Earnings Dynamics and Firm-Level Shocks*

Benjamin Friedrich† Lisa Laun‡ Costas Meghir§ Luigi Pistaferri¶

November 25, 2018

Abstract

In this paper we use matched employer-employee data from Sweden to study the role of the firm in affecting the stochastic properties of the wage process. We consider two channels: time-varying match effects, and the transmission of firm-specific productivity shocks. In both cases we separate temporary from more persistent effects. Our statistical model accounts for endogenous participation and mobility decisions and thus deals explicitly with censoring induced by people quitting into unemployment or changing employer. We document a number of key findings. First, firm-specific permanent productivity shocks transmit to individual wages, but the effect is mostly concentrated among the high-skilled workers; the opposite pattern is found for firm-specific temporary shocks. Moreover, we find only modest updates in match effects over the life of a firm-worker relationship. Finally, we estimate a significant role for permanent individual shocks that stick with the worker. However, a substantial part of the growth in earnings variance over the life cycle for high-skilled workers is driven by firms. By age 55, 44% of the cross-sectional variance is attributable to firm-level shocks passing through wages.

Keywords: Income process, Wage dynamics, Firm dynamics

JEL-codes: H51, H55, I18, J26

*We thank the IFAU for providing the data and supporting the project. We are grateful to participants in seminars at the SED, University of Chicago, Northwestern, NBER, and the Conference in honor of Joe Altonji for valuable comments. We thank Joe Altonji, Peter Arcidiacono, Thibaut Lamadon, Anders Forslund, Greg Kaplan, Magne Mogstad and Emily Nix for fruitful discussions and comments. Costas Meghir thanks the Cowles Foundation and the ISPS for funding. All errors and interpretations are our own.

†Northwestern University, Kellogg School of Management

‡Stockholm University and IFAU

§Yale University, IFS, NBER, IFAU and IZA

¶Stanford University, NBER, CEPR, SIEPR and IZA

1 Introduction

How important is the firm in which a worker is employed in determining his wages? And how much of the wage fluctuations over an individual's career reflect fluctuations in firm productivity? These questions are important for understanding the sources of inequality and of risk facing individuals over the lifecycle.

A number of papers have addressed the former question, starting with Abowd, Kramarz, and Margolis (1999a) as well as more recent papers such as Card, Heining, and Kline (2013). However there is very little work addressing the extent to which fluctuations in the firm's fortunes pass on to wages, in part because of the formidable data requirements. In this paper we address the extent to which the idiosyncratic wage shocks are related to fluctuations in firm-level productivity shocks. This relates directly both to the amount and sources of risk faced by individuals and to the competitiveness of the labor market, making it an issue of first order importance from a number of perspectives.

Related directly to this question are the pay policies of firms in frictional labor markets (Postel-Vinay and Robin, 2002b, Lise, Meghir, and Robin, 2016, for example). This research agenda partly reflects developments in search theory (starting with the seminal models of Burdett and Mortensen (1998) and Mortensen and Pissarides (1994)), which stress departures from perfect competition and the law of one price. The causes of pay heterogeneity also underlie the reasons why workers may, under certain circumstances, share shocks to firm productivity.[1]

In general, workers face multiple sources of risk distinct from their own productivity shocks.[2] Consider for instance fluctuations in the fortunes of the firm, induced by product

[1] Of course pay heterogeneity for the same worker across firms does not require search frictions: complementarities between worker and firm productivities will imply such heterogeneity as in a Becker-style marriage market, although in the absence of search frictions it is hard to understand why we would observe workers moving across the quality distribution of firms, which in practice happens frequently.

[2] Low, Meghir, and Pistaferri (2010) illustrate the importance of such distinctions for understanding the welfare effects of risk.

market shocks. In a competitive labor market workers only bear the risk of shocks to their *own* productivity, which they carry with them wherever they work, and they bear them fully. However, in the presence of search or financial frictions, the response of wages to individual and firm-related shocks may not be straightforward. On the one hand, productivity shocks may not affect wages immediately (for example if firms offer implicit contracts smoothing wages in the face of variable productivity). On the other hand, the worker may have to share some of the shocks to the productivity of the firm itself, since his immediate outside option is unemployment and thus he may not have a credible threat to quit. This issue relates to whether workers share rents with the firm, and an early study in this direction is van Reenen (1996). Indeed, in a recent paper Lamadon, Mogstad, and Setzler (2018) show how the pass-through of firm level shocks to wages reflects wage-setting power.

The existing literature has focused on sorting, and explain wage determination in the absence of firm-related shocks. The transmission of productivity shocks to wages has been examined before by Lise, Meghir, and Robin (2016), who, however, do not use matched employer-employee data. More recently, Lamadon (2016) has developed a structural model with directed job search that offers a theoretical framework for understanding the role of firm level shocks for worker outcomes.

In an earlier paper, Guiso, Pistaferri, and Schivardi (2005) estimate the pass-through of firm level shocks onto wages using Italian matched employer-employee data and interpret the results as estimates of the amount of insurance the firm provides. However, their approach is limited by the fact that they ignore job-to-job mobility and the transitions between employment and unemployment. Such transitions may well hide the impact of firm-level shocks on wages because a worker may quit or switch jobs instead of suffering too large a pay cut, causing wage growth to be censored. At the same time, some of the firm-level adjustments may consist of dismissing workers rather than adjusting their wages, conditional on employment.

In this paper we remain agnostic about the specific structural model that generates the data. We build on the literature modeling the stochastic structure of earnings, see Abowd and Card (1989), MaCurdy (1982), Meghir and Pistaferri (2004), Guvenen (2007) and more recently Altonji, Smith, and Vidangos (2013). We extend the framework considered by these papers by using matched employer-employee data and using information on firm level shocks to explicitly identify the extent to which they can explain individual wage fluctuations. In this way we go beyond the existing literature and identify different sources of risk, including individual productivity as well as firm level shocks. To avoid biases due to censoring we explicitly allow for the endogeneity of transitions between employment and unemployment as well as between jobs.

In a related paper, Low, Meghir, and Pistaferri (2010) find that making job mobility and employment choices endogenous reduces the estimated variance of permanent shocks compared to earlier studies. In their model firms are represented as a fixed matched heterogeneity effect. However, because they do not observe firms they are not able to measure the impact of shocks to firms separately from worker productivity shocks. They do, however, infer indirectly the amount of heterogeneity that can be attributed to the workplace. A related paper is Altonji, Smith, and Vidangos (2013), who specify a model of employment, hours, wages and earnings in order to distinguish between different sources of risk. Selection into employment and between jobs is modeled in a similar way as in Low, Meghir, and Pistaferri (2010). While both studies allow for some firm-related variation in wages, they do not consider the role of firm-level shocks for earnings dynamics, which is the main contribution of the present paper.

Our data are drawn from Swedish administrative records. We have matched these records with data on firm balance sheets. The result is the universe of workers and firms, matched to each other for the years 1997-2008. The data include annual earnings, detailed information of job histories, including the identity of the firm and other important information. However, it does not include hours of work. We thus focus our main analysis on men, who rarely work part-time. We allocate individuals to two education groups: those with some college education and those with less.

We specify a model of wages, employment and job mobility, all of which are interrelated. Specifically, wage shocks drive entry and exit from work, while mobility is allowed to depend on wage improvements between the incumbent and the poaching firm. The stochastic structure of wages includes idiosyncratic effects, reflecting changes in individual productivity and match-specific effects. The latter consist in part of shocks to firm productivity (transitory and permanent) as well as random match effects. As such, it is a particularly rich framework that effectively nests earlier specifications of the stochastic process of income.

We find that firm productivity is quite volatile and that this volatility transmits to wages of high skill workers to a larger extent, particularly when it relates to permanent shocks. It thus turns out that the firm is responsible for a high fraction of cross sectional variance of wages attributable to unobserved components and interpreted as uncertainty. The same is not true for unskilled workers: transitory shocks to productivity transmit to wages, but overall this does not explain a large fraction of the wage variance. We also find that employment is strongly related to wage shocks, consistent with self selection into work and work incentives. Finally, job mobility is highly dependent on wage offers, although other factors lead workers to take wage cuts when they move across workplaces.

To better understand the implications of our main findings, we simulate the model in a number of counterfactual scenarios in which we change the nature of wage variability over the life cycle. In one scenario, we eliminate any pass-through of firm shocks onto wages; in another, we shut down any form of firm influence on wages (both match productivity effects as well as firm shocks pass-through). We find that wage variances over the life cycle decline substantially when eliminating the impact of firm shocks, and less so when match productivity shocks are eliminated (with the effect being particularly relevant for the high skilled). In another set of counterfactual experiments, we eliminate selection by preventing job-to-job moves or quits into unemployment. If workers cannot move or quit (which are extreme forms of labor market frictions), shocks stay with them longer and cannot be avoided, resulting in higher variances over the life cycle. We show that this is mostly due to pass-through of firm-specific shocks. Hence, workers' dynamism (the ability to quit into unemployment or move to alternative employers) represent an implicit form of insurance against labor marker risks.

The paper proceeds as follows. Section 2 presents the model of the income process. Section 3 introduces the dataset and presents descriptive statistics. Section 4 presents the estimation and identification strategy. Section 5 shows the main results for the two-stage estimation procedure and their implications for our understanding of where labor market risks come from. Section 6 concludes.

2 The Stochastic Structure of Earnings

2.1 Overview

The model we set up has various components. At the heart of the specification is a wage equation for each of the two education groups we consider (some college or less). Our focus is on wage growth over the life cycle. We thus allow for a stochastic structure of wages that depends on general productivity shocks, which follow the worker wherever he is employed to the extent that they are persistent. Wages also depend on match-specific effects (relating to the specific worker/firm combination), and possibly on shocks to firm-level productivity, which is the central question of our paper.

Selection into and out of employment and mobility between jobs may be driven, at least in part, by shocks to wages. Ignoring this link may cause a serious bias in the measurement of the impact of firm level shocks, since large adjustments are effectively censored by individual behavior. We thus allow for endogenous employment and mobility and relate this directly to shocks.

2.2 The Statistical Model

Wages We consider a quarterly model for wages, employment and job mobility. The quarterly frequency is designed to capture the effects of job mobility and the associated wage changes. If we were to focus on annual frequencies, there would be too few moves and the model would miss a key source of wage dynamics.

Log wages of individual i in calendar year t who started to work at firm j in period t_0 is given by:

$$\ln w_{i,j(t_0),t} = x'_{i,t}\gamma + P_{i,t} + \varepsilon_{i,t} + v_{i,j(t_0),t}, \tag{1}$$

where x are observable worker characteristics such as age, education, and experience.

We assume that $\varepsilon_{i,t} \sim N\left(0, \sigma_\epsilon^2\right)$ is an i.i.d. transitory productivity shock,[3] and $P_{i,t}$ is permanent productivity, specified as:

$$\begin{aligned} P_{i,t} &= \rho P_{i,t-1} + \zeta_{i,t} \\ &= \rho^t P_{i,0}^{init} + \sum_{s=1}^{t} \rho^{t-s} \zeta_{i,s} \end{aligned} \tag{2}$$

where P^{init} is the initial productivity draw upon entry into the labor market. If $\rho = 1$ we have the standard random walk assumption for the permanent component of wages. The productivity shock is denoted ζ and we make the distributional assumptions

[3]Note that we assume no measurement error because we will use high quality administrative data for estimation. Meghir and Pistaferri (2004) point out the inability to disentangle the variance of the transitory shock, the variance of the measurement error and the parameters of the transitory process in a similar setting. The distinction has economic implications, however, since measurement error is pure noise while transitory shocks reflect uncertainty that may give rise to economic responses. The authors suggest two ways of handling this issue: obtaining bounds for the unidentified variances or using an external estimate of the measurement error (from validation data) to recover the variance of the transitory shock. In practice, if some of the transitory variation in wages that we estimate reflects measurement error, the main effect will be in an overstatement of transitory risk.

$$P^{init} \sim N\left(0, \sigma_P^2\right), \quad \zeta \sim N\left(0, \sigma_\zeta^2\right). \tag{3}$$

The identity of the firm affects wages through the match-specific productivity term $v_{i,j(t_0),t}$. We assume that the match effect evolves stochastically as a result of firm- and match-specific shocks. It is useful to distinguish between a component that reflects permanent (or at least long-run persistent) changes in the value of the worker/firm match, and one that reflects transitory changes. Within that context we will introduce the way the firm affects wage growth. For the periods $t > t_0$ when the worker does not change jobs we assume:

$$v_{i,j(t_0),t} = v^P_{i,j(t_0),t} + v^T_{i,j(t_0),t} \tag{4}$$

The permanent part of the match component follows the law of motion:

$$\begin{aligned} v^P_{i,j(t_0),t} &= v^P_{i,j(t_0),t-1} + \kappa^P \xi^P_{j,t} + \psi^P_{i,j(t_0),t} \\ &= v^P_{i,j(t_0),t_0} + \kappa^P \sum_{s=t_0+1}^{t} \xi^P_{j,s} + \sum_{s=t_0+1}^{t} \psi^P_{i,j(t_0),s} \end{aligned} \tag{5}$$

while the transitory part of the match component equals:

$$v^T_{i,j(t_0),t} = \kappa^T \xi^T_{j,t} + \psi^T_{i,j(t_0),t} \tag{6}$$

The initial draw of the permanent match productivity (at time t_0) in equation (5) is:

$$v^P_{i,j(t_0),t_0} = \tau_j + \psi^{init}_{i,j(t_0)}. \tag{7}$$

Thus we assume that the initial match value of a job is affected by fixed firm characteristics

τ_j and an idiosyncratic match component, which is distributed as follows:

$$\psi_{i,j(t_0)}^{init} \sim N\left(0, \sigma_{\psi^{init}}^2\right).$$

Notice that with this specification we are not modeling sorting because we do not allow the initial match effect to depend on individual characteristics. However, our focus is on wage growth moments and even if there is sorting based on permanent characteristics our estimates are not biased because these effects difference out once we consider log-wage changes. Of course, they may matter to explain wage levels.

