Sorting on Unobserved Skills into New Firms

Knutsson, Polina

2018

Document Version:
Other version

Link to publication

Citation for published version (APA):
Knutsson, P. (2018). *Sorting on Unobserved Skills into New Firms*. (Working Papers ; No. 2018:38).

Total number of authors:
1

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Working Paper 2018:38

Department of Economics
School of Economics and Management

Sorting on Unobserved Skills into New Firms

Polina Knutsson

November 2018

LUND UNIVERSITY

Sorting on Unobserved Skills into New Firms*

Polina Knutsson†

Lund University

November 29, 2018

Abstract

Human capital features prominently in theoretical work on post-entry performance of new firms. Empirical analysis has, however, to a large extent overlooked the unobserved component of human capital focusing on years of education or labor market experience. This paper adds to the literature on worker characteristics and post-entry firm performance by putting the unobserved quality of workers in the center of analysis. I find strong evidence that new firms on average employ workers of lower unobserved quality relative to incumbent firms. Among new firms workers of higher unobserved quality are overrepresented in spin-offs and incorporated new firms. I further show that unobserved quality of workers is important for the post-entry performance of firms as it is a strong predictor of new firm survival.

JEL classification: J24, J60, M13

Keywords: Human capital, occupational choice, sorting, new firms

1 Introduction

Creation of new firms is often seen as a driver of innovation, economic growth, and job creation. Not all new firms, however, innovate or render incumbents' technologies obsolete, not all desire to grow, many new firms fail within the first years after entry (Hurst and Pugsley, 2011; Evans, 1987; Geroski, 1995). In attempt to understand what makes some new firms successful, entrepreneurship researchers have long been interested in assessing the role of human capital. Numerous empirical results have been attributed to the possibility that better performing new firms are better at attracting human capital (Bates, 1990; Colombo and Grilli, 2010; Dahl and Klepper, 2015).

There are however several reasons to expect that individuals of great human capital are less likely to work for new firms. There are in fact both theoretical and empirical arguments suggesting that individuals with greater human capital might find employment in new firms less appealing as compared to incumbent firms. First of all, individuals with greater human capital

*I am grateful to Martin Andersson, Joakim Gullstrand, and Torben Schubert for their constructive feedback. Additionally, I thank David Audretsch, Alessandro Gavazza, Lena Hensvik, Luisa Gagliardi, Kristina Andersen, and seminar participants at Lund University, Indiana University, WIPE workshop in Spain, and DRUID conference in Denmark for their helpful comments.

†Department of Economics and Centre for Innovation, Research and Competence in the Learning Economy (CIRCLE), Lund University. Email: polina.knutsson@nek.lu.se Homepage: https://polinaknutsson.com/ Address: Box 7082 S-22007 Lund, Sweden.

are likely to be better at assessing the employment opportunities (Brüderl et al., 1992). Hence, high exit rate of startups might deter these individuals from accepting employment at new firms. Second, it has been shown that jobs created by new firms are paid less than jobs created by older firms (Brown and Medoff, 2003; Heyman, 2007). New firms are also rarely profitable. If wages and profits are connected through rent sharing, earnings prospects of employees in new firms are likely to be uncertain. Third, new firms are less capital-intensive due to small size (Caves, 1998). If capital and workers' skills are complements, more capital-intensive firms are expected to be prone to hiring more skilled workers and more likely to pay higher wages for their ability (Evans and Leighton, 1989). Finally, incumbent firms are more likely to offer such attractive job opportunities as on-the-job-training and internal promotion, as well as fringe benefits (Shane, 2009). Overall, these findings suggest that individuals with greater human capital might be less likely to select into new firms and more likely to seek greater return on their skills and abilities in incumbent firms.

This paper provides evidence that individuals from the right tail of the human capital distribution are indeed less likely to be employed at new firms relative to incumbent firms. I first provide descriptive evidence that workers of new firms in general have less years in education, less years of labor market experience, and more years out of employment. Next, to obtain a more complete picture of sorting, I go beyond these traditional observed characteristics of workers and estimate the unobserved quality of workers - *unobserved skills* - relying on the framework by Abowd et al. (1999). The measure of unobserved skills in this framework reflects the market value of the workers' unobservable characteristics. I confirm that on average workers in new firms have lower unobserved skills relative to incumbent firms. In other words, individuals with lower values of estimated unobserved skills are overrepresented in new firms as compared to incumbents, while individuals with greater unobserved skills are more likely to be employed at incumbent firms. I find that individuals with higher skills are more likely to work for firms that are older, employ more educated and more skilled workers. Further, this paper confirms that sorting upon the unobserved skills is important for post-entry performance of firms by showing that new firms with workers of greater values of unobserved skills face lower probability of exit and higher growth rate.

These findings are important for a number of reasons. First, our understanding of the differences between new and incumbent firms is incomplete as long as we assume that individuals from the right tail of the ability distribution are as likely to start working for new firms as they are for incumbents. For instance, sorting may be at least partly responsible for wage differentials between startups and incumbents. Secondly, while prior literature has often related poor performance of startups with their inability to attract financial capital (Carpenter and Petersen, 2002b,a), findings presented in this paper point out that inability to attract skilled labor may also be to blame. Thirdly, sorting of workers into firms in general has implications for workers' labor market opportunities, as well as their incomes, implying that better understanding of the match between firms and workers has a potential to contribute to our understanding of broader societal issues (Barth et al., 2018; Lazear and Shaw, 2018).

The framework used in this paper to assess the unobserved dimension of human capital connects individuals' earnings to their observable and unobservable characteristics. Assuming the part of earnings unexplained by the observable factors arises due to the unobserved time-invariant skills, I back out the estimate of fixed effects from the earnings regression. Under a set of assumptions on worker mobility (discussed in Sec. 3.2), the estimated fixed effects capture the permanent component of wages that reflects the market valuation of the individuals' unobservable characteristics. I call this measure *unobserved skills*.[1]

To understand the sorting pattern on unobserved skills, I study the distribution of skills across new and incumbent firms. Next, to assess the extent to which the unobserved skills are important for the performance of new firms, I estimate the probability of firms' exit as a function of the mean unobserved skills of their employees.

This paper is related to several strands of literature. First, it contributes to the literature that studies the compositional differences between new and incumbent firms. This literature remains scant. Using data on Sweden, Nyström and Elvung (2014) show that labor market entrants employed by new firms are more likely to have lower GPA-score and are less likely to have tertiary education than those employed by incumbent firms. Ouimet and Zarutskie (2014) find that employees of startups in the US are on average younger than of incumbents. Bublitz et al. (2017) show that more educated employees are more likely to work for more educated founders. This paper contributes to the literature by showing that differences between the employees of new and incumbent firms extend over the observable factors (age, years of schooling, etc.). The paper confirms that individuals with greater unobserved skills are less likely to be employed at new firms relative to incumbents. Furthermore, the differences are not confined to mean values characterizing employees of new and incumbent firms. In fact, this paper documents differences in the entire *distribution* of workers' skills across new and incumbent firms.

Second, this work is related to the literature connecting particular unobserved characteristics of workers with the performance of new firms. Sauermann (2017) finds that employees of startups have different preferences with respect to job security and these differences can partly explain variation in innovative performance of new firms. Levine and Rubinstein (2017) connect variation in earnings of self-employed individuals to their non-routine cognitive abilities and self-esteem. Ayyagari and Maksimovic (2017) relate cognitive skills of the workforce to growth rates of startups. They also show that cognitive skill levels of employees in startups are lower than of employees in incumbents. This paper complements this literature by documenting that differences between new and incumbent firms, as well as variation in post-entry performance extends to a broader set of unobservable qualities of their workers.

Third, this paper relates to the literature connecting post-entry performance of new firms to their stock of human capital. In fact, the link between human capital and post-entry performance has been extensively covered in the literature (see Unger et al., 2011 for a review). Education and experience has become the customary proxies for human capital in this strand of literature (Bates, 1990; Colombo and Grilli, 2005; Baptista et al., 2014). These proxies do not, however, account for a variety of potentially important characteristics of workers' quality such as innate ability, ability to work in teams, quality of education, conscientiousness, creativity, etc. Moreover,

[1] This empirical approach relates to recent literature studying the distribution of skills across firms and cities (Balsvik, 2011; Combes et al., 2012b; Irarrazabal et al., 2013; Fox and Smeets, 2011; Roca and Puga, 2017). In this literature, the estimates of fixed effects are interpreted as unobserved skills and unobserved skills are in turn understood as a proxy for workers' productivity.

it has been shown that the observable characteristics of worker quality yield low explanatory power in earnings regression (Abowd et al., 2005; Fox and Smeets, 2011). This evidence suggests that a large portion of variation in workers' quality is explained by unobservable factors. In fact, it is the unobserved dimension of human capital that usually features in the literature relating startups' performance with human capital. For instance, in Kogut and Zander (1992) human capital is seen as workers' ability to cooperate and make use of assets in a unique way. Human capital manifests itself as productivity in Åstebro and Bernhardt (2005). Brüderl et al. (1992) argue that greater human capital might take a form of better ability to assess business opportunities. Given that it is skills and abilities that are typically unobserved in data that are emphasized in theory, it is important to go beyond years of education and experience when studying the link between human capital and firm performance. This paper confirms that the stock of the unobserved human capital is an important determinant of firms survival even after controlling for the observable factors.

Finally, this work is related to the literature on the match between firms and workers and its implications for workers (Card et al., 2013; Barth et al., 2018), firms (Balsvik, 2011; Irarrazabal et al., 2013; Bender et al., 2018), and cities (Combes et al., 2012a,b; Roca and Puga, 2017). This paper reveals that sorting on human capital extends to the age of firms.

The paper proceeds as follows. Section 2 presents the data set. Section 3 discusses the empirical strategy implemented to study sorting on unobserved skills and documents sorting patterns between new and incumbent firms. Importance of sorting upon unobserved skill for post-entry performance of new firms is assessed in Section 4. Section 5 concludes.

2 Data

The study makes use of the registry data provided by Statistics Sweden over the period 1993-2010. The data set allows matching firm-level records to the individual-level data covering entire population of Swedish private sector employees. The basis for the employee-employer match is the employee composition of firms as of November each year.

The individual-level data contain information on age, gender, education, employment, and annual earnings. I restrict sample to the employees aged between 18 and 60. The upper threshold of 60 years is imposed to exclude the individuals whose employment decisions are likely to be affected by the anticipation of retirement. Since hours worked are not observed, I exclude individuals with the annual labor earnings below 100,000 SEK in 2000 prices.[2] I further trim the sample by excluding observations with the reported annual labor income exceeding top 1% of earnings distribution to avoid extreme values.

