NBER WORKING PAPER SERIES

EARNINGS DYNAMICS AND FIRM-LEVEL SHOCKS

Benjamin Friedrich
Lisa Laun
Costas Meghir
Luigi Pistaferri

Working Paper 25786
http://www.nber.org/papers/w25786

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2019, Revised September 2021

We thank the IFAU for providing the data and supporting the project. We are grateful to participants in seminars at Chicago, Northwestern, SED, 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Earnings Dynamics and Firm-Level Shocks
Benjamin Friedrich, Lisa Laun, Costas Meghir, and Luigi Pistaferri
NBER Working Paper No. 25786
April 2019, Revised September 2021
JEL No. I18,J24,J31,J63

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. For low-skilled the pass-through is similar for temporary and permanent firm-level shocks and the magnitude smaller. 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. In particular, cross-sectional wage variances by age 55 are roughly one-third higher relative to a scenario with no pass-through of firm shocks onto wages.

Benjamin Friedrich
Kellogg School of Management
Strategy Department
2211 Campus Drive
Evanston, IL 60208
benjamin.friedrich@kellogg.northwestern.edu

Lisa Laun
Institute for Evaluation of Labour Market
and Education Policy (IFAU)
Box 513
SE-751 20 Uppsala
Sweden
Lisa.Laun@ifau.uu.se

Costas Meghir
Department of Economics
Yale University
37 Hillhouse Avenue
New Haven, CT 06511
and IZA
and also NBER
c.meghir@yale.edu

Luigi Pistaferri
Department of Economics
579 Serra Mall
Stanford University
Stanford, CA 94305-6072
and NBER
pista@stanford.edu

1 Introduction

Workers face multiple sources of labor market risks. In the workhorse competitive labor market model, they only bear the risk of shocks to their *own* productivity, which they carry with them wherever they work, and they bear them fully. But labor and credit market frictions weaken this extreme view. A recent literature has argued that job search costs on the two sides of the labor market (or the presence of non-monetary components that workers or employers value, such as job amenities or employee loyalty), underscore a role for firms to affect both the levels and the dynamics of wages over an individual's career. This research agenda also reflects the growing availability of employer-employee data with detailed information on worker and firm characteristics, offering the opportunity to better understand the sources of inequality and of labor market risks that individuals face over the life cycle, and in particular how much of those risks are affected by who we work for.

Most of the literature to date has focused on the extent to which wages of individuals are related to the firm in which they are employed, with various mechanisms being proposed, such as sorting or rent sharing. See, e.g., Abowd, Kramarz, and Margolis (1999) (AKM) as well as more recent papers such as Card, Heining, and Kline (2013) and Card et al. (2018), who provide an excellent review and interpretation of the literature. There is comparatively less research on the extent to which shocks to the firm's fortunes (such as product market or technology shocks) pass onto wages, partly because of more stringent data requirements. As an example, Lise, Meghir, and Robin (2016) do consider the transmission of productivity shocks to wages, but restrict themselves to the implications driven by a specific structural model. Moreover, they do not have access to matched employer-employee data. Balke and Lamadon (2021) develops a structural model with directed job search that offers a theoretical framework for understanding the role of firm-level shocks for worker outcomes and tests some of its implications on Swedish administrative data. Besides job search frictions, the transmission of firm shocks onto wages is often interpreted as reflecting the extent that workers share rents with the firm, and an early study in this direction is Van Reenen (1996). Following a similar line of interpretation, Lamadon, Mogstad, and Setzler (2019) show how the pass-through of firm-level shocks to wages reflects wage-setting power, while Kline, Petkova, Williams, and Zidar (2019) look at how firm shocks originating from the allowance of patent applications impact wages – in both cases using US administrative data. Garin and Silvério (2019) use Portuguese data to measure firm-specific trade shocks during the Great Recession and find that wages particularly respond to firm shocks in less competitive labor markets. 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 partial insurance the firm provides, reflecting imperfections faced by workers in credit and insurance markets. For the US, Juhn et al. (2018) find that top employees' wages are more sensitive to firm shocks than those of rank-and-file workers, consistent with the idea that in certain occupations performance pay acts as a countervailing force to wage insurance.

A common limitation of these papers is that they ignore job-to-job mobility and the transitions between employment and unemployment.[1] Such transitions may well hide the impact of firm-level shocks on wages because a worker may quit or switch jobs instead of suffering too large a pay cut, causing wage growth to be censored. Another element that is missing from these papers is a characterization of how much of the lifetime risk faced by a worker is explained by the firm, how it differs across different type of workers, and how it is affected by movements across employment states and across employers. The goal of our paper is to fill these gaps.

We remain agnostic about the specific structural model that generates the data and build instead on the literature modeling the stochastic structure of earnings.[2] Using matched employer-employee data, we consider a process for earnings that, in addition to individual

[1]Carlsson, Messina, and Skans (2016) distinguish between industry-level and firm-level shocks, but as in Guiso, Pistaferri, and Schivardi (2005) they maintain the focus on stayers at incumbent firms.

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

productivity shocks, allows for match-specific and firm-level productivity shocks passing onto wages. This relates directly 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. With respect to most of the earlier literature on rent sharing, the key innovations are that we distinguish between the nature of the shocks to firms (permanent versus transitory) and whether they translate into permanent shocks to individual wages or just transitory adjustments, which would happen if the workers were to move to another firm or leave employment altogether. This consideration underscores another important element of our framework, which is to allow for endogenous mobility choices and moves between employment and unemployment spells. These choices can have important implications for the measurement of the pass-through of shocks to wages and the identification of different sources of risk. 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.[3] We use our statistical model to examine how wage risks evolve in counterfactual scenarios in which we shut down some sources of variation, such as firm risk or movements across jobs or employment states.

Our data are drawn from Swedish administrative employment records. We match these records with data on firm balance sheets. The result is the universe of workers and firms, matched to each other for the years 1997–2008. The data include annual earnings, detailed information of job histories, including the identity of the firm, and information on value added and total employment. However, it does not include hours of work. We thus focus our

[3] 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)

main analysis on men, who rarely work part-time.[4] 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. Firm productivity also affects the frequency and quality of outside offers. 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) passing onto wages, as well as individual match effects originating from production complementarities. As such, it is a particularly rich framework that effectively nests most of the earlier specifications of the stochastic process of earnings.

We find that firm productivity is quite volatile and that this volatility transmits to wages of high-skill workers to a larger extent than for low-skill workers, 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. We also find that employment is related to wage shocks, consistent with self-selection into work and work incentives, although the implied elasticity is rather low. Finally, job mobility is sensitive to wage improvements, although other factors may lead to pay cuts when moving across workplaces.

To better understand the implications of our findings, we simulate the model in a number of counterfactual scenarios in which we change the nature of wage variability faced by workers 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 (match productivity effects, firm shocks pass-through, and origin of outside offers). We find that wage variances over the life cycle decline substantially when eliminating the impact of firm shocks (with the effect

[4] For the years that overlap with our sample period, Labor Force Survey data show that 89% of men aged 16-64 work full-time (35 hours or more); for women, this is 64%. See Friedrich, Laun, and Meghir (2021) for a gender comparison of earnings dynamics in Sweden.

being particularly relevant for the high-skilled), while eliminating match productivity shocks matters little. Specifically, eliminating transmission of firm-level shocks into wages would result in 30% lower cross-sectional wage variation among high-skill workers by age 55. 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, again, this is mostly due to pass-through of firm-specific shocks. Hence, workers' fluidity (the opportunity to quit into unemployment or move to alternative employers) represents an implicit form of insurance against labor market risks.[5] We further use counterfactuals to illustrate that ignoring this insurance motive by only focusing on stayers (as in most of the literature following Guiso, Pistaferri, and Schivardi (2005)) understates the transmission of firm-level permanent shocks to workers by 40% for high-skill workers and by almost half for low-skill workers.

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 specified separately for each of the two education groups we consider (at least some college, and high school or less). As we shall see, these groups differ quite dramatically in how the distribution of wages (mean and variance)

[5] Work by Davis and Haltiwanger (2014) documents a decline in labor market fluidity for the US, with important welfare implications.

evolve over the life cycle. We 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); these shocks may reflect depreciation of skills, health shocks, etc. 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 do 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 risk suffering large pay cuts as a result of negative 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.

We consider a quarterly model for firm productivity, 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.

2.2 The Statistical Model

Firm Productivity As mentioned above, a potentially key source of wage variation is transmission of firm productivity shocks onto wages. Empirically, we measure firm productivity with value added per worker. We also 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). We thus 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,

$$a_{j,t} = a^P_{j,t} + \xi^T_{j,t} \tag{1}$$

where

$$a^P_{j,t} = a^P_{j,t-1} + \xi^P_{j,t} \tag{2}$$

and

$$\begin{aligned} \xi^P_{j,t} &\sim N\left(0, \sigma^2_{\xi^P}\right) \\ \xi^T_{j,t} &\sim N\left(0, \sigma^2_{\xi^T}\right). \end{aligned}$$

The shocks $\xi^P_{j,t}$ and $\xi^T_{j,t}$ may pass through to wages differently. We now turn to a description of the stochastic process for individual wages.

Wages The education-specific log wage equation in period t for individual i who started to work at firm j in period t_0 is given by:

$$\ln w_{i,j(t_0),t} = x'_{i,t}\gamma + P_{i,t} + \varepsilon_{i,t} + v_{i,j(t_0),t}, \tag{3}$$

where all parameters are education specific, but we leave this dependence implicit to avoid cumbersome notation; x are observable worker characteristics such as age. The remaining terms describe individual ($P_{i,t} + \varepsilon_{i,t}$) and match-specific productivity ($v_{i,j(t_0),t}$). We now describe them in detail. Again all distributions are education specific and we leave this implicit.

Individual productivity is subject to transitory shocks $\varepsilon_{i,t} \sim N(0, \sigma^2_\epsilon)$.[6] $P_{i,t}$ is the accu-

[6]We assume that standard measurement error is negligible. Meghir and Pistaferri (2004) point out the

mulation of persistent idiosyncratic productivity shocks and is specified as:

$$P_{i,t} = \rho P_{i,t-1} + \zeta_{i,t}$$

This persistent productivity evolves starting from the initial productivity draw upon entry into the labor market, $P_{i,0} \equiv f_i^{init} \sim N\left(0, \sigma_f^2\right)$. If $\rho = 1$ we have the standard random walk assumption for the permanent component of wages. The productivity shock is denoted ζ_{it} and we make a flexible distributional assumption:

$$\zeta_{i,t} \sim \text{mixture of Normals}\left(\{\mu_{\zeta_1}, \sigma_{\zeta_1}\};\ \{\mu_{\zeta_2},\ \ \sigma_{\zeta_2}\};\ \lambda_m\right) \tag{4}$$

where $\{\mu_{\zeta_s},\ \sigma_{\zeta_s}\}$ $(s = 1, 2)$ represent the mean and standard deviation of each of the two normals in the mixture, and 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 skewness and kurtosis. The importance of higher order moments in earnings growth has been examined in the context of US data by Guvenen 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 $v_{i,j(t_0),t}$, where $j(t_0)$ is the firm j where the worker i started in period t_0. We assume that the match effect evolves stochastically as a result of firm- and match-specific shocks. We

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.

also find it 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:

$$v_{i,j(t_0),t} = v^P_{i,j(t_0),t} + v^T_{i,j(t_0),t} \tag{5}$$

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

$$\begin{aligned} v^P_{i,j(t_0),t} &= v^P_{i,j(t_0),t-1} + \kappa^P \xi^P_{j,t} + \psi^P_{i,j(t_0),t} \\ &= v^P_{i,j(t_0),t_0} + \kappa^P \sum_{s=t_0+1}^{t} \xi^P_{j,s} + \sum_{s=t_0+1}^{t} \psi^P_{i,j(t_0),s} \end{aligned} \tag{6}$$

where κ^P is the pass-through of the permanent firm-level shock ($\xi^P_{j,t}$) to wages . The transitory part of the match component equals:

$$v^T_{i,j(t_0),t} = \kappa^T \xi^T_{j,t} + \psi^T_{i,j(t_0),t} \tag{7}$$

where κ^T is the transitory shock ($\xi^T_{j,t}$) pass-through rate. The initial value of the permanent match productivity (at time t_0) in equation (6) is drawn from a firm-type specific distribution, reflecting the initial idiosyncratic match value,

$$v^P_{i,j(t_0),t_0} = \psi^{init}_{i,j(t_0)} \sim N\left(\tau_{j,t_0}, \sigma^2_{\psi^{init}}\right). \tag{8}$$

Thus we assume that the initial match value of a job depends on firm characteristics at the time of hiring, τ_{j,t_0}. Specifically, we rank firms by their productivity and classify them into four quartiles that differ by the firm premium $\tau_{j,t_0} = \tau_s$, with $s = \{1, 2, 3, 4\}$ (the "firm type"). Since firm-level shocks will affect firm productivity over time, a firm may change its wage premium over time as it moves up or down in the firm ranking.[7]

Finally, the two ψ shocks are i.i.d. normal and capture idiosyncratic match productivity. Specifically we assume that $\psi^l \sim N\left(0, \sigma^2_{\psi^l}\right)$, for $l = \{P, T\}$.