In equations (5) and (6), the terms $\xi_{j,t}^P$ and $\xi_{j,t}^T$ are permanent and transitory shocks to the productivity of the firm, respectively. The properties of these shocks will be measured directly from the firm level data. The two ψ shocks are i.i.d. normal. Specifically we assume that $\psi^l \sim N\left(0, \sigma_{\psi^l}^2\right)$, for $l = \{P, T\}$. By allowing for these match specific shocks that are unrelated to firm level productivity we guard against the possibility that the productivity shocks just proxy for such effects.

The existence of a match-specific effect has been motivated theoretically within the search and matching framework by, among others, Topel and Ward (1992). Abowd, Kramarz, and Margolis (1999b) use French employer-employee data to show that match-specific effects matter empirically. Most studies on earnings dynamics, however, have not explicitly modeled the firm side. Low, Meghir, and Pistaferri (2010) include a match-specific component in the wage process, but in their paper the match is not allowed to change within the firm-worker relationship and is not subject to shocks that could be related to firm-level productivity. Thus, in their model, wage growth does not depend on the identity of the firm.

These additions to the match-specific component are one of the contributions of our work compared to earlier studies. The other key part is that some of the evolution of the match component (due to "learning" or firm-specific human capital accumulation) may mask rent sharing. In our framework, these two are kept distinct, which is an important deviation from earlier work. Our framework is general enough that it nests previous characterizations of the role of firms in wages. If $\kappa^P = 0$, the persistent part of the match component evolves independently of the firm's fortunes; if $\sigma^2_{\psi^P} = 0$ the match productivity component changes only in response to firm-related permanent shocks.

Employment and Job-to-Job Mobility In thinking about the dynamics of earnings, a key issue is controlling for selection into work and for job mobility, both of which may truncate the distributions of shocks. For example, if there is a large pass-through of firm level shocks onto wages, the worker may actually quit his job rather than suffer the resulting pay cut, which may even be permanent. Similarly, workers with large pay cuts in firms that have had bad productivity shocks may be more likely to accept alternative job offers. Observationally, there may be two workers paid exactly the same - one of whom moves, while the other does not - just because of the different reasons for observing a pay cut. In one case it may be because of an adverse firm level shock, while in the other a negative individual productivity shock that is carried everywhere.[4]

We model the employment decisions E as:

$$E_{i,t} = \mathbf{1}\left\{z'_{i,t}\delta + \phi\left(P_{i,t} + \varepsilon_{i,t} + v_{i,j(t_0),t}\right) + u^E_{i,t} > 0\right\} \tag{8}$$

The decision to work depends on the stochastic component of wages $P_{i,t} + \varepsilon_{i,t} + v_{i,j(t_0),t}$. A more general specification not pursued here would allow a different impact of the transitory and the permanent components because the former only causes substitution effects, while the latter also causes wealth effects (see Blundell, Pistaferri and Saporta-Eksten, 2017). The coefficient ϕ in part reflects the incentive effect of working but also the importance

[4]Positive shocks work in reverse, lowering quits and reducing the likelihood of a move to an alternative employer. We discuss below that allowing for asymmetric effects appears not to affect our findings much.

of unobserved heterogeneity in participation choices.[5] In other words, in the absence of exclusion restrictions that would allow us to distinguish the causal impact of wages from heterogeneity this coefficient captures both. This is sufficient for our purposes of controlling for censoring due to labor market transitions. Other observable wage components (as well as taste shifter variables such as age) are summarized in z.

Similarly, job-to-job mobility is defined as:

$$J_{i,t} = \mathbf{1}\left\{z_{i,t}'\theta + b\left(v_{i,j(t),t}^{init} - v_{i,j(t_0),t}\right) + u_{i,t}^{J} > 0\right\}, \tag{9}$$

and is also affected by a set of variables z, such as age. Job mobility depends only on the difference in match values between new and incumbent firms, $(v_{i,j(t),t}^{init} - v_{i,j(t_0),t})$, and not on the remaining stochastic components, because permanent and transitory productivity shocks do not depend on a particular firm match but are portable characteristics of a worker across different jobs. The importance of wage differences as opposed to worker observable characteristics in determining mobility is captured by the parameter b.

Finally, both the employment and the mobility equation depend on stochastic shocks, respectively $u^E \sim N(0,1)$ and $u^J \sim N(0,1)$. These shocks reflect exogenous job destruction and mobility (or lack thereof) due to unexplained random factors. In other words workers may move to unemployment despite an attractive wage or may move to a job paying less than the current one for unobserved reasons, or indeed may not move despite an excellent alternative offer. The two stochastic components also reflect unobserved tastes for work or job mobility. Finally, the observed characteristics in the two equations also reflect labor market attachment and employment and mobility costs.

[5]By participation we always mean employment versus non-employment. We use the terms interchangeably.

Labor Market Frictions and Job Offers Upon entry in the labor market, workers receive job offers at a rate λ_{init}. In subsequent unemployment spells, job offers are received at an age-dependent rate $\lambda_U = \lambda_{U,0} + \lambda_{U,1} \cdot age$. The age dependency is, of course, testable. Job offers while employed are subsumed into age-dependent mobility preferences in equation (9), since the two cannot be separately identified. If a worker receives a job offer while employed, we also model the origin of the offer to match transition patterns across broad categories. We classify firms according to their sector and size, and we assume that the probability of new offers from any given sector and size group depend on the current job, i.e.:

$$Pr\left(sector,\, size\right)_t = \omega_0 + \omega_1 \mathbf{1}\left\{sect_{j(t)} = sect_{j(t_0)}\right\} + \omega_2 \mathbf{1}\left\{size_{j(t)} = size_{j(t_0)}\right\}. \qquad (10)$$

This specification can capture the empirical fact that two-thirds of job-to-job moves occur within the same sector and 50% across similarly sized employers, see Table 7 for details.

3 Data

Our empirical analysis uses a matched employer-employee data set that combines information from four different data sources, compiled by Statistics Sweden. The first is the Longitudinal Database on Education, Income and Employment (LOUISE) that contains information on demographic and socioeconomic variables for the entire working age population in Sweden from 1990 onward. We use information about age, gender, municipality of residence, number and ages of children, marital status, education level as well as the collection of public transfers such as disability, public pension, sickness, unemployment and parental leave benefits. All variables in LOUISE are available on a yearly basis.

The second data set is the Register-Based Labour Market Statistics (RAMS) that contains information about the universe of employment spells in Sweden from 1985 onward. On the worker side, RAMS registers the gross yearly earnings and the first and last remunerated month for each employment/firm spell, as well as firm and plant identifiers. On the firm side, RAMS registers information about industry and the type of legal entity for all firms with employees.

The third data set is the Structural Business Statistics (SBS), which contains accounting and balance sheet information for all non-financial corporations in Sweden from 1997 onward, and for a subset of corporations during the 1990–1996 period.

The final data set is the Unemployment Register, containing all spells of unemployment registered with the Public Employment Service.

Since the SBS covers all non-financial corporations in Sweden only from 1997 onward, we focus the analysis on the period 1997–2008. We include all firms with the legal entity being limited partnership, limited company other than banking and insurance companies, and exclude sole traders because data for these firms are not available for the entire period. The final sample represents 83 percent of value added and 83 percent of employment in the Swedish private sector over the 1997–2008 period.

Table 1 presents descriptive statistics for the firms in our data set. The data includes 98,630 unique firms and 678,792 firm-year observations. The four sectors construction, manufacturing, retail and services account for 17%, 19%, 24% and 40% of all firms in the sample, respectively. Within sectors, larger firms display, on average, higher revenue per worker. The growth rate of revenue per worker does not follow the same pattern across sectors. For construction, larger firms grow more slowly on average, whereas growth rates are higher for larger firms in the other sectors.

We include all individuals who work at firms in our sample at some point during the 1997–2008 period. We use the data from RAMS together with registrations of unemployment

Table 1: **Summary statistics, firms**

	Firm size: number of employees			
	<20	20–50	50–100	100+
A. Construction				
No. unique firms	11,699	973	1,738	1,143
Value added per worker	494,117	533,599	552,487	566,002
Growth, log va/worker	0.0377	0.0368	0.0338	0.0229
B. Manufacturing				
No. unique firms	11,654	2,696	9,971	1,166
Value added per worker	528,331	578,780	622,656	1,017,031
Growth, log va/worker	0.0282	0.0211	0.0136	0.0133
C. Retail Trade				
No. unique firms	20,016	2,244	555	402
Value added per worker	528,697	625,053	634,861	756,280
Growth, log va/worker	0.0303	0.0235	0.0260	0.0182
D. Services				
No. of unique firms	32,938	3,847	1,010	826
Value added per worker	575,102	682,203	861,547	775,826
Growth, log va/worker	0.0382	0.0375	0.0430	0.0302

Note: Revenue per worker is reported in real SEK for base year 2007.

at the Public Employment Service to define employment on a quarterly basis. We use daily unemployment records to measure the exact length of employment spells. For individuals with multiple jobs during a quarter we keep the main employment, defined as the employment that accounts for the largest share of quarterly earnings. We define a worker as employed if he is working at least 2 months for any employer during the quarter. In each quarter, we record if an individual is a job mover, a job stayer or an entrant from non-employment. Average monthly earnings are recorded based on the yearly earnings and the number of remunerated months as registered in the RAMS data.

We exclude individuals until the last year that they receive public study grants (typically, young workers at the beginning of their working life who are still completing their formal education). We also exclude individuals from the first year that they receive disability benefits, occupational pension or public pension benefits (typically, workers at the end of their working life). We further exclude individuals when they move to a workplace that is not in the firm sample (typically, these are moves to the public sector, a financial corporation, or self-employment). Importantly, however, we keep all the records of non-employment that are in connection with employment spells at the firms in our sample.

In this paper we focus on men only. Results for women are much harder to interpret given that earnings variation reflects changes in both hours and productivity.[6] We estimate the model separately for each of two education groups: workers with at most high school education ("low skill") and workers with at least some college education ("high skill"). We take as given education choices and restrict our estimation sample to individuals age 26-55 for both education groups.

Table 2 presents summary statistics for each group of workers. Workers with lower education are on average slightly older, which reflects changes in years of schooling across cohorts. Workers with lower education are also less likely to have children living at home. The employment rate increases with education, but the fraction of employed workers who remain at their current job each quarter is fairly constant across groups. More educated workers are more likely to move from job to job, and less likely to enter a new job from non-employment. The data indicate that job-to-job mobility and transitions between employment and non-employment are fairly common. Each quarter, 2–3 percent of employed workers change jobs and 2–4 percent enter employment after a period of non-employment.

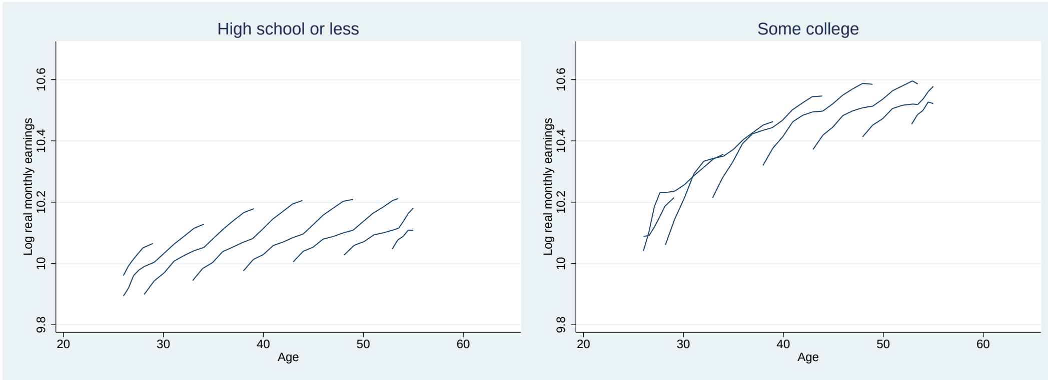
Life-cycle earnings Table 2 also reveals some important differences in earnings across education groups. We take a more detailed look at life-cycle earnings profiles in Figure 1,

[6]In 1997 (our first sample year), the part-time employment rate (defined as the fraction of employed workers who work less than 30 hours per week in their main job) was 6.5% for men and 23% for women; in 2008, the rates were 10% and 20%, respectively (source: OECD). For women of child-bearing age, it is also more frequent to observe shifts from full-time to part-time work and *vice versa*, making the analysis of earnings volatility in administrative data much more challenging. In an earlier working paper version of the paper, we documented that wage variances for women exhibit a hump-shaped pattern over the life cycle. Given these differences, we defer the study of women's earnings dynamics to future work.

Table 2: **Summary statistics, Male workers**

	≤High school	College
No. unique workers	1,290,727	458,622
No. worker-quarter obs.	34,130,603	10,900,427
Monthly earnings,	24,562	36,095
(2008 SEK)	(7,899)	(17,097)
Age	38.16	39.00
Married	0.5181	0.6102
Having children	0.4196	0.4979
Employed, of which	0.8671	0.9063
Job stayer	0.9491	0.9496
Job mover	0.0248	0.0317
Re-entrant	0.0261	0.0186
Industry		
Construction	0.1518	0.0580
Manufacturing	0.4052	0.3602
Retail Trade	0.1878	0.1394
Services	0.2552	0.4424
Firm size		
≤20	0.3156	0.2611
20–50	0.1406	0.1214
50–100	0.1016	0.0884
100+	0.4421	0.5291

Figure 1: Log real monthly earnings for five-year cohort groups against age, 1997-2008

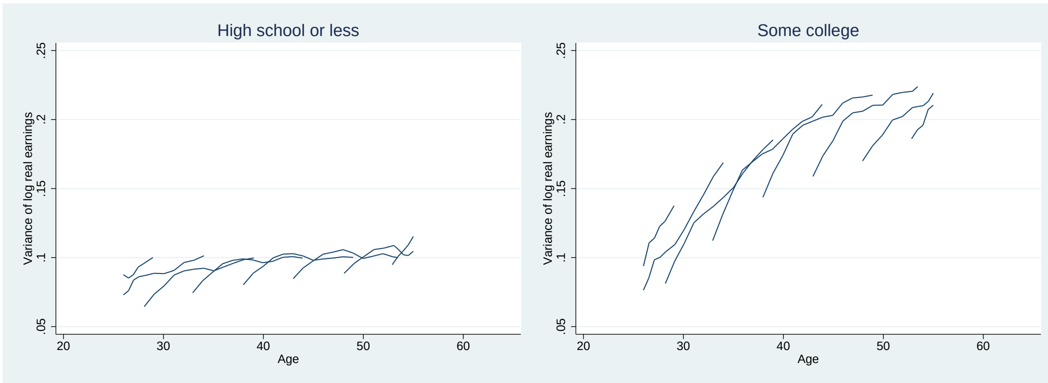


using observations for different birth cohorts in the data. In particular, for each education group we construct five-year cohort groups and separately plot their average log earnings over the age span in which we observe this particular cohort. The vertical distance between earnings of different cohort groups at a given age can then be interpreted as cohort effects, while the overall slope of the profile can be interpreted as reflecting age effects (ignoring for simplicity the usual age, time, cohort identification issues).