The population of firms includes all the firms that have reported their activity to the Tax Agency (Skatteverket) for the personnel registered in Sweden. For each firm the industry affiliation and the number of employees are reported. To separate entrants from incumbents I use the FAD data set. The data set identifies new and continuing firms by tracking the flow of employees over time, ensuring that new firms reflect true entries and not the nominal changes

[2] Antelius and Björklund (2000) show that exclusion of observations with the low annual labor income serves as a good approximation of full-time employment in Sweden. Akerman et al. (2013); Ahlin et al. (2014) are examples of studies using Swedish data that implement the same strategy to approximate full-time employment. As a robustness check, I have replicated the analysis for other thresholds as well. See Appendix B.

Table 1: Worker Characteristics

	Incumbents		New Firms	
	Mean	s.d.	Mean	s.d.
Age	39.19	11.53	36.07	10.89
Female	0.35	0.48	0.31	0.46
Years of schooling	11.92	2.08	11.80	2.07
Primary school	0.15	0.36	0.17	0.37
High school	0.59	0.49	0.60	0.49
College	0.10	0.30	0.08	0.28
University	0.15	0.36	0.13	0.34
GPA	13.46	2.95	13.20	3.03
Work experience	11.47	5.15	8.76	5.21
Not in employment	2.10	2.60	3.33	3.11
No. observations	892764		16759	

Note: Mean values and standard deviation (s.d.) of observable worker characteristics for 2006. *Primary school*, *High school*, *College* and *University* are indicator variables taking value 1 if the corresponding educational level has been achieved. *GPA* is high-school grade measured on measured on 0-20 scale. *Work experience* refers to the years of full-time work. *Not in employment* includes years when a worker was not working full-time or was unemployed.

in firm identification number. As in Eriksson and Kuhn (2006) and Andersson and Klepper (2013), to fulfill the definition of new firms, firms are required to have no affiliation to the existing firms. Further, the number of employees in the first year must be between 2 and 10. This restriction excludes self-employed individuals, as well as larger entrants that are likely to be divested establishments of the existing firms rather than truly independent new firms. Incumbents are defined as firms that have been on the market for at least eight years.[3]

Table 1 reports a snapshot of the demographic characteristics for the full population of employees in new and incumbent firms for year 2006. Employees of new firms appear to be somewhat younger than the employees of incumbents. They have spent slightly less time in education, have lower GPA (high-school grades), have less years of full-time work experience and have spent more years out of full-time employment. Hence, employees of new firms do on average have lower values for the observed characteristics which are commonly associated with greater human capital. Snapshots of worker characteristics for other years point out to the same pattern.

[3]The idea here is to separate more mature firms from new entrants. Defining incumbents as firms with shorter history on the market (6 years) does not lead any noticeable change in the sorting patterns. See Appendix B.

3 Sorting on skills

3.1 Measure of unobserved skills

This section introduces the approach used to study sorting of workers. The intention is to go beyond the standard proxies for human capital summarized in Table 1. Measures of worker quality that are based on years of education or labor market experience do not account for a variety of potentially important characteristics of workers' ability such as innate ability, ability to work in teams, conscientiousness, creativity, etc. Moreover, the observable characteristics of worker quality tend to yield low explanatory power in wage regression, indicating that a large portion of variation in wages is explained by unobservable factors (Abowd et al., 2005; Fox and Smeets, 2011).

I make use of the panel structure of the data to address the unmeasured component of worker's quality. I build on the framework developed by Abowd, Kramarz, and Margolis (1999) - hereafter AKM - and decompose earnings of individual i in year t into individual- and firm-specific components in the following way:

$$Earnings_{it} = \alpha_i + x'_{it}\beta + \Psi_{\boldsymbol{J}_{(i,t)}} + \nu_{it} \tag{1}$$

In this decomposition, α_i is an individual-specific time-invariant pay component. $x'_{it}\beta$ is a set of time-varying factors that affect i's productivity across different jobs. It includes the effects of education and experience (discussed in greater detail in Sec. 3.2). Together, α_i and $x'_{it}\beta$ represent the individuals-specific component of earnings. $\Psi_{\boldsymbol{J}_{(i,t)}}$ gives the identity of a firm where individual i is employed at time t. $\Psi_{\boldsymbol{J}_{(i,t)}}$ is then a firm-specific component of earnings that can be interpreted as firm-specific wage premium or discount to its workers. Finally, ν_{it} is the error component of earnings.

The decomposition allows purging earnings from the firm-specific component ($\Psi_{\boldsymbol{J}_{(i,t)}}$) and the observable workers' characteristics ($x'_{it}\beta$) that affect individual's wage setting. The remaining individual-specific component α_i is then the combination of time-invariant skills and other factors that raise or lower individual i's earnings regardless of where i is employed. Put differently, α_i should reflect market valuation of the portable component of workers' skills which in principle should incorporate such attributes of workers as their innate ability, creativity, ambition, analytic reasoning, social skills, etc. Since workers in the sample are observed for many years (mean amount of observations per worker is 9 years) and potentially across many employers (mean amount of employees is 3), the estimate of α_i is based on the valuation across employers. The estimate of α_i is in the focus of this paper.

Theoretical motivation behind the idea to use workers' earnings to infer the distribution of their skills builds on the assumption that workers' earnings reflect their marginal productivity. Even in the presence of market frictions, wages should be highly correlated with workers' productivity (Irarrazabal et al., 2013). Differences in wages are used to approximate unobserved skills in Eeckhout and Kircher (2011) to compare skills across cities and in Balsvik (2011) to compare workers in multinational corporations with those of domestic firms. Ability is measured as the residual from the fully saturated Mincerian equation in Dal Bó et al. (2017) who study selection of individuals into politics. Most recent applications of AKM decomposition in the context of worker skills include Bender et al. (2018), who study the relationship between management quality and firm productivity as well as Card et al. (2015, 2018) whose focus is on labor market inequality. The empirical approach adopted in this paper is closely related to the studies that derive the estimates of worker fixed effects to examine sorting on ability. For instance, Combes et al. (2012b) as well as Roca and Puga (2017) use estimated worker fixed effects to explain urban wage premium. The estimated worker fixed effects are used in Fox and Smeets (2011) and Irarrazabal et al. (2013) to explain productivity differences across firms.

3.2 AKM estimates

In this section I discuss the empirical issues related to the estimation of worker fixed effects. Abowd et al. (1999) show that the unbiased estimates of α_i (and Ψ_j) can be obtained using ordinary least squares. As shown by Abowd et al. (2002), worker and firm fixed effects can be separately identified only for the observations in the connected set. A connected set is a set of individuals connected by worker mobility. The set includes all the individuals ever employed by a firm in the set of all the firms that any individual in the set has ever been employed at. I restrict the analysis to the largest connected set, which covers 98.2% of the sample.

Another issue related to the identification of worker fixed effects in AKM decomposition is the specification of age effects. In the traditional Mincerian equation, age effects are typically controlled for by including a polynomial in age or in potential experience. At the same time, to control for economic trends, a set of year indicators is usually included. This gives rise to an identification challenge due to the collinearity between age and year count (Card et al., 2013). To deal with this challenge, Abowd et al. (1999) use actual labor market experience instead of the potential, as gaps in actual experience mitigate the collinearity of experience with calendar year. A limitation of this approach in the context of this paper is that in my data set actual experience is fully available only for a subsample of workers. Furthermore, as argued by Card et al. (2013), the gaps are likely to be exogenous even after controlling for worker effects and hence might confound the estimated of firm fixed effects. I follow the solution implemented by Card et al. (2013). To deal with this identification issue they include a third-order polynomial in age by education group together with year indicator variables and impose a linear restriction on age effects.[4]

An additional complication in estimating worker fixed effects arises due to the way new firms are defined. By definition, new firms are observed in the data up to three years. That is, new firms have on average fewer observations over time than incumbents. The estimates of firm fixed effects can still be obtained for firms observed for only a few data points, as long as the observations are in the connected data set. Still, the estimated firm fixed effects for the firms observed only a few periods will systematically use less variation than the estimates for firms observed for more periods. To make sure that the discrepancy in the period under observation does not systematically affect the estimates of interest, I proceed as follows. I split the available data into two periods: a longer pre-sample (1993-2005) and a shorter main sample (2006-2010).

[4] A linear restriction is achieved by restricting the age effects to be equal to zero at age of 40. That is, a linear term is omitted for each educational group. The choice of age 40 is motivated by the evidence that wage profile tends to be relatively flat around this age (Card and Cardoso, 2012).

I perform AKM decomposition for the pre-sample.[5] The pre-sample is sufficiently long for the largest connected set to cover a large portion of observations (98.2% as discussed above). Next, I obtain the estimated worker fixed effects for all the individuals in the pre-sample and match with the individuals in the main sample (recall that the estimates are time-invariant). All the subsequent analysis is performed for the main sample. This approach ensures that the estimates of worker fixed effects for the employees of new firms are not systematically biased due to the way new firms are defined. Moreover, the split ensures that the estimated worker fixed effects are not endogenous to the subsequent performance of the firms. A consequence of this approach is that the analysis is restricted only to the individuals who are observed both in the sample and in the pre-sample.[6]

Finally, a discussion on the assumptions of AKM decomposition is due. AKM relies on the assumption that worker mobility is not correlated to the error component of earnings ν_{it} conditional on the observable and unobservable worker characteristics (as known as "exogenous mobility assumption"). Put differently, the residual component is assumed to be uncorrelated with the sequence of employer identifiers in worker's employment record. While this assumption appears strong, it commonly finds support in the data. For instance, Card et al. (2013) conclude that AKM assumptions are justified for German data, Card et al. (2015) and Macis and Schivardi (2016) reach the same conclusion for correspondingly Portugal and Italy. Following the event study analysis developed by Card et al. (2013) and Card et al. (2015), I document several empirical facts suggesting that the exogenous mobility assumption is also substantially met in the Swedish register data (see Appendix A).

3.3 Analysis of sorting

Recent literature on sorting has emphasized importance of comparing the entire distributions of variables of interest. Focus on mean values might be misleading if the distributions differ in tails and in the overall variance of values. In principle, distributions can be compared compared graphically. Combes et al. (2012a) offer a more formal way of assessing sorting.[7] The method is designed to compare the distribution of one variable across two groups. It assesses how well the first distribution can approximate the second one if the latter is shifted by a parameter A and dilated with a parameter D. An additional parameter in the assessment is truncation of the distributions. For simplicity, I restrict the exposition to the discussion based only on the shift and dilation. For a detailed overview of the method see Combes et al. (2012a). As it will become evident in the Sec. 3.4, the estimated truncation is negligible. In the context of sorting into new and incumbent firms, the distribution of worker fixed effects in new firms $F_{New}(\alpha)$ is approximated by shifting the distribution of worker fixed effects in incumbent firms $F_{Inc}(\alpha)$ by

[5] Choosing a different year for sample split does not change main conclusions of the paper (see Appendix B)

[6] It is worth noting that this empirical strategy restricts the analysis to the individuals who are observed for at least three years in the data (at least two years in the pre-sample plus at least one in the sample). Hence, workers in this analysis are older and have more labor market experience relative to the full sample of workers. This restriction has a meaningful interpretation in the context of this paper. It has been argued that workers with greater labor market experience are less likely to switch jobs and their current employment is, therefore, more likely to be the result of a match that satisfies both the employer and the employee (Topel and Ward, 1992). Put differently, this empirical strategy restricts the analysis to individuals whose set of skills is more likely to match the demand of their employer.

