EARNINGS DYNAMICS AND FIRM-LEVEL SHOCKS

By

Benjamin Friedrich, Lisa Laun, Costas Meghir, and Luigi Pistaferri

April 2019

COWLES FOUNDATION DISCUSSION PAPER NO. 2175

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Earnings Dynamics and Firm-Level Shocks*

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

April 20, 2019

Abstract

We use matched employer-employee data from Sweden to study the role of the firm in affecting the stochastic properties of wages. Our model accounts for endogenous participation and mobility decisions. We find that firm-specific permanent productivity shocks transmit to individual wages, but the effect is mostly concentrated among the high-skilled workers; firm-specific temporary shocks mostly affect the low-skilled. The updates to worker-firm specific match effects over the life of a firm-worker relationship are small. Substantial growth in earnings variance over the life cycle for high-skilled workers is driven by firms accounting for 44% of cross-sectional variance by age 55.

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, IFAU, NBER Summer Institute 2016, 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
[‡]IFAU
[§]Yale University, IFS, NBER, IFAU, CEPR and IZA
[¶]Stanford University, NBER, CEPR, SIEPR and IZA

1 Introduction

How important is the firm in which a worker is employed in determining 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 that individuals face over the lifecycle.

A number of papers have addressed the former question, starting with Abowd, Kramarz, and Margolis (1999) 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 study how 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, individual 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 their 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 explains 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 and restrict themselves to the implications driven by a specific structural model. 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.[3] Such transitions may well hide the impact of firm-level

[3]Subsequent work by Carlsson, Messina, and Skans (2016) has distinguished industry-level and firm-level shocks, but maintains the focus on stayers at incumbent firms.

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.

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.[4] We extend the framework considered by the earlier 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 and match-specific 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

[4]See Abowd and Card (1989), MaCurdy (1982), Meghir and Pistaferri (2004), Guvenen (2007) and more recently Altonji, Smith, and Vidangos (2013).

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 earnings, 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 individual 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 market 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

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. Our administrative data does not measure hours of work and thus we do not distinguish between earnings and wages, terms we use interchangeably.

An important feature of careers is mobility between employment and out-of-work as well as between jobs. 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: individuals who may suffer large pay cuts as a result of productivity shocks, may quit into unemployment or are more likely to accept alternative job offers. We thus allow for endogenous employment and mobility and relate this directly to wage 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 period 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,[5] and $P_{i,t}$ is

[5]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 an overstatement of transitory risk.

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

$$
\begin{aligned}
&\mathrm{P}^{init} \sim N\left(0, \sigma_P^2\right) \\
&\zeta \sim \text{mixture of Normals}(\mu_{\zeta_1},\, \sigma_{\zeta_1};\, \mu_{\zeta_2},\, \sigma_{\zeta_2};\lambda_m)
\end{aligned}
\tag{3}
$$

where μ_{ζ_s}, σ_{ζ_s} $s = 1, 2$ represent the mean and standard deviation of each of the two normals in the mixture, while λ_m is the mixing parameter. By allowing for a mixture of normals we are able to fit higher order moments of the distribution of wage growth, such as the observed kurtosis. The importance of higher order moments in earnings growth has been examined in the context of US data by Guvenen, Karahan, Ozkan *et al.* (2019). Earlier papers that consider a mixture of normals for income processes include Geweke and Keane (2000) and Bonhomme and Robin (2009). One interpretation of the mixture is that on occasion workers draw a large wage change, possibly representing promotions or other important changes; another is that a non-negligible fraction of workers experience no wage growth from one period to the next. These features of the model turn out to be important empirically.

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:

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

Notice that with this specification we are not modeling sorting on permanent worker and firm 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 for explaining wage levels.

In equations (5) and (6), the terms $\xi^P_{j,t}$ and $\xi^T_{j,t}$ 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 $\psi^l \sim N\left(0, \sigma^2_{\psi^l}\right)$, 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 (1999) 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 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.[6]

We model the employment decisions E as:

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

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 of unobserved heterogeneity in participation choices.[7] 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

[6] 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.

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

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.

Labor Market Frictions and Job Offers Upon entry in the labor market, workers receive job offers at a rate λ_{entry}. 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.:

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

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 8 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), containing information about the universe of employment spells in Sweden from 1985 onward. On the worker side, RAMS includes 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 includes 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. The sample includes all firms with the legal entity being limited partnership or limited company (other than banking and insurance

Table 1: **Summary statistics, firms**

	Firm size: number of employees			
	5–20	20–50	50–100	100+
A. Construction				
No. unique firms	15,527	984	195	142
Value added per worker	486,027	528,201	558,381	576,954
Growth, log V.A./worker	0.0363	0.0372	0.0390	0.0247
B. Manufacturing				
No. unique firms	14,373	2,705	1,080	1,166
Value added per worker	515,661	577,966	621,752	1,018,796
Growth, log V.A./worker	0.0290	0.0208	0.0130	0.0123
C. Retail				
No. unique firms	27,013	2,245	554	403
Value added per worker	507,697	624,140	633,776	760,339
Growth, log V.A./worker	0.0291	0.0245	0.0260	0.0206
D. Services				
No. of unique firms	45,637	3,931	1,015	832
Value added per worker	553,601	654,343	841,577	771,384
Growth, log V.A./worker	0.0368	0.0399	0.0439	0.0327

Note: Value added per worker is in real SEK for base year 2008.

companies), and we exclude sole traders because data for these firms are not available for the entire period. The final sample represents 84 percent of value added and 86 percent of employment in the Swedish non-financial private sector over the 1997–2008 period.

Table 1 presents descriptive statistics for the firms in our data set. The data includes almost 120,000 unique firms and 920,000 firm-year observations. The four sectors construction, manufacturing, retail and services account for 15%, 18%, 27% and 40% of all firms in the sample, respectively. Within sectors, larger firms display, on average, higher value added per worker. For construction and manufacturing, larger firms grow more slowly on average, whereas growth rates are more similar across firm size 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 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.[8] 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.

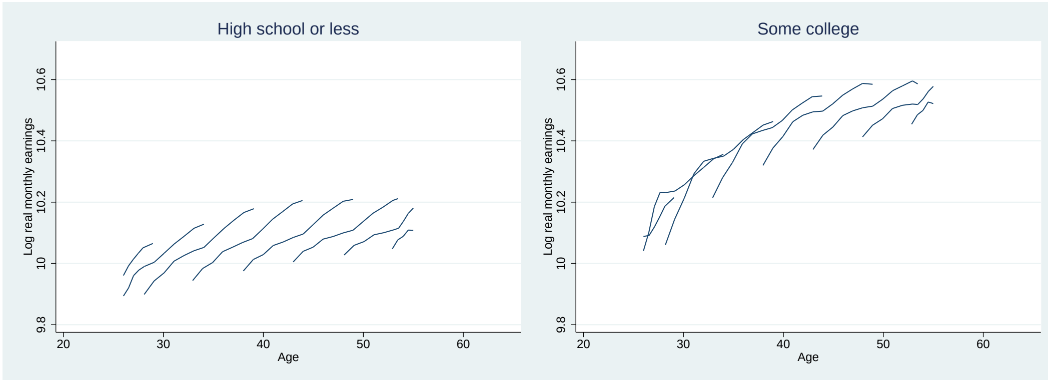
[8]In 1997 (our first sample year), the part-time employment rate (defined as the fraction of employed workers who work less than 30 hours per week in their main job) was 6.5% for men and 23% for women; in 2008, the rates were 10% and 20%, respectively (source: OECD). For women of child-bearing age, it is also more frequent to observe shifts from full-time to part-time work and *vice versa*, making the analysis of wages volatility in administrative data much more challenging. In an earlier working paper version of the paper, we documented that earnings variances for women exhibit a hump-shaped pattern over the life cycle (unlike the growing pattern documented below for men). 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	1,152,933	464,616
No. worker-quarter obs.	31,091,423	11,188,448
Monthly earnings	24,960	35,930
(2008 SEK)	(8,009)	(17,115)
Age	40.31	39.07
Married	0.5797	0.6112
Having children	0.4569	0.4985
Employed, of which	0.8764	0.9040
Job stayer	0.9547	0.9493
Job mover	0.0233	0.0319
Re-entrant	0.0220	0.0188
Industry		
Construction	0.1535	0.0582
Manufacturing	0.4054	0.3554
Retail Trade	0.1853	0.1409
Services	0.2558	0.4455
Firm size		
≤20	0.3264	0.2752
20–50	0.1355	0.1191
50–100	0.0994	0.0869
100+	0.4387	0.5188

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 around 2 percent enter employment after a period of non-employment.