Overall, we observe the familiar life-cycle earnings profile increasing quite rapidly early in the career and then flattening or slightly decreasing towards the end of the life-cycle. Level-differences show the absolute gain from achieving a higher level of education. There seem to be some modest, but positive cohort effects (with new cohorts being more productive than older cohorts at each point of the life cycle).

The first moment of wages may give only a partial description of the life cycle evolution of earnings. Figure 2 presents the evolution of the variance of residual log real earnings, obtained after removing year and age effects. The patterns here display striking differences between education groups. While for the higher education group the variance increases by age, as has often been noted in US data (Meghir and Pistaferri, 2004), for lower education men the variance is either flat or increases at a very low rate. The lifecycle variance profile for those with some college is consistent with a random walk (or possibly heterogeneous

Figure 2: The variance of log real monthly earnings for five-year cohort groups against age, 1997-2008



age profiles). However the profile for those with high school or less is more consistent with stationary wages over the life cycle. Hence within-group inequality is increasing among the higher educated, but not among the lower educated.

Participation and job transitions The top-left graph in Figure 3 presents the employment rate by age for each education group. In our sample employment rates are above 75% for all age groups. The lower the achieved level of education, the lower is participation at young ages. Interestingly, there is an increase in participation from the beginning of individuals' careers until their mid-50s for high-school graduates, whereas participation for workers with some college education quickly levels off at around 90%. The figure also shows a substantial drop in employment after age 55 for both education groups.

The bottom panels of Figure 3 shows that young workers across both education groups have high quarterly job separation and re-entry rates when out-of-work. The entry rate from non-employment is rapidly falling with age and comparable across education groups, but the respective separation rates are higher for low-educated workers. As a result, the share of unemployed workers differs across groups. As the employment, separation and reentry rates illustrate, transitions in and out of employment are an important feature of the labor market.

The top-right panel in Figure 3 presents the quarterly job-to-job transition rates by age

Figure 3: Quarterly employment and job transition rates by age and education

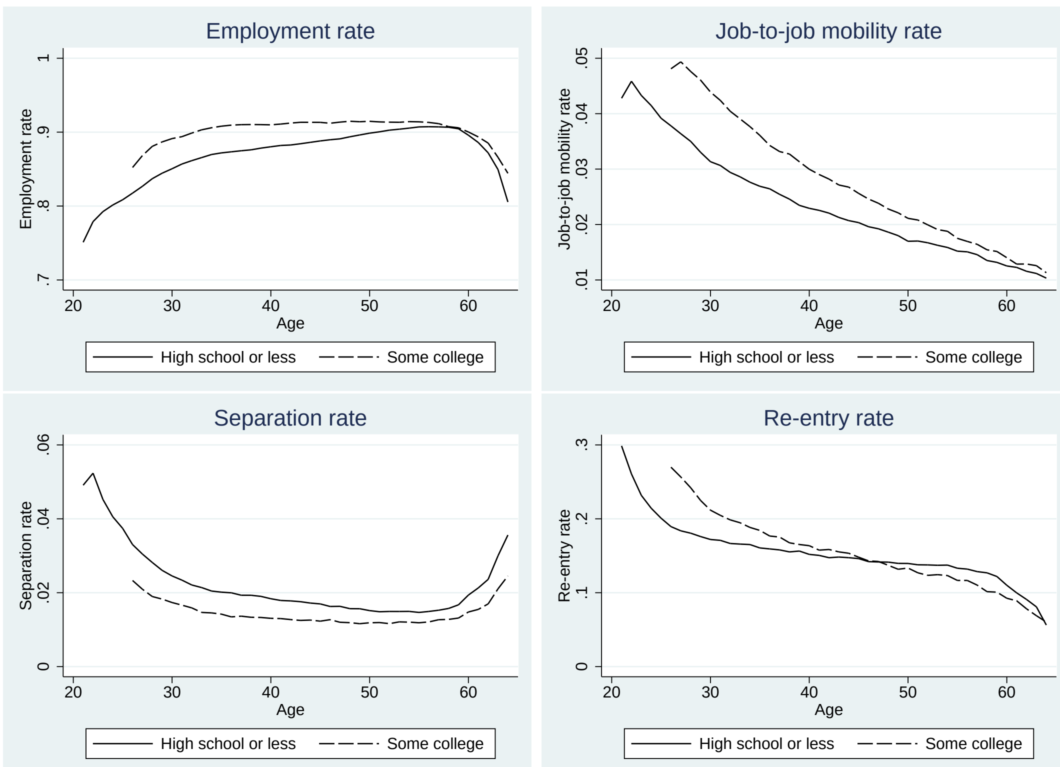
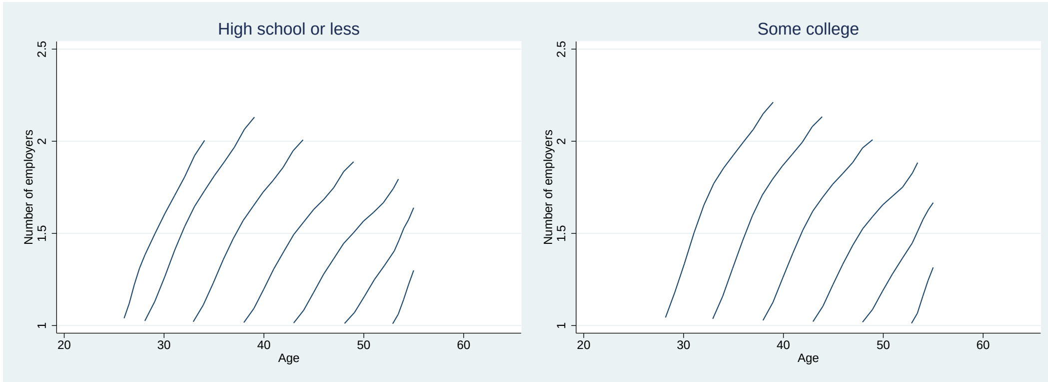


Figure 4: Number of jobs by age and education in five-year cohort groups



for each education group. The frequency of job to job transitions is particularly high at younger ages. Workers with at least some college switch employers more frequently than less educated workers.

Finally, Figure 4 presents the average number of employers for each cohort that we observe over the sample period from 1997 to 2008, conditioning on age in 1997. The figure confirms that job-to-job mobility is an important feature of the labor market. Individuals aged 20 to 25 in 1997 had on average more than 2 employers in the period between 1997 and 2008.

Table 3 reports the amount of wage variance that can be attributed to differences between firms.[7] The results show that most of the variance of wages is, in fact, within firms. For low skill workers this remains stable over time. However, for high skill workers the share of between firm variance is increasing over time. This increase is in line with recent findings by Card, Heining, and Kline (2013) for Germany and motivates the investigation of the role of firms for wage inequality and wage dynamics in our paper.

Mobility and wages In Table 4 we describe mobility patterns between firms sorted by the average wage they pay, and describe the way wages change between jobs when mobility

[7]We obtain these values by a standard decomposition of the total wage variance into between- and within-firm contributions.

Table 3: **Share of between firm wage variance**

	At Least Some College		High School or Less	
Year	Log Earnings	Residual	Log Earnings	Residual
1997	37.09%	38.44%	38.74%	35.57%
2000	38.96%	39.95%	37.19%	35.01%
2004	42.32%	40.64%	38.28%	36.21%
2008	42.06%	40.70%	38.70%	36.60%

The proportion of the cross sectional variance attributable to variation between firms. Residual refers to the variance after controlling for age and cohort effects, within each education group.

does not involve an unemployment spell in between jobs, separately for low- and high-skill workers. We compare wage growth in the year before the job move to the year after the job move, conditional on no other transition happening in this three-year window. Among job-to-job movers about 48% of low skill workers and 46% of high skill workers move to a firm of the same wage quartile level; in both groups, slightly less than 30% move to a higher-paying firm, and about a quarter to a lower-paying one.

Table 4: **Job Mobility and Wage Growth**

		Low skill workers *Departing firm quartile*											
		Share of transitions				Log wage growth				Share wage cuts			
		1	2	3	4	1	2	3	4	1	2	3	4
Arriving firm quartile	1	0.145	0.071	0.038	0.014	0.063	0.026	–0.004	–0.021	0.337	0.410	0.482	0.475
	2	0.068	0.110	0.059	0.016	0.116	0.048	0.035	0.008	0.261	0.346	0.401	0.452
	3	0.041	0.077	0.144	0.042	0.168	0.083	0.051	0.034	0.199	0.277	0.324	0.391
	4	0.019	0.026	0.048	0.081	0.184	0.127	0.092	0.066	0.198	0.232	0.283	0.322
		High skill workers *Departing firm quartile*											
		Share of transitions				Log wage growth				Share wage cuts			
		1	2	3	4	1	2	3	4	1	2	3	4
Arriving firm quartile	1	0.114	0.050	0.033	0.016	0.099	0.056	0.044	–0.020	0.294	0.341	0.363	0.465
	2	0.060	0.107	0.067	0.022	0.134	0.093	0.083	0.055	0.211	0.246	0.294	0.360
	3	0.036	0.080	0.152	0.061	0.177	0.114	0.085	0.064	0.187	0.218	0.275	0.354
	4	0.023	0.033	0.062	0.084	0.186	0.153	0.141	0.100	0.217	0.215	0.232	0.311

On average wage movers experience positive wage growth, unless they move from the very top firms to the very bottom ones (in terms of wage quartile). However, this average experience is masking a very large number of wage cuts: for both groups of workers between 20%-50% experience some wage cut when moving from one firm to another. The size of the wage cut depends very much on the direction of the move. Our model allows for such wage cuts: the motive for changing jobs, expressed in equation 9, trades-off wage improvements to other observed and unobserved reasons for mobility. However, many search models do not allow for wage cuts: The Burdett-Mortensen wage posting model excludes them, unless one rigs the model to force some random transitions. The model by Postel-Vinay and Robin (2002a) does allow for wage cuts: the worker may choose to move to a firm where the match surplus is higher; he may wish to pay for this move in terms of a lower upfront wage because of the option value of future wage increases. Finally, in Lise, Meghir, and Robin (2016) wage movers are either improving their match or are moving away from a firm that has suffered a productivity shock. This formulation allows for a much more flexible relationship between wage changes and mobility. The large prevalence of wage cuts surrounding job-to-job mobility is an indicator of the importance of such shocks in determining mobility and our model allows us to assess this.

4 Estimation Strategy

The estimation of the model is complex because of the combination of dynamics, endogenous selection into work and mobility and the unobserved factor structure. To address these complexities, we proceed in three steps. First, we estimate the stochastic process of firm-level productivity and treat the results as an input into the model estimation. Second, we estimate wage residuals from a selection model that takes into account the quarterly frequency of job mobility and the annual frequency of earnings. Finally, we estimate the full model using simulated method of moments based on the wage residuals, quarterly transition rates and firm-level shocks.

4.1 Firm Productivity Shocks

The source of stochastic variation that we are directly interested in are the productivity shocks to firms. We distinguish between permanent and transitory shocks because we can expect them to have very different impacts on wages. For example in a world with adjustment costs on either wages or employment we can expect the firm to smooth over transitory shocks but consider adjustments in response to a permanent change (see also Guiso, Pistaferri, and Schivardi (2005)).

The key point is that by observing data on firms we are able to measure shocks to their productivity directly (instead of relying on proxies such as employment, which may be subject to inaction bias due to the presence of adjustment costs). We can then relate these shocks to wages, since firms are also matched to individual records. Our measure of productivity is log value added (VA) per worker, which is observed annually.

We first run a regression of log VA per worker controlling for industry, municipality, firm size and year effects and save the residuals of this regression. Thus the shocks to firm level productivity that we use are purely idiosyncratic and do not include aggregate, regional, scale or industry effects. In Table 5 we show the autocovariance structure of firm productivity across all firms and separately by industry. From these results it seems that a random walk with an i.i.d. transitory component is a good approximation of the stochastic structure of VA per worker because the second and third order autocovariances for productivity growth in the data are close to zero for all sectors.[8]

Based on this empirical pattern, we assume that the stochastic process of log productivity for firm j observed in year t and quarter q, denoted a_{j,t_q}, can be decomposed into permanent and transitory components,

$$a_{j,t_q} = a^P_{j,t_q} + \xi^T_{j,t_q} \tag{11}$$

[8] While some of these autocovariances are statistically significant, they are economically negligible (in all cases considered, second- and third-order autocovariances are an order of magnitude smaller than first-order autocovariances).