[7] The context in which the method is explicated in Combes et al. (2012a) is comparison of TFP distributions. For the application of the method in the context of unobserved worker characteristics see Combes et al. (2012b); Roca and Puga (2017).

parameter A and dilating it with a parameter D. For two identical distributions the shift has to be 0 and the estimated dilation has to be 1. The parameters of interest are estimated by minimizing the mean quantile difference between the two distributions. Let $M(\hat{A}, \hat{D})$ denote the total mean quantile difference between the first distribution and the second distribution transformed by shifting it by parameter $\hat{A}$ and dilating it with a parameter $\hat{A}$. Then the measure of fit can be defined as $R^2 = 1 - M(\hat{A}, \hat{D})/M(0, 1)$. Intuitively, this measure of fit is a fraction of the difference in the two distributions which is explained by shift and dilation. Standard errors of the estimated parameters are obtained by bootstrapping. The observations are drawn for the full sample size with replacement and the parameters are re-estimated for each bootstrap replication.[8]

In general, when firms of type f and g are compared with respect to the unobserved skills of their workers, a positive shift in the distribution of unobserved skills in type f relative to g implies that workers employed in firms type f have greater unobserved skills relative to g. That is, to approximate the distribution of skill in g, skills in f need to be decreased. That distribution of f is dilated relative to g means that workers are overrepresented in the both tails of the distribution.

3.4 Main results: Sorting on skills

Figure 1 Panel A plots the distribution of worker fixed effects estimated from Equation 1 on the pre-sample. The formal analysis of the shift in the distribution is presented in Table 2, row 1.[9] The worker fixed effects are demeaned so that the mean value for the whole distribution of workers in the sample is zero.

Figure 1 Panel A reveals that the distribution of worker fixed effects of new firms is shifted to the left relative to the incumbent firms. The shift implies the over-representation of less skilled workers (as measured by the fixed effects) in new firms relative to incumbents. This result is further confirmed in Table 2, top row, where the estimated shift is statistically significant. The estimated shift parameter Â is -0.03, indicating that to approximate the distribution of worker fixed effects in new firms, the distribution of worker fixed effects in incumbent firms needs to be shifted 0.03 units to the left (magnitude of the shift corresponds approximately to 10% of the standard deviation in worker fixed effects). Hence, the average estimated worker fixed effects are lower in new firms relative to incumbents. The estimated dilation parameter D̂ is very close to 1 indicating that the distribution of unobserved skills in new firms is only slightly dilated as compared to incumbent firms. The modest dilation indicates that there is no considerable over-representation of high- and low-skilled workers in the tails of the distributions. Estimated truncation is negligible as indicated by parameter Ŝ. Together, shift, dilation, and truncation account for 92.3% mean-squared quantile difference between the distributions of worker effects in new and incumbent firms.

Focus of this paper is on the unobserved qualities of workers. To complement the findings

[8] I use Stata module *estquant* by Kondo et al. (2017) to perform the estimation.

[9] The results presented in Figure 1 and in Table 2 are based on the observation for year 2006. The reason for presenting the results only for one year is that workers move across firms over time and to make a meaningful comparison of the distributions it is necessary to take a snapshot on a specific date. Plots for other years look very similar, even though the number of the individuals matched from the pre-sample and sample declines. Signs of the coefficients in the analytical results are the same, magnitudes of the coefficients are very similar. See Appendix B.

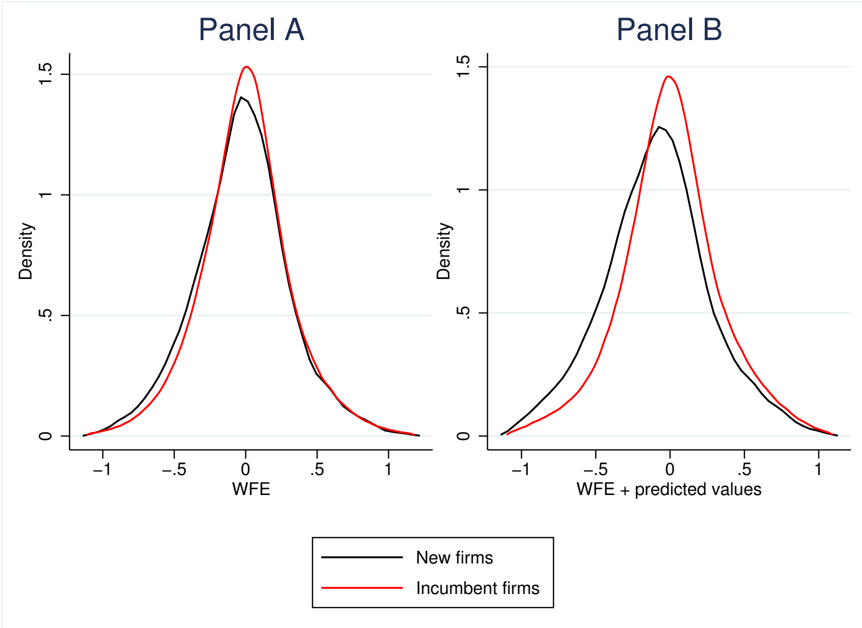

Figure 1: Worker fixed effects in new and incumbent firms (Panel A). Sum of worker fixed effects and predicted values in new and incumbent firms (Panel B)

presented in Panel A, Panel B of Figure 1 plots the distribution of the sum of worker fixed effects and predicted values from Eq. (1). In AKM decomposition, the sum of worker fixed effects and predicted values can be interpreted as human capital (Abowd et al., 2005), as it is the sum of the observable and unobservable qualities of workers. As Panel B reveals, the difference between employees of new and incumbent firms is even more stark with this measure. The estimated shift in the distribution of the measure of human capital in incumbents to approximate new firms is -0.1, indicating that the estimated human capital is on average higher for workers in the incumbent firms. Again, the distribution for workers of new firms is slightly more dilated and the truncation parameter is negligible.

3.5 Further results

Panel A of Figure 1 establishes that workers with greater unobserved skills are underrepresented in new firms as compared to incumbents. This section explores the extent to which this pattern reflects the compositional differences between new and incumbent firms. In particular, I first

Table 2: Worker fixed effects in new and incumbent firms

	Â	D̂	Ŝ	R^2	Obs.1	Obs.2
α_{Inc} to α_{New}	-0.0287***	1.048***	0.001***	0.923	1538732	20868
	(0.002)	(0.006)	(0.000)			
$(\alpha + X\beta)_{Inc}$ to $(\alpha + X\beta)_{New}$	-0.100***	1.104***	0.007***	0.986	1536594	20757
	(0.003)	(0.007)	(0.001)			

The estimated parameters are shift Â, dilation D̂, and truncation Ŝ. In the first row the distribution the distribution of worker fixed effects in incumbent firms denoted by α_{Inc} is compared to the distribution of worker fixed effects in new firms denoted by α_{New}. In the second row, the distributions of the sum of worker fixed effects and predicted values $(\alpha + X\beta)$ are compared. Obs. 1 and Obs. 2 denote number of observations in the first and the second group of firms. Bootstrapped standard errors in parentheses (100 iterations). ***p<0.01, **p<0.5, *p<0.1.

Table 3: Worker fixed effects: displaced workers; workers by age group

	Â	D̂	Ŝ	R^2	Obs.1	Obs.2
$\alpha_{Inc_{(D)}}$ to $\alpha_{New_{(D)}}$	-0.0597^{***}	0.907^{***}	-0.005	0.940	13901	2892
	(0.007)	(0.020)	(0.004)			
$\alpha_{Inc_{<35}}$ to $\alpha_{New_{<35}}$	0.0146^{***}	1.066^{***}	0.006^{***}	0.868	434621	7468
	(0.004)	(0.012)	(0.002)			
$\alpha_{Inc_{[35;45)}}$ to $\alpha_{New_{[35;45)}}$	-0.0115^{**}	1.086^{***}	0.001	0.873	494542	7553
	(0.004)	(0.011)	(0.001)			
$\alpha_{Inc_{[45;55)}}$ to $\alpha_{New_{[45;55)}}$	-0.0210^{***}	1.125^{***}	0.001	0.857	375524	4051
	(0.006)	(0.013)	(0.001)			
$\alpha_{Inc_{>55}}$ to $\alpha_{New_{>55}}$	-0.0440^{***}	1.103^{***}	0.005	0.853	201106	1534
	(0.010)	(0.025)	(0.003)			

The estimated parameters are shift Â, dilation D̂, and truncation Ŝ. In the first row the distribution the distribution of worker fixed effects for the displaced workers moving to incumbents ($\alpha_{Inc_{(D)}}$) is compared to the corresponding contribution in new firms ($\alpha_{New_{(D)}}$). In rows 2-5, worker fixed effects in incumbents and new firms are compared across four age categories: workers aged below 35; aged from 35 to 45, form 45 to 55, and above 55. Obs. 1 and Obs. 2 denote number of observations in the first and the second group of firms. Bootstrapped standard errors in parentheses (100 iterations). ***p<0.01, **p<0.5, *p<0.1.

explore the composition of workers and then the composition of firms.

3.5.1 Tenure

While Panel A of Figure 1 is informative about the distribution of the unobserved skill across new and incumbent firms, it may reflect the differences in the possible tenure for employees in startups and incumbents. Indeed, the largest possible length of employment for employees in new firms by definition is three years, while it is unlimited for the employees of incumbent firms. Put differently, employees of new firms are, by construction, the individuals who have recently changed jobs. Whereas employees of incumbents might have been employed at the same firm for many years. If there are some systematic differences in unobserved skills between the individuals leaving one job for another and those who stay at the same job, the sorting pattern presented in Figure 1 is a reflection of differences in the tenure of workers. By comparing only individuals that have just started their employment at new or incumbent firms, we can find out whether sorting is present in the very decision to accept the job or sorting resides in the difference in tenures. To do this, I focus on the displaced workers - workers who were forced to search for new jobs due to the closure of the plants where they were employed in the previous period.[10] By restricting the analysis to only displaced workers, I purge the unobserved factors that might be related to the selection into a voluntary change of job. In this I follow a voluminous literature that interprets plant closures as quasi-natural experiment in a sense that workers are forced to search for new jobs for reasons exogenous to their performance (Eliason and Storrie, 2006; Hijzen et al., 2010; Eliason, 2011; Browning and Heinesen, 2012).