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

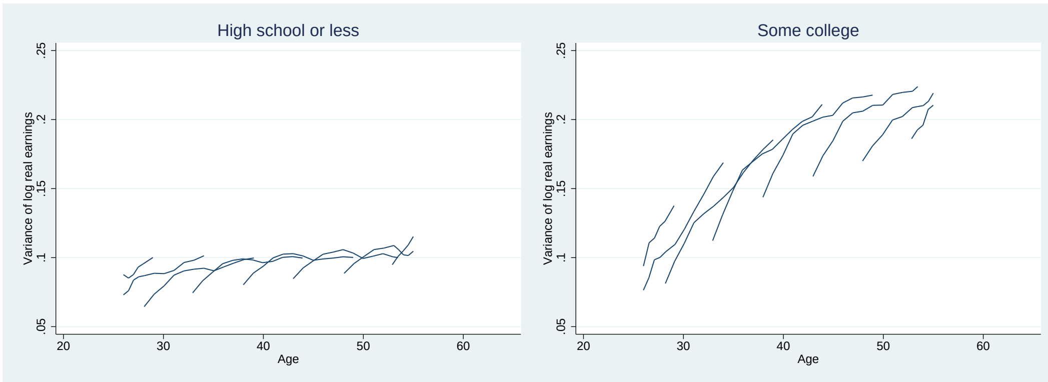


Life-cycle earnings Table 2 also reveals some important differences in earnings across education groups. We take a more detailed look at life-cycle earnings profiles in Figure 1, using observations for different birth cohorts in the data. In particular, for each education group we construct five-year 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 earnings 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

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



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 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, which justifies our sample selection choice of focusing on workers younger than 55.

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. Low-educated workers face higher separation rates and lower re-entry rates at young ages. The entry rate from non-employment is rapidly falling with age and comparable across education groups around

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

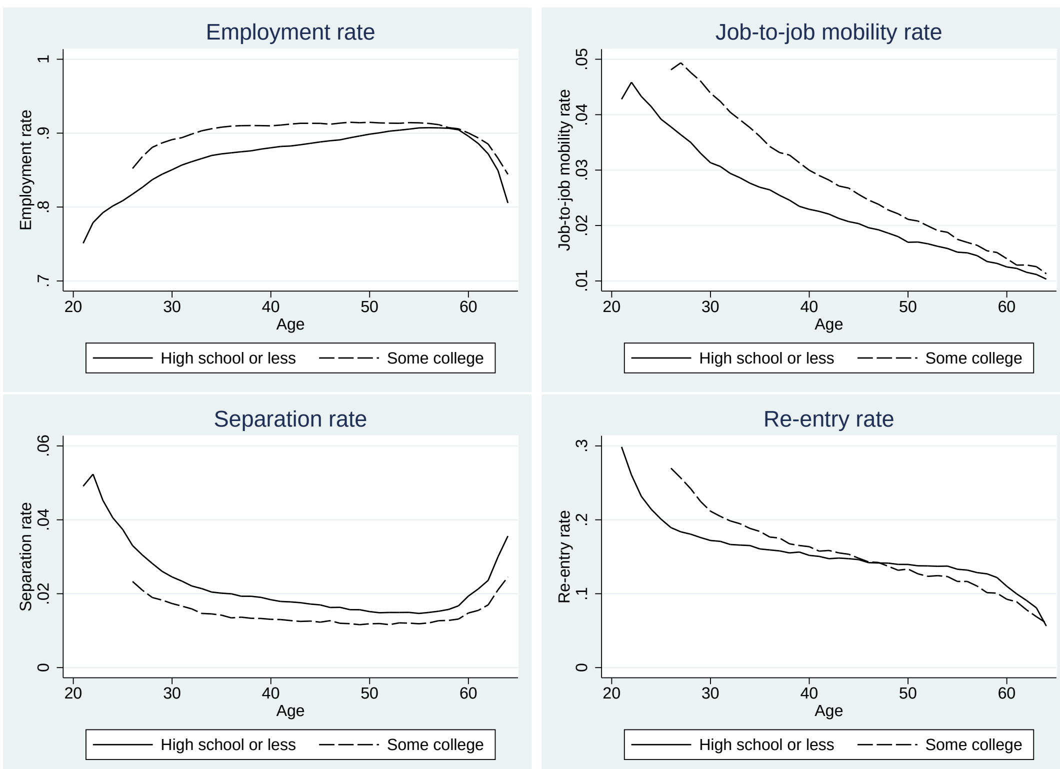


Table 3: **Share of between firm wage variance**

	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%

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.

age 35, 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 re-entry 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 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.

Table 3 reports the amount of earnings variance that can be attributed to differences between firms.[9] The results show that most of the variance of earnings 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

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

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	0.409	0.482	0.475
	2	0.068	0.110	0.058	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	0.278	0.323	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	0.343	0.363	0.464
	2	0.060	0.107	0.067	0.022	0.134	0.094	0.083	0.056	0.211	0.246	0.294	0.360
	3	0.036	0.080	0.152	0.061	0.177	0.113	0.085	0.064	0.187	0.218	0.275	0.354
	4	0.023	0.033	0.062	0.085	0.186	0.153	0.140	0.100	0.217	0.216	0.232	0.310

Note: Firms sorted based on average wage paid

On average, 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 based on a model that accounts for selection into employment and that allows for the fact that we take job mobility at a quarterly frequency but observe earnings annually, unless there is a change in employer. 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 fixed 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 economy-wide, regional, scale or industry effects. This ensures, that what we estimate to be a transmission of firm level shocks to individual wages, will not be confounded by common shocks to workers in the same industry. 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.[10]

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

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

[10]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
$\mathrm{Var}(\Delta A_t)$	0.1791	0.1603	0.1469	0.1698	0.2078
	(0.0008)	(0.0019)	(0.0017)	(0.0015)	(0.0015)
$\mathrm{Cov}(\Delta A_t, \Delta A_{t-4})$	-0.0537	-0.0587	-0.043	-0.0487	-0.0602
	(0.0004)	(0.0010)	(0.0008)	(0.0008)	(0.0008)
$\mathrm{Cov}(\Delta A_t, \Delta A_{t-8})$	-0.0041	-0.0005	-0.0045	-0.0049	-0.0048
	(0.0003)	(0.0007)	(0.0006)	(0.0005)	(0.0005)
$\mathrm{Cov}(\Delta A_t, \Delta A_{t-12})$	-0.0022	-0.0036	-0.0023	-0.0011	-0.0024
	(0.0003)	(0.0007)	(0.0006)	(0.0006)	(0.0006)

