): "Job Mobility and the Careers of Young Men," *The Quarterly Journal of Economics*, 107(2), 439–79.

VERAMENDI, G. (2011): "Labor Market Dynamics: A Model of Search and Human Capital Accumulation," unpublished manuscript.

YAMAGUCHI, S. (2010): "Job Search, Bargaining, and Wage Dynamics," *Journal of Labor Economics*, 28(3), pp. 595–631.