Row 1 of Table 3 presents the results of analytical comparison of the distributions of unobserved skills of the displaced workers. The estimated shift is -0.06, which is larger in absolute

[10] Closures of plants are defined using FAD data set. I only include closure of the plants employing at least 50 employees to minimize the possibility that a closure is endogenous to the unobserved skills of a particular employee.

Table 4: Worker fixed effects by firm size

	$\hat{A}$	$\hat{D}$	$\hat{S}$	R^2	Obs.1	Obs.2
$\alpha_{Inc_{(1;10]}}$ to $\alpha_{New_{(1;10]}}$	-0.0148^{***}	0.979^{***}	-0.000	0.938	395951	19536
	(0.002)	(0.006)	(0.001)			
$\alpha_{Inc_{(10;20]}}$ to $\alpha_{New_{(10;20]}}$	-0.0249	1.069^{***}	0.000	0.750	204817	1095
	(0.018)	(0.054)	(0.029)			
$\alpha_{Inc_{>20}}$ to $\alpha_{New_{>20}}$	-0.0413^{*}	1.111^{***}	0.007	0.891	943239	246
	(0.019)	(0.081)	(0.012)			

The estimated parameters are shift Â, dilation D̂, and truncation Ŝ. The distribution the distribution of worker fixed effects are compared across firm size groups: firms with 10 employees or less, from 10 to 20 and above 10 employees. Obs. 1 and Obs. 2 denote number of observations in the first and the second group of firms. Bootstrapped standard errors in parentheses (100 iterations). ***p<0.01, **p<0.5, *p<0.1.

terms than the result for the whole sample of workers. Out of the workers who start new jobs due to being displaced from the previous workplaces, more skilled ones are less prone to start at new firms. That is, high-skilled workers are underrepresented among the new hires of new firms, when the sample is restricted to the displaced workers only. Neither dilation, nor truncation appear to be major contributors to the differences in the two distributions.

3.5.2 Age of workers

Rows 2-5 of Table 3 explore the heterogeneity in the estimates of unobserved skills across age groups. As coefficients reveal, the sorting pattern with less skilled workers sorting into new firms persists among all age groups, except for the youngest workers. The estimated shift coefficient is positive and significant for workers below 35 years (Row 2), suggesting that among workers younger than 35 years, the sorting pattern is reversed relative to the main findings and more skilled younger workers sort into new firms. What could be driving this result? As such, younger workers are more likely to change jobs than older employees as they learn about their own productivity relative to coworkers and other jobs (Topel and Ward, 1992). They might also be less risk averse when making decision about joining firms (Ouimet and Zarutskie, 2014). If risk tolerance is correlated with the market value of workers' unobserved skills, it can explain the sign of the shift coefficient, suggesting that less risk averse individuals whose unobserved skills has been highly valued by their previous employers are more likely to take a risk of joining a new firm. At the same time, for older groups of workers the sorting pattern is as before negative. The fact that the difference in worker fixed effects increase with age might reflect the discrepancy in wages paid by new and incumbent firms (Brown and Medoff, 2003).

3.5.3 Firm size

New firms are by definition small at year of entry. While some firms experience a more than tenfold increase in the number of full-time employees within their first years in the market, the majority of firms remains small. Mean size of new firms in the sample is 4.5 employees and maximum size at the age of three is 88 employees. To find out whether sorting upon unobserved skills persists even among firms of similar size, I split the firms into three size groups: firm

Table 5: Worker fixed effects by sector group

	Â	D̂	Ŝ	R^2	Obs.1	Obs.2
$\alpha_{Inc_{(HT)}}$ to $\alpha_{New_{(HT)}}$	-0.0290^{*}	0.933^{***}	-0.016	0.690	371505	1190
	(0.012)	(0.045)	(0.017)			
$\alpha_{Inc_{(LT)}}$ to $\alpha_{New_{(LT)}}$	-0.0404^{***}	1.049^{***}	-0.001	0.893	148283	692
	(0.011)	(0.044)	(0.010)			
$\alpha_{Inc_{(KIBS)}}$ to $\alpha_{New_{(KIBS)}}$	0.0074	1.026^{***}	0.001	0.783	217709	5263
	(0.005)	(0.010)	(0.001)			
$\alpha_{Inc_{(OS)}}$ to $\alpha_{New_{(OS)}}$	-0.0434^{*}	1.087^{***}	0.004	0.854	938548	245
	(0.004)	(0.012)	(0.003)			

The estimated parameters are shift Â, dilation D̂, and truncation Ŝ. In all the rows worker fixed effects in incumbent firms are compared to new firms for a one of four broadly defined sectors. HT denotes High-tech, LT - low-tech manufacturing, KIBS - knowledge-intensive business services, OS - Other Services. Obs. 1 and Obs. 2 denote number of observations in the first and the second group of firms. Bootstrapped standard errors in parentheses (100 iterations). ***p<0.01, **p<0.5, *p<0.1.

employing 10 or less employees, from 10 to 20, and firms with above 20 employees.[11] Row 1 of Table 4 reveals that sorting on unobserved skills is persistent also for micro firms. That is, incumbents with 10 or less employees have on average more skilled workers than entrants of the same size. Shift remains negative, but becomes insignificant for the second size group firms (Rows 2). At the same time, shift, dilation and truncation of the distribution of worker fixed effects in micro incumbent firms does not describe the distribution in new firms as well as in the main analysis: the measure of fit is only 75%. The shift is negative and weakly significant for the largest size group. Overall, the analysis by size groups reveals no considerable heterogeneity in the sorting pattern.

3.5.4 Sectoral composition

Figure 1 does not account for the possible difference in sectoral composition of new and incumbent firms. To examine sorting in closer detail, I group all the industries into four broad groups: high-tech manufacturing, low-tech manufacturing, knowledge-intensive business services and other services. Results reported in Table 5 suggest that on average employees of new firms are less skilled relative to incumbents in high-tech manufacturing, low-tech manufacturing, and in other services. These three sectors contribute to the underrepresentation of high-skilled individuals in new firms relative to incumbents, as reported in the main results. The shift in the distributions for KIBS is positive, but is not statistically significant, implying that at least for KIBS we cannot conclude that high skilled workers are underrepresented in new firms relative to incumbents. This finding might reflect correlation between unobserved skills of workers and professional expertise which firms in KIBS build on.

Overall, neither worker nor firm composition seem to mask any considerable heterogeneity in the sorting pattern found in the main analysis, suggesting that the sorting pattern documented in the main analysis holds in various contexts.

[11] The choice of size threshold is arbitrary. Larger threshold for the last group (e.g. above 50 employees) result in a low amount of observation, thus yielding a low fit of the distributions. The estimated shift becomes insignificant, but the coefficient is still negative.

Table 6: Determinants of sorting into new firms

	$\hat{\alpha}$	$\hat{\alpha}$	$\hat{\alpha}$	$\hat{\alpha}$
Coworkers' $\hat{\alpha}$			0.304***	0.294***
			(0.009)	(0.009)
Coworkers' educ		0.022***	0.018***	0.017***
		(0.001)	(0.001)	(0.001)
Firm age	0.007***	0.006***	0.005***	0.005***
	(0.000)	(0.000)	(0.000)	(0.000)
Firm size	-0.032***	-0.036***	-0.027***	-0.040***
	(0.004)	(0.007)	(0.005)	(0.006)
Firm size, sq.	0.000	0.001	0.000	0.003*
	(0.001)	(0.002)	(0.001)	(0.001)
Share females	-0.229***	-0.238***	-0.185***	-0.183***
	(0.005)	(0.005)	(0.004)	(0.004)
Spinoff				0.031***
				(0.002)
Incorporated				0.047***
				(0.003)
Constant	0.030	-0.238***	-0.217***	-0.181**
	(0.036)	(0.039)	(0.031)	(0.066)
N	340151	318560	318560	316900
Adj.R^2	0.09	0.09	0.11	0.11

Notes: Dependent variable is worker fixed effects ($\hat{\alpha}$), it is centered around zero. All specifications include industry and year indicator variables. Estimation covers the period from 2002 to 2010, only firms whose entry is observed in the dataset are included. Firm size is log of number of employees. Standard errors are clustered at the firm level. ***p<0.01, **p<0.5, *p<0.1.

3.6 Determinants of sorting

To provide more insights into the drivers of sorting, this section discusses the results of estimating a series of simple regressions relating worker fixed effects to firm-level characteristics.

Section 3.4 has established that workers sort into new firms upon the unobserved skills. But how does sorting upon unobserved skills take place among market entrants? To provide a more detailed description of the sorting patterns across new firms, I have estimated several regression models that relate worker fixed effects to firm-level variables for the sample of new firms. To have more variation in age of entrants, for this analysis I follow cohorts of firms whose entry I observe in the data set, restricting the firms to be at most ten years old.[12] The analysis is performed on individual level.

Table 6 summarizes the results. All the specifications control for industry and year effects. Column 1 confirms that older firms have higher skilled workers. Column 1 reports a baseline model with only firm age, size, size squared, share of females as controls. Firm size as measured the number of employees in logs is negatively correlated with the worker fixed effects. This

[12] To be more consistent with the analysis in Section 3.4, as a check I also run the analysis where I follow only entrants up till their third year on the market. This sample restriction results in lower amount of observations and significance of the age coefficient in the first specification disappears. Overall, the direction and the significance of the effects remains the same. The results are available upon request.

Table 7: Comparison of worker fixed effects distributions

	$\hat{A}$	$\hat{D}$	$\hat{S}$	R^2	Obs.1	Obs.2
$\alpha_{New_{(I)}}$ to $\alpha_{New_{(N)}}$	0.0820***	1.009***	-0.001	0.986	3918	16744
	(0.006)	(0.017)	(0.001)			
$\alpha_{New_{(Sp)}}$ to $\alpha_{New_{(NSp)}}$	0.0674***	0.896***	-0.007**	0.988	15573	5304
	(0.005)	(0.015)	(0.003)			

The estimated parameters are shift $\hat{A}$, dilation $\hat{D}$, and truncation $\hat{S}$. In the first row $\alpha_{New_{(I)}}$ and $\alpha_{New_{(N)}}$ denote correspondingly the distribution of worker fixed effects in new incorporated and non-incorporated firms. In the second row, spin-offs are compared with all new firms not classified as spin-offs. Obs. 1 and Obs. 2 denote number of observations in the first and the second group of firms. Bootstrapped standard errors in parentheses (100 iterations). ***p<0.01, **p<0.5, *p<0.1.

result might reflect the fact that larger firms are more likely to perform a broader set of tasks, which in turn requires a broader variety of skills. Specification in Column 2 includes education of coworkers, which is a simplest measure of the observed human capital of the coworkers. Education of coworkers appears to be positively correlated with worker fixed effects. In Column 3 worker fixed effects of coworkers enter as a covariate. Coefficient of coworkers' fixed effects is positive and significant, suggesting that estimated worker fixed effects is a strong determinant of sorting on unobserved skill.