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

where

$$a^P_{j,t} = a^P_{j,t-1} + \xi^P_{j,t}$$
$$\xi^P_{j,t} \sim N\left(0, \sigma^2_{\xi^P}\right)$$
$$\xi^T_{j,t} \sim N\left(0, \sigma^2_{\xi^T}\right).$$

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}, where t refers to the first quarter of the relevant year, and the annual measure can be related to the underlying quarterly measure by

$$e^{A_t} = e^{a_t} + e^{a_{t+1}} + e^{a_{t+2}} + e^{a_{t+3}}$$

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. 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

$$A_t = a^P_{t-1} + \log\left[\sum_{k=0}^{3} exp\left(\xi^T_{t+k} + \sum_{s=0}^{k}\xi^P_{t+s}\right)\right].$$

Analogously, the value in the second year is

$$A_{t+4} = a^P_{t-1} + \sum_{k=0}^{3}\xi^P_{t+k} + \log\left[\sum_{k=0}^{3} exp\left(\xi^T_{t+4+k} + \sum_{s=0}^{k}\xi^P_{t+4+s}\right)\right]$$

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

$$A_{t+1} - A_t = \sum_{k=0}^{3}\xi^P_{t+k} + \log\left[\sum_{k=0}^{3} exp\left(\xi^T_{(t+4+k} + \sum_{s=0}^{k}\xi^P_{t+4+s}\right)\right] - \log\left[\sum_{k=0}^{3} exp\left(\xi^T_{t+k} + \sum_{s=0}^{k}\xi^P_{t+s}\right)\right]$$

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.[11] These estimates will be used to draw firm shocks in the simula-

[11] If we shut down the transitory shock, the annualized standard deviation of the permanent shock is 21.2%. Similarly, the annualized standard deviation of the transitory shock is 24.6%.

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 wages 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 selection 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 the binary indicator $E_{t_q} = 1$ 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 ($\lambda^M_{t_q}$) 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

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

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.[12] 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-

[12]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 as well as 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. Industry by time effects are included to control for aggregate and industry 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.[13] 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

[13]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 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.[14] The approach requires us to simulate wages and career paths, including transitions between employment and unemployment and between jobs.

Conditional on a guess for the parameter vector, we simulate life-cycle behavior and wages for 40,000 workers in the model. Specifically, 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

[14] 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 (ρ, σ^2_P, μ_{ζ_1}, $\sigma^2_{\zeta_1}$, μ_{ζ_2}, $\sigma^2_{\zeta_2}$, λ_m, σ^2_{ε}, $\sigma^2_{\psi^P}$, $\sigma^2_{\psi^T}$, $\sigma^2_{\psi^{init}}$), the job arrival rate coefficients (λ_{entry}, $\lambda_{U,0}$, $\lambda_{U,1}$) and the coefficients determining the source of outside offers (ω_0, ω_1, ω_2).

equal probability. The population of workers within a firm then receives a permanent and transitory firm shock drawn from the distributions that 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 observed spatial correlation of wages within a firm to identify the transmission coefficients.

In the simulation some workers may receive an offer immediately after education and others do not. The model includes a probability of this event as a parameter, which is estimated by matching it to the actual proportions in the data. The initial source of the job offer reflects the sectoral distribution of employees by education group in Table 2. However, in this version of the paper we do not allow the stochastic properties of wages to differ systematically by sector.

Once we simulate these career paths we compute moments from the simulated data to match them to those from the actual matched employer-employee dataset. In doing this we aggregate data from a quarterly to an annual frequency whenever needed to match the observed data. The wages in the data are the residuals we constructed earlier.[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 structure in the simulation, we use the empirical age distribution by education group as weights to compute the simulated moments from the model. We repeat this full life-cycle simulation procedure for 20 independent samples of workers and firms to further increase precision.

The full set of moments is described in the section below.

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

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.[16] These help identifying 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 as 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 association between wages and work decisions (ϕ).[17]

Quarterly job separations are endogenous and directly relate to transitory and permanent wage shocks. To distinguish "general" from "match-specific" wage shocks, we add annual moments related to wages. 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

[16] The age groups we use are 26-30, 31-35, 36-40, 41-45, 46-50, 51-55.

[17] This coefficient will be a function of both the causal impact of wages on participation and of the covariance of the errors, reflecting a composition effect on employment. Without exclusion restrictions these two effects cannot be disentangled. However, it does allow us to deal with censoring due to employment, whatever the interpretation of the coefficient.

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, skewness and kurtosis of wage growth to identify the contribution of more persistent shocks (including the parameters characterizing the mixture of normals).

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 often do not know the exact timing for job-to-job mobility. We therefore choose not to use wage information for these years and instead use mover information by looking at residual wage growth across years 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 wage 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 wages.

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, firm affiliation. This intra-class 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 and $\Delta\bar{e}$ is average residual wage growth across all firms and workers. 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 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.[18] 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:

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

[18]Wage moments that are calculated across the entire age distribution are weighted by a factor of 6 to give them equal importance as the job transition moments we compute separately by 6 age groups.

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:

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

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 $\hat{\beta}_{MCMC}$.[19]

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.[20]

To estimate standard errors we use the sandwich formula. Normally, the variance of the MCMC chain would provide an estimate of the variance of the parameters if the weights used in the method of moments criterion function were the optimal ones. But we use a diagonally weighted approach. The estimated covariance matrix has the form

[19] 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% (Roberts, Gelman, and Gilks (1997)).

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

$$\hat{V}(\hat{\beta}) = (G'(\hat{\beta})\Omega G(\hat{\beta}))^{-1} G'(\hat{\beta})\Omega \hat{E}\left[(g(\hat{\beta}) - \hat{g})(g(\hat{\beta}) - \hat{g})'\right]\Omega G(\hat{\beta})(G'(\hat{\beta})\Omega G(\hat{\beta}))^{-1}$$

where Ω is the weight matrix used in the estimation, $G\left(\hat{\beta}\right)$ is the gradient matrix evaluated at the estimated parameter vector $\hat{\beta}$. Finally, $\hat{E}$ denotes an estimated expected value.

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

We also need to compute $\hat{E}\left[(g(\hat{\beta}) - \hat{g})(g(\hat{\beta}) - \hat{g})'\right]$, which turns out to be the most complex component: this is because of the combination of serial and spatial correlation combined with the large number of observations and the huge combination of workers that can find themselves in a particular firm. While it is relatively straightforward to deal with either spatial correlation or serial correlation, doing both is intractable. We thus decided to simplify. For all moments other than the spatial correlation we allow only for within individual serial correlation, which is likely to be a very important source of dependence; in our calculation of the standard errors we ignore the within firm spatial correlation of residuals; allowing for both sources would have been straightforward with the bootstrap, but the estimation procedure is far too slow for this to be feasible. For the spatial correlation coefficient we assume all variation is between firms. While the simplification may underestimate our standard errors, the size of our data set is so large that this shortcut is unlikely to make much of a difference. The standard errors we compute are very small in general. We show in an appendix the details of the derivation of our covariance matrix, which draws from Hansen (1982).

5 Results

5.1 Model Fit

Our model is overidentified and consequently considering the fit of targeted moments can be informative on the performance of the model. In Figure 4 we plot the actual and fitted cross sectional variance of wages, conditional on fixed effects that have been differenced out when we constructed the wage residuals. These are replicated extremely well, showing a growing variance over the lifecycle for the high education group and a flat one for the lower one.