Overall, the match value $v_{i,j(t_0),t}$ reflects two important ways that firms may affect workers' pay: (a) systematic differences in the match quality, and (b) transmission of firm-level shocks to wages over time. Conceptually, firm shocks are shared by all workers within the firm (at least within a broad skill group) and are associated to changes in wage policy, contractual arrangements, etc. In contrast, match-specific shocks are idiosyncratic and are associated to changes in production complementarities, learning, individual performance evaluation, etc. By allowing for match specific shocks that are unrelated to firm-level productivity we guard against the possibility that the productivity shocks just proxy for worker-firm pair effects. Whether they matter in practice is an empirical question.

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, that remains unchanged over the duration of the match and thus is not subject to shocks that could be related to firm-level productivity. As a result, 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

[7] In practice, τ_s is the average residual wage at firm of type s at the time of hiring.

productivity component changes only in response to firm-related permanent shocks.

Sorting. In this paper, wages may depend on fixed individual and firm characteristics in a log-separable way as in Abowd, Kramarz, and Margolis (1999). We do not take a stand on whether there is assortative matching in the labor market (i.e., a correlation between these effects). But note that even if there is sorting based on permanent worker or firm characteristics (as some papers have found), we assume that the mechanism through which firm-level productivity shocks affect wages is common across all firms and hence there is no effect of sorting on wage growth. While this is restrictive, it is still more general than earlier models of wages estimated on matched data because of the richer structure of the 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 the job rather than suffer the resulting pay cut, which may even be permanent (within the firm). Similarly, workers with large pay cuts in firms that have had bad productivity shocks may be more likely to accept alternative job offers.[8]

We model the employment E as:

$$E_{i,t} = \mathbb{1}\left\{z'_{i,t}\delta + \phi\left(P_{i,t} + \varepsilon_{i,t} + v_{i,j(t_0),t}\right) + u^E_{i,t} > 0\right\}. \tag{9}$$

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 (2016)). The coefficient ϕ in part reflects the incentive effect of working but also the im-

[8]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.

portance of unobserved heterogeneity in participation choices, including unobserved factors explaining persistence in the value of employment.[9] Other observable determinants of employment (such as age) are summarized in z.

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

$$J_{i,t} = \mathbb{1}\left\{z_{i,t}'\theta + b\left(v_{i,j(t),t}^{init} - v_{i,j(t_0),t}\right) + u_{i,t}^{J} > 0\right\}, \tag{10}$$

and is also affected by a set of variables z, such as age. Job mobility also depends only on the difference in match values between new and incumbent firms, $(v_{i,j(t),t}^{init} - v_{i,j(t_0),t})$, and not on the remaining stochastic components, because permanent and transitory productivity shocks do not depend on a particular firm match but are portable characteristics of a worker across different jobs. The importance of wage differences as opposed to worker observable characteristics in determining mobility is captured by the parameter b. Note that, relative to the case in which the initial draw of the match component is random, in our case a higher initial draw is also a signal of better career/wage growth opportunities (due to the "homophily" in the arrival rate of outside offers, see below).

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 particular unobserved tastes for work or job mobility. 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. As usual, identification of parameters in (9) and (10) is only up to scale, and hence the variances of u^E and u^J are both normalized to 1. Finally, the observed characteristics in the two equations also reflect labor market attachment and employment and mobility costs.

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

Labor Market Frictions and Job Offers Upon entry in the labor market, workers receive job offers at a rate λ_{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 (10), since the two cannot be separately identified.

These arrival rates create an asymmetry between the probability of entry and exit in employment. Given this flexibility we can interpret the employment equation (9) as reflecting both the decision to work or not and exogenous job destruction. Thus there is no presumption that exit from employment reflects purely endogenous individual decisions.

If a worker receives a job offer while employed, we also model the origin of the offer as a function of the firm type s of the current employer j,

$$Pr\left(\text{offer from firm type } k \mid \text{current firm type } s\right) = \omega_s \cdot \frac{exp(-\omega_{dist} \cdot |k-s|)}{\sum_v exp(-\omega_{dist} \cdot |v-s|)}. \tag{11}$$

This process allows for different offer arrival rates by firm type, ω_s,[10] and further characterizes the offer origin as a function of the current firm type. The latter is based on firm productivity as measured by the wage premium τ_s (see the discussion following equation (8)). In particular, the specification allows for the possibility that workers face higher offer probabilities from firm types that are more "similar" to the current firm type; this is done in order to match the empirical pattern that most job-to-job mobility occurs between similar firms, see Table 3 for details. Hence, equation (11) provides a third channel for firms to affect worker careers, in addition to the firm premium and the transmission of shocks: The identity of the current employer and their performance over time also affect the frequency and quality of outside offers for their employees and will hence affect workers' ability to switch to alternative positions and accumulate search capital over the life cycle. If there is some form of "homophily" in the origin of offers, being in a growing firm may increase the

[10] This is subject to a normalization because we also include a constant term in mobility preferences. We choose to define ω_s relative to always receiving an offer at the highest ranked firms.

magnitude and quality of outside offers, while being in a shrinking firm worsens both.

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), containing information on demographic and socioeconomic variables for the entire working age population in Sweden from 1990 onward. We use information on 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 companies), and we exclude sole proprietors because data for these entities are not available

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.

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 construction, manufacturing, retail and service sectors 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 (Friedrich, Laun, and Meghir, 2021).[11] 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. Heterogeneity by skills reflect exposure to different types of shocks or contractual arrangements.

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

[11] 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 wage 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**

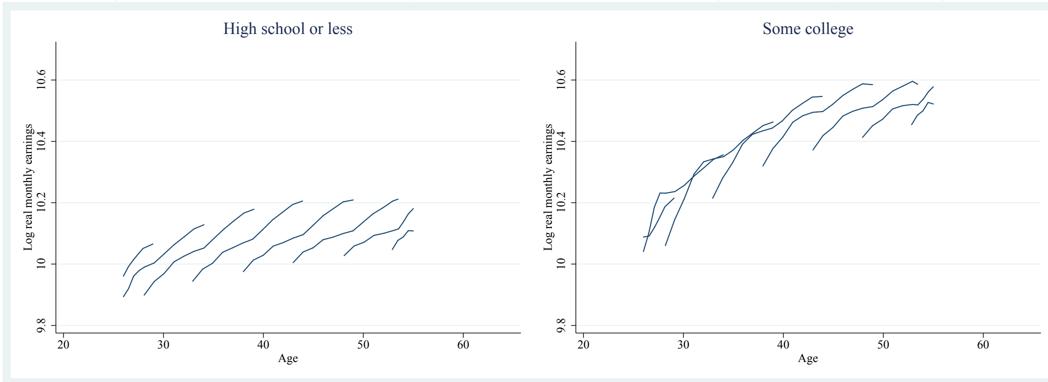
	≤High school	College
No. unique workers	1,152,933	464,616
No. worker-quarter obs.	31,091,423	11,188,448
Monthly earnings	24,960	35,930
(2008 SEK)	(8,009)	(17,115)
Age	40.31	39.07
Married	0.5797	0.6112
Having children	0.4569	0.4985
Employed, of which	0.8764	0.9040
Job stayer	0.9547	0.9493
Job mover	0.0233	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

at their current job each quarter is fairly constant across groups. More educated workers are more likely to experience a job-to-job move, 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.

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 log earnings over the age span in which we observe this particular cohort.[12] The vertical distance between earnings of different cohort groups at a given age can then be interpreted as cohort effects, while the

[12]We use average monthly earnings, obtained dividing annual earnings by the number of months worked.

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

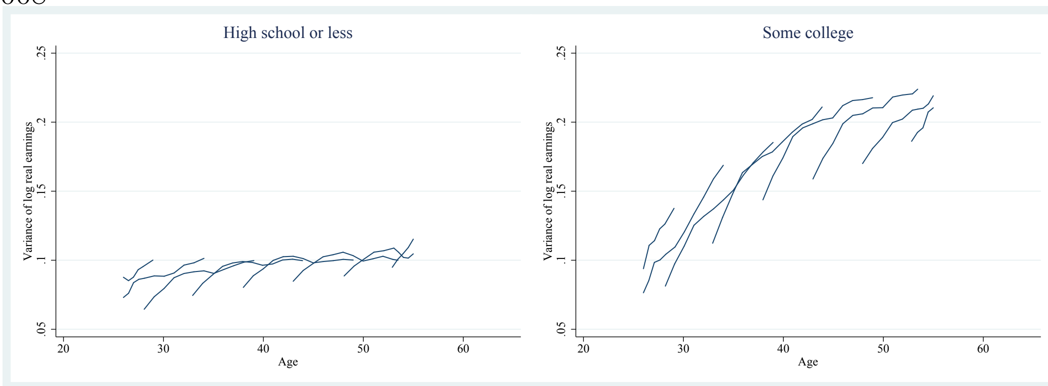


overall slope of the profile can be interpreted as reflecting age effects ignoring for simplicity the usual age, time, cohort identification issues (see for example Gosling, Machin, and Meghir, 2000, for alternatives ways to address these).

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 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.

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

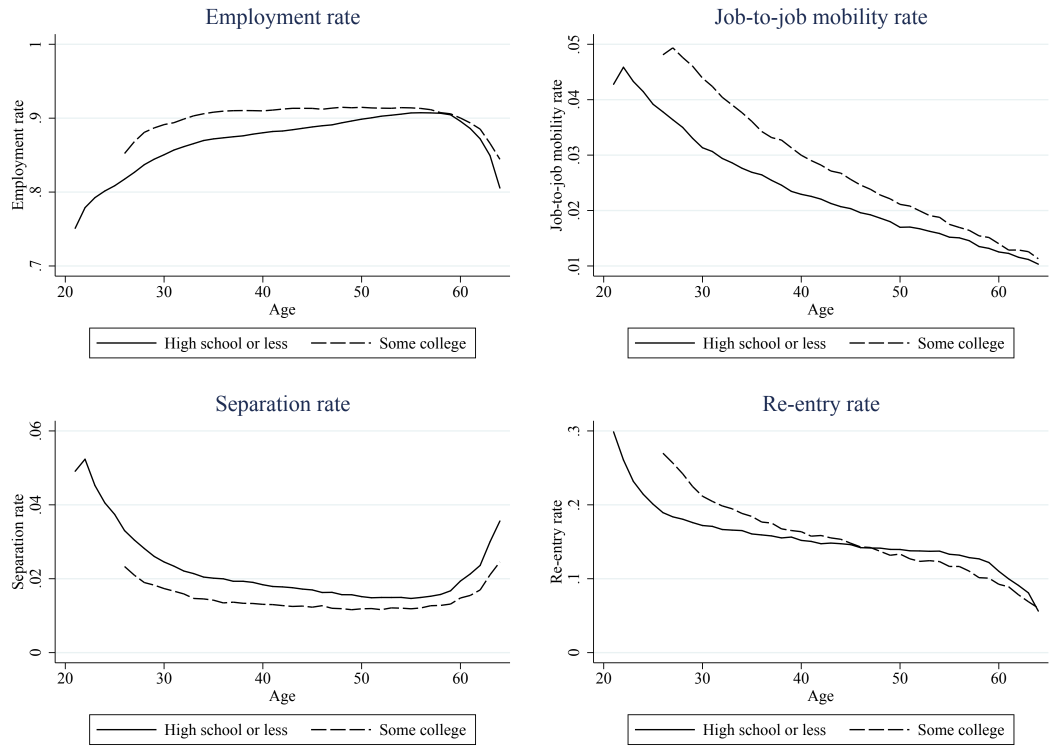


Participation and job transitions The top-left graph in Figure 3 presents the employment rate by age for each education group. In our sample employment rates are above 75% for all age groups. The lower the achieved level of education, the lower is participation at young ages. Interestingly, there is 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 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. In addition, the top-right panel in Figure 3 presents substantial quarterly job-to-job transition rates by age for each education group, in particular at younger ages. Workers with at least some college switch employers more frequently than less educated workers.

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



Mobility, Firms, and Wages It is well documented that a substantial and increasing share of wage inequality is driven by differences across firms (see Card, Heining, and Kline (2013) for Germany, Song et al. (2018) for the U.S., and Akerman et al. (2013) for Sweden).[13] This motivates the investigation of the role of firms for wage inequality and wage dynamics in our paper and suggests that understanding worker mobility across firms may be crucial for our analysis.

In Table 3 we describe mobility patterns between firms sorted by the average wage they pay, and describe the way (residual) wages change between jobs when mobility does not involve an unemployment spell in between jobs, separately for low- and high-skill workers. We compare residual wage growth in the year that precedes and in the year that follows a job move, conditional on no other transition happening in this three-year window. Among job-to-job movers about 43% of low-skill workers and 41% of high-skill workers move to a

[13]We find for our sample that the increase in the role of firms over time is stronger for high-skill workers, see Appendix D.