In column 4 an indicator variable for incorporated firms[13] and an indicator variable for spin-offs[14] are included. These indicator variables are supposed to capture new firms with greater economic potential. Incorporation status of a startup has often been interpreted as a proxy for startup potential (La Porta and Shleifer, 2014; Guzman and Stern, 2015; Levine and Rubinstein, 2017), as costs of incorporation (stricter reporting rules, upfront cash payment), as well as its benefits (limited liability) are arguably more important for businesses that plan to undertake innovative risky projects. At the same time, spin-offs have often been found to outperform other new firms in terms of survival and growth rates (Thompson, 2005; Eriksson and Kuhn, 2006; Andersson and Klepper, 2013). As Table 6, Column 4 reveals, higher skilled individuals are more likely to be employed at incorporated firms and at spin-offs.

The fact that incorporated startups and spin-offs employ more skilled individuals is further illustrated in Figure 2 and confirmed analytically in Table 7. Indeed, the distribution of worker fixed effects in incorporated new firm is shifted to the right as compared to the distribution in new firms that start non-incorporated. A similar pattern is observed among spin-offs and other new firms. These results are important, as they complement the literature that studies how these types of startups - incorporated startups and spin-offs - are different from the other new firms. It turns out that these startups are more likely to employ individuals with greater human capital as approximated by worker fixed effects. The finding that new firms whose superior economic potential has been documented in the literature are more likely to employ more skilled workers indicates that sorting on unobserved skill is linked to firm performance.

Overall, the estimates presented in this section provide some insights about firm-level characteristics of startups that are associated with higher estimated worker fixed effects. It turns

[13]Incorporated firms are limited liability companies with a separate legal identity. Incorporated firms are subject to stricter reporting and auditing procedures and face minimum capital requirements.

[14]Spin-offs are firms with the majority of their initial employees coming from the same firm.

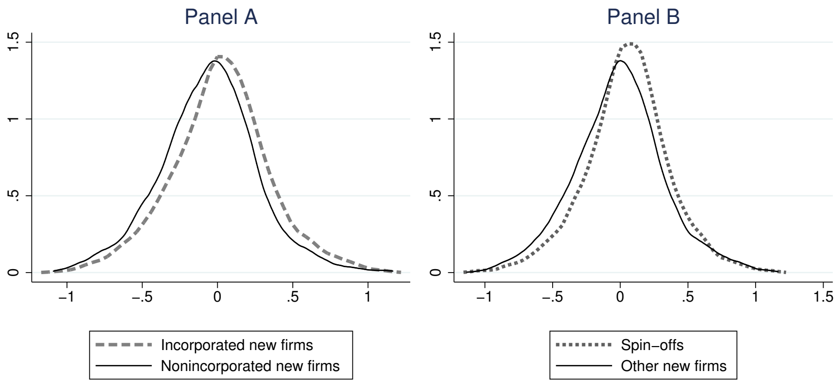

Figure 2: Worker fixed effects in new firms. In Panel A new firms are divided into two groups according to their incorporation status. In Panel B new firms are divided according to the entry mode into spin-offs and other new firms.

out that more skilled workers are more likely to work for new firms with high skilled employees. This strong positive correlation between fixed effects of a worker and its coworkers is stable across all the specifications. It also holds when the correlation exercise is performed for all the firms, i.e. when incumbents are included in the sample (see Appendix, Table D.1).

Overall, the analysis of the determinants of sorting into new firms enriches our understanding of the sorting pattern observed in the main analysis. Age appears to be an important determinant of sorting even among younger firms. Moreover, individuals with greater estimated unobserved skills are more likely to work for market entrants which employ other workers with high unobserved skills. Incorporated startups are more likely to host individuals with greater unobserved skills relative to non-incorporated ones. Similarly, employees with greater unobserved skills are overrepresented in spin-offs as compared to other new firms.

3.7 Alternative measures to study sorting

A limitation of the measure of unobserved skills used in the main analysis is that the measure can be obtained only for full-time employees, thus ignoring part-time employees and individuals who were previously unemployed. Moreover, main analysis discards the observations that find no match in the pre-sample. To ensure that the sorting pattern described in the main analysis holds when all the employees are included in the sample, I use a proxy for unobserved skill - high-school grades (GPA hereafter), as GPA is often argued to be highly correlated with ability and ambition (Grogger and Eide, 1995; Miller, 1998). GPA is measured on 0-20 scale which was introduced in Sweden in 1994. For individuals who finished high school before 1994 grades are obtained using the conversion table provided by antagning.se.

Panel A of Figure 3 illustrates kernel density plots of GPA for employees of new and incumbent firms. It is evident from the plot, that GPA of employees of new firms is lower across the distribution.[15]

To complement main analysis I use yet another proxy for unobserved skills - earnings of

[15] Due to lumpiness of GPA, the bandwidth for kernel density was chosen above the "optimal" one to facilitate the comparison of plots. Larger bandwidth means that the figure illustrates smoothed density. With smaller bandwidth the sorting pattern remains the same with an evident shift to the right in GPA of incumbents, but the density is more lumpy. The results are available upon request.

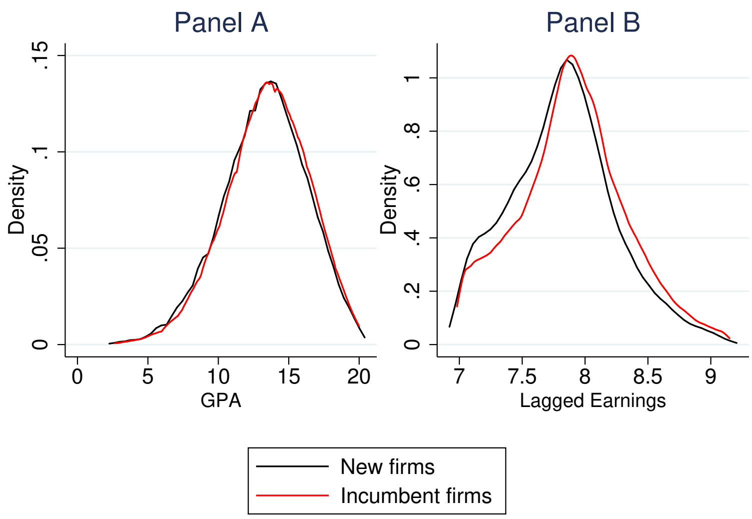

Figure 3: GPA of all the workers and lagged log Earnings of new hires.

new hires (former displaced workers). AKM estimation utilizes annual earnings and backs out the estimate of unobserved skill relying on a set of assumptions. At the same time, monetary reimbursement itself has been used in literature to compare worker quality. Using wages to approximate workers unobserved skills can be motivated by the assumption that workers' wages should reflect their marginal productivity (Balsvik, 2011; Irarrazabal et al., 2013; Eeckhout et al., 2014). Even if markets are not competitive and workers' wages do not fully correspond to their marginal productivity, they are likely to be highly correlated with workers' productivity (Fox and Smeets, 2011; Irarrazabal et al., 2013). When comparing workers of new and incumbent firms upon unobserved skill, it is important to ensure that the measure of unobserved skill is not driven by the performance of the firm where an individual is currently employed. Therefore, I use earnings (in logs) two years before starting at new or incumbent firms. Lag of two years is chosen due to the annual character of the data at use to ensure that these are the earnings at the previous workplace.

Panel B of Figure 3 plots the distributions of lagged earnings for the displaced workers hired by new and incumbent firms. It shows a clear differences between employees of new firms and incumbents, with new employees of new having on average lower lagged earnings. This result corroborates the results with the main analysis.

4 Unobserved skills and firms outcomes

4.1 Probability of exit

To assess the importance of sorting upon unobserved skills for new firms, I proceed by examining the link between the probability that a new firm exits and the unobserved skills of its workers. Probability of firm exit is a natural choice for assessing firm success given high exit rate of new firms documented across different countries and time periods (Audretsch and Mahmood, 1995; Evans, 1987; Thompson, 2005). I start by relating the probability that a firm entering in 2006 exits before 2010 to the measure of unobserved skills and a set of firm-level characteristics at

Table 8: Probability of exit of new firms

	(1) Logit	(2) LPM
Mean $\hat{\alpha}$	-0.084**	-0.084**
	(0.027)	(0.027)
Mean Educ	-0.009	-0.009
	(0.005)	(0.005)
Firm size	-0.056***	-0.055***
	(0.015)	(0.014)
Share females	-0.044	-0.044
	(0.024)	(0.024)
Baseline prob. of exit	0.337	0.337
Observations	4045	4045
Pseudo-R^2	0.02	
Adjusted-R^2		0.02

Notes: Column 1 reports the marginal effects at mean values from a logistic regression. Coefficients from linear probability model are reported in column 2. Both specifications include industry controls. Robust standard errors are in the parentheses. Baseline probability of exit is the probability of exit at mean values of dependent variables. ***p<0.01, **p<0.5, *p<0.1.

year of entry. In appendix I proceed by estimating a duration model including all the entrants and incorporating time-variant covariates, which yields very similar results (see Appendix C).

I estimate the following model:

$$Pr(Exit_f^{t_0;T}) = \Phi(\mathbf{W'}_{ft_0}\beta_1 + \mathbf{X'}_{ft_0}\beta_2 + d_s) \tag{2}$$

$\mathbf{W}_{ft_0}$ is measured as the average of worker fixed effects for all the workers employed in $t_0 = 2006$. $\mathbf{X}_{ft_0}$ includes firm size (number of employees in logs), mean education of employees, share of female employees. Firm size is a standard control in the analysis of firm survival and has been motivated by work of Mata and Portugal (1994); Audretsch and Mahmood (1994); Dunne et al. (1988). Mean education of employees and share of female employees are included to control for differences in the worker composition. d_s are industry effects which capture unobservable industry characteristics that influence the probability of exit. The dependent variable takes value 1 for those entrants who have exited before 2010.[16]

4.2 Results: probability of exit and unobserved skills

Table 8 Column 1 reports the marginal effects at mean values from a logistic regression of firm exit probability as specified in Eq. (2). The second column reports the estimates from the liner probability model. Marginal effects from both models are very similar. The probability of exit is decreasing in mean unobserved skills of workers. An increase in one standard deviation on mean worker fixed effect is associated with a decline in probability of exit of 2.6% (-0.084× 0.31,

[16] Firms that restructure due to mergers and acquisitions are not considered as exiting.

see Appendix D.2 for summary statistics). This is a relevant magnitude, given that the baseline probability of exit is 33.7%. The estimated marginal effect is of the same magnitude if linear probability model (LPM) is used to estimate the probability of exit (Column 2).

To sum up, Table 8 reveals that new firms with greater average unobserved skills are less likely to exit. Table D.3 in Appendix shows that the probability of exit is decreasing in unobserved skills of its workers also if alternative ways to aggregate worker fixed effects across employees are used (median, percentiles, maximum). Interestingly, in this specification, the coefficient of mean unobserved skills is significant even after controlling for mean education of workers, which is a common control for human capital stock in firms, indicating that both measures capture variation in different aspects of worker characteristics. Overall, I conclude that the unobserved skills are an important determinant of firm survival.