Then in Table 7 we show the dynamic moments of wage growth; overall the fit is excellent. Some autocovariances show sign reversals, but they are all very close to zero and this is inconsequential.[21] When it comes to the moments relating 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 eliminates the effects of measurement error in the exact data of the transition. 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 $Var(\tilde{e}_{t+1} - \tilde{e}_{t-1}|E_{t-1} = 1, E_{t+1} = 1, J_t = 1)$) are reproduced very accurately by the model. We also consider the covariance between an employment residual (from a linear probability model) and the wage residual, separately for stayers and movers. They help capture the selection effect of employment decisions on

[21] All units are in logs.

Table 7: Model Fit for a Selection of Moments on Wage Dynamics

	At least some College		High School or Less	
	Data	Model	Data	Model
Residual wage growth moments for job stayers				
$Var(\Delta\tilde{e}_t \mid E_{t-1}=1, E_t=1, J_t=0)$	0.0344	0.0275	0.0250	0.0227
$Cov(\Delta\tilde{e}_t, \Delta\tilde{e}_{t-1} \mid J_t=0)$	-0.0047	0.0030	-0.0035	-0.0009
$Skewness(\Delta\tilde{e}_t \mid E_{t-1}=1, E_t=1, J_t=0)$	0.0154	0.0175	0.1925	0.1943
$Kurtosis(\Delta\tilde{e}_t \mid E_{t-1}=1, E_t=1, J_t=0)$	6.0813	6.0949	6.5075	6.5038
Residual wage growth moments for job movers				
$\mathbf{E}(\tilde{e}_{t+1}-\tilde{e}_{t-1} \mid E_{t-1}=1, E_t=1, J_t=1)$	0.0400	0.0394	0.0266	0.0259
$Var(\tilde{e}_{t+1}-\tilde{e}_{t-1} \mid E_{t-1}=1, E_t=1, J_t=1)$	0.0668	0.0689	0.0537	0.0565
Covariance between wage growth and employment residuals				
$Cov(\tilde{u}_t, \tilde{e}_t \mid E_t=E_{t-1}=1, J_t=0)$	0.0003	-0.0004	-0.0002	-0.0002
$Cov(\tilde{u}_t, \tilde{e}_t \mid E_t=E_{t-1}=1, J_t=1)$	0.0192	0.0211	0.0031	0.0060
Common shocks at the firm level				
Spatial correlation coefficient (for stayers)	0.1822	0.1883	0.1783	0.1785
$Cov(\mathbf{E}_j[\Delta\tilde{e}_t], \mathbf{E}_j[\Delta\tilde{e}_{t-1}] \mid J_t=0)$	-0.0015	0.0004	-0.0010	-0.0021

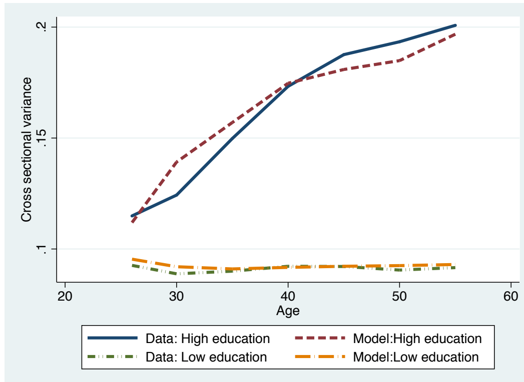
V: Variance, C: Covariance, $\mathbf{E}$: average, $\mathbf{E}_j$: average within firm j. $\tilde{e}_t$ is the estimated wage residual at age t. $\tilde{u}_t$ is a residual from a linear probability regression for employment. $E_t = 1$ indicates employment, $J_t = 1$ denotes a job mover between period $t-1$ and t. Kurtosis and Skewness drop top and bottom 1% of wage growth observations.

wages.

Among the moments we consider are the skewness and kurtosis of wage growth for stayers. These two moments capture the possibility of non normality, one interpretation of which is that most wage adjustments are small but occasionally we see big changes, say because of a promotion or an important adverse effect on productivity; we will discuss this below when we look at the estimated parameters. Skewness is close to zero, but kurtosis is relatively high. Both these moments are fitted very well by the model.

In the last panel of Table 7 we show two moments designed to capture the co-movement of wage growth among stayers in a firm; these moments identify the transmission coefficients and are thus of central importance. These are the spatial correlation of wage shocks and the autocovariance of average wage growth. Since we measure these moments using residual wage *growth* they are unlikely to reflect correlation in wages due to sorting of similar workers into a firm. Rather, they reflect how changes in wages are correlated across

Figure 4: Cross sectional variance by education over the lifecycle - model fit



individuals, reflecting the influence of firm-level shocks. This spatial correlation is quite high: 0.18 for both education groups and this is closely reproduced by the model. Similarly, the model accurately matches the autocovariance of average wage growth within firms ($Cov(\mathbf{E}_j[\Delta\tilde{e}_t], \mathbf{E}_j[\Delta\tilde{e}_{t-1}]|J_t = 0)$). 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 and job mobility (including the heterogeneity by education), which are all declining over the life cycle. It also fits the non-employment rate exactly for the lower education group and slightly under predicts it for people with more education over 40, although not in an economically consequential way. Importantly, the model replicates the increasing participation over the lifecycle. It does, however, overpredict job separation for the youngest higher education group. 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 as shown in the lower part of Table 8.

Table 8: Model Fit for Moments on Labor Market Transitions

	Age	At least some college Data	At least some college Model	High School or Less Data	High School or Less 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

Note: All transitions are quarterly

5.2 Parameter estimates

Transitions We start by presenting results in Table 9 for the decisions to work and to move to another firm.[22] 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.[23]

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		Estimate	s.e.	Estimate	s.e.
		Employment			
δ_0	Constant	0.251	(0.004)	1.839	(0.001)
δ_{age}	Age	0.807	(0.002)	-0.015	(0.0004)
δ_{age^2}	Age squared	-0.080	(0.0002)	0.016	(0.0001)
ϕ	Wage residual	0.763	(0.003)	0.268	(0.0012)
Marginal Effect of 10% wage change (%)		0.273		0.133	
		Job-to-job Mobility			
θ_0	Constant	-0.829	(0.013)	-2.579	(0.006)
θ_{age}	Age	-0.514	(0.006)	0.448	(0.0023)
θ_{age^2}	Age squared	0.058	(0.0007)	-0.074	(0.0003)
b	Wage improvement	2.865	(0.054)	1.501	(0.041)
Marginal Effect of 10% wage improvement (%)		1.711		0.871	

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.

[22] 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.

[23] 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 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		Estimate	Estimate
		(s.e.)	(s.e.)
	Job arrival rate		
λ_{entry}	Arr. rate at entry	0.977	0.744
		(0.0012)	(0.0005)
$\lambda_{U,0}$	Arr. rate, subs. spells	0.367	0.221
		(0.0005)	(0.0002)
$\lambda_{U,age}$	Arr. rate, subs. spells (age shift)	0.0047	0.0016
		(0.0000)	(0.0000)
	Origin of offer		
Parameter	Description	Estimate	Estimate
		(s.e.)	(s.e.)
ω_0	Different firm size & sector	0.210	0.213
		(0.0002)	(0.0001)
$\omega_0 + \omega_1$	Different size, same sector	0.284	0.266
		(0.0002)	(0.0001)
$\omega_0 + \omega_2$	Same size, different sector	0.144	0.124
		(0.0002)	(0.0001)
$\omega_0 + \omega_1 + \omega_2$	Same size, same sector	0.362	0.397
		-	-

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 random exogenous shocks. This is important when we consider structural models of mobility because it suggests that wage concerns are only a part of the story driving job changes.

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}, implying a faster integration in the labor market post education. The arrival rate of job offers over the life-cycle implies that at age 30, one job is 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 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.[24]

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.