Table 3: **Data: Job Mobility and Residual Wage Growth by Employment-Wage Quartiles**

		Low-skill workers											
		Departing firm quartile											
		Share of transitions				Residual wage growth				Share wage cuts			
		1	2	3	4	1	2	3	4	1	2	3	4
Arriving firm quartile	1	.134	.072	.043	.017	.063	−.021	−.058	−.11	.414	.575	.626	.669
	2	.066	.095	.063	.018	.104	.006	−.024	−.077	.355	.511	.576	.643
	3	.051	.079	.13	.04	.145	.036	−.001	−.035	.277	.441	.532	.575
	4	.023	.032	.066	.071	.189	.081	.044	.018	.247	.328	.393	.474

		High-skill workers											
		Departing firm quartile											
		Share of transitions				Residual wage growth				Share wage cuts			
		1	2	3	4	1	2	3	4	1	2	3	4
Arriving firm quartile	1	.095	.051	.034	.016	.079	.001	−.022	−.102	.402	.504	.551	.603
	2	.056	.095	.073	.025	.104	.021	−.005	−.049	.353	.457	.53	.578
	3	.042	.096	.137	.065	.138	.046	.032	.001	.297	.404	.451	.5
	4	.023	.037	.077	.078	.19	.094	.071	.031	.257	.332	.368	.444

Note: Employees are sorted on the basis of average residual wages paid at their firm to construct quartiles.

firm of the same wage quartile level; in both groups, about one-third of all moves are to a higher-paying firm, and about a quarter to a lower-paying one.

On average, movers experience positive wage growth when moving to higher ranked firms, especially when leaving firms in the bottom wage quartile. In contrast, mobility to lower ranked firms goes along with zero or negative wage growth, and the decline is larger if the new firm ranks further below the previous employer. These patterns are consistent with the results from Portugal reported in Card et al. (2018). In general, these average changes mask wide dispersion in pay changes as evidenced by the share of wage cuts: for both groups of workers between 25%–67% experience some wage cut when moving from one firm to another. As expected, the share of wage cuts is inversely related to the direction of the move and rank distance between previous and new firm. Our model allows for such wage cuts: the motive for changing jobs, expressed in equation (10), trades off wage improvements to other observed and unobserved reasons for mobility. Understanding the underlying theoretical model that would replicate such intricate patterns would be particularly interesting. Many search models do not allow for wage cuts: the basic Burdett-Mortensen wage posting model excludes them. On the other hand the model by Postel-Vinay and Robin (2002) does allow for wage cuts: the worker may choose to move to a firm where the match surplus is higher; they 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.[14] 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.[15]

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 measure job mobility at a quarterly frequency but observe earnings annually, unless there is a change in employer. Finally, we estimate the remaining parameters of the full model using the simulated method of moments (McFadden, 1989; Pakes and Pollard, 1989) 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. Our measure of firm productivity is log value added (VA) per worker, which is observed annually. By observing data on firm performance we are able to measure firm productivity shocks directly (instead of relying on proxies such as employment, which

[14] Bonhomme and Jolivet (2009) and Sorkin (2018) consider search models in which amenities are valued in utility, also rationalizing the possibility of job-to-job transitions with wage cuts.

[15] Recall, however, that these shares are computed with respect to residual wage growth. Hence, what we call a "pay cut" may simply reflect below-average wage growth.

may be subject to inaction bias due to the presence of adjustment costs). Some studies of how firm productivity affects wages have difficulty establishing causality because of sorting or other mechanisms that confound the causal impact with selection. However, with our approach, which eliminates permanent TFP components as well as permanent effects from wages (and allows for selective mobility and employment), we can better argue that we eliminate confounding factors and we identify causal impacts.

We use a statistical procedure to identify idiosyncratic firm shocks. An alternative approach would be to identify explicitly events that might have affected firm productivity. While this adds transparency and may be attractive in some cases, it has also some important limitations. If the events are external economic shocks, including say access to new infrastructure, then they will have general equilibrium effects and it will not be possible to clearly identify the impact of idiosyncratic effects on productivity. If instead the events are truly specific to the firm, such as an innovation or a new large procurement contract, they offer a more promising identification avenue. However, such events may not capture many of the shorter and longer run changes in the fortunes of firms. These events are almost by design large and salient (or else it would be hard for the researchers to record them), and hence seem more appropriate for testing whether there is pass-through of productivity shocks to wages in the presence of large, identifiable shocks. However, they do not measure the effect of the myriad of events that lead to changes in firm productivity and how these affect the stochastic process of wages. Along the same lines, our distinction of permanent and transitory shocks, while statistical in nature, points to an important direction, which the explicit measurements of shocks may miss: permanent shocks are associated with longer term re-positioning of the firm that may lead to larger changes in the pay structure, while we can reasonably expect transitory shocks to be smoothed over. “Assigning” an explicit productivity event to be permanent or temporary is inherently arbitrary and it may miss the important distinction between shocks of different duration, with implications for the economic interpretation of the pass-through.[16]

One important data issue is time aggregation. While our model is quarterly, in the data we can only construct annual productivity, which means (among other things) that we cannot identify an MA component within the year. To match the data frequency, we apply simulation-based estimation to estimate the stochastic components of the quarterly firm-shock process from data on annual growth in log VA per worker. As we derive in detail in the Appendix, the linear structure for quarterly productivity implies tractable expressions for annual productivity growth. We use the variance and first-order autocovariance of annual changes in firm log productivity as auxiliary moments to estimate the parameters of the productivity process that minimize the distance between these moments in the model and in the data.

We start by running a regression of log VA per worker on the interaction of industry, municipality, firm size, and year, and save the residuals of this regression. Thus the shocks to firm-level productivity that we use are purely idiosyncratic and do not depend on economy-wide, regional, scale or industry effects. This ensures that the transmission of firm-level shocks to individual wages is not confounded by common shocks to workers in the same industry/region/firm scale bins (which may reflect changes in outside options, etc.).

In Panel A of Table 4, we show the autocovariance structure of firm productivity across all firms. The results indicate that a random walk with an i.i.d. transitory component (1)-(2) 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.[17] We find relatively similar empirical autocovariance patterns across industries (see Appendix Table B1), and therefore focus on the pooled sample.

Panel B of Table 4 reports the estimation results for the standard deviations of shocks on a quarterly basis. The implied process for quarterly value added per worker shows sizable

[16]In practice, we see the two approaches as complementing each other, rather than being substitutes.

[17]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 4: **Autocovariance of log Value Added per Worker: Data**

Panel A: Empirical Autocovariance of log Value Added per Worker			
Var (Δa_τ)	Cov $(\Delta a_\tau, \Delta a_{\tau-1})$	Cov $(\Delta a_\tau, \Delta a_{\tau-2})$	Cov $(\Delta a_\tau, \Delta a_{\tau-3})$
0.1791	–0.0537	–0.0041	–0.0022
(0.0008)	(0.0004)	(0.0003)	(0.0003)

Panel B: Structural Estimates of Quarterly Productivity Process			
σ_{ξ^T}	0.4758	σ_{ξ^P}	0.1303
	(0.0016)		(0.0007)

Note: a_τ denotes log of annual productivity in year τ. See the Appendix for details about estimation of the variance of quarterly productivity shock variances.

transitory shocks, which suggest 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.[18] These estimates will be used to draw firm shocks in the simulation estimation below.

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.[19]

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 (3). Based on this first stage, we can then use the wage residuals $\tilde{u}^w_{i,t} = (P_{i,t} + \varepsilon_{i,t} + v_{i,j(t_0),t})$ as the relevant input into the model estimation.

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

[18] 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%. These results are similar when estimating separate stochastic processes by industry, see Appendix Table B2.

[19] Note that firm survivorship bias in the estimation of the productivity shock variances is not an issue since we observe the firm up to the year in which it exits (if any).

tion into work and for the discrepancy in data frequency between model and data.[20]. 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,

$$w_{i,t} = \frac{\sum_{q=1}^{4} E_{i,t_q} \times w_{i,t_q}}{\sum_{q=1}^{4} E_{i,t_q}},$$

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

As we discuss in more detail in the Appendix, aggregating individual wage information at annual frequency even though wages are determined at a higher frequency may lead to aggregation bias akin to the bias due to individual heterogeneity in Blundell, Reed, and Stoker (2003) when analyzing aggregate wages. In addition, seasonality of participation decisions represents a second source of aggregation bias. To address these challenges, we first estimate a discrete choice model for employment (E_{i,t_q}) for each individual at quarterly frequency and construct the Mills ratio ($\lambda^M(.)$) for each of these periods. To make the model consistent with the data, we then aggregate these quarterly selection correction terms in the annual wage model,

$$\log w_{i,t} = x'_{i,t}\gamma + \log \left[\sum_{q=1}^{4} E_{i,t_q} \times e^{\eta \lambda^M (z'_{i,t_q}\delta)} / \sum_{q=1}^{4} E_{i,t_q} \right] + u^w_{i,t} \tag{12}$$

Equation (12) is estimated separately for our two broad education groups (less than college and some college or more). We control for a fourth-order polynomial in age, as well as marital status, dummies for children in different age groups, and region-fixed effects. Industry×year 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

[20]For reasons we explain below, we construct the relevant Mills ratios using more observables than those we assume entering the "structural" employment equation (9).

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.[21] In order to address this type of measurement error, we directly include controls for parental leave and sickness benefits. The presence of these sources of heterogeneity is also the reason for our choice of estimating the wage equation "outside" of the indirect inference step: there is simply too much heterogeneity in wage data that we can ever hope to replicate in simulations of wage profiles by age or education.

The same set of control variables used in the wage equation are also included in the participation choice equation (with the exception of industry, which is missing for non-workers), but we use municipality×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 the community level and the cost of living, in particular housing and rental prices, differ widely across the 290 municipalities 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, $\tilde{u}^{w}_{i,t}$ (excluding the contribution of the selection correction) to construct some key moments for identification (detailed next).

[21] 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).

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.[22] 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 overlapping cohorts of workers and a fixed number of firms in the model.[23] 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-estimate, to determine the evolution of firm types and transmission of firm shocks to current employees.

When entering the labor market after completing their education, some workers receive an offer immediately and others do not. The model includes a probability of this event as a parameter (λ_{entry}), which is estimated by matching it to the actual proportions working in the data. Throughout their career, workers draw job offers according to the offer arrival process in unemployment or on-the-job depending on their current firm type (see equation (11)). For those workers who get an offer, the offer origin (firm type) is drawn based on the offer arrival process for the current or most recent firm type, and randomly if a worker never had a job before.

For previously employed workers, firm shocks (and idiosyncratic match shocks) are real-

[22]We list these here for convenience: the parameters determining participation (δ and ϕ), job-to-job mobility (θ and b), the transmission of firm-related shocks (κ^P and κ^T), the parameters of the stochastic processes determining wage dynamics ($\rho,\ \sigma^2_f,\ \mu_{\zeta_1},\ \sigma^2_{\zeta_1},\ \mu_{\zeta_2},\ \sigma^2_{\zeta_2},\ \lambda_m,\ \sigma^2_{\varepsilon},\ \sigma^2_{\psi^P},\ \sigma^2_{\psi^T},\ \sigma^2_{\psi^{init}}$), wage premia by firm type (τ_s, with $s = \{1, 2, 3, 4\}$), the job arrival rate coefficients ($\lambda_{entry},\ \lambda_{U,0},\ \lambda_{U,1}$) and the coefficients determining the source of outside offers by firm type ($\omega_s,\ \omega_{dist}$).

[23]A simulated economy consists of 4 overlapping cohorts with 6,000 individuals per cohort followed over their entire life cycle and who are matched with 80 firms. We repeat this simulation procedure for 5 independent samples of workers and firms to further increase precision.

ized and affect the current match value. Co-workers at the same firm experience the same firm-level shocks; this will allow us to use the observed spatial correlation of wage growth within a firm to identify the transmission coefficients. Workers compare available offers (if any) to their current job, determine the best option, and decide whether to switch jobs (if applicable) and whether to participate. We then allocate entrants or movers with equal probability to one of the firms in the firm-type group from which their offer originated. We keep track of employment status, firm type, individual productivity, and current match value, because all these factors affect future offers, participation and mobility decisions.