5 Conclusion

This paper starts by arguing that new firms might be less appealing as workplaces to the individuals with greater human capital. I first provide descriptive evidence suggestive that individuals employed by new firms are on average less educated, have lower market experience and were more likely to be have gaps in their full-time employment history. I then argue that these observable characteristics that measure amount of years in education or on labor market mask an important portion of human capital, as they do not capture ambition, creativity, innate ability, etc. To overcome this limitation of the traditional proxies for human capital, I employ a framework that aims at estimating the unobserved dimension of human capital - a composite measure which I refer to as unobserved skills. Using the estimated worker fixed effects as a measure of unobserved skills, I show that on average new firms employ less skilled workers relative to incumbent firms. Put differently, workers sort into new firms upon unobserved skills with more skilled individuals working in older firms. With a few exceptions, these findings hold across economic sectors, firms of different size groups, workers of different age and tenure.

The additional results reveal that sorting upon unobserved skills is also present among new firms. In particular, the results suggest that workers in incorporated new firms have on average higher skills than workers in non-incorporated new firms. Spin-offs are also found to employ more skilled workers relative to other new firms. In addition, workers with greater unobserved skills as measured by their fixed effects are more likely to work for establishments that are older and have more employees with high values of unobserved skill.

Finally, sorting appears to be important for post-entry performance of firms. In particular, I show that startups with greater average workers' unobserved skills face lower probability of exit. The measure of unobserved skills remains a significant predictor of firm survival even after controlling for workers' education, indicating that both measures capture variation in different aspects of worker characteristics. Put differently, the measure of workers' unobserved skills gives a more nuanced picture of the link between workers' characteristics and firm longevity.

These findings inform the discussion on policies aimed at facilitating the ability of younger firms to attract and retain workers. A recent example on the urgency of this discussion comes from Sweden, where in 2016 several Swedish startups publicly expressed stark concerns about their ability to hire competent individuals. Swedish government responded by introducing changes to the taxation of stock options - a tool often used by new firms to attract and retain skilled employees.[17] The results presented in this paper are in line with the notion that new firms struggle with attracting skilled individuals. Furthermore, given that workers' skills were also found to be systematically related to firm survival rates, inability to attract skilled individuals may be critical to post-entry firm performance. This suggests that policies aimed at spurring new firm creation by means of facilitating access to financial capital, may miss out another important type of asset - human capital.

The findings presented in this paper give rise to a number of questions for future research. To start with, what are the mechanisms behind the sorting pattern? There are several possible non-mutually exclusive explanations. First, the existence of sorting is in line with the argument that new firms might seem less appealing to more skilled workers. A few potential reasons that have been highlighted in earlier literature are the uncertainty surrounding the startups' ability to ensure employment along with stable salary and non-monetary benefits (Brown and Medoff, 2003; Heyman, 2007; Shane, 2009). The finding that new firms whose superior economic potential has been documented in the literature (incorporated firms and spin-offs) are more likely to employ more skilled workers indicates that sorting on unobserved skill is linked to firm performance. These "superior" new firms might be more appealing to the individuals from the right tail of human capital distribution already when they hire their first employees. Second, individuals with lower unobserved skills may join new ventures because they cannot obtain employment at older firms due to, for instance, complex hiring procedures. Third, the existing sorting patterns may be shaped by startups' labor demand. The findings presented in this paper accord with the idea that firms try to match new hires to the skills of their existing employees and, therefore, seek only a certain set of skills when hiring (Kremer, 1993). More skilled workers are more valuable for more efficient firms, whereas less efficient firms do not gain by hiring the most skilled employees. Similar arguments are presented in Dahl and Klepper (2015), who maintain that startups base their hiring choices on their assessment of their economic prospects. Hence, more promising startups target more skilled workers. At the same time, the sorting pattern is in line with the argument that entrants on average perform tasks that require lower skill content. All these potential explanations deserve further exploration.

Further research on employment practices of new firms is of great importance, given the substantial role of new firms in job creation (Haltiwanger et al., 2013) and turbulent economic context that seems to favor larger and older businesses more than smaller and younger ones (Decker et al., 2014; Ayyagari and Maksimovic, 2017). As such, better understanding of how individuals match with firms has a potential to influence both policy-making and business practice.

[17] Henrekson (2005) argues that taxation of stock options in Sweden has been discouraging new ventures from entry. Stock options is a way of attracting employees to new ventures. That stock options in Sweden were taxed as income implies that gains from accepting the options were low relative to other countries, such as the US.

References

Abowd, J. M., Creecy, R. H., Kramarz, F., et al. (2002). Computing person and firm effects using linked longitudinal employer-employee data. Technical report, Center for Economic Studies, US Census Bureau.

Abowd, J. M., Haltiwanger, J., Jarmin, R., Lane, J., Lengermann, P., McCue, K., McKinney, K., and Sandusky, K. (2005). The relation among human capital, productivity, and market value: Building up from micro evidence. In *Measuring capital in the new economy*, pages 153--204. University of Chicago Press.

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2):251--333.

Ahlin, L., Andersson, M., and Thulin, P. (2014). Market thickness and the early labour market career of university graduates: an urban advantage? *Spatial Economic Analysis*, 9(4):396--419.

Akerman, A., Helpman, E., Itskhoki, O., Muendler, M.-A., and Redding, S. (2013). Sources of wage inequality. *American Economic Review*, 103(3):214--19.

Allison, P. D. (1982). Discrete-time methods for the analysis of event histories. *Sociological methodology*, 13:61--98.

Andersson, M. and Klepper, S. (2013). Characteristics and performance of new firms and spinoffs in sweden. *Industrial and corporate change*, 22(1):245--280.

Antelius, J. and Björklund, A. (2000). How reliable are register data for studies of the return on schooling? an examination of swedish data. *Scandinavian Journal of Educational Research*, 44(4):341--355.

Åstebro, T. and Bernhardt, I. (2005). The winner's curse of human capital. *Small Business Economics*, 24(1):63--78.

Audretsch, D. B. and Mahmood, T. (1994). The rate of hazard confronting new firms and plants in us manufacturing. *Review of Industrial organization*, 9(1):41--56.

Audretsch, D. B. and Mahmood, T. (1995). New firm survival: new results using a hazard function. *The Review of Economics and Statistics*, pages 97--103.

Ayyagari, M. and Maksimovic, V. (2017). Fewer and less skilled? human capital, competition, and entrepreneurial success in manufacturing.

Balsvik, R. (2011). Is labor mobility a channel for spillovers from multinationals? evidence from norwegian manufacturing. *The review of economics and statistics*, 93(1):285--297.

Baptista, R., Karaöz, M., and Mendonça, J. (2014). The impact of human capital on the early success of necessity versus opportunity-based entrepreneurs. *Small Business Economics*, 42(4):831--847.

Barth, E., Davis, J., and Freeman, R. B. (2018). Augmenting the human capital earnings equation with measures of where people work. *Journal of Labor Economics*, 36(S1):S71--S97.

Bates, T. (1990). Entrepreneur human capital inputs and small business longevity. *The review of Economics and Statistics*, pages 551--559.

Bender, S., Bloom, N., Card, D., Van Reenen, J., and Wolter, S. (2018). Management practices, workforce selection, and productivity. *Journal of Labor Economics*, 36(S1):S371--S409.

Bonhomme, S., Lamadon, T., Manresa, E., et al. (2015). A distributional framework for matched employer employee data. *Unpublished manuscript, University of Chicago*.

Brown, C. and Medoff, J. L. (2003). Firm age and wages. *Journal of Labor Economics*, 21(3):677--697.

Browning, M. and Heinesen, E. (2012). Effect of job loss due to plant closure on mortality and hospitalization. *Journal of health economics*, 31(4):599--616.

Brüderl, J., Preisendörfer, P., and Ziegler, R. (1992). Survival chances of newly founded business organizations. *American sociological review*, pages 227--242.

Bublitz, E., Nielsen, K., Noseleit, F., and Timmermans, B. (2017). Entrepreneurship, human capital, and labor demand: a story of signaling and matching. *Industrial and Corporate Change*, page dtx027.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge university press.

Card, D. and Cardoso, A. R. (2012). Can compulsory military service raise civilian wages? evidence from the peacetime draft in portugal. *American Economic Journal: Applied Economics*, 4(4):57--93.

Card, D., Cardoso, A. R., Heining, J., and Kline, P. (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, 36(S1):S13--S70.

Card, D., Cardoso, A. R., and Kline, P. (2015). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics*, 131(2):633--686.

Card, D., Heining, J., and Kline, P. (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics*, 128(3):967--1015.

Carpenter, R. E. and Petersen, B. C. (2002a). Capital market imperfections, high-tech investment, and new equity financing. *The Economic Journal*, 112(477).

Carpenter, R. E. and Petersen, B. C. (2002b). Is the growth of small firms constrained by internal finance? *Review of Economics and statistics*, 84(2):298--309.

Caves, R. E. (1998). Industrial organization and new findings on the turnover and mobility of firms. *Journal of economic literature*, 36(4):1947--1982.

Colombo, M. G. and Grilli, L. (2005). Founders' human capital and the growth of new technology-based firms: A competence-based view. *Research policy*, 34(6):795--816.

Colombo, M. G. and Grilli, L. (2010). On growth drivers of high-tech start-ups: Exploring the role of founders' human capital and venture capital. *Journal of business venturing*, 25(6):610--626.

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D., and Roux, S. (2012a). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, 80(6):2543--2594.

Combes, P.-P., Duranton, G., Gobillon, L., and Roux, S. (2012b). Sorting and local wage and skill distributions in france. *Regional Science and Urban Economics*, 42(6):913--930.

Dahl, M. S. and Klepper, S. (2015). Whom do new firms hire? *Industrial and corporate change*, 24(4):819--836.

Dal Bó, E., Finan, F., Folke, O., Persson, T., and Rickne, J. (2017). Who becomes a politician? *The Quarterly Journal of Economics*, 132(4):1877--1914.

Decker, R., Haltiwanger, J., Jarmin, R., and Miranda, J. (2014). The role of entrepreneurship in us job creation and economic dynamism. *The Journal of Economic Perspectives*, 28(3):3--24.

Dunne, T., Roberts, M. J., and Samuelson, L. (1988). Patterns of firm entry and exit in us manufacturing industries. *The RAND journal of Economics*, pages 495--515.

Eeckhout, J. and Kircher, P. (2011). Identifying sorting---in theory. *The Review of Economic Studies*, 78(3):872--906.

Eeckhout, J., Pinheiro, R., and Schmidheiny, K. (2014). Spatial sorting. *Journal of Political Economy*, 122(3):554--620.