Table 11: The stochastic process of individual productivity

		At least some College		High School or Less	
Parameter		Estimate	st. error	Estimate	st. error
σ_ϵ	Transitory shock, wages	0.051	(0.0013)	0.068	(0.0007)
ρ	AR(1) coefficient	0.971	(0.0001)	0.962	(0.0001)
σ_P	Initial perm. productivity, wages	0.335	(0.0002)	0.303	(0.0003)
	Mixture of normals for persistent productivity shocks				
μ_{ζ_1}	mean of distribution 1	0.0008	(0.0000)	-0.0014	(0.0000)
σ_{ζ_1}	standard dev. of distribution 1	0.0006	(0.0003)	0.011	(0.0007)
μ_{ζ_2}	mean of distribution 2	-0.007	-	0.017	-
σ_{ζ_2}	standard dev. of distribution 2	0.280	(0.0007)	0.281	(0.0006)
λ_m	Probability of distribution 1	0.897	(0.0007)	0.924	(0.0007)

Wages at labor market entry show a remarkable amount of dispersion (as measured by σ_P). Thereafter the shocks are quite persistent. However, recall that the autocorrelation coefficient ρ is quarterly, which implies that wages are not a random walk for either of the two groups. For example, after 10 years only 30% of a shock to high education workers remains; for the low education group 21% of the shock remains after that amount of time.

[24]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.

A feature of the wage data is heavy tails; one interpretation of this is that workers occasionally obtain large wage increases, possibly reflecting promotions, while otherwise there are small fluctuations reflecting small adjustments to pay possibly because of sickness or other events. To capture this we allow the distribution of individual productivity shocks to be a mixture of Normals, which allows for very general structure of moments. As we showed in the section on the model fit we are indeed able to match the observed kurtosis of wages. In Table 11 we show the estimated parameters of the mixture (μ_{ζ_s}, σ_{ζ_s}, s=1,2 and λ_m). The key feature here is that with some low probability the individual draws a shock from a distribution with a very high standard deviation. Thus for the higher education group with a probability of 0.1 ($1-\lambda_m$) the individual draws an idiosyncratic productivity shock with a standard deviation of 0.28. Otherwise the shock standard deviation is very small (0.0006). The result is very similar for the lower education group, except that the standard deviation in the more frequent regime is higher (0.011). 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.

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	0.1031	(0.0014)	0.1931	(0.0002)
κ^P	Permanent firm shock, match value	0.3105	(0.0013)	0.0810	(0.0017)
σ_{ψ^T}	Transitory idiosyncratic shock, match value	0.0009	(0.0004)	0.0071	(0.0008)
σ_{ψ^P}	Permanent idiosyncratic shock, match value	0.0003	(0.0001)	0.0000	(0.0001)
$\sigma_{\psi^{init}}$	Permanent initial shock, match value	0.0056	(0.0006)	0.0775	(0.0010)

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.

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 10% of a transitory shock is transmitted to workers. This is is not large but still substantial and given the size of our data set the impact is highly significant.

Permanent shocks, on the other hand, are transmitted to a much larger extent, with 31% 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). The implication of this result is of considerable firm level market power, allowing the firm to adjust wages to reflect its fortunes. 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. This is also a strong result because it points to mechanisms of rent sharing, rather than the results of credible renegotiation as in Lise, Meghir, and Robin (2016). In that model an improvement in the productivity of the firm should not lead to increased wages because workers do not have a credible threat to quit: if they were happy with their wage before they should continue to be so following the improvements of the firm's fortunes.

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 19% transmission coefficient) but much less so in response to permanent shocks, where the effect is only 8%. 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, which does not increase, in contrast to that of the higher skill workers.

The remaining coefficients in Table 12 relate to the idiosyncratic match value. This is a component of wage variation that relates to the specific worker firm match, but is purely idiosyncratic to the pair and is not shared in equal measure by similar workers within the firm (unlike the "rent sharing" component we commented on above). In settings where information on firm performance is missing, this distinction is lost, while it plays an important role here and can be separately identified from the impact of firm level shocks.[25]

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 a larger role for an initial variance of offers (0.0775), but beyond that idiosyncratic match effects are effectively zero, playing no role in the variance of wage growth. The important point that emerges from these results is that a large fraction of "match 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 spatial correlation of wages between workers in a firm. Allowing for idiosyncratic match value is not particularly important: 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. Such non competitive behavior seems to be much more important for workers with higher

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

education.

5.3 Simulations

The identity of the firm in which one works appears to have a substantial impact on the evolution of wages over the lifecycle, pointing to non-competitive behavior. 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. This is beyond the issues of sorting that other authors have identified and relates to the level of wages and firm productivity. 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 40,000 individuals. To start off the lifecycle, individuals receive an initial offer with probability λ_{entry} in the first period. 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 wage 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 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 above.

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 0.13, compared to the full variance of 0.19. In other words, permanent firm level shocks, which are transmitted to wages, explain 32% 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

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	0.1104	0.8906	0.0345	0.0954	0.7909	0.0230
35	0.1531	0.9059	0.0249	0.0912	0.8727	0.0252
45	0.1721	0.9220	0.0224	0.0922	0.8901	0.0175
55	0.1869	0.9060	0.0238	0.0930	0.9122	0.0099
			Panel B: No Firm Shocks			
Age	Var(Earnings)	Participation	Mobility	Var(Earnings)	Participation	Mobility
26	0.1089	0.8902	0.0330	0.0939	0.7906	0.0224
35	0.1306	0.8996	0.0247	0.0877	0.8726	0.0229
45	0.1333	0.9142	0.0240	0.0876	0.8893	0.0169
55	0.1344	0.8927	0.0257	0.0881	0.9123	0.0093
			Panel C: No Idiosyncratic Match Effects			
Age	Var(Earnings)	Participation	Mobility	Var(Earnings)	Participation	Mobility
26	0.1104	0.8906	0.0343	0.0894	0.7910	0.0219
35	0.1530	0.9059	0.0248	0.0850	0.8709	0.0258
45	0.1722	0.9221	0.0223	0.0856	0.8886	0.0182
55	0.1871	0.9058	0.0237	0.0863	0.9112	0.0099
			Panel D: No Firm Shocks, No Match Effects			
Age	Var(Earnings)	Participation	Mobility	Var(Earnings)	Participation	Mobility
26	0.1089	0.8903	0.0330	0.0879	0.7905	0.0209
35	0.1306	0.8995	0.0249	0.0813	0.8711	0.0237
45	0.1332	0.9139	0.0240	0.0812	0.8879	0.0171
55	0.1344	0.8924	0.0258	0.0881	0.9109	0.0092

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). For the lower skill workers, switching off transmission has a much smaller effect, but again in the same direction: wages would have been slightly less dispersed if it were not for firm level shocks.