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.[24]

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

4.3.2 Data Moments and Identification

This section describes the choice and computation of the data moments that we use to estimate the model (summarized in Table 5). 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 employment shares and job-to-job mobility rates by age group.[25] These help identify the deterministic part of the participation and

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

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

Table 5: **Moments used in estimation**

Quarterly moments		Annual moments	
(a)	Employment shares (by age group)	(a)	Residual wage level variances (by age)
(b)	Job-to-job mobility rates (by age group)	(b)	Residual wage growth variance (stayers)
(c)	Job creation rates (by age group)	(c)	Residual wage growth autocov. (stayers)
(d)	Job separation rates (by age group)	(d)	Residual wage growth skewness (stayers)
(e)	Job-to-job mobility rates (by firm type)	(e)	Residual wage growth kurtosis (stayers)
		(f)	Average wage growth (movers)
		(g)	Variance residual wage growth (movers)
		(h)	Cov. wage growth and empl. residuals (stayers)
		(i)	Cov. wage growth and empl. residuals (movers)
		(j)	Spatial correlation (stayers)
		(k)	Autocov. average wage growth (stayers)
		(l)	Average wage growth for job movers (by firm type)

job-to-job transition equations (the vectors δ and θ). The shift over the life cycle of mobility rates is also crucial for estimating the impact of differences in firm-specific matches on the probability of a job-to-job move (the parameter b in equation (10)). The second set of moments includes quarterly job creation rates (fractions moving into work from unemployment) and job separation rates (fractions moving from employment to unemployment) for the same age groups as above. 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_{enter}). Moreover, we use job-to-job flows by firm type (see Table 3) to characterize the offer arrival process as a function of the current employer's characteristics (ω_s and ω_{dist}). Separation rates by age groups relate to participation preferences and shock variances (see equation 9), but also partly reflect the lack of available outside offers to switch jobs in the event of adverse shocks. Another set of moments we use is the covariance between wage residuals and participation residuals (obtained as described in section 4.2), which pins down the association between wages and work decisions (ϕ).[26]

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, both levels and growth rates. Since the model assumes quarterly

[26] 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.

processes for all shocks, all simulation outcomes are quarterly as well. As a result, we need to aggregate simulated outcomes such as firm shocks and wages within each year to make the simulation comparable to the observed moments.

A first set of wage moments are the variances of wage residuals at selected points of the life cycle (age 26, 30, 35, 40, 45, 50, and 55; see Figure 4). The level of residual wage variance at the beginning of the life cycle pins down the variance of initial productivity (σ_f^2). The size of the autocorrelation coefficient in permanent productivity (ρ) is instead identified through the life-cycle pattern of the variance of residual wages. We also 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 of a job-to-job move. 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 $(\tilde{e}_{t+1} - \tilde{e}_{t-1})$. 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 ($\sigma^2_{\psi^{init}}$). Moreover, we target average wage growth by type of job transition, defined as pairs of firm types of the previous and new employer. These differences in wage growth are closely related to differences in wage premia τ_s that different firm types offer.

Some of the key structural parameters are the pass-through coefficients 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 (or spatial) correlation of wage growth is defined as:

$$\rho_{\Delta\tilde{e}} = \frac{\sum_{\text{firm } j} \sum_{\text{worker } k\in j} \sum_{\text{worker } l\in j, k\neq l} (\Delta\tilde{e}_{kt} - \Delta\bar{e})(\Delta\tilde{e}_{lt} - \Delta\bar{e})}{\sum_{\text{firm } j} \sum_{\text{worker } k\in j} (\Delta\tilde{e}_{kt} - \Delta\bar{e})^2}$$

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.[27]

4.3.3 MCMC Estimation

We use a Markov-Chain Monte Carlo method (MCMC) to estimate the model. This derivative-free estimation method only requires many evaluations of the objective function at different parameter guesses. This is computationally attractive because the simulated moments may not be smooth. The method can deal with large parameter spaces and multiple local minima quite well; see also the discussion in Chernozhukov and Hong (2003b). We describe our procedure in detail in Appendix C.

5 Results

5.1 Model Fit

Our model is overidentified and consequently considering the fit of targeted moments can be informative about the performance of the model. In Figure 4 we plot the actual and fitted

[27] Neglecting for simplicity selection issues, one can notice that – among stayers – $\rho_{\Delta\tilde{e}} \times var(\Delta\tilde{e}_{kt}) = (\kappa^P)^2\sigma^2_{\xi^P} + 2(\kappa^T)^2\sigma^2_{\xi^T}$, and that the autocovariance of average wage growth is: $C_{\Delta\tilde{e},\Delta\tilde{e}_{-1}} = -(\kappa^T)^2\sigma^2_{\xi^T}$. Since firm data are used to identify the variance of firm shocks, $\sigma^2_{\xi^P}$ and $\sigma^2_{\xi^T}$, these two moments form a system of two equations in two unknowns that can be used to identify the pass-through coefficients κ^P and κ^T. Identification is predicated on having eliminated sources of spatial correlation other than being exposed to common firm shocks (which is accomplished by appropriate controls in the wage equation), and on the absence of other reasons for observing a spatial correlation (i.e., peer effects in labor supply, etc.).

Table 6: **Model Fit for Selected Wage Dynamics Moments**

	At least some College		High School or Less	
	Data	Model	Data	Model
Residual wage growth moments for job stayers				
$\text{Var}(\Delta\tilde{e}_t \mid E_{t-1} = E_t = 1, J_t = 0)$	0.0350	0.0256	0.0250	0.0218
$\text{Cov}(\Delta\tilde{e}_t, \Delta\tilde{e}_{t-1} \mid J_t = 0)$	–0.0047	0.0021	–0.0035	0.0009
$Skewness(\Delta\tilde{e}_t \mid E_{t-1} = E_t = 1, J_t = 0)$	0.0154	0.0187	0.1925	0.1910
$Kurtosis(\Delta\tilde{e}_t \mid E_{t-1} = E_t = 1, J_t = 0)$	6.0813	6.0809	6.5075	6.5299
Residual wage growth moments for job movers				
$\mathbb{E}(\tilde{e}_{t+1} - \tilde{e}_{t-1} \mid E_{t-1} = E_{t+1} = 1, J_t = 1)$	0.0386	0.0427	0.0266	0.0284
$\text{Var}(\tilde{e}_{t+1} - \tilde{e}_{t-1} \mid E_{t-1} = E_{t+1} = 1, J_t = 1)$	0.0675	0.0681	0.0537	0.0558
Covariance between wage growth and employment residuals				
$\text{Cov}(\tilde{u}_t, \tilde{e}_t \mid E_t = E_{t-1} = 1, J_t = 0)$	0.0003	–0.0003	–0.0002	–0.0002
$\text{Cov}(\tilde{u}_t, \tilde{e}_t \mid E_t = E_{t-1} = 1, J_t = 1)$	0.0189	0.0209	0.0031	0.0034
Common shocks at the firm level				
Spatial correlation coefficient (for stayers)	0.1816	0.1827	0.1783	0.1757
$\text{Cov}(\mathbb{E}_j[\Delta\tilde{e}_t], \mathbb{E}_j[\Delta\tilde{e}_{t-1}] \mid J_t = 0)$	–0.0015	–0.0003	–0.0010	-0.0012

Var: Variance, Cov: Covariance, $\mathbb{E}$: average, $\mathbb{E}_j$: 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 are computed excluding the top and bottom 1% of wage growth observations.

cross sectional variance of (residual) wages. The variances are replicated extremely well by the model, showing a growing dispersion of wages over the life cycle for the high-skill group and little to no age effects for the low-skill group.

In Table 6 we next show the dynamics of wage growth. As for wage levels, the overall fit is excellent. Some autocovariances show sign reversals, but they are all very close to zero and this is inconsequential.[28] 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 date of the transition and the associated measurement error in earnings changes between jobs.[29] The relevant statistics (the conditional mean $\mathbb{E}(\tilde{e}_{t+1} - \tilde{e}_{t-1} \mid E_{t-1} = E_{t+1} = 1, J_t = 1)$, and the conditional variance $\text{Var}(\tilde{e}_{t+1} - \tilde{e}_{t-1} \mid E_{t-1} = E_{t+1} = 1, J_t = 1)$) are reproduced very accurately by the model. We also consider the covariance between the employment residual (obtained from estimating a linear probability model) and the

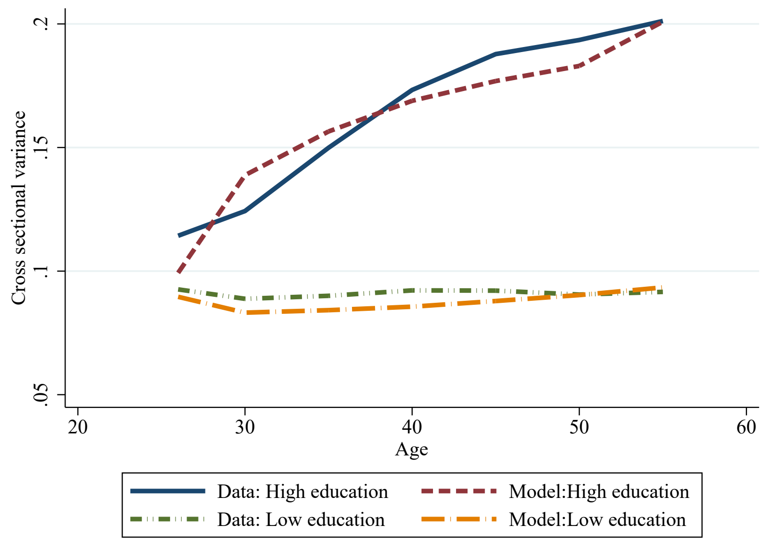
[28]All units are in logs.

[29]No bias arises from this process because the model and the data moments are constructed using the same rules.

wage residual, separately for stayers and movers. They help capture the selection effect of employment decisions on wages.

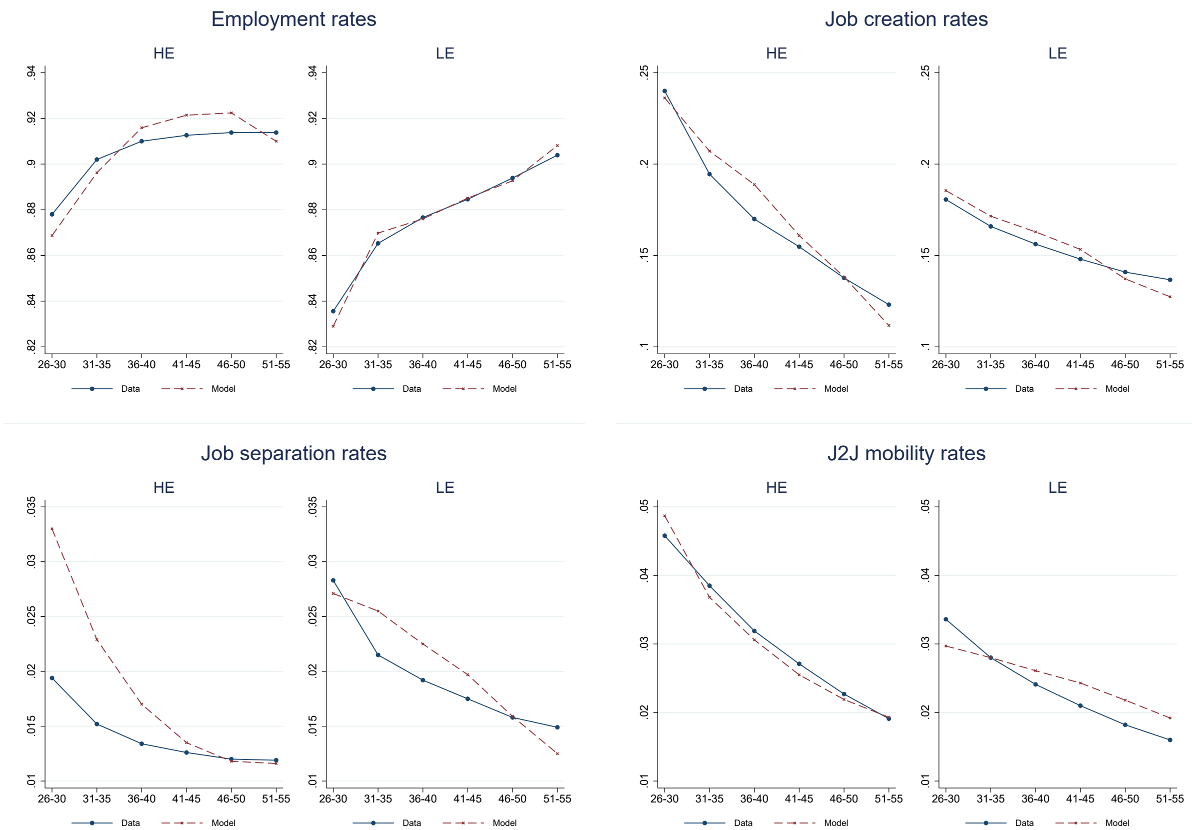
Among the moments we consider are the skewness and kurtosis of wage growth for stayers. These two moments capture the possibility of non-normality, one interpretation of which is that most wage adjustments are small but occasionally we see big changes, say because of a promotion or an important adverse 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.

Figure 4: **Model Fit for Cross sectional variance by education over the life cycle**



In the last panel of Table 6 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 individuals within a firm, 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 ($Cov(\mathbb{E}_j[\Delta\tilde{e}_t], \mathbb{E}_j[\Delta\tilde{e}_{t-1}]|J_t = 0)$). In sum, the model captures rather well the way wages of workers in the same workplace move together from period to period.