Table 5: Autocovariance of log Value Added per Worker: Data

	Value Added per Worker: Data				
	All firms	Construction	Manufacturing	Retail	Services
$\mathrm{Var}\left(\Delta A_t\right)$	0.2295	0.1950	0.1822	0.2182	0.2754
	(0.0017)	(0.0032)	(0.0033)	(0.0032)	(0.0031)
$\mathrm{Cov}\left(\Delta A_t, \Delta A_{t-4}\right)$	-0.0666	-0.0709	-0.0520	-0.0607	-0.0768
	(0.0009)	(0.0020)	(0.0019)	(0.0017)	(0.0017)
$\mathrm{Cov}\left(\Delta A_t, \Delta A_{t-8}\right)$	-0.0045	-0.0012	-0.0052	-0.0047	-0.0053
	(0.0005)	(0.0010)	(0.0009)	(0.0008)	(0.0010)
$\mathrm{Cov}\left(\Delta A_t, \Delta A_{t-12}\right)$	-0.0022	-0.0034	-0.0028	-0.0014	-0.0018
	(0.0005)	(0.0011)	(0.0008)	(0.0008)	(0.0010)

A denotes log of annual productivity. Time is measured in quarters. So t and $t-4$ are one year apart.

where

$$
\begin{aligned}
a^P_{j,t_q} &= a^P_{j,t_{(q-1)}} + \xi^P_{j,t_q} \\
\xi^P_{j,t_q} &\sim N\left(0, \sigma^2_{\xi^P}\right) \\
\xi^T_{j,t_q} &\sim N\left(0, \sigma^2_{\xi^T}\right).
\end{aligned}
$$

where the subscript on the time index t denotes the quarter of that year and varies from 1 to 4.[9] In the data, we can only construct annual productivity (while our model is quarterly) which means we cannot identify an MA component within year. We denote annual productivity by e^{A_t}, which can be be related to the underlying quarterly measure by

$$
e^{A_t} = e^{a_{t_1}} + e^{a_{t_2}} + e^{a_{t_3}} + e^{a_{t_4}}
$$

where we drop the firm subscript j for convenience.

We apply simulation-based estimation to estimate the quarterly firm-shock process. Given the parametric assumptions of the quarterly shock process, we make guesses about the parameter vector $\left\{\sigma^2_{\xi^T},\ \sigma^2_{\xi^P}\right\}$ and simulate firm productivity for a set of hypothetical firms.

[9] If $q = 1$ then $q - 1$ refers the last quarter of the previous year.

We then aggregate these simulated shocks to replicate the structure of the actual data. The quarterly shock process for log VA per worker is additive. As a result, the annual log VA per worker can be written as

$$A_t = a^P_{(t-1)_4} + \log\left[\sum_{k=1}^{4} exp\left(\xi^T_{t_k} + \sum_{s=1}^{k} \xi^P_{t_s}\right)\right].$$

Analogously, the value in the second year is

$$A_{t+1} = a^P_{(t-1)_4} + \sum_{k=1}^{4} \xi^P_{t_k} + \log\left[\sum_{k=1}^{4} exp\left(\xi^T_{(t+1)_k} + \sum_{s=1}^{k} \xi^P_{(t+1)_s}\right)\right]$$

and the analytical expression for annual growth in log VA per worker is

$$A_{t+1} - A_t = \sum_{k=1}^{4} \xi^P_{t_k} + \log\left[\sum_{k=1}^{4} exp\left(\xi^T_{(t+1)_k} + \sum_{s=1}^{k} \xi^P_{(t+1)_s}\right)\right] - \log\left[\sum_{k=1}^{4} exp\left(\xi^T_{t_k} + \sum_{s=1}^{k} \xi^P_{t_s}\right)\right]$$

The important point is that the initial conditions drop out of the expression. To estimate the parameters of the productivity process we define a set of auxiliary moments that can be easily computed in the data as well as from the simulation. We choose the structural parameters that minimize the distance between these moments in the model and in the data. In particular, we identify the underlying parameters of the shock process from the variance and first-order autocovariance for the annual change in firm productivity.

Table 6 reports the estimation results for the standard deviations of the shocks on a quarterly basis. The implied process for quarterly value added per worker shows sizable transitory shocks, which are similar across industries: this implies considerable mean reversion. However, the permanent shocks are also substantial, implying quite volatile firm level productivity. This in itself is an important result and consistent with what Guiso, Pistaferri, and Schivardi (2005) find.[10] These estimates will be used to draw firm shocks in the simula-

[10]If we shut down the transitory shock the annualized standard deviation of the permanent shock is 0.212.

Table 6: Results: Quarterly Firm-Shock Process

	All firms	Construction	Manufacturing	Retail	Services
σ_{ξ^T}	0.4758	0.4804	0.4335	0.4598	0.5021
	(0.0016)	(0.0034)	(0.0032)	(0.0029)	(0.0027)
σ_{ξ^P}	0.1303	0.1003	0.1199	0.1319	0.1442
	(0.0007)	(0.0019)	(0.0012)	(0.0012)	(0.0012)

Standard errors obtained using the bootstrap.

tion estimation procedure below. There are some interesting differences between industries, with services being most volatile.

One issue concerns measurement error. It is not possible to distinguish measurement error from the variance of the transitory shock. This means that we may well be overstating the variance of the transitory component. This will imply understating the transmission of the transitory shocks to wages. Under the assumption of orthogonality of transitory and permanent shocks however, the pass-through coefficient for permanent shocks is unaffected.

4.2 Wage Residuals

In the next step, we use individual-level earnings and labor market participation to estimate the effects of individual characteristics on wages (γ) in equation (1). Based on this first stage, we can then use the wage residuals $\tilde{e}_t = (P_{i,a,t} + \varepsilon_{i,a,t} + v_{i,j,a,t})$ as the relevant input into the model estimation. In what follows we use interchangeably earnings and wages. The administrative data does not include information on hours, so to the extent that some fluctuations reflect changes in hours of work during the work spell we will not be able to distinguish this from other sources of fluctuations. This point may be particularly pertinent for women, which is a reason why we do not model their income process in this paper and focus only on men.

The estimation applies a modified Heckman two-step procedure that accounts for selec-

Similarly the annualized standard deviation of the transitory shock is 24.6%.

tion into work and for the discrepancy in data frequency between model and data. In the model, we assume that all decisions of individuals and firms happen at a quarterly frequency. Yet, in the data we only observe wages as an annual average over all quarters. As a result, our observed outcome variable in levels is the average quarterly wage for those who have worked at least one quarter,

$$w_t = \frac{\sum_{q=1}^{4} E_{t_q} \times w_{t_q}}{\sum_{q=1}^{4} E_{t_q}}.$$

where t_q is the q quarter in year t and $E_{t_q} = 1$ denotes working in that quarter.

We start by estimating a discrete choice model for employment (E_{t_q}) for each individual at a quarterly frequency and construct the Mills ratio (λ^M) for each of these periods. To make the model consistent with the data, we aggregate these quarterly selection correction terms in the annual wage model. If the error term follows a log normal distribution, the log of the conditional expectation of observed average quarterly wages is given by

$$\log \mathbb{E}\left[w_t \,|\, x_t, z_t,\, E_{t_q} = 1\, \forall q = 1, ..., 4\right] = x_t'\gamma + \log\left[\sum_{q=1}^{4} E_{t_q} \times e^{\rho\lambda^M(z_{t_q}'\delta)} \Big/ \sum_{q=1}^{4} E_{t_q}\right] + \frac{\sigma_v^2}{2}, \quad (12)$$

where again we omit the firm subscript j and we take the x characteristics as constant within the year (for simplicity). The last term in this equation explicitly shows the bias from aggregating individual wage information at annual frequency, even though wages are determined at a higher frequency.[11] The additional variance term $\frac{\sigma_v^2}{2}$ will be absorbed by the constant term in the regression. The second term is a nonlinear function of quarterly Mills ratios $\lambda^M\left(z_{t_q}'\delta\right)$. This term implies that seasonality of participation decisions can introduce a second bias when running a simple linear specification of log wages on individual characteristics, even when controlling for selection. If some of the decision criteria for partic-

[11]This aggregation bias term is reminiscent of the bias due to individual heterogeneity in Blundell, Reed and Stoker (2003) when analyzing aggregate wages.

ipation z_{t_q} change at quarterly frequency, a nonlinear specification is needed that accounts for seasonal changes in participation when aggregating employment choices to the annual level. The estimation approach based on equation (12) then controls for these two sources of aggregation bias that occur because of data availability and can be used to get consistent estimates of γ.

Equation (12) is estimated separately for our two broad education groups (less than college and some college or more). Within each category there are more detailed educational levels (i.e., grades completed) and we control for these. We also include industry dummies and a fourth-order polynomial in age. Since our selection equation also includes demographic characteristics, which we do not wish to use as exclusion restrictions, we also include marital status and dummies for children in different age groups as well as region-fixed effects. Time dummies are used to control for aggregate trends. Finally, we acknowledge the role of measurement error in employment. For example, it is quite common for individuals in Sweden to receive some payments from their employers while on parental leave. If these payments are sufficiently high, then those individuals will be falsely considered employed and will appear as particularly bad working types in the data even though they should be considered out of work during that period. These cases would lead to overestimating the amount of low-productivity types in the labor market and will bias the estimation results.[12] In order to address this type of measurement error, we directly include controls for parental leave and sickness benefits.

The same set of control variables used in the wage equation are also included in the participation choice equation, but we use region-time fixed effects in the quarterly participation equation as excluded instruments to estimate the selection effect. These instruments are motivated by the fact that income taxes in Sweden are determined at a community level

[12]Note that the familiar result of consistent estimates despite measurement error in the dependent variable does not apply for the participation equation because we estimate a nonlinear model. See Hausman (2001) for more details.

and the cost of living, in particular housing or rental prices, differs widely across regions and over time. As a consequence, the opportunity cost of work differs across regions and time. However, we assume that the labor market is integrated and that, other than fixed regional effects and time effects, the interactions can be excluded (see for example Blundell, Duncan, and Meghir (1998)). We use the residual from the estimated participation regression, $\tilde{u}_t$ to construct some key moments for identification (detailed next).

4.3 Full Model Estimation

4.3.1 Simulation

We now estimate the remaining parameters defining individual careers and wages using the simulated method of moments (McFadden, 1989, Pakes and Pollard, 1989). Each set of parameters is estimated for the lower and higher education groups separately.[13] The approach requires us to simulate wages and career paths, including transitions between employment and unemployment and between jobs.

We draw from the distribution of idiosyncratic shocks to determine the stochastic evolution of individual productivity (which is estimated simultaneously with the entire model) and from the distribution of permanent and transitory firm level shocks, which we pre-estimated. To construct the firm level shocks, such that a large number of workers receive the same ones (because they work together) we need to allocate workers to firms in the simulation. To do so, we create two firm identities for each size/sector bin. The model generates offers for each of these bins, based on the probability model (10). We then allocate the individual randomly to one of the firms with equal probability. The population of workers within a firm then receives a permanent and transitory firm shock drawn from the distributions that

[13]We list these here for convenience: the parameters determining participation (δ and ϕ), job-to-job mobility (θ and b), the transmission of firm-related shocks (κ^P and κ^T), the parameters of the stochastic processes determining wage dynamics ($\rho,\ \sigma^2_P,\ \sigma^2_\zeta,\ \sigma^2_\varepsilon,\ \sigma^2_{\psi^P},\ \sigma^2_{\psi^T}\ \sigma^2_{\psi init}$), the job arrival rate coefficients ($\lambda_{init},\ \lambda_{U,0},\ \lambda_{U,1}$) and the coefficients determining the source of outside offers ($\omega_0,\ \omega_1,\ \omega_2$).

were estimated in advance before, as described. The key point is that we have groups of workers with the same shocks; this will allow us to use the spatial correlation of wages within a firm to identify the transmission coefficients. Once we estimate these career paths we compute moments from the simulated data to match them to those from the actual matched employer-employee dataset. The wages in the data are the residuals we constructed earlier. The full set of moments is described in the section below.

4.3.2 Data Moments and Identification

This section describes the choice and computation of the data moments to estimate the model. In particular, we emphasize challenges because of different data frequencies. Since different moments simultaneously contribute to pin down the structural parameters, the identification discussion in this section is naturally informal.

The first set of moments we use are quarterly participation and job mobility rates by age group (26-30, 31-35, 36-40, 41-45, 46-50, 51-55). These help identify the deterministic part of the participation and job-to-job transition equations (δ and θ). The second set of moments includes quarterly job creation rates (fractions moving into work from unemployment) and job destruction rates (fractions moving from employment to unemployment) for the same age groups above. Moreover, we use job to job flows towards firms of similar size and industry. The job creation rate relates to the arrival rate of offers by age ($\lambda_{U,0}$ and $\lambda_{U,1}$) and the distribution of initial offers (λ_{init}). Quarterly job transition rates across sectors and firm size groups are directly related to the on-the-job offer probabilities (ω_0, ω_1 and ω_2). The shift over the life cycle of job mobility flows is crucial in estimating the impact of differences in firm-specific matches on the probability of a job-to-job move (the parameter b in equation (9)). The covariance between wage residuals and participation residuals (obtained as described in section 4.2) pins down the impact of wages on the decision to work (ϕ).

Quarterly job separations are endogenous and directly relate to transitory and permanent shock wage shocks. To distinguish "general" from "match-specific" wage shocks, we add annual moments related to earnings. Since the model assumes quarterly processes for all shocks, all simulation outcomes are quarterly as well. As a result, we need to aggregate simulated outcomes such as firm shocks and wages within each year to make the simulation comparable to the observed moments. Specifically, we use the variance and autocovariance of wage growth for stayers. The first-order autocovariance pins down the contribution of transitory fluctuations, leaving the variance of wage growth to identify the contribution of more persistent shocks.

We further distinguish match-specific and individual-specific shocks by comparing average wage growth for stayers and movers. Wage information in transition years is not very reliable because we do not know the exact timing for job-to-job mobility.[14] We therefore choose to not use wage information for these years and instead use mover information by looking at residual wage growth across periods before and after the switch occurred. We focus on workers with only one job move between periods $t-1$ and $t+1$, i.e. we compute, $\{\epsilon\}_{jj} = \tilde{e}_{t+1} - \tilde{e}_{t-1}$. We then use this residual wage growth measure to determine average wage growth and the variance of wage growth for movers, which in turn will be informative about the variance of match-specific effects ($\sigma^2_{\psi^{init}}$).