In Panel C we switch off match-specific effects and Panel D eliminates both firm-level shocks and idiosyncratic match components. As we would expect from the parameter estimates, standard match effects do not contribute to the cross-sectional variance of high-skilled workers. Perhaps surprisingly, these shocks do not explain much of the overall participation

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	0.1105	0.8905	0.0000	0.0955	0.7908	0.0000
35	0.1567	0.9016	0.0000	0.0923	0.8712	0.0000
45	0.1860	0.9140	0.0000	0.0933	0.8884	0.0000
55	0.2090	0.8942	0.0000	0.0941	0.9110	0.0000
			Panel B: No Job-to-Job Mobility, No Firm Shocks			
Age	Var(Earnings)	Participation	Mobility	Var(Earnings)	Participation	Mobility
26	0.1089	0.8902	0.0000	0.0940	0.7905	0.0000
35	0.1306	0.8996	0.0000	0.0877	0.8714	0.0000
45	0.1332	0.9141	0.0000	0.0872	0.8878	0.0000
55	0.1344	0.8926	0.0000	0.0877	0.9112	0.0000
			Panel C: Full Participation			
Age	Var(Earnings)	Participation	Mobility	Var(Earnings)	Participation	Mobility
26	0.1108	1.0000	0.0395	0.0953	1.0000	0.0306
35	0.1556	1.0000	0.0268	0.0909	1.0000	0.0286
45	0.1768	1.0000	0.0221	0.0921	1.0000	0.0189
55	0.2000	1.0000	0.0218	0.0939	1.0000	0.0101
			Panel D: Full Participation, No Firm Shocks			
Age	Var(Earnings)	Participation	Mobility	Var(Earnings)	Participation	Mobility
26	0.1092	1.0000	0.0376	0.0939	1.0000	0.0296
35	0.1331	1.0000	0.0279	0.0865	1.0000	0.0262
45	0.1358	1.0000	0.0266	0.0868	1.0000	0.0186
55	0.1383	1.0000	0.0288	0.0871	1.0000	0.0097

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. This is true for both education groups. The implication of these results is that the low skill labor market looks much more competitive and closer to the standard paradigm than that of the high skill group. This is consistent with other studies with US data, following a different methodology and data, such as Lise, Meghir, and Robin (2016).

In Table 14 we explore the role of mobility and participation choices for overall wage 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 illus-

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	0.014	0.015	0.015	0.016	0.016	0.015
35	0.172	0.200	0.169	0.040	0.052	0.051
45	0.291	0.396	0.302	0.053	0.070	0.061
55	0.391	0.555	0.446	0.056	0.073	0.078

trates 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 12% for high-skilled workers but only 1.2% for low-skilled workers. Job mobility is thus an important earnings (and hence consumption) smoothing mechanism for high skill workers.

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, firm level shocks account for about 55% of the cross sectional variance, a much larger share than when workers are allowed to change jobs. For low skill workers this number is about 7%, so although much less important, it is still a substantial amount.[26]

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 4.5% for the high skill group. However, it has little to no impact for the lower skill group. Switching off the transmission of firm level shocks eliminates this impact.

[26] 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 wage variance and participation from Panel C of Table 13 apply to the case of no mobility as well.

The various results are summarized in Table 15. The results emphasize the crucial role of the firm in determining careers and how mobility and participation choices can mitigate some of these effects. 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.

The role of selection To illustrate this point further, 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) (GPS) and most of the literature that followed. The table shows how the estimated firm contribution to the variance of wage growth differs in our model and in that of GPS. For example, our model implies that the standard deviation of wage growth for stayers attributable to permanent firm shocks is 0.055. The GPS model implies a much smaller contribution of 0.033. Overall for high skilled workers the firms contribute 32% of the variance of wage growth, while the implied number in GPS is almost half at 18%. Thus selection plays a very strong role in censoring the impact of the firm. For lower skill workers this is less of an issue because the role of the firm is much reduced.

Thus, focusing on stayers gives the impression of a much lower transmission rate of firm shocks to wages, which in turn implies a bias towards more competitive labor markets. 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 a larger 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, Guiso, and Pistaferri (2017, 2018) use this insight to study

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
$sd(\Delta w\ firm\ trans)$	0.0248	0.0072	0.0479	0.0261
$sd(\Delta w\ firm\ perm)$	0.0546	0.0334	0.0059	0.0258
$sd(\Delta w\ firm)$	0.0601	0.0342	0.0483	0.0367
Share firm shocks	0.324	0.184	0.306	0.232

how exogenous permanent firm shocks passing through wages impact household savings and portfolio choices, respectively.

6 Conclusion

The extent to which the firm in which a person works is a factor in their wages and their fluctuations is an important question both from the perspective of understanding the degree of labor markets competitiveness and to identify the sources and nature of uncertainty that individuals face. 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, directly addressing this question. Our model accounts for endogenous participation and mobility decisions and thus deals with the potential truncation in the impact of productivity shocks 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 firm productivity shocks is 0.31. In other words firms pass a third of their permanent change in their fortunes to wages. However transitory (i.i.d.) shocks have no impact on the wages of high skill workers. They do however affect the wages of the low skill workers with an elasticity of 0.19; yet, this does not have a large impact on wage profiles. We find that the variance of wages increases over the life-cycle for high skill workers. By age 55 about 39% 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. On the other hand, for lower skill workers we find that the firm has a much lower impact, with wages not depending much either on firm level shocks or even on idiosyncratic match specific effects. For them about 5.6% of the cross sectional variance can be attributed to the firm by age 55.

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 is transferred across jobs. It varies over time due to transitory and permanent components - for example because of short-lived spells of sickness or long-lasting 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 from them. This is consistent with union protection being more important for these workers. Indeed, one way of interpreting the results is that the wages of low-skill workers are close to the minimum wage thresholds set in collective bargaining agreements, reducing the transmission of negative firm-level shocks onto wages.[27]

References

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

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). '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.

Bonhomme, S. and Robin, J.-M. (2009). 'Assessing the Equalizing Force of Mobility Using Short Panels: France, 1990-2000.' *Review of Economic Studies*, vol. 76(1):pages 63–92.

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.

Carlsson, M., Messina, J., and Skans, O. N. (2016). 'Wage Adjustment and Productivity Shocks.' *The Economic Journal*, vol. 126(595):pages 1739–1773.

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

[27]Saez, Schoefer, and Seim (2017) argue that the union minimum wage floors mostly bind for new, young employees.

Fagereng, A., Guiso, L., and Pistaferri, L. (2017). 'Firm-Related Risk and Precautionary Saving Response.' *American Economic Review Papers and Proceedings*, vol. 107(5):pages 393–397.

Fagereng, A., Guiso, L., and Pistaferri, L. (2018). 'Portfolio choice, firm shocks and uninsurable wage risk.' *Review of Economic Studies*, vol. 85(1):pages 437–474.

Geweke, J. and Keane, M. (2000). 'An empirical analysis of earnings dynamics among men in the PSID: 1968-1989.' *Journal of Econometrics*, vol. 96:pages 293–356.

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.

Guvenen, F., Karahan, F., Ozkan, S., and Song, J. (2019). 'What Do Data on Millions of U.S. Workers Reveal About Life-Cycle Earnings Dynamics?' *University of Minnesota mimeo*.

Hansen, L. P. (1982). 'Large Sample Moments of Generalized Method of Moments Estimators.' *Econometrica*, vol. 50:pages 1029–1054.

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.

Roberts, G. O., Gelman, A., and Gilks, W. R. (1997). 'Weak convergence and optimal scaling of random walk Metropolis algorithms.' *The Annals of Applied Probability*, vol. 7(1):pages 110–120.

Saez, E., Schoefer, B., and Seim, D. (2017). 'Payroll Taxes, Firm Behavior, and Rent Sharing: Evidence from a Young Workers' Tax Cut in Sweden.' NBER Working Paper No. 23976.

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.

A Wage Residuals

The results for the first-stage estimation are presented in Table 17. For readability, we suppress region FE, time FE and region-time interactions in the participation equation, and region effects as well as industry-time FE in the wage regression.

First, consider the results for participation choices in Table 17. Column (1) reports the results for workers with high school education or less. Column (3) reports the results for workers with at least some college. The results are probit estimates, and we focus on their sign patterns. For both groups, having children up to three years of age significantly decreases the probability of participating in the labor market, but older children increase participation.