Figure 5: **Model Fit for Employment and Mobility Moments**



Note: HE and LE denote high and low education, respectively.

In Figure 5 we report the fit of the model for employment rates and various labor market transitions (Table D1 in the Appendix reports the actual figures). The model fits the employment rate exactly for the lower education group and slightly over predicts it for people with more education in their 40s, although not in an economically consequential way. Importantly, the model replicates the increasing participation over the life cycle. It also fits accurately the age profile of entry and job-to-job mobility (including the heterogeneity by education), which are all declining over the life cycle. It does, however, overpredict job separation for the youngest higher education group.

Since we allow job offer arrival rates to differ by firm type of the current employer, we further target transition frequencies and average wage growth for movers across firm types. The model captures extremely well these transition moments as we document in Figure 6 (see Appendix Table D3 for the specific numbers).

Finally, we show that the model replicates well features that we do not target explicitly: the share of workers experiencing pay cuts when moving across firms with different levels of productivity. This is shown in Figure 7.

5.2 Parameter estimates

Transitions We start by presenting results in Table 7 for the decisions to work and to move to another firm.[30] 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.[31]

To interpret the size of the coefficient, we report at the bottom of the table the marginal effect of a 10 percent wage increase on employment for workers aged 40. 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.

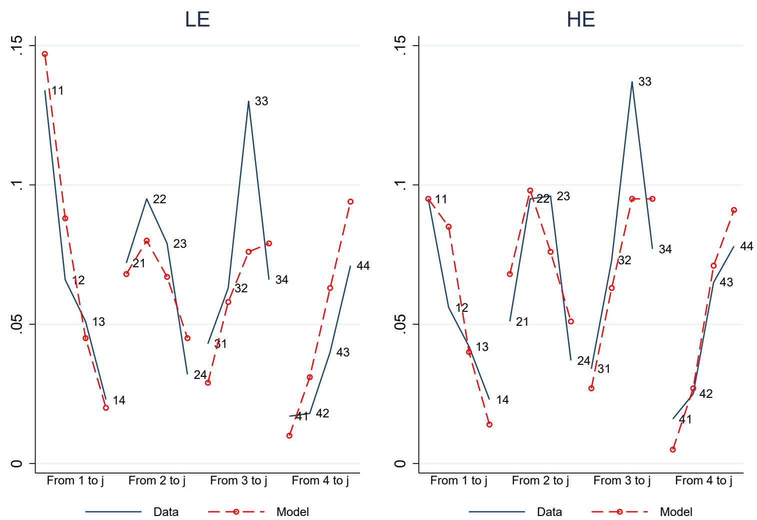
In the bottom part of Table 7 we look at the determinants of job-to-job mobility. We find that transitions across firms are decreasing in age, matching what we see in the data (e.g., the top-right panel of Figure 3). The coefficient b is estimated to be large and positive for both education levels, which shows that mobility choices are influenced by the wage difference between incumbent and poaching firm. This sensitivity limits the ability of the incumbent firm to lower wages as a result of shocks (conditioning on the flow of outside

[30] The results of estimating equation (3) to obtain estimates of the effects of individual characteristics on wages (γ) and the wage residuals ($\tilde{e}$) are presented in the Appendix.

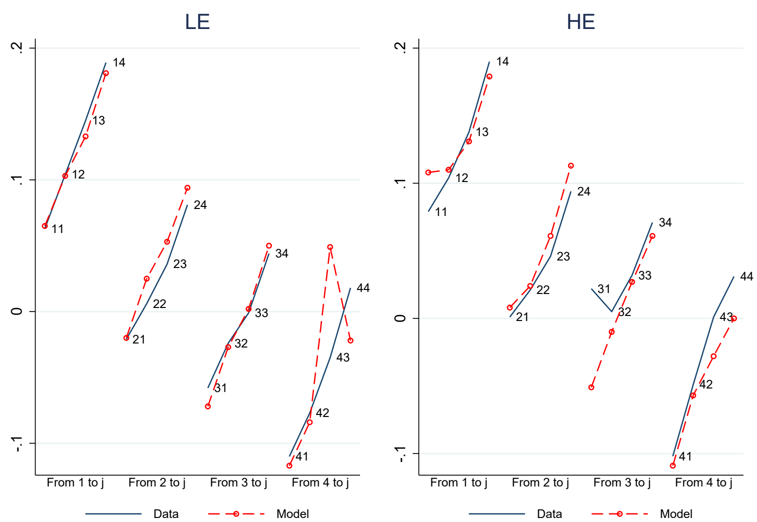
[31] As noted earlier, this is a mix of a selection and an incentive effect and in this context we cannot distinguish the two since we do not have appropriate exclusion restrictions. Nevertheless this is not a threat to the identification of the stochastic process of wages, which is the central focus of this study.

Figure 6: **Model Fit for Transition Moments by Firm Type**

Fractions moving across firm types

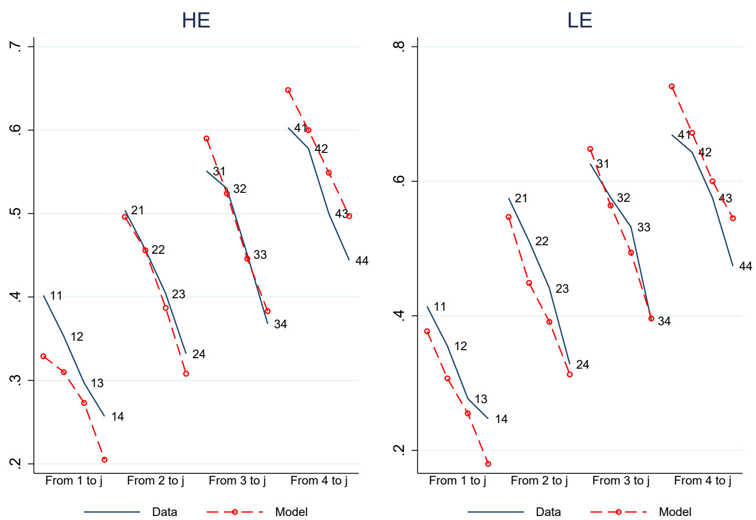


Avg. wage growth, movers across firm types



Note: HE and LE denote high and low education, respectively. ij denotes a transition from firm type i to firm type j (with 1 and 4 being the lowest and highest productivity type, respectively).

Figure 7: **Model Validation: Shares with Pay Cuts when Moving Across Firm Types**



Note: HE and LE denote high and low education, respectively. Each point in the figure, ij, denotes a transition from firm type i to firm type j (with 1 and 4 being the lowest and highest productivity type, respectively).

Table 7: **Results: Participation and job mobility**

		At least some College		High School or Less	
Parameter		Estimate	s.e.	Estimate	s.e.
		Employment			
δ_0	Constant	0.202	(0.004)	2.072	(0.001)
δ_{age}	Age	0.811	(0.002)	–0.150	(0.0004)
δ_{age^2}	Age squared	–0.080	(0.0002)	0.034	(0.0001)
ϕ	Wage residual	0.645	(0.003)	0.156	(0.0012)
Marginal effect of 10% wage increase (%)		0.246		0.082	
		Job-to-Job Mobility			
θ_0	Constant	–0.914	(0.013)	–1.998	(0.006)
θ_{age}	Age	–0.270	(0.006)	0.087	(0.0023)
θ_{age^2}	Age squared	0.018	(0.0007)	–0.018	(0.0003)
b	Wage improvement	3.351	(0.054)	1.938	(0.041)
Marginal effect of 10% wage increase (%)		3.12		1.18	

Table 8: **Results: Arrival Rate of Offers**

		At least some College	High School or Less
Parameter	Description	Estimate (s.e.)	Estimate (s.e.)
	Job arrival rate		
λ_{entry}	Arr. rate at entry	0.847	0.617
		(0.0010)	(0.0004)
$\lambda_{U,0}$	Arr. rate, subs. spells	0.378	0.259
		(0.0003)	(0.0001)
$\lambda_{U,1}$	Arr. rate, subs. spells (age shift)	0.0050	0.0025
		(0.0000)	(0.0000)
	Origin of offer		
ω_1	Offer rate of type 1 firms	0.031	0.823
		(0.0002)	(0.0001)
ω_2	Offer rate of type 2 firms	0.912	0.900
		(0.0002)	(0.0001)
ω_3	Offer rate of type 3 firms	0.952	0.975
		(0.0002)	(0.0001)
ω_{dist}	Relative frequency of offers	2.123	1.339
	by type distance	(0.0040)	(0.0015)

offers received). 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 8 presents the results on job offer arrival rates. 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 2.8 quarters for the high-skilled and every 4 quarters for lower skill workers. These rates decrease in frequency as workers age, but very moderately.

In the bottom half of Table 8, the coefficients ω_k ($k = \{1, 2, 3\}$) show the probability of on-the-job offers from different firm types (defined on the basis of average wage premium, or productivity). We normalize the arrival rate for the highest-ranked firms (type 4) to one because we cannot separately identify overall arrival frequency and mobility preferences.

The large ω_{dist} implies that most job offers originate from similar firms (74% of offers for low-skilled, 88% for high-skilled workers). In addition, the results reveal monotonically lower offer arrival rates on the job for workers at lower ranked firms.

Stochastic process of individual productivity The stochastic process of wages contains the contribution of firm-specific components as well as components unrelated to firms and which the worker carries from job to job. The latter are shown in Table 9. There are clear similarities across education groups, but also some important differences as we would expect when considering the life-cycle patterns of log earnings dispersion in Figure 2.

Table 9: **Results: The stochastic process of individual productivity**

		At least some College		High School or Less	
Parameter		Estimate	s.e.	Estimate	s.e.
σ_f	Initial perm. productivity, wages	0.299	(0.0002)	0.304	(0.0001)
ρ	AR(1) coefficient	0.973	(0.0000)	0.952	(0.0000)
σ_ϵ	Transitory shock, wages	0.055	(0.0003)	0.018	(0.0002)
Mixture of normals for persistent productivity shocks					
μ_{ζ_1}	mean of distribution 1	0.0005	(0.0000)	–0.0013	(0.0000)
σ_{ζ_1}	standard dev. of distribution 1	0.0172	(0.0001)	0.0063	(0.0000)
μ_{ζ_2}	mean of distribution 2	–0.0052	–	0.0145	–
σ_{ζ_2}	standard dev. of distribution 2	0.2723	(0.0001)	0.2762	(0.0001)
λ_m	Probability of distribution 1	0.907	(0.0002)	0.918	(0.0000)

Wages at labor market entry show a remarkable amount of dispersion (as measured by σ_f). 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 33% of a shock to high education workers remains; for the low education group 14% of the shock remains after that amount of time.

A feature of the wage data is heavy tails; one interpretation of this is that workers occasionally obtain large wage increases, possibly reflecting promotions, or large negative productivity shocks reflecting say an accident, while otherwise there are small fluctuations reflecting small adjustments to pay. To capture this we allow the distribution of individual productivity shocks to be a mixture of Normals, which allows for a 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 9 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 occasionally the individual draws a shock from a distribution with a very high standard deviation. Thus for the higher education group there is a 9.3% probability ($(1-\lambda_m)$) that the idiosyncratic productivity shock would be drawn from a distribution with a standard deviation of 0.27; while in the vast majority of circumstances the draw is from a distribution with a much smaller standard deviation (0.0172). The findings are qualitatively similar for the lower education group, except that the standard deviation of shocks in the more frequent regime is lower (0.0063). 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 10: **Results: Shocks and their transmission**

		At least some College		High School or Less	
Parameter	Description	Estimate	s.e.	Estimate	s.e.
τ_4	Wage premium, type-4 firms	0.0970	(0.0002)	0.0948	(0.0001)
τ_3	Wage premium, type-3 firms	–0.0181	(0.0002)	0.0136	(0.0002)
τ_2	Wage premium, type-2 firms	–0.0354	(0.0001)	–0.0149	(0.0001)
τ_1	Wage premium, type-1 firms	–0.0435	(-)	–0.0935	(-)
$\sigma_{\psi^{init}}$	Permanent initial shock, match value	0.0628	(0.0001)	0.0239	(0.0001)
σ_{ψ^T}	Transitory idiosyncratic shock, match value	0.0103	(0.0000)	0.0294	(0.0001)
σ_{ψ^P}	Permanent idiosyncratic shock, match value	0.0098	(0.0000)	0.0181	(0.0000)
κ^T	Transitory firm shock, match value	0.1320	(0.0003)	0.1633	(0.0001)
κ^P	Permanent firm shock, match value	0.2606	(0.0002)	0.1632	(0.0002)

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 10 we show the key parameters for our study, related to the impact of firms onto wages.

The initial match is quite disperse across firm types, with the average premium of joining the highest productivity firm (as opposed to the lowest, the difference between τ_4 and τ_1) ranging from 14 percentage points for the high-skilled to 19 percentage points for the low-skilled, albeit with a lower variance.[32] These differences persist over time (absent shocks to the match) also because, as we shall see, workers are more likely to receive outside offers from similar firm types.