Eliason, M. (2011). Income after job loss: the role of the family and the welfare state. *Applied Economics*, 43(5):603--618.

Eliason, M. and Storrie, D. (2006). Lasting or latent scars? swedish evidence on the long-term effects of job displacement. *Journal of Labor Economics*, 24(4):831--856.

Eriksson, T. and Kuhn, J. M. (2006). Firm spin-offs in denmark 1981--2000---patterns of entry and exit. *International Journal of Industrial Organization*, 24(5):1021--1040.

Evans, D. S. (1987). The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries. *The journal of industrial economics*, pages 567--581.

Evans, D. S. and Leighton, L. S. (1989). Some empirical aspects of entrepreneurship. *The American Economic Review*, 79(3):519--535.

Fox, J. T. and Smeets, V. (2011). Does input quality drive measured differences in firm productivity? *International Economic Review*, 52(4):961--989.

Geroski, P. A. (1995). What do we know about entry? *International Journal of Industrial Organization*, 13(4):421--440.

Grogger, J. and Eide, E. (1995). Changes in college skills and the rise in the college wage premium. *Journal of Human Resources*, pages 280--310.

Guzman, J. and Stern, S. (2015). Where is silicon valley? *Science*, 347(6222):606--609.

Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2013). Who creates jobs? small versus large versus young. *Review of Economics and Statistics*, 95(2):347--361.

Henrekson, M. (2005). Entrepreneurship: a weak link in the welfare state? *Industrial and Corporate change*, 14(3):437--467.

Heyman, F. (2007). Firm size or firm age? the effect on wages using matched employer--employee data. *Labour*, 21(2):237--263.

Hijzen, A., Upward, R., and Wright, P. W. (2010). The income losses of displaced workers. *Journal of Human Resources*, 45(1):243--269.

Hurst, E. and Pugsley, B. W. (2011). What do small businesses do? *Brookings Papers on Economic Activity*, pages 73--143.

Irarrazabal, A., Moxnes, A., and Ulltveit-Moe, K. H. (2013). Heterogeneous firms or heterogeneous workers? implications for exporter premiums and the gains from trade. *Review of Economics and Statistics*, 95(3):839--849.

Kogut, B. and Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization science*, 3(3):383--397.

Kondo, K. et al. (2017). Estquant: Stata module to implement quantile approach by combes et al. *Statistical Software Components*.

Kremer, M. (1993). The o-ring theory of economic development. *The Quarterly Journal of Economics*, 108(3):551--575.

La Porta, R. and Shleifer, A. (2014). Informality and development. *Journal of Economic Perspectives*, 28(3):109--26.

Lazear, E. P. and Shaw, K. L. (2018). Introduction: Firms and the distribution of income: The roles of productivity and luck. *Journal of Labor Economics*, 36(S1):S1--S12.

Levine, R. and Rubinstein, Y. (2017). Smart and illicit: who becomes an entrepreneur and do they earn more? *The Quarterly Journal of Economics*, 132(2):963--1018.

Macis, M. and Schivardi, F. (2016). Exports and wages: Rent sharing, workforce composition, or returns to skills? *Journal of Labor Economics*, 34(4):945--978.

Mata, J. and Portugal, P. (1994). Life duration of new firms. *The journal of industrial economics*, pages 227--245.

Miller, S. R. (1998). Shortcut: High school grades as a signal of human capital. *Educational Evaluation and Policy Analysis*, 20(4):299--311.

Nyström, K. and Elvung, G. Z. (2014). New firms and labor market entrants: Is there a wage penalty for employment in new firms? *Small Business Economics*, 43(2):399--410.

Ouimet, P. and Zarutskie, R. (2014). Who works for startups? the relation between firm age, employee age, and growth. *Journal of financial Economics*, 112(3):386--407.

Roca, J. D. L. and Puga, D. (2017). Learning by working in big cities. *The Review of Economic Studies*, 84(1):106--142.

Sauermann, H. (2017). Fire in the belly? employee motives and innovative performance in start-ups versus established firms. *Strategic Entrepreneurship Journal*.

Shane, S. (2009). Why encouraging more people to become entrepreneurs is bad public policy. *Small business economics*, 33(2):141--149.

Thompson, P. (2005). Selection and firm survival: evidence from the shipbuilding industry, 1825--1914. *The review of economics and statistics*, 87(1):26--36.

Topel, R. H. and Ward, M. P. (1992). Job mobility and the careers of young men. *The Quarterly Journal of Economics*, 107(2):439--479.

Unger, J. M., Rauch, A., Frese, M., and Rosenbusch, N. (2011). Human capital and entrepreneurial success: A meta-analytical review. *Journal of business venturing*, 26(3):341--358.

Appendix

A Exogenous mobility

Endogenous mobility and its consequences
OLS estimates of coefficients in AKM decomposition are biased if worker mobility is correlated with the residual component of earnings. Put differently, AKM coefficients are biased if there is sorting into firms based on the value of the worker-firm match component. To facilitate the discussion of the plausibility of the assumption, Card et al. (2013) decompose ν_{it} into a match component, drift component and the transitory error term ($\nu_{it} = \eta_{i\boldsymbol{J}_{(i,t)}} + \zeta_{it} + \varepsilon_{it}$). For OLS estimates to be biased it suffices for any of the components to be correlated with the employment history. There are in particular three forms of endogenous mobility that could violate the exogenous mobility assumption. The first form of endogenous mobility is sorting based on the match component $\eta_{i\boldsymbol{J}_{(i,t)}}$, that is workers receive different wage premiums at the same employer depending on the quality of the match. In presence of such sorting, wage gains for workers moving from one establishment to another should be quite different in absolute terms from the wage losses of workers moving in the opposite direction. If, however, there is no sorting on match component, gains and losses should be symmetric. The second form of endogenous mobility is mobility based on the drift component ζ_{it}, where workers' abilities are revealed to the employer over time and induce movement to another firm. Workers who displayed higher productivity are expected to get a raise at the current employer and be more likely to move to higher-wage firms, whereas workers who displayed lower productivity are expected to experience a decline in current wages and become more likely to move to low-wage firms. If this type of mobility systematically occurs in the data, we should observe systematic upward trends in wages of movers to the higher-wage firms before the move and downward trends to those moving to lower-wage firms. The firm fixed effects will be overstated due to the positive correlation between the drift component and the change of the employer. Finally, the third form of endogenous mobility arises if the transitory shock ε_{it} is correlated to the mobility history. This means that workers' mobility is related to the transitory shocks experienced by firms (e.g. workers move from firms experiencing negative shocks to those experiencing positive shocks).

Event-study analysis
To test for these forms of sorting I closely follow Card et al. (2013, 2015) and compare wages of moving workers. The idea is to compare wage development of movers from different wage quartiles. To compare wages of moving workers, I assign each worker into four quartiles based on the estimated firm fixed effects. I then restrict the analysis to movers who have been employed both in the origin firm (before the move) and in the destination firm (after the move) for at least two years. Since I do not observe the date when an employee moves and, therefore, cannot assign yearly earnings to two different employers, I exclude the year when move occurred from the analysis. I classify the move in one of sixteen groups defined in accordance with the combination of the wage quartiles of the origin and the destination firms (movers from quartile 1 to quartile 1, movers from quartile 1 to quartile 2, etc.). Figure A.1 illustrates the results. For clarity, Panel A illustrates the wage profiles for movers from the first and the fourth quartiles, in Panel B

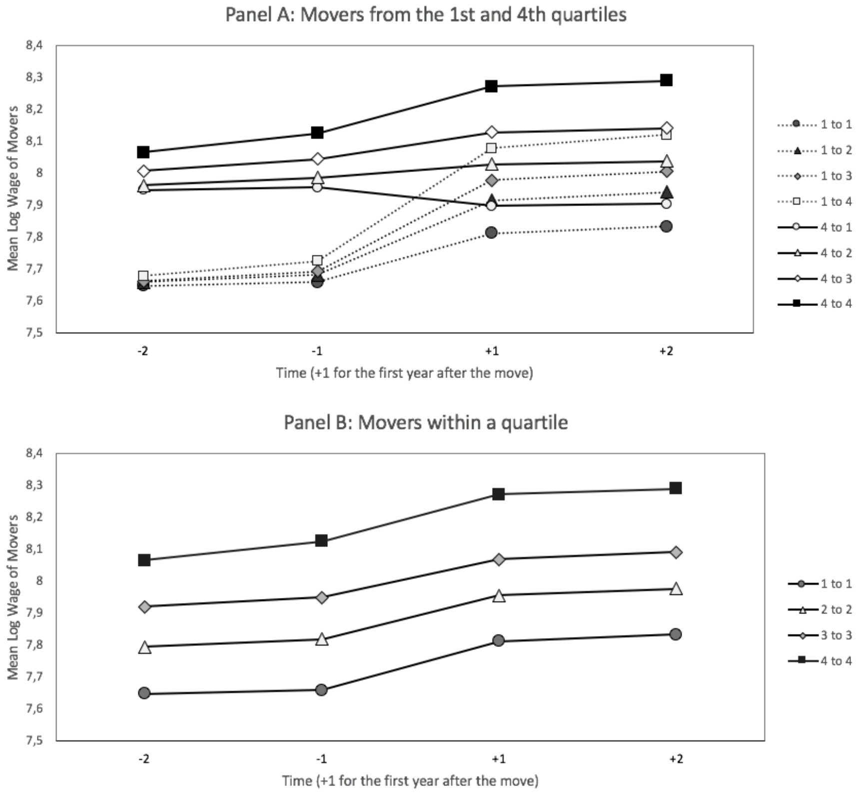

Figure A.1: Earnings of job changers by direction of move between the earning quartiles

the wage profiles for movers who stay within the same wage quartile are illustrated. All the directions of the move are presented in Table A.1. In general, the figures resemble those obtained in Card et al. (2013, 2015); Macis and Schivardi (2016).

Evidence from wage profiles of movers

The analysis reveals the following. First of all, gains and losses from moving between the quartiles are approximately symmetric. Recall that AKM implies that firms offer proportional wage premiums to their employees. Hence, workers moving from firms paying lower wages to firms with higher wages should experience wage rises, while those moving in the opposite direction should experience wage losses in earnings. Moreover, these losses and benefits should be symmetric. Workers who move to firms that pay similar wage premiums should not experience substantial wage changes. Indeed, wage profiles for movers within the same quartile are close to being parallel to each other without any jumps. Overall, the symmetry in wage profiles is at odds with the models that link firm-worker sorting to the match component of wages (Eeckhout and Kircher, 2011) and predict that wages should increase irrespective of the direction of the move. Importantly, this symmetry is at odds with the sorting on match component. Secondly, although wage profiles are overall increasing over time, there are no substantial drops or rises in wages before the move, as it would be expected in the presence of the mobility based on the drift component. Finally, absence of substantial trends to wage profiles before the move also suggests that there is little evidence that transitory wage fluctuations predict workers' mobility, which in turn means that any correlation between ε_{it} and mobility history is small.