Temporary absence is facilitated by the Swedish system of parental leave benefits that offers 80% of previous wages for up to 13 months with a very generous cap. The full benefit

Table 17: First-Stage Results: Participation and Log wages

	High School of less		Some College	
	Participation	Log wages	Participation	Log wages
age	0.4066	0.3288	0.6457	0.7442
	(0.019)	(0.004)	(0.033)	(0.010)
age^2	-0.3814	-0.1997	-0.7058	-0.4223
	(0.025)	(0.004)	(0.045)	(0.010)
age^3	0.1655	0.0617	0.2967	0.1259
	(0.013)	(0.002)	(0.023)	(0.005)
age^4	-0.0229	-0.0072	-0.0413	-0.0148
	(0.002)	(0.000)	(0.004)	(0.001)
child 0-3 yrs	-0.0492	-0.0344	0.0087	-0.0114
	(0.003)	(0.000)	(0.005)	(0.001)
child 4-6 yrs	0.0234	-0.0021	0.0626	0.0278
	(0.003)	(0.000)	(0.004)	(0.001)
child 7-10 yrs	0.0192	-0.0047	0.0598	0.0208
	(0.003)	(0.000)	(0.005)	(0.001)
child 11-17 yrs	0.0677	0.0107	0.1211	0.0373
	(0.003)	(0.000)	(0.005)	(0.001)
married	0.3236	0.0996	0.2089	0.1334
	(0.003)	(0.001)	(0.005)	(0.001)
parental leave	0.0184	-0.0369	0.0309	-0.0357
	(0.001)	(0.000)	(0.001)	(0.000)
sickness benefits	-0.0933	-0.0739	-0.1010	-0.1023
	(0.000)	(0.000)	(0.001)	(0.001)
Mills ratio		0.4966		1.0987
		(0.007)		(0.021)
Mills ratio * age		-0.1941		-0.4552
		(0.006)		(0.020)
Mills ratio * age^2e		0.0459		0.0849
		(0.002)		(0.006)
Year FE	Yes	Yes	Yes	Yes
Industry-Year FE	No	Yes	No	Yes
Region FE	Yes	Yes	Yes	Yes
Region-Year FE	Yes	No	Yes	No
Observations	31,091,423	7,114,874	11,188,448	2,643,040
R-squared	0.074	0.162	0.041	0.173
Wald test [df=220]	19425.29		5881.99	
Wald test [p-value]	0.0000		0.0000	

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

Note: Robust standard errors in parentheses. Wald tests report test statistics and p-values for the exclusion restriction of region-time interactions in each specification. We use a Probit model for participation and report the Pseudo R-Squared.

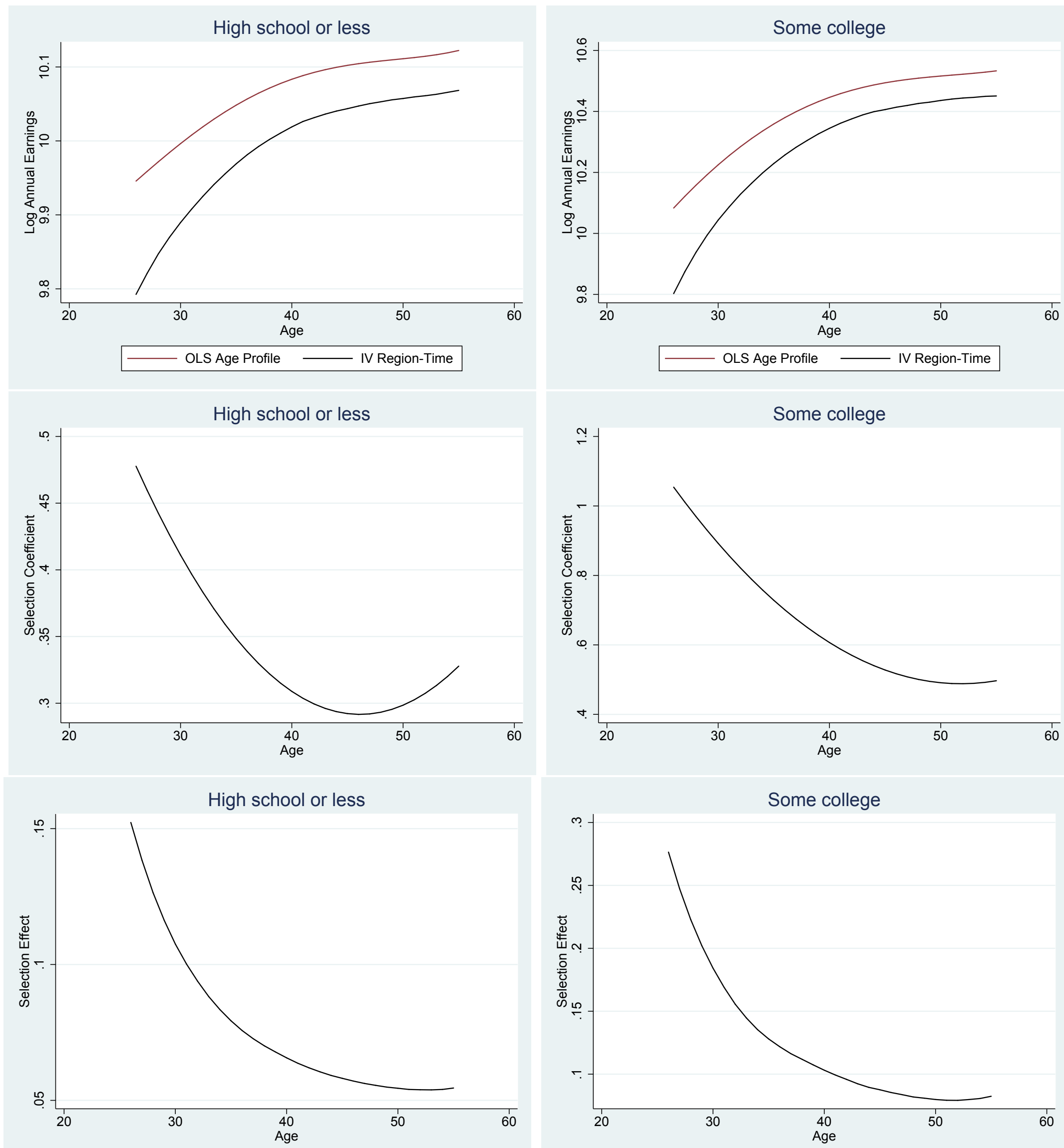
period only applies if the father also stays with the child for some time, which is consistent with the lower participation probability for men with young children. Interestingly, married men are more likely to work in general.

The coefficients on parental leave and sickness benefits confirm the measurement problems in employment status described above. In particular, parental leave payments increase the probability of being employed. 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. The coefficient for sickness benefits is negative and significant for both 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.

Next, consider the results for wages in columns (2) and (4) of Table 17 respectively. The results confirm the familiar concave life-cycle profile of wages. The predicted wage profiles across the lifecycle are illustrated graphically in the top row of Figure 5. As we can see from the comparison with simple OLS wages profiles, the model predicts that selection has an effect on the slope of the wages 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 wages is underestimated.

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

Figure 5: Wage Profiles and Selection



corresponding to the regression results in Table 17 can be found in the second row of Figure 5. For both education groups, 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. These patterns directly mirror the results for wages profiles taking selection into account. Overall, the wage regression implies a positive and significant selection effect for both samples. As the average selection effects by age in the third row of Figure 5 suggest, wage differences because of selection are in the range of 0-20%, where these effects are higher for highly educated workers.