We target the level of residual earnings variance at the beginning of the life cycle to identify the variance of initial productivity (σ^2_P). The size of the autocorrelation coefficient in permanent productivity (ρ) is identified through the life-cycle pattern of the variance of residual earnings.

Some of the key structural parameters are the pass-through of firm-level shocks onto wages. To identify these parameters, we measure the share of variation in wage growth that is due to variation across firms, i.e. the share of wage growth explained by a common factor,

[14] While the dates are recorded in one data set, another records continuing payments to the worker well after he was recorded as having moved, within that year.

firm affiliation. This intraclass correlation of wage growth is defined as:

$$\rho_{\Delta\tilde{e}} = \frac{\sum_{\text{firms } j}\sum_{\text{worker } k\in j}\sum_{l\in j, k\neq l}(\Delta\tilde{e}_{kt} - \Delta\bar{e})(\Delta\tilde{e}_{lt} - \Delta\bar{e})}{Var(\Delta\tilde{e}_{it})\sum_j n_j(n_j - 1)}$$

where $\Delta\tilde{e}$ is residual wage growth. We complement this moment with the autocovariance of average wage growth among stayers to capture the mean reversion of transitory firm-level shocks. These two moments are closely related to the structural pass-through parameters κ^P and κ^T.

4.3.3 Model Simulation

Conditional on a guess for the parameter vector β, we simulate life-cycle behavior and wages for 800,000 workers in the model. We estimate the share of workers who receive an initial offer immediately when entering the labor market as an additional parameter. The initial source of the job offer reflects the sectoral distribution of employees by education group in Table 2.

In the data, employment and job mobility are observed at a quarterly frequency, whereas wages and firm shocks are only observed annually. We want to exploit the additional variation in choices and transition probabilities in order to estimate quarterly variances of permanent and transitory shocks to the worker. At the same time we acknowledge the coarser structure of the data with respect to wages in order to estimate the wage equation for example. In practice, we will assume quarterly processes for all shocks in the simulation, but in order to compute moments corresponding to the outcome in the real data, we need to aggregate firm shocks and wages within each year.[15]

The moments simulated from the model mimic the moments we compute from the data and hence any sample selection is controlled for. In order to exactly replicate the data

[15]This aggregation step requires aggregating in levels and then taking logs to maintain the properties of the wage shock process.

structure in the simulation, we use the empirical age distribution by education group as weights to compute the simulated moments from the model.

4.3.4 MCMC Estimation

We maximize the GMM objective function

$$L_n(\beta) = -\frac{n}{2}(g_n(\beta))' W_n(\beta)(g_n(\beta))$$

where $g_n(\beta) = \frac{1}{n}\sum_{i=1}^{n} m_i(\beta)$ and $m_i(\beta)$ is a vector of differences between simulated moments $\Gamma^S(\beta)$ and data moments Γ^D such that

$$E[m_i(\beta_0)] = E\left[\Gamma^D - \Gamma^S(\beta_0)\right] = 0.$$

The concerns raised by Altonji and Segal (1996) are particularly pertinent for our context, where we are estimating variances. As a result we use an equally weighted distance criterion, which we minimize to obtain our parameter estimates. Since the simulated moments may not be smooth, we use a Laplace-type estimator (LTE) following Chernozhukov and Hong (2003) to obtain this minimum. The main computational advantage of the LTE approach is that it uses functions of the criterion function that can be computed by Markov Chain Monte Carlo methods (MCMC). In particular, we use the Metropolis-Hastings algorithm with uniform priors. We transform the objective function $L_n(\beta)$ into a quasi-posterior:

$$p_n(\beta) = \frac{e^{L_n(\beta)}}{\int_{\beta \in B} e^{L_n(\beta)} d\beta}$$

and evaluate this function at the current parameter guess $\beta^{(j)}$ and at an alternative draw χ from a multivariate normal distribution. The parameter guess is then updated according to:[16]

$$\beta^{(j+1)} = \begin{cases} \chi & \text{with probability} \quad \pi\left(\beta^{(j)}, \chi\right) \\ \beta^{(j)} & \text{with probability} \quad 1 - \pi\left(\beta^{(j)}, \chi\right) \end{cases}$$

where

$$\pi(x, y) = \min\left(\frac{p_n(y)}{p_n(x)}, 1\right) = \min\left(e^{L_n(y) - L_n(x)}, 1\right).$$

Our estimator follows as the quasi-posterior mean

$$\hat{\beta} = \int_{\beta \in B} \beta p_n(\beta)\, d\beta,$$

which in practice can be computed as the average over all N_S elements of the converged Markov chain

$$\hat{\beta}_{MCMC} = \frac{1}{N_S} \sum_{j=1}^{N_S} \beta^{(j)}.$$

In practice, we estimate 100 chains of 40,000 elements per education group and we use the last 20,000 elements to compute $\hat{\beta}_{MCMC}$. [17]

This estimation strategy is a good fit for our problem because MCMC only requires many function evaluations $L_n(\beta)$ at different parameter guesses. The method is derivative-free and can deal with large parameter spaces and multiple local minima quite well.[18]

[Insert paragraph on standard errors HERE. [Is it the one below???]]

The method also allows us to easily compute standard errors for the estimates using the GMM sandwich formula $J^{-1}IJ^{-1}$. First note that I is given by

$$I = G(\beta_0)' W_n \Omega(\beta_0) W_n G(\beta_0)$$

[16] This acceptance procedure is similar in spirit to simulated annealing.

[17] The first 10,000 elements of the chain are computed based on a preset error variance. For the subsequent chain, we use adaptive MCMC to target the asymptotically optimal acceptance rate of 23.4% (**?**).

[18] See the discussion in Chernozhukov and Hong (2003) for more details.

where W_n is the weight matrix used in the estimation, $G(\beta_0)$ is the gradient at the true parameter vector β_0 and Ω is the variance covariance matrix of the moment function at the true parameter vector. We can easily estimate these objects at $\hat{\beta}_{MCMC}$. Finally, we can also compute the numerical Hessian to get an estimate for the inverse Hessian J^{-1}.

We obtain estimates for G and Ω through simulation. We first calculate each element j of the numerical gradient vector at the parameter estimate $\hat{\theta}$ as

$$\hat{G}_j = \frac{g\left(\hat{\beta} + h_j\right) - g\left(\hat{\beta} - h_j\right)}{0.02\hat{\beta}_j}$$

where g is the vector of moments that we evaluate at $\hat{\beta} + h_j$ and $\hat{\beta} - h_j$ respectively, in our case the vector of participation rates, mobility rates, wage growth moments, spatial correlation of wage growth etc. Lastly, h_j is a vector of zeros with one positive element at the j-th position equal to 1% of the parameter value $\hat{\theta}_j$, the j-th element of the vector of parameter estimates. In the next step, we calculate an estimate for Ω as:

$$\hat{\Omega} = n \cdot g\left(\hat{\theta}\right) g\left(\hat{\theta}\right)'.$$

Note that J^{-1} is the inverse Hessian that we estimate as:

$$\hat{J}^{-1} = \left(\hat{G} W_n \hat{G}'\right)^{-1}.$$

5 Results

5.1 Model Fit

Our model is overidentified and consequently considering the within sample fit can be informative on the performance of the model. In Table 7 we start with the stochastic properties of wage growth, which are of central importance for our purposes. The model replicates quite closely wage variation for both low and high education groups.

We report the variance and covariance properties of wage residuals. Despite the parsimonious stochastic structure of the model, we can fit dynamics of both levels and growth rates of wages quite well. The model generates life-cycle earnings profiles that are close but slightly steeper than in the data. At the same time, the variance of wage growth for stayers ($J = 0$) in the model is below the data moment. When it comes to job movers ($J = 1$), we only consider the growth in wages that occurs between the year before the move and the year after the move, as explained above. This reduces the effects of measurement error, induced by the fact that the earnings record is for the entire calendar year. The relevant statistics (the conditional mean $E(\tilde{e}_{t+1} - \tilde{e}_{t-1}|E_{t-1} = 1, E_{t+1} = 1, J_t = 1)$, and the conditional variance $V(\tilde{e}_{t+1} - \tilde{e}_{t-1}|E_{t-1} = 1, E_{t+1} = 1, J_t = 1))$ are presented at the bottom of the table. They are reproduced very accurately by the model. The last two moments are the covariance between an employment residual (from a linear probability model) and the wage residual, separately for stayers and movers. They help capturing the selection effect of employment decisions on wages.

One of the most important moments for our purposes is the spatial correlation of wage shocks. Since we measure this moment using wage growth residuals, it is unlikely to reflect correlation in wages due to sorting of similar workers into a firm. Rather, it reflects how changes in wages are correlated across individuals, most likely due to the fact that wages of all workers may be changing because of firm-level shocks. This correlations is quite high (0.18) and is closely reproduced by the model. Similarly, the model accurately matches the autocovariance of average wage growth within firms. In sum, the model captures rather well the way wages of workers in the same workplace move together from period to period.

In Table 8 we report the fit of the model for labor market transitions. Again, the model does a good job of capturing the age profile of entry, exit and job mobility (including the

Table 7: Stochastic Properties: Data and Simulation

	At least some college		High School or Less	
	Data	Model	Data	Model
$V(\tilde{e}_{age26})$	0.1149	0.1094	0.0926	0.0944
$V(\tilde{e}_{age30})$	0.1243	0.1298	0.0926	0.0912
$V(\tilde{e}_{age35})$	0.1505	0.1475	0.0926	0.0887
$V(\tilde{e}_{age40})$	0.1733	0.1655	0.0922	0.0901
$V(\tilde{e}_{age45})$	0.1876	0.1858	0.0922	0.0912
$V(\tilde{e}_{age50})$	0.1934	0.1988	0.0922	0.0927
$V(\tilde{e}_{age55})$	0.2008	0.2118	0.0916	0.0937
$V(\Delta\tilde{e}_t \mid E_{t-1}=1, E_t=1, J_t=0)$	0.0344	0.0328	0.0250	0.0240
$C(\Delta\tilde{e}_t, \Delta\tilde{e}_{t-1} \mid J_t=0)$	-0.0047	-0.0081	-0.0035	-0.0041
$E(\tilde{e}_{t+1}-\tilde{e}_{t-1} \mid E_{t-1}=1, E_t=1, J_t=1)$	0.0400	0.0383	0.0258	0.0248
$V(\tilde{e}_{t+1}-\tilde{e}_{t-1} \mid E_{t-1}=1, E_t=1, J_t=1)$	0.0668	0.0666	0.0537	0.0543
$C(E_j[\Delta\tilde{e}_t], E_j[\Delta\tilde{e}_{t-1}] \mid J_t=0)$	-0.0015	0.0007	-0.0010	-0.0021
$C(\tilde{u}_t, \tilde{e}_t \mid E_t=E_{t-1}=1, J_t=0)$	0.0003	-0.0004	-0.0002	-0.0002
$C(\tilde{u}_t, \tilde{e}_t \mid E_t=E_{t-1}=1, J_t=1)$	0.0192	0.0158	0.0031	0.0033
Spatial correlation coefficient (for stayers)	0.1822	0.1818	0.1783	0.1799

heterogeneity by education), which are all declining over the life cycle. Moreover, while the model replicates qualitatively the declining age pattern of unemployment, it predicts slightly higher participation rates than we observe in the data. Since in equation (10) we let job-to-job transition probabilities differ according to firm size and sector of origin, we add moments that capture such heterogeneity, namely the proportion of movers to a different industry, different firm size type, or both. The model captures extremely well such transitions.

5.2 Parameter estimates

Transitions We start by presenting results for the decisions to work and to move to another firm in Table 9.[19] Starting with employment, we find the expected increasing concave pattern in age (the δ parameters). The association of wages with participation is given by the coefficient ϕ in the table. The coefficient is positive and significant, with a notably higher value for high skill workers.[20]

[19] The results from the first step to obtain estimates of the effects of individual characteristics on wages (γ) and the wage residuals ($\tilde{e}$) are presented in the appendix.

[20] As noted earlier, this is a mix of a selection and an incentive effect and in this context we have no way of distinguishing the two, because we do not have appropriate exclusion restrictions. Nevertheless this is not

Table 8: Transitions: Data and Simulation

		At least some college		High School or Less	
	Age	Data	Model	Data	Model
Unemployment frequency	26-30	0.1220	0.1229	0.1644	0.1610
	31-35	0.0980	0.1027	0.1347	0.1325
	36-40	0.0900	0.0827	0.1234	0.1231
	41-45	0.0874	0.0752	0.1154	0.1155
	46-50	0.0862	0.0777	0.1061	0.1050
	51-55	0.0862	0.0921	0.0961	0.0984
Job creation frequency	26-30	0.2400	0.2226	0.1806	0.1768
	31-35	0.1945	0.2016	0.1659	0.1689
	36-40	0.1699	0.1816	0.1562	0.1594
	41-45	0.1548	0.1591	0.1480	0.1510
	46-50	0.1377	0.1348	0.1409	0.1425
	51-55	0.1231	0.1135	0.1367	0.1345
Job separation frequency	26-30	0.0194	0.0326	0.0283	0.0285
	31-35	0.0152	0.0220	0.0215	0.0253
	36-40	0.0134	0.0157	0.0192	0.0219
	41-45	0.0126	0.0128	0.0175	0.0193
	46-50	0.0120	0.0118	0.0158	0.0164
	51-55	0.0119	0.0125	0.0149	0.0143
Job mobility frequency	26-30	0.0458	0.0472	0.0336	0.0345
	31-35	0.0385	0.0347	0.0280	0.0278
	36-40	0.0319	0.0261	0.0241	0.0232
	41-45	0.0271	0.0212	0.0210	0.0193
	46-50	0.0227	0.0193	0.0182	0.0167
	51-55	0.0191	0.0182	0.0160	0.0146
Pr(E-to-E to new industry)		0.3573	0.3572	0.3372	0.3376
Pr(E-to-E to new firm size)		0.5066	0.5062	0.4779	0.4803
Pr(E-to-E to new industry and new size)		0.2144	0.2137	0.2118	0.2119

To interpret the size of the coefficient we report at the bottom of the table the marginal effect of a wage increase on employment. This turns out to be much higher for higher educated workers than the rest, implying a stronger combined effect of self-selection and incentives for the higher skilled group.