The next set of coefficients in Table 10 (σ_{ψ^P}, σ_{ψ^T}) 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 discuss below). 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.[33]

The results indicate a substantial role for initial heterogeneity in idiosyncratic match effects for higher skill workers: the variance of the initial match value is about one-fifth as large as initial heterogeneity in permanent productivity at labor market entry. This match component also implies that wage-induced moves from one job to another are not only driven by an accumulation of bad shocks in the job of origin, but also by a location of a much improved opportunity. The transitory and permanent shocks to this initial match value (σ_{ψ^T} and σ_{ψ^P}) are much smaller than idiosyncratic shocks to individual productivity in Table 9.

When we turn to lower skill workers there seems to be a smaller role for the initial variance of offers, but a larger role for idiosyncratic changes in the match value over time. The important point that emerges from these results is that a large fraction of permanent "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.

The final set of parameters in Table 10 are the key ones, and relate to the transmission of firm-related shocks onto wages. For workers with higher education 13% of a transitory shock

[32] As a normalization, we set average firm premia equal to zero.

[33] Although we cannot rule out the possibility that these idiosyncratic match effects reflect heterogeneity in the pass-through of firm-related shocks.

is transmitted to workers. This 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 a 26% pass-through coefficient. 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 economically significant 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).[34] 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 slightly more in response to transitory shocks in the firm's value added (16% transmission coefficient), but at the same time we find a substantially lower transmission of permanent shocks (16% transmission) than for high-skilled workers. 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 and is consistent with a lower share of variable compensation compared to high-skilled workers, see also Juhn et al. (2018). From an econometric point of view this result may be traced back to the fact that overall permanent shocks are less

[34]In principle, we may even underestimate the amount of rent sharing for workers with executive-level positions, since they are more likely to have part of their compensation in the form of stock options or firm ownership shares, which induce an even tighter link with the firm's fortunes.

important for low-skill workers, as implied by the descriptive analysis of their life-cycle variance, which does not increase, in contrast to that of the higher skill workers.

Card et al. (2018) offer a comprehensive survey of earlier estimates of rent sharing elasticities. Our elasticities cannot be readily compared to those in the literature they cite both because we distinguish between permanent and transitory shocks (a distinction that is absent in most of the studies they cite, with the exception of Guiso, Pistaferri, and Schivardi (2005) and the studies that replicate their approach) and because we report estimates separately for high- and low-skill workers (while most studies ignore this form of heterogeneity). Nevertheless our elasticity to permanent shocks for workers with some college is at the high end of their list and indeed quite close to that reported by Van Reenen (1996), who finds an elasticity of wages with respect to firm profits of 0.29. Finally, the only study we are aware of that uses data from Sweden is Carlsson, Messina, and Skans (2016). The authors use a different methodology than the one used here and focus on manufacturing firms. They find a pass-through elasticity of approximately 0.1 when considering a measure of productivity close to ours. Thus, while our estimates are comparable, they tend to be higher, especially for workers with some college education. Besides differences in samples and methodologies, a possible explanation for these differences is that controlling for endogenous mobility increases the estimates, since large downward wage adjustments are likely to induce worker departures, either to unemployment or to other firms. We examine this explanation by applying the Guiso, Pistaferri, and Schivardi (2005) approach to our data in section 5.3.

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 does not appear particularly important: match specificity mainly 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 education.

5.3 Simulations

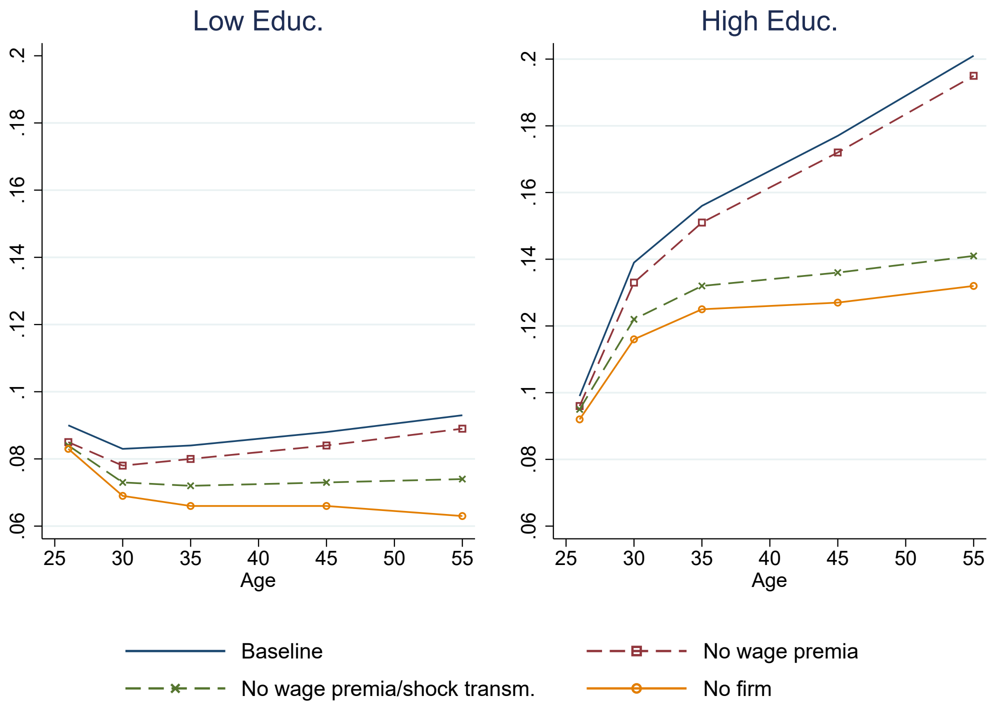
The identity of the firm in which one works appears to have a substantial impact on the evolution of wages over the lifecycle, pointing to non-competitive behavior. Given that we are looking at innovations to wages and productivity, our conclusion is that a substantial amount of uncertainty faced by individuals has its origins in the fluctuating fortunes of their firm. This is beyond the issues of sorting that other authors have identified and that relate 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 life-cycle profiles. For simplicity, we report statistics for five selected points in the life cycle: age 26, 30, 35, 45 and 55.

In Figure 8 we plot the baseline variance of log earnings over the life cycle, as well as counterfactual profiles obtained ruling out (sequentially) the key aspects of the firm's impact on careers: (a) initial wage premia, (b) transmission of firm productivity shocks, and (c) firm-worker pair specific shocks (together with "homophily" of outside wage offers). The right panel is for the low-skilled; the left panel for the high-skilled (with identical scale). Table E1 in the Appendix shows the actual numbers behind these pictures.

The "baseline" profile in Figure 8 is the life-cycle profile of variances of log wages for the full model, featuring endogenous participation and mobility choices. As we expect from the data (Figure 2), the predicted cross sectional variance of earnings increases over time for the higher skilled while the growth for the low-skilled is minimal (despite starting values being similar). We target these life-cycle patterns in the estimation, and the levels closely match the data as shown above.

The first counterfactual experiment is to eliminate differences in initial pay across firm types (the lines labeled "No wage premia"). This reduces the cross-sectional variance at young ages by about 5% and this share slightly declines with age for both education groups. It is important to note, however, that this channel only captures how firms' initial pay offers

Figure 8: **The variance of log earnings over the life cycle**



affect earnings dispersion. Any subsequent evolution in pay and career opportunities that depends on firm characteristics will further contribute to the overall role of the firm.

Hence, we next investigate the impact of firm shock transmission to workers' life-cycle earnings variance. In the next experiment (lines labeled "No wage premia/shock transm.") we switch off both initial differences in wage premia and the contribution of firm-level shocks (i.e., set the pass-through parameters $\kappa^P = \kappa^T = 0$). We now find large effects as well as substantial differences across groups. While firm shock transmission explains a small share of log earnings variance at young ages, we find that by age 55, the cross sectional variance for the high-skilled is only 0.147, compared to the full variance of 0.201. In other words, permanent firm-level shocks, which are transmitted to wages, increase the cross-sectional dispersion of wages for 55-year-old workers with at least some college education by 25% relative to a world where there is no shock pass-through onto wages. This effect is important because, as documented in Table 10, 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). Models that ignore the firm and this shock transmission structure will estimate a variance of permanent shocks that aggregates firm- and person-specific components and hence, to the extent that transmission of firm shocks can be avoided by moves out of work or into new firms, exaggerate the extent of economic risk faced by workers. For the lower skill workers, switching off transmission has a smaller effect, but again in the same direction: wages would be 15% less dispersed at age 55 in the absence of transmission of firm-level shocks. Taken together, initial wage premia and firm shock transmission account for 30% (21%) of earnings dispersion at age 55 among the high (low) skilled, with the lion's share of the decline accounted for by ruling out transmission of firm productivity shocks to wages. Yet these channels still ignore the impact of firms on career opportunities of workers through job mobility and that of shocks to the idiosyncratic match effects. The final counterfactual experiment rules out these channels as well (on top of eliminating initial heterogeneity in wage premia and transmission of firm shocks). This is achieved by (a) assuming equal offer frequency for workers at all firm types, and (b) imposing that $\sigma_{\psi^T} = \sigma_{\psi^P} = 0$ in the wage model. Since these counterfactuals effectively eliminate any role for firms to influence the variance of wages and its growth over the life cycle, we label the line in Figure 8 that represent them the "No firm" case.[35]

In contrast to the substantial role of firm-related shocks, for highly educated workers idiosyncratic match effects account for a stable and relatively small share of earnings variation. For low-skill workers, the dispersion in idiosyncratic match components grows over the life-cycle and accounts for a similar share of variation as firm shock transmission.

In sum, firm-level and idiosyncratic match components account for about one-third of earnings dispersion for each education group. Yet the underlying sources of this dispersion differ across groups, with a larger role for the firm among high-skill workers and a larger role

[35] For completeness, in Table E1 in the Appendix we look at the impact of the four channels (wage premia, shock transmission, "homophily" of offers, and idiosyncratic match shocks) separately instead of sequentially. These exercises confirm that the quantitatively most relevant channel is the ruling out of firm shock transmission.

for idiosyncratic match values among low-skill workers.

The role of selection Our results point to a larger role of firm-shock transmission to workers' wages than most previous studies. One key distinctive feature of our analysis is that we take endogenous participation and mobility choices into account, rather than focusing only on stayers. The latter choose to stay with the firm, and this selection may lead to underestimating the role of the firm, in the sense that extreme firm shocks may trigger moves (perhaps not immediately due to frictions), and hence stayers are a sample of workers with presumably lower impact of firm shocks.

To illustrate this point further, in Table 11 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. Our model implies that the standard deviation of wage growth for stayers attributable to permanent firm shocks is 0.0464. The GPS model implies a much smaller contribution of 0.0144. This lowers the contribution of permanent firm shocks to cross-sectional wage growth variation among stayers in the data from 1/4 to 1/10. Thus selection plays an important role in censoring the impact of the firm. For lower skill workers, permanent shock transmission is similarly underestimated by GPS, even though this is less of an issue overall because the role of the firm is smaller.

In sum, focusing on stayers gives the impression of a much lower transmission rate of firm shocks to wages, which in turn implies a bias towards concluding that labor markets are more competitive than they actually are. 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 life cycle variances than transitory firm shocks. This has important considerations for an evaluation of life cycle 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 how exogenous permanent firm shocks passing through

Table 11: **Simulations: Comparison with a Stayer-Only** Model

	At least some College		High School	
	Full Model	Stayers	Full Model	Stayers
$sd(\Delta w\ firm\ perm)$	0.0464	0.0144	0.0291	0.0149
Share perm firm shocks	0.250	0.111	0.184	0.094

"Share perm firm shocks" is the ratio of $sd(\Delta w\ firm\ perm)$ to the cross-sectional standard deviation of wage growth among stayers in the data. This dispersion is 0.1855 for workers with at least some college, 0.1581 for workers with high school education or less.

wages impact household savings and portfolio choices, respectively.

Participation and mobility as insurance mechanisms Building on the previous insights, we now shed further light on the importance of participation and mobility choices to avoid the transmission of firm-level shocks. We conduct additional simulations in which we (i) exclude the option of non-participation and instead assume full employment, (ii) rule out the arrival of new job offers while employed (OTJ offers), and (iii) combine restrictions (i) and (ii).

Intuitively, while scenario (i) eliminates the option of leaving the current job into non-participation in response to large negative shocks, scenario (ii) rules out job-to-job mobility (implying that workers can join new firms only after an unemployment spell). The effect of no on-the-job offers on earnings dispersion is a priori ambiguous because it prevents both shock mitigation and accumulating search capital through moving to opportunity. Finally, the combination of these constraints on participation and mobility implies that workers are fully tied to the fortunes of their first employer.

For each of these scenarios, we simulate the full model and a counterfactual without firm-level shock transmission. The resulting difference between the profiles of earnings variance over the lifecycle will shed light on the role that firm shocks play in overall earnings dispersion when excluding specific insurance mechanisms.