B Deriving Standard Errors

Define an outcome k relevant for period t and individual i as y_{kit}. This could be the log wage or the log wage squared or the log wage in t multiplied by the log wage in period $t-1$. The expected value of this moment given the model is denoted by $E(y_{kit}) = g_k(\theta)$. This is a function of the p parameters of the model θ. The empirical counterpart for g_k is

$$\hat{g}_k = \frac{1}{\Sigma_i^{N_k} T_{ki}} \Sigma_i^{N_k} \Sigma_t^{T_{ki}} y_{kit}$$

where T_{ki} the number of observations over time used for moment k for the case of individual i, N_k is the number of individuals used in computing moment k.

The model counterpart is

$$\widehat{g_k(\theta)} = \frac{1}{\Sigma_i^{N_k} T_{ki}} \Sigma_i^{N_k} \Sigma_t^{T_{ki}} g_{kit}(\theta)$$

where $g_{kit}(\theta)$ is a function defined by the model and predicting an individual level outcome such as participation or mobility. The $\widehat{}$ denotes the fact that this is a simulated object. Given the data for each individual we can use many simulations to improve the approximation and mitigate simulation error. We henceforth drop the $\widehat{}$ for simplicity of notation and assume that there are enough simulations to make simulation errror negligible.

We associate a weight with each moment. Denote the $k \times k$ weight matrix by Ω with diagonal element ω_k. The average of these predictions is the finite sample model counterpart of the moment we are fitting as defined above.

We only take diagonal weight matrices here. The criterion to be minimized is

$$D = \frac{1}{2} min_\theta [\Sigma_{k=1}^{K} \omega_k (g_k(\theta) - \hat{g}_k)^2]$$

Define the $k \times 1$ vector of moments as $g(\theta)$ and the $k \times p$ matrix of first derivatives by $G(\theta)$. The k-th row is denoted by $g_k'(\theta)$ and is a $1 \times p$ vector.

The first order conditions for minimizing D are

$$\frac{\partial D}{\partial \theta} \equiv \Sigma_{k=1}^{K} \omega_k (g_k(\theta) - \hat{g}_k) \frac{\partial g_k(\theta)}{\partial \theta} = 0$$

Approximating the first order conditions around the true value θ^0 we get

$$\frac{\partial D}{\partial \theta^0} + \frac{\partial^2 D}{\partial \theta^0 \partial \theta^{0\prime}} (\hat{\theta} - \theta^0) = 0$$

which gives

$$\hat{\theta} - \theta^0 \simeq - (\frac{\partial^2 D}{\partial \theta^0 \partial \theta^{0\prime}})^{-1} \times \frac{\partial D}{\partial \theta^0}$$

Hence the variance of the method of moments estimator is

$$Var(\hat{\theta}) = (\frac{\partial^2 D}{\partial \theta^0 \partial \theta^{0\prime}})^{-1} E(\frac{\partial D}{\partial \theta^0} \times \frac{\partial D}{\partial \theta^{0\prime}})(\frac{\partial^2 D}{\partial \theta^0 \partial \theta^{0\prime}})^{-1}$$

Taking each component in turn and evaluating it at the estimated $\hat{\theta}$ and taking plims we have that

$$plim[\frac{\partial^2 D}{\partial\theta^0\partial\theta^{0'}}] = \Sigma_k\omega_k[plim(g_k - \hat{g}_k)\frac{\partial^2 g_k}{\partial\hat{\theta}\partial\hat{\theta}'} + plim\frac{\partial g}{\partial\hat{\theta}} \times \frac{\partial g}{\partial\hat{\theta}'}] = \Sigma_k\omega_k[plim\frac{\partial g}{\partial\hat{\theta}} \times \frac{\partial g}{\partial\hat{\theta}'}] = G'\Omega G$$

where G is the $k \times p$ matrix of first derivatives of the moments. The k-th row contains the derivatives of of the k-th moment with respect to all parameters.

We can write the first order conditions as

$$\frac{\partial D}{\partial\hat{\theta}} = G'\Omega(g(\hat{\theta}) - \hat{g})$$

with $g(\hat{\theta})$ being the vector of moments from the model evaluated at the estimated parameters $\hat{\theta}$ and $\hat{g}$ being their data counterparts. Hence the covariance matrix for the estimated parameters is given by

$$\hat{V}(\hat{\theta}) = (G'(\theta)\Omega G(\theta))^{-1}G'(\theta)\Omega E\left[(g(\hat{\theta}) - \hat{g})(g(\hat{\theta}) - \hat{g})'\right]\Omega G(\theta)(G'(\theta)\Omega G(\theta))^{-1}$$

To estimate $E\left[(g(\hat{\theta}) - \hat{g})(g(\hat{\theta}) - \hat{g})'\right]$ for arbitrary heteroskedasticity and serial correlation we need to express each element of $(g(\hat{\theta}) - \hat{g})$ as

$$g_k(\hat{\theta}) - \hat{g}_k = \frac{1}{\Sigma_{i=1}^N T_{ki}}\Sigma_{i=1}^N\Sigma_{t=1}^{T_{ki}}(g_{kit}(\hat{\theta}) - y_{kit}) \equiv \frac{1}{\Sigma_{i=1}^N T_{ki}}\Sigma_{i=1}^N\Sigma_{t=1}^{T_{ki}}v_{kit}$$

For variables such as frequency of unemployment at age a we have that

$$v_{kia} = y_{kia} - g_{kia}(\hat{\theta})$$

where y_{kia} is the value of the outcome (say unemployed or not) for person i in period t when their age is a and all other variables that enter the moment are evaluated at the value for person i in period when they are age a. If a is an interval say 26-30 then the person will appear five times, possibly with other conditioning variables (if present) taking on different values each time. While a will not change the other predictive variables may change. For variables such as $V(\Delta\tilde{e}_t|E_{t-1}=1, E_t=1, J_t=0)$ we will get

$$v_{kit} = (\tilde{e}_{it} - \tilde{e}_{it-1})^2 - (\text{predicted amount for this object by model for person } i \text{ in period } t)$$

This will be operative for the periods where the conditions are true and this will define T_{kit}. Note that $plim_{N\to\infty}E\left[(g(\hat{\theta})-\hat{g})(g(\hat{\theta})-\hat{g})'\right]=0$ so long as there is because once we have imposed independence across individuals the numerator will be of order N while the denominator of order N^2.

So the (k,s) element of $E\left[(g(\hat{\theta})-\hat{g})(g(\hat{\theta})-\hat{g})'\right]$ can be written as

$$E\left[(g(\hat{\theta})-\hat{g})(g(\hat{\theta})-\hat{g})'\right]_{k,s} = plim_{N\to\infty}\left[\frac{1}{\Sigma_{i=1}^N T_{ki}\Sigma_{i=1}^N T_{si}}\Sigma_{i=1}^N\Sigma_{t=1}^{T_{ki}}\Sigma_{q=1}^{T_{si}}v_{kit}v_{siq}\right]$$

A more complex issue is the variance related to the spatial correlation

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

One approach would be to assume that all the independent variation comes from between firms. Then denoting

$$\rho_{\Delta\tilde{e}} - g_\rho(\hat{\theta}) = \Sigma_{j=1}^M v_j$$

where M is the number of firms. Then the variance for this residual will be

$$Var(\rho_{\Delta\tilde{e}} - g_\rho(\hat{\theta})) \simeq \Sigma_{j=1}^M v_j^2$$

Similarly the covariance of $\rho_{\Delta\tilde{e}} - g_\rho(\hat{\theta})$ with the other elements of $g(\hat{\theta})-\hat{g}$.