Table 9: Participation and job mobility

		At least some college		High School or Less	
Parameter	Description	Estimate	s.e.	Estimate	s.e.
Employment					
δ_0	Constant, participation	-0.1363	(0.0013)	1.5454	(0.0011)
δ_{age}	Age, participation	1.0158	(0.0007)	0.1309	(0.0006)
δ_{age^2}	Age squared, participation	-0.1070	(0.0001)	-0.0019	(0.0001)
ϕ	Wage residual	0.6379	(0.0021)	0.1761	(0.0010)
Marginal Effect of 10% wage change		0.219		0.088	
Job-to-job Mobility					
θ_0	Constant, mobility	-0.3507	(0.0030)	-1.1629	(0.0047)
θ_{age}	Age, mobility	-0.6557	(0.0018)	-0.2880	(0.0026)
θ_{age^2}	Age squared, mobility	0.0626	(0.0003)	0.0175	(0.0003)
b	Wage improvement	1.7185	(0.0141)	1.1877	(0.0059)
Marginal Effect of 10% wage improvement		0.982		0.598	

In the bottom part of Table 9 we look at the determinants of job-to-job mobility. We find that transitions across firms are decreasing in age, matching what we see in the data. The coefficient b is estimated to be large and positive, which shows that mobility choices are influenced by the wage difference between incumbent and poaching firm; this is true for both education levels. This limits the ability of the incumbent firm to lower wages as a result of shocks. However, mobility is not driven by wages only. Mobility costs that vary by age also matter, as do exogenous shocks.

Table 10 presents information on the transition process between jobs and sectors. High skilled workers have a substantially higher probability of job offers at labor market entry, λ_{entry}. The arrival rate of job offers over the life-cycle implies that at age 30, one job is

a threat to the identification of the stochastic process of wages, which is the central focus of this study.

Table 10: Estimation Results

		At least some college	High School or Less
Parameter	Description	Estimate (s.e.)	Estimate (s.e.)
	Job arrival rate		
λ_{init}	Arr. rate at entry	0.9723	0.7571
		(0.0012)	(0.0005)
$\lambda_{U,0}$	Arr. rate, subs. spells	0.3625	0.2337
		(0.0003)	(0.0001)
$\lambda_{U,age}$	Arr. rate, subs. spells (age shift)	-0.0047	-0.0018
		(0.0000)	(0.0000)
	Origin of offer		
Parameter	Description	Estimate (s.e.)	Estimate (s.e.)
ω_0	Different firm size & sector	0.2141	0.2113
		(0.0001)	(0.0001)
$\omega_0 + \omega_1$	Different size, same sector	0.2930	0.2688
		(0.0002)	(0.0001)
$\omega_0 + \omega_2$	Same size, different sector	0.1417	0.1251
		(0.0001)	(0.0001)
$\omega_0 + \omega_1 + \omega_2$	Same size, same sector	0.3512	0.3948
		-	-

sampled approximately every 2.9 quarters for the high skilled and every 4.9 quarters for lower skill workers. These rates decrease in frequency as workers age, but very moderately. However, there is an age profile in labor force participation, induced by the the age profile shown in Table 9.

In the bottom half of Table 10, the coefficients ω_k ($k = \{0, 1, 2\}$) show how offers to different firm types and industries vary. They imply that sampling jobs from other sectors is smaller than from the same sector.[21]

Stochastic process of individual productivity We first consider the stochastic process of wages that is unrelated to firms and which the worker carries from job to job. This is shown in Table 11. There are clear similarities across education groups, but also some important differences as we would expect when considering Figure 2. Wages at labor market entry show a remarkable dispersion. Thereafter the shocks are persistent, but much more so for the higher education group for which they are almost a random walk. The quarterly AR(1) coefficient is substantially smaller for low skilled workers, which leads to faster depreciation of initial conditions and past shocks. In particular, after 30 years in the labor market at age 55, worker productivity still contains 30% of the initial productivity draw for a college graduate, but only 3% for a high school graduate. Individual productivity shocks are only a part of the story driving wage fluctuations. The next key component are firm level shocks, to which we now turn.

Match value and transmission of shocks In Table 12 we show the key parameters for our study, namely the transmission of firm-related shocks onto wages. For workers with higher education transitory shocks are transmitted to workers at a rate of 0.09%. This is quite

[21] In the current version of the paper we have not explored the implications of such persistence because we have not allowed wage growth to depend on firm size or sector. We intend to consider this issue in future work in more detail.

Table 11: The stochastic process of individual productivity

		At least some college		High School or Less	
Parameter	Description	Estimate	s.e.	Estimate	s.e.
σ_ϵ	Transitory shock, wages	0.1814	(0.0004)	0.1225	(0.0002)
σ_ζ	Permanent shock, wages	0.0469	(0.0002)	0.0637	(0.0001)
ρ	AR(1) coefficient	0.9902	(0.0001)	0.9717	(0.0000)
σ_P	Initial perm. productivity, wages	0.3138	(0.0002)	0.2960	(0.0001)

low and consistent with earlier studies that find firms fully absorbing firm-level idiosyncratic shocks. It is worth pointing out that our data set is extremely large, which allows us to pick up even small effects. Permanent shocks, on the other hand, are transmitted to a much larger extent, with 35% of a firm permanent shock being transmitted to wages. Thus when the fortunes of firms change permanently, they change the wages of high skill workers permanently (or at least until job separation), implying a high degree of rent sharing. This result is qualitatively consistent with Guiso, Pistaferri, and Schivardi (2005) (see below for a more quantitative comparison highlighting the importance of accounting for job mobility and periods out of work). We also experimented with allowing for an asymmetric impact of shocks, depending on whether they were positive or negative, but we were not able to detect any difference.

The story is quite different for lower skill workers. Their wages fluctuate quite substantially in response to transitory shocks in the firm's value added (a 20% transmission coefficient) but much less so in response to permanent shocks, where the effect is only 10%. This may indicate a stronger level of competition in the lower skill market, as well as wages closer to reservation values, which do not allow for large reductions without workers quitting. It may also reflect more union protection against structural revisions in pay. From an econometric point of view this result may be traced back to the fact that overall permanent shocks are less important for low skill workers, as implied by the descriptive analysis of their lifecycle variance.

Table 12: Shocks and their transmission

		At least some college		High School or Less	
Parameter	Description	Estimate	s.e.	Estimate	s.e.
κ^T	Transitory firm shock, match value	0.0949	(0.0016)	0.2006	(0.0003)
κ^P	Permanent firm shock, match value	0.3487	(0.0010)	0.0959	(0.0016)
σ_{ψ^T}	Transitory idiosyncratic shock, match value	0.0811	(0.0007)	0.0117	(0.0019)
σ_{ψ^P}	Permanent idiosyncratic shock, match value	0.0071	(0.0002)	0.0163	(0.0002)
$\sigma_{\psi^{init}}$	Permanent initial shock, match value	0.0200	(0.0003)	0.0665	(0.0006)

Note: The standard deviation of the transitory firm-level shock is 0.4758; the standard deviation of the permanent firm-level shock is 0.1303.

The remaining coefficients in Table 12 relate to the idiosyncratic match value. This is a component of wage variation that is obviously related to the type of firm one is working for, but that is not shared in equal measure by similar workers within the firm (unlike the "rent sharing" component we commented on above). In settings in which information on firm performance is missing, this distinction is lost, while it plays an important role here.[22]

The results here indicate a relatively small role for initial heterogeneity in idiosyncratic match effects for higher skill workers: the variance of the initial match value is very small, which also implies that wage induced moves from one job to another are primarily driven by an accumulation of bad shocks in the job of origin, rather than a location of a much improved opportunity. The permanent shocks to this initial match value (σ_{ψ^P}) are much smaller than permanent productivity shocks (see Table 11). Finally, transitory shocks to the matched value (σ_{ψ^T}) although substantial, are small compared to idiosyncratic shocks to individual productivity in Table 11. When we turn to lower skill workers there seems to be larger permanent heterogeneity in idiosyncratic match values (standard deviation 0.02). This heterogeneity increases over time, as the permanent shocks to this initial value (σ_{ψ^P}) are substantial.

The important point that emerges from these results is that a large fraction of "match

[22]We cannot rule out that these idiosyncratic match effects reflect heterogeneity in the pass-through of firm -related shocks.

effects" on wage variability is explained by shocks to firm productivity rather than more idiosyncratic components reflecting, say, learning or wage improvements due to between-firm competition for workers.

To summarize, our results are not driven by omitted match specific effects, but by the firm level shocks that are observed and by the correlation of wages between workers. Allowing for idiosyncratic match value is also important because it accounts for changes in wages across firms for the same worker, even without allowing for firm level productivity shocks. Thus match specificity originates from productivity shocks and essentially relates to non-competitive behavior in the labor market that allows both for rent sharing and a pass-through of negative fluctuations.

5.3 Simulations

The identity of the firm in which one works appears to have a substantial impact on the evolution of earnings over the lifecycle. Given that we are looking at innovations to wages and productivity, our conclusion is that a substantial amount of uncertainty faced by individuals has its origins in the fluctuating fortunes of their firm. In order to better understand the implications of these results we carry out a number of simulations of actual and counterfactual lifecycle profiles.

We simulate the life-cycle for 20,000 individuals, of which 85% receive an initial offer in the first period, while 15% enter into unemployment. We allocate these offers across individuals according to the cross-sectional distribution of workers in different sector and firm-size bins as reported in Table 2. We then analyze earnings dispersion, participation and mobility over the life-cycle for the full model and in counterfactual scenarios in which we shut down different types of shocks subsequently. For simplicity, we report statistics for four points in the life cycle: age 26, 35, 45 and 55.

In Panel A of Table 13 we consider the baseline model with endogenous participation

Table 13: Simulations: Firm and Match-Specific Shocks

	At least some college			High School		
Panel A: Full Model						
Age	Var(Earnings)	Participation	Mobility	Var(Earnings)	Participation	Mobility
26	0.1095	0.8813	0.0472	0.0944	0.7952	0.0299
35	0.1481	0.9055	0.0267	0.0888	0.8706	0.0206
45	0.1870	0.9255	0.0189	0.0913	0.8889	0.0166
55	0.2138	0.9001	0.0164	0.0938	0.9042	0.0126
Panel B: No Firm Shocks						
Age	Var(Earnings)	Participation	Mobility	Var(Earnings)	Participation	Mobility
26	0.1078	0.8821	0.0446	0.0925	0.7953	0.0283
35	0.1150	0.9020	0.0272	0.0841	0.8706	0.0208
45	0.1193	0.9181	0.0216	0.0850	0.8884	0.0163
55	0.1206	0.8879	0.0172	0.0868	0.9040	0.0121
Panel C: No Firm Shocks, No Match Effects						
Age	Var(Earnings)	Participation	Mobility	Var(Earnings)	Participation	Mobility
26	0.1062	0.8823	0.0440	0.0881	0.7953	0.0278
35	0.1119	0.9022	0.0258	0.0761	0.8699	0.0204
45	0.1156	0.9181	0.0208	0.0748	0.8880	0.0169
55	0.1166	0.8880	0.0168	0.0746	0.9028	0.0116

and mobility choices. As we expect from the data, the cross sectional variance of earnings increases over time for the higher skilled and is flat for the low skilled. We target these life-cycle patterns in the estimation, and the levels closely match the data as shown in Table 7. In Panel B we switch off firm level shocks (i.e., set the pass-through parameters $\kappa^P = \kappa^T = 0$). By the age of 55, the cross sectional variance for the high skilled is only 0.12, compared to the full variance of 0.21. In other words, permanent firm level shocks, which are transmitted to wages, explain 44% of the cross-sectional dispersion of wages for 55-year-old workers with at least some college education. This effect is important because, as documented in Table 12, it is the permanent shocks that are transmitted, and these accumulate over the life-cycle to a much larger extent than transitory ones (at least so long as people stay with the firm). Since the share of earnings variance that can be accounted for by firm-level shocks is a key statistic, we report it separately in Table 15 both for the baseline model and for some of the counterfactual models we discuss next.

In Panel C we switch off both firm-level shocks as well as match component effects. As we would expect from the parameter estimates, standard match effects contribute less than 2% to the cross-sectional variance of high-skilled workers. Perhaps surprisingly, these shocks do not explain much of the overall participation or mobility rates by age; this is despite the fact that both of these decisions depend on the wage and the wage gains from moving, respectively.

The results differ significantly for workers with no more than high school education. First, the share of wage variation explained by firm-level shocks at age 55 is only 7.4%. Second, as illustrated in Table 12, there is substantial variation in initial match values and in permanent match shocks for low-skilled workers. Panel C shows that these match effects plays an important role for the overall wage variance of these workers; comparing Panel B and C, match-specific idiosyncratic shocks account for 5-13% of the cross-sectional earnings variance over the life-cycle.

In Table 14 we explore the role of mobility and participation choices for overall earnings variation. Panel A simulates the model without allowing for job-to-job mobility (implying that workers can join new firms only after an unemployment spell). The first column illustrates the increase in overall earnings variance compared to the results of the full model in Panel A of Table 13. If workers cannot switch jobs, either to move to opportunity or to leave a sinking ship, the simulated cross-sectional earnings variance at age 55 increases by 14% for high-skilled workers and 8.5% for low-skilled workers. Job mobility is thus an important earnings (and hence consumption) smoothing mechanism.