Table 12 presents the results. We find that the impact of firm-level shock transmission on life cycle earnings dispersion increases substantially in the absence of the option to leave a sinking ship. Losing the options of either moving to a different firm or choosing non-

Table 12: **Simulations: Participation and Mobility as Insurance**

Panel A: Earnings variance among workers with high school or less							
Counterfactual	Baseline	Full participation		No OTJ Offers		Full part., No OTJ Offers	
	w/ firm	w/ firm	w/o firm	w/ firm	w/o firm	w/ firm	w/o firm
Age	shocks	shocks	shocks	shocks	shocks	shocks	shocks
26	0.090	0.089	0.088	0.090	0.089	0.089	0.088
30	0.083	0.087	0.080	0.086	0.079	0.091	0.081
35	0.084	0.088	0.078	0.090	0.079	0.102	0.083
45	0.088	0.099	0.082	0.099	0.082	0.132	0.096
55	0.093	0.109	0.085	0.110	0.085	0.163	0.108

Panel B: Earnings variance among workers with at least some college							
Counterfactual	Baseline	Full participation		No OTJ Offers		Full part., No OTJ Offers	
	w/ firm	w/ firm	w/o firm	w/ firm	w/o firm	w/ firm	w/o firm
Age	shocks	shocks	shocks	shocks	shocks	shocks	shocks
26	0.099	0.100	0.099	0.100	0.098	0.101	0.099
30	0.139	0.141	0.127	0.138	0.124	0.146	0.126
35	0.156	0.163	0.135	0.159	0.133	0.178	0.136
45	0.177	0.191	0.140	0.183	0.137	0.234	0.142
55	0.201	0.222	0.149	0.212	0.144	0.307	0.155

Panel A shows simulation results for life-cycle earnings variance among low-skill workers, Panel B for high-skill workers. Each panel compares four mobility and participation scenarios: (i) "Baseline" allows for both endogenous choices; "Full Participation" rules out non-participation; "No OTJ Offers" rules out job-to-job mobility by excluding on-the-job offers; "Full Part, No OTJ Offers" excludes mobility and non-participation. For each scenario that limits individual choices, we consider two alternatives, with and without transmission of firm-level shocks, respectively.

participation has similar impact on the results: Compared to the main model simulation without firm shocks but allowing for both endogenous worker choices, both scenarios increase the role of the firm in cross-sectional earnings variation substantially, from 15% to 22% for low-skill workers and from 27% to 33% for high-skill workers.

Ruling out both insurance mechanisms in the third counterfactual yields a one-third smaller earnings variance for low-skill workers at age 55 in the absence of firm shocks. For high-skill workers, the difference is even more dramatic, with firm shocks accounting for half of the wage variance at age 55.

In sum, these results emphasize the key role that participation and mobility choices play in responding to firm-level shocks. Accounting for these endogenous responses is crucial not only in estimating the transmission of firm-level shocks to wages but also in assessing the extent of risk that different workers are exposed to through firm-level shocks.

6 Conclusion

The extent to which the firm in which individuals work matter for explaining the level and fluctuations of their wages is an important question, both from the perspective of understanding the degree of labor markets competitiveness as well as to better characterize 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 firm-specific permanent productivity shocks transmit to individual wages for high-skill workers: the elasticity of wages with respect to permanent firm productivity shocks is 0.26. In other words firms pass roughly one-fourth of their permanent change in their fortunes to wages. Transitory (i.i.d.) shocks have an impact on the wages of high-skill workers that is half as large. For low-skill workers the elasticity is 0.16 and there is no meaningful difference between less and more persistent firm shocks; moreover, this transmission of shocks does not have a large impact on wage profiles. We find that for the high-skilled the variance of wages doubles between age 25 and age 55. Eliminating transmission of firm shocks alone would generate a much more modest 48% growth. Hence, without transmission of firms shocks (especially permanent ones) the growth in wage variances over the life cycle would be approximately half of what it actually is (at least for highly educated workers). For these workers, match-specific effects, other than those that are common to all workers in the same firm, do not play a substantial role. For lower skill workers the effects are qualitatively similar, but the role of the firm is much less relevant because the dispersion in wages does not increase much over the life cycle (12%).

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 firm-worker pair 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 a component (reflecting rent sharing or partial insurance) 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 appear insulated from them. This finding 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.[36]

[36]Saez, Schoefer, and Seim (2019) argue that in Sweden the union minimum wage floors mostly bind for new, young employees.

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.

Akerman, A., Helpman, E., Itskhoki, O., Muendler, M.-A., and Redding, S. (2013). 'Sources of Wage Inequality.' *American Economic Review*, vol. 103(3):pages 214–19.

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.

Balke, N. and Lamadon, T. (2021). 'Productivity Shocks, Long-Term Contracts and Earnings Dynamics.' *American Economic Review*. (forthcoming).

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

Blundell, R., Pistaferri, L., and Saporta-Eksten, I. (2016). 'Consumption Inequality and Family Labor Supply.' *American Economic Review*, vol. 106(2):pages 387–435.

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 Jolivet, G. (2009). 'The pervasive absence of compensating differentials.' *Journal of Applied Econometrics*, vol. 24(5):pages 763–795.

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.

Card, D. E., Cardoso, A. R., Heining, J., and Kline, P. M. (2018). 'Firms and Labor Market Inequality: Evidence and Some Theory.' *Journal of Labor Economics*, vol. 36(S1):pages S13–S70.

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

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

Davis, S. J. and Haltiwanger, J. (2014). 'Labor Market Fluidity and Economic Performance.' Working Paper 20479, National Bureau of Economic Research.

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.

Friedrich, B., Laun, L., and Meghir, C. (2021). 'Income Dynamics in Sweden 1985-2016.' Working Paper 28527, National Bureau of Economic Research.

Garin, A. and Silvério, F. (2019). 'How Responsive are Wages to Demand within the Firm? Evidence from Idiosyncratic Export Demand Shocks.' Working Papers w201902, Banco de Portugal, Economics and Research Department.

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.

Gosling, A., Machin, S., and Meghir, C. (2000). 'The changing distribution of male wages in the UK.' *Review of Economic Studies*, vol. 67:pages 635–666.

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.

Hausman, J. (2001). 'Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left.' *Journal of Economic Perspectives*, vol. 15(4):pages 57–67.

Juhn, C., McCue, K., Monti, H., and Pierce, B. (2018). 'Firm Performance and the Volatility of Worker Earnings.' *Journal of Labor Economics*, vol. 36(S1):pages S99–S131.

Kline, P., Petkova, N., Williams, H., and Zidar, O. (2019). 'Who Profits from Patents? Rent-Sharing at Innovative Firms.' *The Quarterly Journal of Economics*, vol. 134(3):pages 1343–1404.

Lamadon, T., Mogstad, M., and Setzler, B. (2019). 'Imperfect Competition, Compensating Differentials and Rent Sharing in the U.S. Labor Market.' Working Paper 25954, National Bureau of Economic Research.

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.

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. (2002). 'Equilibrium W age Dispersion with Worker and Employer Heterogeneity.' *Econometrica*, vol. 70(6):pages 2295–2350.

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. (2019). 'Payroll Taxes, Firm Behavior, and Rent Sharing: Evidence from a Young Workers' Tax Cut in Sweden.' *American Economic Review*, vol. 109(5):pages 1717–63.

Song, J., Price, D. J., Guvenen, F., Bloom, N., and von Wachter, T. (2018). 'Firming Up Inequality*.' *The Quarterly Journal of Economics*, vol. 134(1):pages 1–50. ISSN 0033-5533.

Sorkin, I. (2018). 'Ranking Firms Using Revealed Preference*.' *The Quarterly Journal of Economics*, vol. 133(3):pages 1331–1393.

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.

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.

Not For Publication Appendix

A Wage Residuals

If the error term follows a log normal distribution, the log of the conditional expectation of observed average quarterly wages is given by

$$\log \mathbb{E}\left[w_t \mid x_t, z_t,\, E_{t_q} = 1, \forall q = 1, ..., 4\right] = x_t'\gamma + \log\left[\sum_{q=1}^{4} E_{t_q} \times e^{\eta\lambda^M(z_{t_q}'\delta)} / \sum_{q=1}^{4} E_{t_q}\right] + \frac{\sigma_v^2}{2}, \quad (13)$$

where we omit both the individual i and 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. This aggregation bias term is reminiscent of the bias due to individual heterogeneity in Blundell, Reed, and Stoker (2003) when analyzing aggregate wages. 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 variables z_{t_q} affecting the decision to work 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 (13) 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.

The results for the first-stage estimation are presented in Table A1. 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 A1. 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

Table A1: 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^2		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.

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 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 A1 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 A1. 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 corresponding to the regression results in Table A1 can be found in the second row of Figure A1. 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 A1 suggest, wage differences because of selection are in the range of 0-20%, where these effects are higher for highly educated workers.

B Firm Productivity Estimation

In Table B1, we compare the autocovariance structure of firm productivity across all firms to separate patterns by industry. This evidence supports the approximation of the stochastic structure of VA per worker by a random walk with an i.i.d. transitory component because the second and third-order autocovariances for productivity growth in the data are close to zero for all sectors. 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).

To derive the analytical expressions for annual productivity growth in the model, we proceed in two steps. First, we assume that firm productivity can be measured by value added per worker. Firm productivity (in levels) in year τ is the sum of productivity in quarters $q = \{1, 2, 3, 4\}$:

$$A_\tau = A_{\tau_1} + A_{\tau_2} + A_{\tau_3} + A_{\tau_4}$$

Figure A1: Wage Profiles and Selection

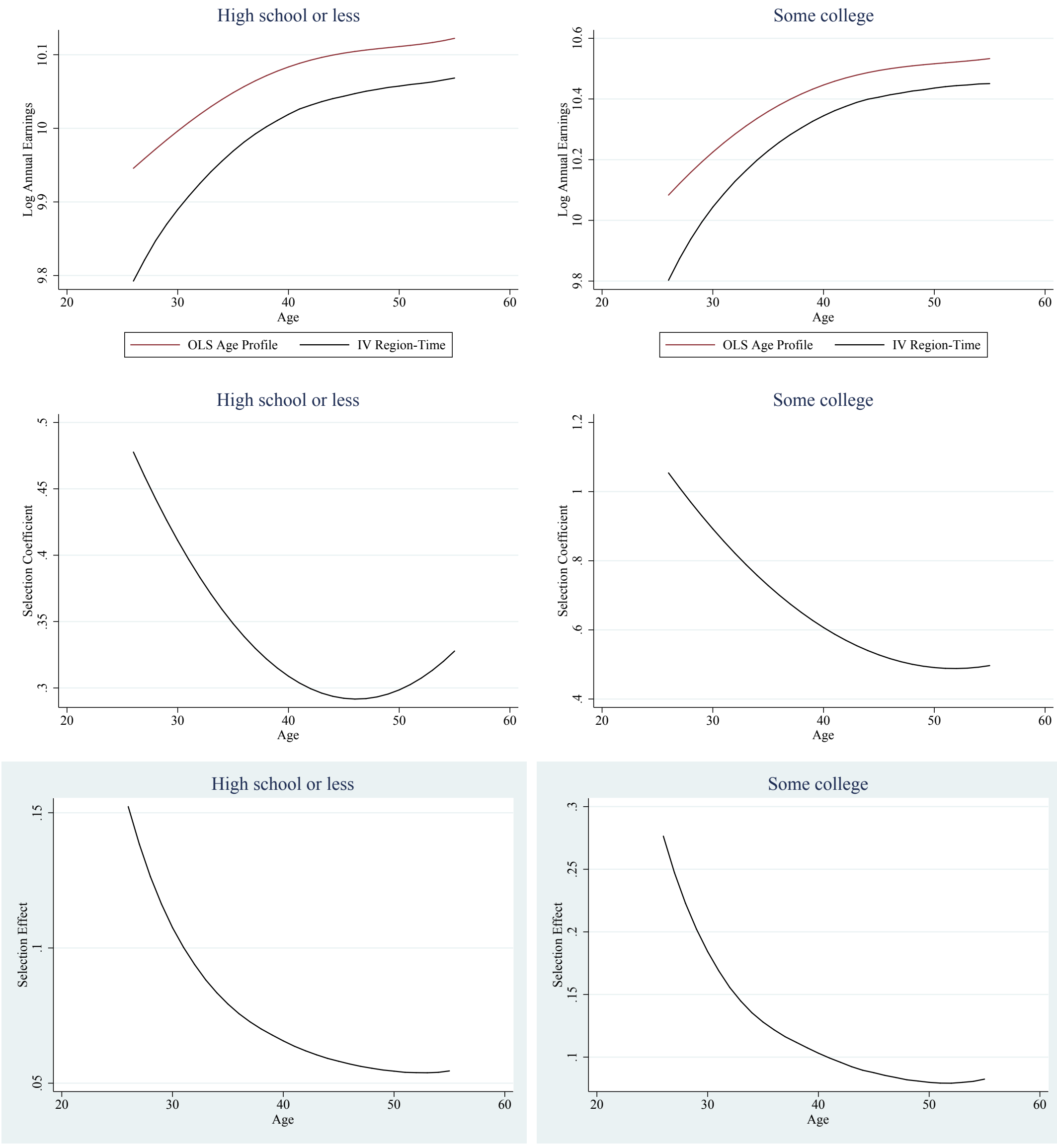


Table B1: Autocovariance of log Value Added per Worker: Data by Industry

	Value Added per Worker: Data				
	All firms	Construction	Manufacturing	Retail	Services
$\mathrm{Var}\left(\Delta a_{\tau}\right)$	0.1791	0.1603	0.1469	0.1698	0.2078
	(0.0008)	(0.0019)	(0.0017)	(0.0015)	(0.0015)
$\mathrm{Cov}\left(\Delta a_{\tau}, \Delta a_{\tau-1}\right)$	-0.0537	-0.0587	-0.043	-0.0487	-0.0602
	(0.0004)	(0.0010)	(0.0008)	(0.0008)	(0.0008)
$\mathrm{Cov}\left(\Delta a_{\tau}, \Delta a_{\tau-2}\right)$	-0.0041	-0.0005	-0.0045	-0.0049	-0.0048
	(0.0003)	(0.0007)	(0.0006)	(0.0005)	(0.0005)
$\mathrm{Cov}\left(\Delta a_{\tau}, \Delta a_{\tau-3}\right)$	-0.0022	-0.0036	-0.0023	-0.0011	-0.0024
	(0.0003)	(0.0007)	(0.0006)	(0.0006)	(0.0006)

Note: a_{τ} denotes log of annual productivity in year τ.