Panel B shows that a large part of the increase in wage variance is due to the role of mobility in mitigating exposure to negative firm-level shocks. Without the option of switching jobs, a higher share of the cross-sectional earnings variation is accounted for by firm-level shocks; at age 55, this share is now 50.4% (an increase of 6.8 p.p.). The same pattern holds for low-skill workers, where the share of earnings variation explained by firm-

Table 14: Simulations: Mobility and Participation Choices

	At least some college			High School		
Panel A: No Job-to-Job Mobility						
Age	Var(Earnings)	Participation	Mobility	Var(Earnings)	Participation	Mobility
26	0.1098	0.8808	0.0000	0.0949	0.7951	0.0000
35	0.1627	0.8999	0.0000	0.0923	0.8699	0.0000
45	0.2106	0.9186	0.0000	0.0971	0.8879	0.0000
55	0.2445	0.8897	0.0000	0.1018	0.9033	0.0000
Panel B: No Job-to-Job Mobility, No Firm Shocks						
Age	Var(Earnings)	Participation	Mobility	Var(Earnings)	Participation	Mobility
26	0.1079	0.8819	0.0000	0.0928	0.7953	0.0000
35	0.1152	0.9015	0.0000	0.0864	0.8701	0.0000
45	0.1198	0.9178	0.0000	0.0886	0.8876	0.0000
55	0.1213	0.8875	0.0000	0.0915	0.9036	0.0000
Panel C: Full Participation						
Age	Var(Earnings)	Participation	Mobility	Var(Earnings)	Participation	Mobility
26	0.1088	1.0000	0.0556	0.0936	1.0000	0.0394
35	0.1533	1.0000	0.0296	0.0906	1.0000	0.0244
45	0.1957	1.0000	0.0192	0.0962	1.0000	0.0181
55	0.2291	1.0000	0.0170	0.1023	1.0000	0.0132
Panel D: Full Participation, No Firm Shocks						
Age	Var(Earnings)	Participation	Mobility	Var(Earnings)	Participation	Mobility
26	0.1069	1.0000	0.0529	0.0916	1.0000	0.0373
35	0.1148	1.0000	0.0307	0.0850	1.0000	0.0247
45	0.1196	1.0000	0.0232	0.0878	1.0000	0.0179
55	0.1213	1.0000	0.0193	0.0917	1.0000	0.0128

Table 15: Simulations: Summary

Share of earnings variance accounted for by firm-level shocks						
	At least some college			High School		
Age	Full Model	No Mobility	Full Participation	Full Model	No Mobility	Full Participation
26	0.015	0.017	0.017	0.020	0.041	0.021
35	0.223	0.292	0.251	0.054	0.221	0.062
45	0.362	0.431	0.389	0.068	0.263	0.088
55	0.436	0.504	0.470	0.074	0.271	0.104

level shocks increases to 27.1% (a large 20 p.p. increase) if workers do not have the option of job mobility.[23]

Finally, Panel C and D consider the role of non-participation. Intuitively, if workers do not have the option of leaving their current job into non-participation in response to large negative shocks, the role of firm-level shocks in explaining overall earnings variation will also increase substantially compared to the baseline. Indeed quitting and searching for another job can mitigate the rise in wage inequality over the lifecycle: forcing all individuals to work increases the variance at age 55 by about 6.5% for the high skill group and about 9% for the lower skill one. Switching off the transmission of firm level shocks eliminates this impact for the former, but for the latter the importance of simple match specific effects means that the cross sectional variance increases by 23% when the transitions between employment and unemployment are shut down.

In sum, these results emphasize the crucial role of mobility and participation choices to mitigate the effect of permanent firm-level and match-specific shocks. These endogenous choices mask the high transmission of firm-level shocks to workers' wages. This is a crucial insight that helps explain the larger transmission effects that we find compared to the previous literature. Focusing only on the set of workers who choose not to adjust along these two margins systematically underestimates the role of firms for earnings variation. To

[23]Note that if the match value is entirely fixed, job mobility does not matter for the earnings variance because individual productivity shocks are carried over to any other job as well. This means the results on earnings variance and participation from Panel C of Table 13 apply to the case of no mobility as well.

Table 16: Simulations: Mobility and Participation Choices

	At least some College		High School	
	Full Model	Stayers	Full Model	Stayers
$sd(\Delta w \mid J = 0)$	0.1855	0.1855	0.1581	0.1581
$sd(\Delta w\ firm\ trans)$	0.0228	0.0050	0.0482	0.0268
$sd(\Delta w\ firm\ perm)$	0.0614	0.0328	0.0165	0.0266
$sd(\Delta w\ firm)$	0.0656	0.0332	0.0511	0.0378
Share firm shocks	0.354	0.179	0.323	0.239

illustrate this point quantitatively, in Table 16 we compare results obtained in the baseline model ("Full model") with those obtained in a counterfactual model where we only focus on stayers, similar to Guiso, Pistaferri, and Schivardi (2005) and most of the literature that followed. Due to selection, focusing on stayers gives the impression that firm shocks matter much less for overall wage variation than it actually does. The downward bias is particularly large for the high educated since for this group the transmission of permanent firm shocks is higher and these shocks have large cumulative effect on lifecycle variances than transitory firm shocks. This has important considerations for an evaluation of lifecycle risks faced by workers, since most firm-level shocks are not under the control of the agent. Fagereng et al. (2017, 2018) use this insight to study how exogenous permanent firm shocks passing through wages impact household savings and portfolio choices, respectively.

6 Conclusion

In this paper we use rich matched employer-employee data from Sweden to estimate the stochastic properties of the wage process for individuals and the way it may be impacted by productivity shocks to the firm. Our model accounts for endogenous participation and mobility decisions and thus deals with the potential truncation in the impact of productivity on wages that is induced by people quitting into unemployment or changing employer.

The key finding is that permanent productivity shocks transmit to individual wages for high skill workers: the elasticity of wages with respect to permanent productivity shocks is 0.38. In other words firms pass a part of their permanent good or bad fortune to wages. However transitory (i.i.d.) shocks have no impact on the wages of the high skill workers. They do however affect the wages of the low skill workers. We find that the variance of wages increases over the life-cycle because of a permanent individual shock that sticks with the worker. However, by age 55 about 44% of the cross sectional variance of wages for high skill workers is attributable to firm level shocks. For these workers, match specific effects, other than those that are common to all workers in the same firm, do not play a substantial role. For lower skill workers, random match specific effects do play a role in driving the variance of wages.

Our paper emphasizes that there are three sources of stochastic variation in wages that are often confounded (mostly due to imperfect data). The first is purely idiosyncratic to the worker and can be transferred across jobs. It varies over time due to transitory and permanent components - for example because of short-lived spells of sickness or long-lived skill depreciation. The second is specific to the match and can potentially also vary over the life of the worker-firm relationship, due again to short-term or long-term developments (such as learning or between firm competition for talents). Finally, there is an insurance or rent-sharing component that depends on how much the fortunes of a firm make their way onto the workers' wages. By its very nature, this component induces correlation across wages of similar workers within the firm. It would be unimportant in settings in which labor markets were perfectly competitive. It would also be absent in settings in which institutional features (such as union contracts) prevent wages from absorbing firm-side fluctuations (while allowing for industry-wide developments to matter, say). Our results show that the firm-level component plays a more important role than the match component (which only explains initial heterogeneity of job offers among the low skilled). They also provide evidence that this affects the wages of workers of different skills differently. Highly skilled workers partake of the structural changes occurring in the firm's fortunes, while low-skilled workers are insulated against them. If the main interpretation of the transmission coefficient is that it reflects the degree of insurance implicitly provided by firms, it is an intuitive result as long as one believes that self-insurance opportunities are more limited among the low educated. It is also consistent with union protection being more important for these workers. Indeed, one way of interpreting the results is that low-skill workers' wages are close to the minimum wage thresholds set in collective bargaining agreements, preventing the transmission of negative firm-level shocks onto wages.[24]

[24] Forslund et al. (2012) and Saez et al. (2017) argue that the union minimum wage floors mostly bind for new, young employees.

A Wage Residuals

The results for the first-stage estimation are presented in Tables 17 and 18. For readability, we suppress region effects, time effects and region-time interactions in the participation equation and time and region effects in the wage regression. Instead we only report the coefficients for personal characteristics.

First, consider the results for participation choices in Table 17. The table reports probit estimates, and we focus on the sign patterns of the results. For men, having children up to three years of age significantly decreases the probability of participating in the labor market, but older children increase participation. Women with at most high school education are less likely to work if they have children and the relationship is stronger the younger the children are. Highly educated women often seem to combine having children and a career. They are likely to temporarily leave the workforce when the child is very young but they reenter soon afterwards and are more likely to participate than women without children. This behavior is consistent with high-productivity types achieving higher education and being more likely to work. Temporary absence is facilitated by the Swedish system of parental leave benefits that offers 80% of previous earnings for up to 13 months with a very generous cap. The full benefit period only applies if the father also stays with the child for some time, which is consistent with the lower participation probability for men with young children. Interestingly, married men are more likely to work, but the same is true to a lesser extent for women as well.

The coefficients on parental leave and sickness benefits confirm the measurement problems in employment status described above. For example, parental leave payments increase the probability of being employed for men. The reason is that men usually only take out parental leave benefits for a few months. Yet employers are likely to add some bonus payments during this time, which makes these fathers appear working at low wages. For women the relationship is negative, but relatively small. Hence these results suggest that as for men,

Table 17: First-Stage Results: Participation Equation

	Male		Female	
	High School	Some College	High School	Some College
constant	0.1940***	0.9833***	0.3825***	1.0144***
	(0.008)	(0.019)	(0.013)	(0.026)
age	0.9059***	0.4661***	0.7258***	0.3402***
	(0.005)	(0.010)	(0.007)	(0.012)
age^2	-0.7315***	-0.5345***	-0.4293***	-0.3274***
	(0.004)	(0.011)	(0.006)	(0.014)
age^3	0.2609***	0.2313***	0.1456***	0.1445***
	(0.001)	(0.004)	(0.002)	(0.006)
age^4	-0.0314***	-0.0326***	-0.0187***	-0.0210***
	(0.000)	(0.001)	(0.000)	(0.001)
child 0-3 yrs	-0.0576***	-0.0062***	-0.2948***	-0.1838***
	(0.001)	(0.002)	(0.001)	(0.002)
child 4-6 yrs	0.0264***	0.0525***	-0.0995***	0.0022
	(0.001)	(0.002)	(0.001)	(0.002)
child 7-10 yrs	0.0262***	0.0495***	-0.0815***	-0.0017
	(0.001)	(0.002)	(0.001)	(0.002)
child 11-17 yrs	0.0341***	0.0983***	-0.1007***	0.0108***
	(0.001)	(0.002)	(0.001)	(0.002)
married	0.3060***	0.2279***	0.1476***	0.1532***
	(0.001)	(0.001)	(0.001)	(0.002)
parental leave	0.0205***	0.0318***	-0.0853***	-0.0583***
	(0.000)	(0.001)	(0.000)	(0.000)
sickness benefits	-0.0900***	-0.0987***	-0.0979***	-0.0872***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	38,824,193	11,653,681	17,881,403	6,463,989
Wald test [df=220]	19425.29	4102.50	5881.99	1623.23
Wald test [p-value]	0.0000	0.0000	0.0000	0.0000
Pseudo R-squared	0.0744	0.0406	0.1132	0.0688

*** p<0.01, ** p<0.05, * p<0.1

Note: All specifications include region, year and region-year fixed effects. Standard errors in parentheses. Wald tests report test statistics and p-values for the exclusion restriction of region-time interactions in each specification.

some women will be considered employed at low wages during their parental leave. The coefficient for sickness benefits is negative and significant for all gender-education groups, but a similar caveat applies: Short time sickness benefits will make individuals appear to be working nevertheless, but at a lower average wage.

Table 18: Wage equation

	Male		Female	
	High School	Some College	High School	Some College
constant	9.6343***	9.9937***	9.1751***	9.9701***
	(0.0075)	(0.0097)	(0.0103)	(0.0075)
age	0.6124***	0.7015***	0.7811***	0.4164***
	(0.003)	(0.008)	(0.005)	(0.009)
age^2	-0.3627***	-0.3701***	-0.4204***	-0.2532***
	(0.002)	(0.007)	(0.004)	(0.009)
age^3	0.1000***	0.1095***	0.1066***	0.0730***
	(0.001)	(0.003)	(0.001)	(0.003)
age^4	-0.0102***	-0.0133***	-0.0108***	-0.0086***
	(0.000)	(0.000)	(0.000)	(0.000)
child 0-3 yrs	-0.0373***	-0.0169***	-0.1694***	-0.1351***
	(0.000)	(0.001)	(0.001)	(0.002)
child 4-6 yrs	-0.0059***	0.0234***	-0.0888***	-0.0514***
	(0.000)	(0.001)	(0.001)	(0.001)
child 7-10 yrs	-0.0069***	0.0167***	-0.0785***	-0.0569***
	(0.000)	(0.001)	(0.001)	(0.001)
child 11-17 yrs	0.0006*	0.0230***	-0.0644***	-0.0525***
	(0.000)	(0.001)	(0.001)	(0.001)
married	0.0838***	0.1425***	-0.0213***	0.0388***
	(0.001)	(0.001)	(0.001)	(0.001)
parental leave	-0.0421***	-0.0402***	-0.0916***	-0.0830***
	(0.000)	(0.000)	(0.000)	(0.001)
sickness benefits	-0.0698***	-0.1006***	-0.0855***	-0.1005***
	(0.000)	(0.001)	(0.000)	(0.001)
Mills ratio	0.2161***	0.9770***	0.4693***	0.4431***
	(0.006)	(0.021)	(0.008)	(0.023)
Mills ratio * age	-0.1113***	-0.5110***	-0.0534***	0.3696***
	(0.004)	(0.017)	(0.005)	(0.015)
Mills ratio * age^2	0.0148***	0.0896***	0.0116***	-0.0654***
	(0.001)	(0.004)	(0.001)	(0.004)
Observations	9,010,548	2,796,200	3,921,223	1,514,611
R-squared	0.199	0.168	0.275	0.251

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Note: All specifications include region and year fixed effects. Standard errors in parentheses. Wald tests report test statistics and p-values for the exclusion restriction of region-time interactions in each specification.

Next, consider the results for wages in Table 18. The results confirm the familar concave life-cycle profile of wages. These profiles are illustrated graphically in Figure 5. As we can see from the comparison with simple OLS earnings profiles in Figure 5, the model predicts that selection has an effect on the slope of the earnings profile. Positive selection into the labor market is stronger at early ages, which means that without selection correction, wage growth at the beginning of the life-cycle will be underestimated by looking at cross-sectional worker data as lower ability individuals enter the labor force later. This is an important finding that needs to be taken into account for analyses of wage inequality for example. Furthermore, we find increasing positive selection at the end of workers' careers again. One explanation could be early retirement based on disability, which is very common in Sweden and is more likely to be chosen by low-ability types. As a result, the wage decrease in the life-cycle of earnings is underestimated. Finally, the selection patterns for women are consistent with higher-educated women having children later in their lives, thereby leading to peak positive selection in their early to mid-thirties as illustrated in Figure 7 below.

To illustrate selection patterns across the lifecycle, we allow for a fairly flexible specification of the Mills ratio in the wage regression. The overall selection coefficients by age corresponding to the regression results in Table 18 can be found in Figure 6. For male workers, selection is highest early in the life-cycle and decreases over time as lower-productivity types enter the labor market. Finally selection increases again as workers get closer to retirement age. The same qualitative pattern holds for women with at most high-school education even though selection increases again quite strongly after the age of 45. Highly educated females are the exception here, they display increasing selection in their 30s and 40s as lower productivity types are more likely to decide to stay out of work to bring up children for example. These patterns directly mirror the results for earnings profiles taking selection into account in Figure 5.

Overall, the wage regression implies a positive and significant selection effect for all samples. As Figure 7 suggests, wage differences because of selection are in the range of 0-20

Figure 5: Predicted Wages controlling for parental leave and sickness benefits (f0 = male, e0 = low-educated)

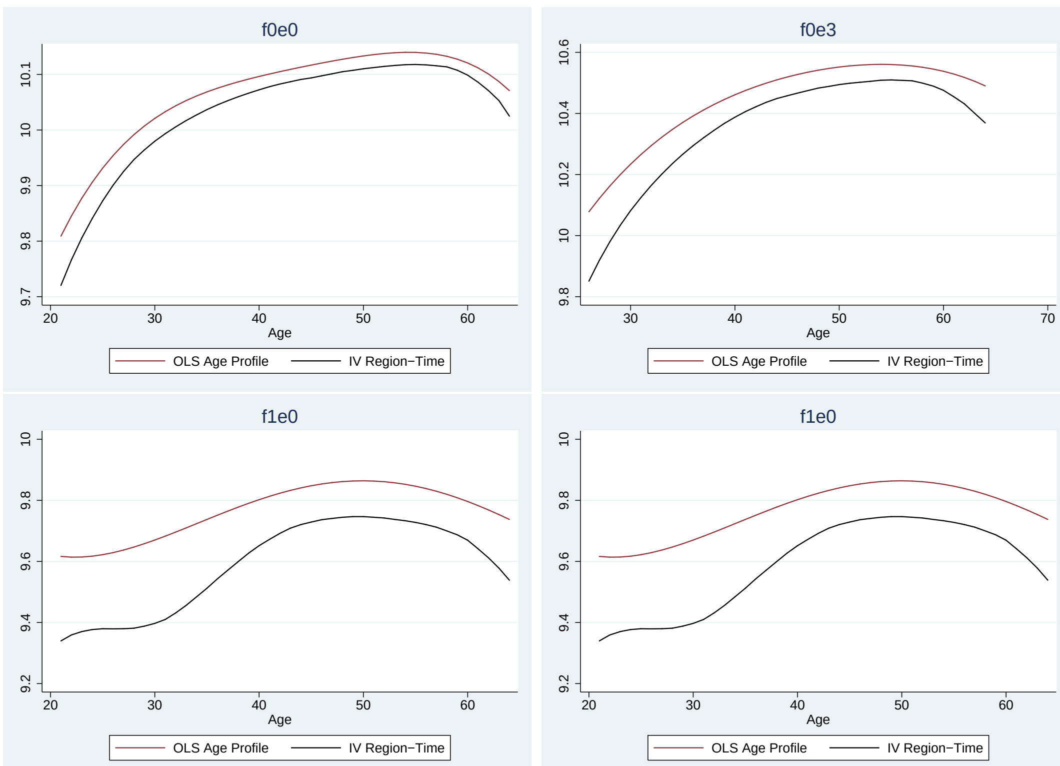


Figure 6: Selection Coefficient by Age (f0 = male, e0 = low-educated)

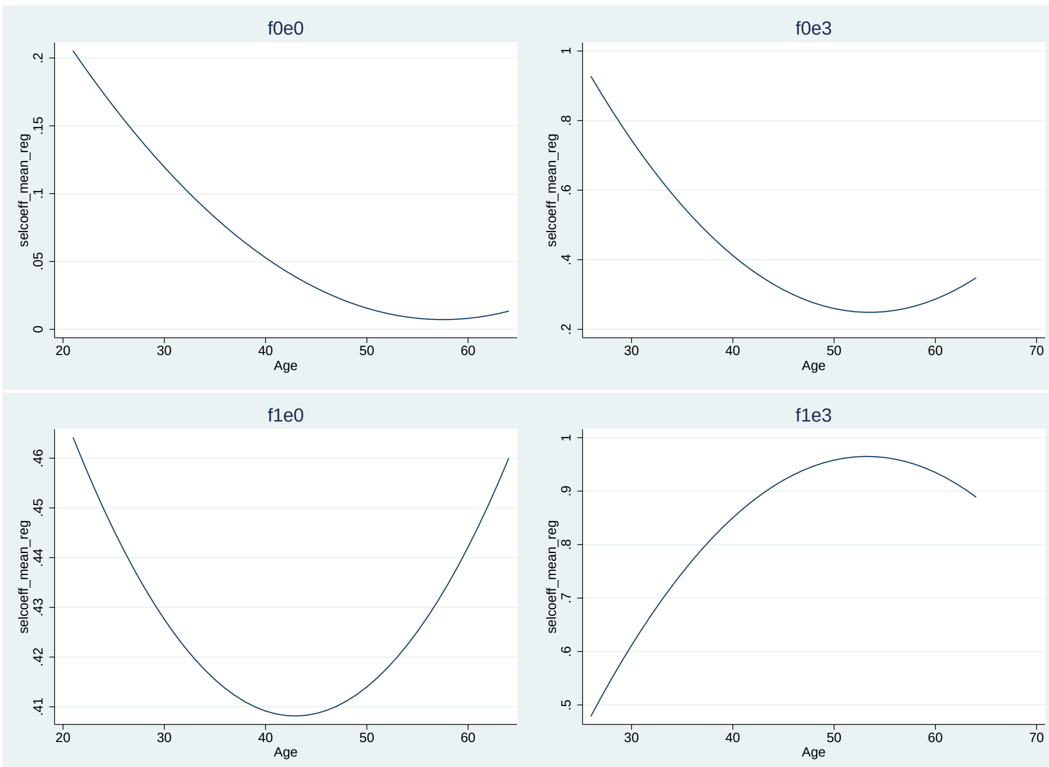
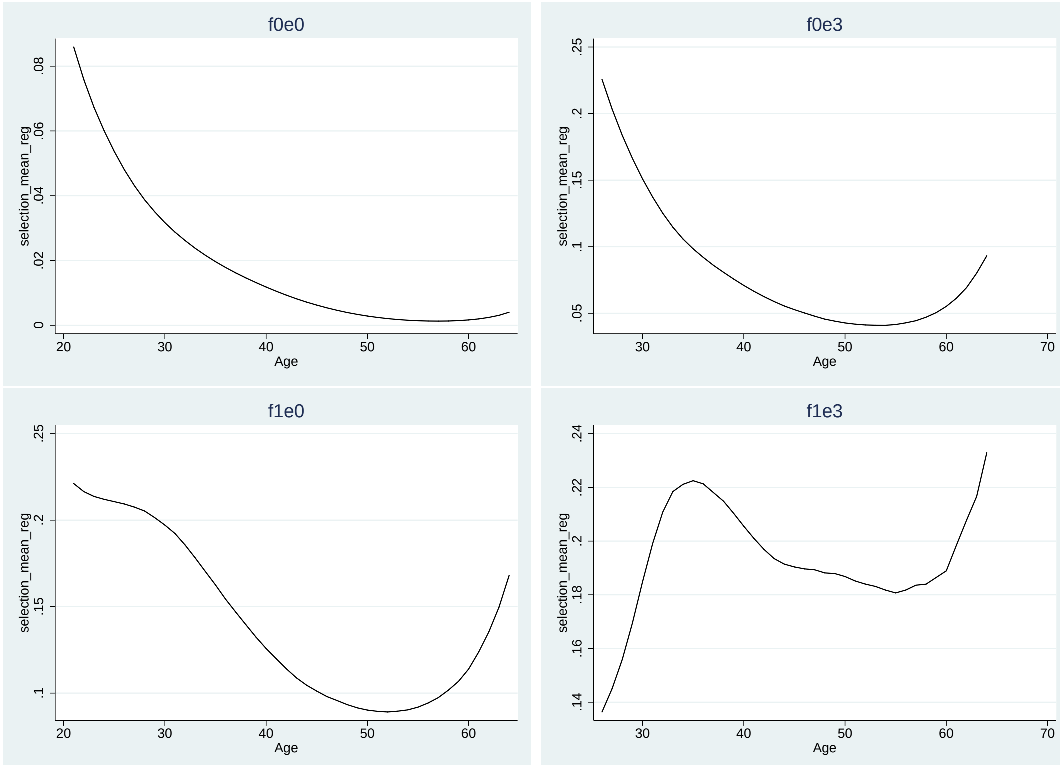


Figure 7: Average Selection Effects by Age (f0 = male, e0 = low-educated)



log points, where these effects are higher for groups with higher education. Interestingly, selection among women tends to be larger than for their male counterparts conditional on education group.

References

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

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999a). 'High Wage Workers and High Wage Firms.' *Econometrica*, vol. 67(2):pages 251–334.

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999b). 'High Wage Workers and High Wage Firms.' *Econometrica*, vol. 67(2):pages 251–333.

Altonji, J. G. and Segal, L. M. (1996). 'Small-Sample Bias in GMM Estimation of Covariance Structures.' *Journal of Business & Economic Statistics*, vol. 14(3):pages 353–66.

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

Blundell, R., Duncan, A., and Meghir, C. (1998). 'Estimating Labor Supply Responses Using Tax Reforms.' *Econometrica*, vol. 66(4):pages 827–862.

Blundell, R., Reed, H., and Stoker, T. M. (2003). 'Interpreting Aggregate Wage Growth: The Role of Labor Market Participation.' *American Economic Review*, vol. 93(4):pages 1114–1131.

Burdett, K. and Mortensen, D. T. (1998). 'Wage Differentials, Employer Size, and Unemployment.' *International Economic Review*, vol. 39:pages 257–273.

Card, D. E., Heining, J., and Kline, P. M. (2013). 'Workplace Heterogeneity and the Rise of West German Wage Inequality.' *The Quarterly Journal of Economics*, vol. 128(3):pages 967–1015.

Chernozhukov, V. and Hong, H. (2003). 'An MCMC approach to classical estimation.' *Journal of Econometrics*, vol. 115(2):pages 293–346.

Guiso, L., Pistaferri, L., and Schivardi, F. (2005). 'Insurance within the Firm.' *Journal of Political Economy*, vol. 113(5):pages 1054–1087.

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

Lamadon, T. (2016). 'Productivity Shocks, Long-Term Contracts and Earnings Dynamics.' University of Chicago mimeo.

Lamadon, T., Mogstad, M., and Setzler, B. (2018). 'Imperfect Competition and Rent Sharing in the U.S. Labor Market.' University of Chicago mimeo.

Lise, J., Meghir, C., and Robin, J.-M. (2016). 'Matching, Sorting and Wages.' *Review of Economic Dynamics*, vol. 19:pages 63–87.

Low, H., Meghir, C., and Pistaferri, L. (2010). 'Wage Risk and Employment Risk over the Life Cycle.' *American Economic Review*, vol. 100(4):pages 1432–67.

MaCurdy, T. E. (1982). 'The Use of Time Series Processes to Model the Error Structure of Earnings in Longitudinal Data Analysis.' *Journal of Econometrics*, vol. 18:pages 82–114.

McFadden, D. (1989). 'A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration.' *Econometrica*, vol. 57(5):pages 995–1026.

Meghir, C. and Pistaferri, L. (2004). 'Income Variance Dynamics and Heterogeneity.' *Econometrica*, vol. 72(1):pages 1–32.

Mortensen, D. and Pissarides, C. (1994). 'Job Creation and Job Destruction in the Theory of Unemployment.' *The Review of Economic Studies*, vol. 61:pages 397–415.

Pakes, A. and Pollard, D. (1989). 'Simulation and the Asymptotics of Optimization Estimators.' *Econometrica*, vol. 57(5):pages 1027–1057.

Postel-Vinay, F. and Robin, J.-M. (2002a). 'Equilibrium W age Dispersion with Worker and Employer Heterogeneity.' *Econometrica*, vol. 70(6):pages 2295–2350.

Postel-Vinay, F. and Robin, J.-M. (2002b). 'Equilibrium Wage Dispersion with Worker and Employer Heterogeneity.' *Econometrica*, vol. 70(6):pages 2295–2350.

van Reenen, J. (1996). 'The Creation and Capture of Rents: Wages and Innovation in a Panel of U. K. Companies.' *The Quarterly Journal of Economics*, vol. 111:pages 195–226.

Topel, R. H. and Ward, M. P. (1992). 'Job Mobility and the Careers of Young Men.' *The Quarterly Journal of Economics*, vol. 107(2):pages 439–79.