We consider a stochastic model for quarterly log productivity:

$$\begin{aligned} \log A_{\tau_q} &\equiv a_{\tau_q} = a^P_{\tau_q} + \xi^T_{\tau_q} \\ a^P_{\tau_q} &= a^P_{\tau_q} + \xi^P_{\tau_q} \end{aligned}$$

(with the understanding that $\tau_q = \tau - 1_4$ if $q = 1$). So log annual productivity is

$$\begin{aligned} \log A_{\tau} &\equiv a_{\tau} \\ &= \log\left(A_{\tau_1} + A_{\tau_2} + A_{\tau_3} + A_{\tau_4}\right) \\ &= \log\left(e^{a_{\tau_1}} + e^{a_{\tau_2}} + e^{a_{\tau_3}} + e^{a_{\tau_4}}\right) \end{aligned}$$

Using the stochastic process for log productivity, note that:

$$a_{\tau_q} = a^P_{\tau-1_4} + \sum_{s=1}^{q} \xi^P_{\tau_s} + \xi^P_{\tau_q}$$

for $q = \{1, 2, 3, 4\}$, and hence:

$$
\begin{aligned}
a_\tau &= \log\left(e^{a_{\tau 1}} + e^{a_{\tau 2}} + e^{a_{\tau 3}} + e^{a_{\tau 4}}\right) \\
&= a^P_{\tau-1_4} + \log\left(\sum_{q=1}^{4} \exp\left(\sum_{s=1}^{q} \xi^P_{\tau_s} + \xi^P_{\tau_q}\right)\right)
\end{aligned}
$$

Moreover, by the same token, in year $\tau + 1$:

$$
\begin{aligned}
a_{\tau+1} &= \log\left(e^{a_{\tau+1_1}} + e^{a_{\tau+1_2}} + e^{a_{\tau+1_3}} + e^{a_{\tau+1_4}}\right) \\
&= a^P_{\tau-1_4} + \log\left(\sum_{q=1}^{4} \exp\left(\sum_{s=1}^{4} \xi^P_{\tau_s} + \sum_{s=1}^{q} \xi^P_{\tau+1_s} + \xi^T_{\tau+1_q}\right)\right)
\end{aligned}
$$

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

$$
\Delta a_{\tau+1} = \log\left(\sum_{q=1}^{4} \exp\left(\sum_{s=1}^{4} \xi^P_{\tau_s} + \sum_{s=1}^{q} \xi^P_{\tau+1_s} + \xi^T_{\tau+1_q}\right)\right) - \log\left(\sum_{q=1}^{4} \exp\left(\sum_{s=1}^{q} \xi^P_{\tau_s} + \xi^P_{\tau_q}\right)\right)
$$

The important point is that the initial conditions drop out of the expression and this remain only a function of productivity shocks.

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. 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 B2: Standard Deviations of Quarterly Firm-Shocks

	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.

C MCMC Estimation

We maximize the GMM objective function

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

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

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

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.[37] Since the simulated moments may not be smooth, we use a Laplace-type estimator (LTE) following Chernozhukov and Hong (2003a) 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:

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

[37] 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 $\hat{\beta}_{MCMC}$.[38]

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, see the discussion in Chernozhukov and Hong (2003) for more details.

Standard Errors 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

[38] 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)).

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

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 below the details of the derivation of our covariance matrix, which draws from Hansen (1982).

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 $\hat{}$ 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 $\hat{}$ 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

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

Hence the variance of the method of moments estimator is

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

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

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

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

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

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

D Goodness of Fit

This section provides goodness-of-fit statistics for the model estimates from section 5. We simulate the model and report the patterns of data vs. model.

We start with Table D1 (employment and labor market transitions by age group). Next, we look at transitions and wage growth of movers across different firm types (Tables D2 and D3). For the latter exercise, we group firms into four bins based on their average residual wages paid. We consider wage growth for movers between one year before the move and one year after the move analogous to our measure in the data. Since the model analyzes residual wages, we provide the empirical patterns for job mobility and residual wage growth

Table D1: **Model Fit for Moments on Labor Market Transitions**

	Age Group	At least some college		High School or Less	
		Data	Model	Data	Model
Unemployment frequency	26–30	0.1220	0.1313	0.1644	0.1710
	31–35	0.0980	0.1038	0.1347	0.1302
	36–40	0.0900	0.0841	0.1234	0.1240
	41–45	0.0874	0.0786	0.1154	0.1149
	46–50	0.0862	0.0776	0.1061	0.1073
	51–55	0.0862	0.0900	0.0961	0.0919
Job creation frequency	26–30	0.2400	0.2362	0.1806	0.1855
	31–35	0.1945	0.2071	0.1659	0.1715
	36–40	0.1699	0.1888	0.1562	0.1629
	41–45	0.1548	0.1609	0.1480	0.1533
	46–50	0.1377	0.1383	0.1409	0.1372
	51–55	0.1231	0.1117	0.1367	0.1274
Job separation frequency	26–30	0.0194	0.0330	0.0283	0.0271
	31–35	0.0152	0.0229	0.0215	0.0255
	36–40	0.0134	0.0170	0.0192	0.0225
	41–45	0.0126	0.0135	0.0175	0.0197
	46–50	0.0120	0.0118	0.0158	0.0159
	51–55	0.0119	0.0116	0.0149	0.0125
Job mobility frequency	26–30	0.0458	0.0487	0.0336	0.0297
	31–35	0.0385	0.0368	0.0280	0.0280
	36–40	0.0319	0.0306	0.0241	0.0261
	41–45	0.0271	0.0255	0.0210	0.0243
	46–50	0.0227	0.0219	0.0182	0.0218
	51–55	0.0191	0.0193	0.0160	0.0192

Note: All transitions are quarterly.

for comparison below in Table D2 (which reproduces Table 3 in the main text).

Table D2: **Data: Job Mobility and Wage Growth**

		Low-skill workers *Departing firm quartile*											
		Share of transitions				Residual wage growth				Share wage cuts			
		1	2	3	4	1	2	3	4	1	2	3	4
Arriving firm quartile	1	.134	.072	.043	.017	.063	−.021	−.058	−.11	.414	.575	.626	.669
	2	.066	.095	.063	.018	.104	.006	−.024	−.077	.355	.511	.576	.643
	3	.051	.079	.13	.04	.145	.036	−.001	−.035	.277	.441	.532	.575
	4	.023	.032	.066	.071	.189	.081	.044	.018	.247	.328	.393	.474
		High-skill workers *Departing firm quartile*											
		Share of transitions				Residual wage growth				Share wage cuts			
		1	2	3	4	1	2	3	4	1	2	3	4
Arriving firm quartile	1	.095	.051	.034	.016	.079	.001	−.022	−.102	.402	.504	.551	.603
	2	.056	.095	.073	.025	.104	.021	−.005	−.049	.353	.457	.53	.578
	3	.042	.096	.137	.065	.138	.046	.032	.001	.297	.404	.451	.5
	4	.023	.037	.077	.078	.19	.094	.071	.031	.257	.332	.368	.444

Note: Firms are sorted on the basis of average residual wages paid.

Next, we report the corresponding results from the model simulation in Table D3. The results show the expected pattern of most mobility among similar firms, higher wage growth when moving to a higher ranked firm, and substantial shares of residual wage cuts after job mobility. Yet, the gradient of these patterns across firm types is steeper in the data than in the simulation. The last matrix in Tables D2 and D3 report the results of the validation exercise in Figure 7, since we do not target these moments explicitly.

Table D3: **Model Simulation: Job Mobility and Wage Growth**

		Low-skill workers *Departing firm quartile*											
		Share of transitions				Residual wage growth				Share wage cuts			
		1	2	3	4	1	2	3	4	1	2	3	4
Arriving firm quartile	1	.147	.068	.029	.010	.065	−.020	−.072	−.117	.377	.547	.648	.741
	2	.088	.080	.058	.031	.103	.025	−.027	−.084	.307	.449	.564	.672
	3	.045	.067	.076	.063	.133	.053	.002	−.049	.255	.391	.494	.600
	4	.020	.045	.079	.094	.181	.094	.050	−.022	.180	.313	.396	.545
		High-skill workers *Departing firm quartile*											
		Share of transitions				Residual wage growth				Share wage cuts			
		1	2	3	4	1	2	3	4	1	2	3	4
Arriving firm quartile	1	.095	.068	.027	.005	.108	.008	−.051	−.109	.329	.496	.590	.648
	2	.085	.098	0.063	0.027	.110	.024	−.010	−.057	.310	.456	.524	.600
	3	.040	.076	.095	.071	.131	.061	.027	−.028	.273	.387	.446	.549
	4	.014	.051	.095	.091	.179	.113	.061	.000	.205	.308	.383	.497

Note: Firms are sorted on the basis of average residual wages paid.

Finally, we conduct a simple variance decomposition exercise on the simulated economy.

The goal is to measure the share of residual wage variation that is accounted for by variation across firms, analogous to the previous literature. Table D4 reports the share of the empirical earnings variance that can be attributed to differences between firms for each education group in our sample. We obtain these values by a standard decomposition of the total wage variance into between- and within-firm contributions. 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. Applying an analogous decomposition to our simulated economy, we find that variation across firms explains 11% of residual wage variation for low-skill workers and 32% for high-skill workers, only slightly below the 40% share we find for this group in the data.

Table D4: **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.

E Counterfactual Experiments

In this section we report the results of conducting counterfactual experiments designed to demonstrate the role of firms in impacting wage risk. We measure the latter with the variance of log wages over the life cycle. Table E1 reports the variance in the full model at selected ages (26, 30, 35, 45, 55), column (1), and then that obtained ruling out, one at a time, the four channels through which firms impact careers: initial differences in wage premia (column 2), transmission of firm shocks onto wages (column 3), search capital (column 4),

Table E1: **Simulations: The Role of Firms over the Life-Cycle**

Panel A: High school or less

Age	Full model	No firm premium	No firm shocks	No offer diff.	No idios. match	No firm prem. or shocks	No firm
	(1)	(2)	(3)	(4)	(5)	(6)=(2)+(3)	(7)=(6)+(4)+(5)
26	0.090	0.085	0.088	0.090	0.089	0.084	0.083
30	0.083	0.078	0.078	0.083	0.080	0.073	0.069
35	0.084	0.080	0.076	0.084	0.079	0.072	0.066
45	0.088	0.084	0.077	0.088	0.080	0.073	0.066
55	0.093	0.089	0.079	0.093	0.082	0.074	0.063

Panel B: At least some college

Age	Full model	No firm premium	No firm shocks	No offer diff.	No idios. match	No firm prem. or shocks	No firm
	(1)	(2)	(3)	(4)	(5)	(6)=(2)+(3)	(7)=(6)+(4)+(5)
26	0.099	0.096	0.098	0.099	0.096	0.095	0.092
30	0.139	0.133	0.128	0.137	0.133	0.122	0.116
35	0.156	0.151	0.137	0.153	0.149	0.132	0.125
45	0.177	0.172	0.142	0.172	0.167	0.136	0.127
55	0.201	0.195	0.147	0.194	0.189	0.141	0.132

and idiosyncratic shocks to the worker-firm pair effects (column 5). Columns 6 shows the combined effects of the first two channels; and column 7 the combined effect of all four channels at once.