Additive separability

Table A.1: Earnings of job changers by direction of move between the earning quartiles, all directions of move

(1) Quartiles	(2) N. obs.	(3) t-2	(4) t-1	(5) t+1	(6) t+2
1 to 1	19974	7.647	7.659	7.812	7.833
1 to 2	13366	7.660	7.683	7.916	7.942
1 to 3	7223	7.663	7.692	7.979	8.006
1 to 4	5448	7.678	7.725	8.078	8.121
2 to 1	11825	7.770	7.781	7.849	7.862
2 to 2	17098	7.794	7.819	7.956	7.976
2 to 3	13742	7.829	7.857	8.041	8.063
2 to 4	9129	7.838	7.888	8.144	8.175
3 to 1	6016	7.864	7.871	7.889	7.901
3 to 2	13979	7.871	7.893	7.981	7.996
3 to 3	18131	7.921	7.950	8.069	8.092
3 to 4	15379	7.966	8.020	8.206	8.231
4 to 1	3494	7.948	7.956	7.898	7.904
4 to 2	7343	7.962	7.986	8.028	8.038
4 to 3	14531	8.007	8.044	8.129	8.141
4 to 4	33487	8.065	8.126	8.272	8.290

Direction of the move in the first column: from origin to destination quartile. Quartiles based on the estimated firm fixed effects. Number of observations in column 2. Columns 3-6 report average log real earnings for job movers before and after the move, with move taking place at time *t*. Sample limited to job movers observed for at least 2 years before a job change and 2 years after.

Another concern with the AKM is the additive separability assumption, i.e. AKM assumes that firm wage premium (or discount) is common for all its employees. This assumption has also received empirical support in various economic settings. Most importantly for the context of this paper, Bonhomme et al. (2015) show that additive model exhibits a very good approximation for Swedish data. In particular, they find that allowing for interactions between employee and employer effects yields only a negligible improvement in the explained wage variance. Similar conclusions about the fit of the additive separability models have been reached for Germany (Card et al., 2013), Portugal (Card et al., 2015), and Italy (Macis and Schivardi, 2016).

Overall, I conclude that the worker fixed effects estimated using AKM decomposition can be considered as reasonable measures of the workers' unobserved skills.

B Robustness of main results

This sections shows that the main result that employees of new firms on average have lower unobserved skills holds if alternative decisions are made when constructing the data set and defining variables. The purpose of this exercise is to see whether the estimated direction of sorting is robust to the above-mentioned alterations to empirical strategy. In particular, in Table B.2 Row 1, I replicate main estimation results corresponding to Table 2, rows 1 if the threshold for approximating full-time workers is 125,000 SEK instead of 100,000 in 2000 prices. In Row 2 the threshold was chosen as 150,000 SEK. In Row 3 the data set is split into pre-sample and main sample in year 2008 (pre-sample 1993-2008 and main sample 2009-2010 instead of 1993-2005 and 2006-2010). Row 4 displays the results if new firms are defined as firms not older than 5 years instead of 3 as in the main analysis. In Row 5 incumbents are defined as firms with at least 6 years on the market, instead of 8. Rows 6 and 7 show the results when the main analysis is performed for years 2007 and 2008 instead of 2006.

It is evident from Table B.2 that these alterations do not change the main conclusion: workers sort into new firms upon unobserved skills with less skilled workers being overrepresented in new firms.

C Duration analysis

In this section I report results of duration analysis relating the probability that a new firm exits to the unobserved skills of its workers. While the analysis reported in the main part (Section 4) might be more intuitive for its simplicity, the results reported here are meant to complement the findings.

To examine whether longevity of new firms increases with workers' skills, I analyze survival patterns of new firms as a function of firm-level covariates. Since the duration in the data set is measured annually, I use a discrete-time model. It is common to assume that in the discrete-time models the hazard rate is described by the logistic function (Allison, 1982). The underlying logistic discrete-time hazard function is as follows:

$$P_{ft} = 1/[1 + exp(-c(t) - \theta' z_{ft})] \tag{3}$$

where P_{ft} is a probability that firm f exits at time t, *c(t)* is the baseline hazard function which summarizes the pattern of duration dependence, and z_{ft} is a vector of firm-level variables. It includes the mean of workers' skills, mean of workers' education, firm size, mean age of employees and its square, and share of females. Eq. (3) can be rewritten in logit form:

$$\log[P_{ft}/(1 - P_{ft})] = c(t) + \theta' z_{ft} \tag{4}$$

To estimate Eq. (4) I specify the baseline hazard *c(t)* as a piecewise constant function. This specification implies that the risk of exit varies across the defined intervals, but is constant within these intervals. I choose two intervals corresponding to the firms' age: 1-2 and 3-4 years.

A known weakness of duration models is the sensitivity of the estimated parameters to the unobserved heterogeneity. The unobserved heterogeneity introduces bias to the estimates even if the heterogeneity is independent of the covariates, in contrast to linear regression where it is simply absorbed into the error term (Cameron and Trivedi, 2005, p.611). A standard way of incorporating unobserved heterogeneity in duration models is to add to the model a unit-specific random effect ν with mean unity and finite variance. This yeilds a so-called frailty model:

$$\log[P_{ft}/(1 - P_{ft})] = c(t) + \theta' z_{ft} + \log(\nu_f) \tag{5}$$

where $\log(\nu_f)$ is assumed to be normally distributed.

C.1 Estimation results

The estimation results are reported in Table C.3. The estimates in the table are marginal effects at mean values. The dependent variable is a binary variable taking value zero when a firm continues operating and one when it exits. The reason for exit is recorded in the data set, allowing to single out exits due to mergers and acquisitions. Such exits are censored. The workers' skills are measured by the mean of the worker fixed effects. A set of explanatory variables is included to account for firm level characteristics that might have impact on firm survival. Controls include mean education of the employees (logged), current firm size as measured by the number of employees (logged), share of females, year and industry dummies. I also include indicator variables for whether an entrant is a spin-off and whether an entrant is incorporated, since these variables are expected to be important determinants of firm survival. The specifications include the firm age intervals as specified by the piecewise constant function.

Similar to the main analysis, probability of exit if decreasing in mean unobserved skills of employees. In Column 1 the average marginal effect of mean workers' skills is -0.025. The estimated marginal effect is -0.019 in Column 2 where unobserved heterogeneity is allowed for. The likelihood ratio test suggests statistically significant frailty, indicating that frailty model is better fit for data. Overall, estimates suggest that the probability of exit declines with workers' skills

D Tables and graphs

Table B.2: Worker fixed effects in new and incumbent firms

	$\hat{A}$	$\hat{D}$	$\hat{S}$	R^2	Obs.1	Obs.2
Trim at wage ≤ 125	-0.0241***	1.0543***	0.001***	0.882	1495375	19227
	(0.003)	(0.007)	(0.000)			
Trim at wage ≤ 150	-0.0193***	1.0608***	0.000***	0.819	1443685	17469
	(0.003)	(0.008)	(0.000)			
Split sample at 2008	-0.0367***	1.047***	0.001***	0.930	1649889	23519
	(0.002)	(0.006)	(0.000)			
Define new firms ≤ 5y	-0.0238***	1.049***	0.001***	0.913	1538732	34272
	(0.002)	(0.005)	(0.000)			
Define incumbents ≥ 6y	-0.0286***	1.048***	0.001***	0.923	1555320	20876
	(0.002)	(0.006)	(0.000)			
Snapshot 2007	-0.0367***	1.058***	0.002***	0.950	1510973	22323
	(0.002)	(0.006)	(0.000)			
Snapshot 2008	-0.0365***	1.0607***	0.002***	0.944	1486241	22395
	(0.002)	(0.002)	(0.002)			

The estimated parameters are shift $\hat{A}$, dilation $\hat{D}$, and truncation $\hat{S}$. Obs. 1 and Obs. 2 denote number of observations in the first and the second group of firms. Bootstrapped standard errors in parentheses (100 iterations). ***p<0.01, **p<0.5, *p<0.1.

Table C.3: Duration analysis

	(1) Logit	(2) Frailty
Mean $\hat{\alpha}$	-0.025***	-0.019***
	(0.004)	(0.003)
Firm size	-0.036***	-0.026***
	(0.002)	(0.002)
Mean educ	-0.002*	-0.001**
	(0.001)	(0.001)
Share females	-0.010**	-0.007**
	(0.004)	(0.003)
Spinoff	-0.012***	-0.009***
	(0.003)	(0.002)
Incorporated	-0.021***	-0.017***
	(0.003)	(0.002)
Firm age 1-2	0.034***	0.005
	(0.007)	(0.006)
Firm age 3-4	0.021**	0.007
	(0.007)	(0.005)
N	50958	50958
pseudo-R^2	0.03	
Log-likelihood	-13842.24	-13003.23
Rho		0.38
LR test of rho=0		0.00

Notes: All specifications control for year and industry effects. ***, **, and * correspond to significance level at 1, 5, and 10 percent. Standard errors in the logistic model are clustered at the firm level, robust standard errors in the frailty model. Rho is the ratio between heterogeneity variance and one plus the heterogeneity variance. ***p<0.01, **p<0.5, *p<0.1.

Table D.1: Determinants of sorting into firms

	$\hat{\alpha}$	$\hat{\alpha}$	$\hat{\alpha}$
Coworkers' $\hat{\alpha}$			0.546***
			(0.005)
Coworkers' educ		0.040***	0.022***
		(0.001)	(0.000)
Firm age	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)
Firm size	-0.029***	-0.025***	-0.013***
	(0.002)	(0.002)	(0.001)
Firm size, sq.	0.004***	0.003***	0.001***
	(0.000)	(0.000)	(0.000)
Share females	-0.239***	-0.262***	-0.146***
	(0.003)	(0.003)	(0.002)
Constant	0.016*	-0.450***	-0.288***
	(0.006)	(0.018)	(0.012)
N	7575614	7203316	7203316
Adj.R^2	0.08	0.09	0.13

Notes: Dependent variable is worker fixed effects ($\hat{\alpha}$), it is centered around zero. All specifications include industry and year indicator variables. Estimation covers the period from 2002 to 2010, only firms whose entry is observed in the dataset are included. Firm size is log of number of employees. Standard errors are clustered at the firm level. ***p<0.01, **p<0.5, *p<0.1.

Table D.2: Firm characteristics for entrants in 2006

	Mean	s.d.
Mean $\hat{\alpha}$	-0.04	0.32
Firm size	0.63	0.52
Mean education	2.46	0.15
Share females	0.25	0.37
Exit	0.34	0.47
No. observations	4045	

Notes: Sample includes firms that entered in 2006.

Table D.3: Probability of exit of new firms

	(1) Exit	(2) Exit	(3) Exit
Median $\hat{\alpha}$	-0.083**		
	(0.027)		
75th perc $\hat{\alpha}$		-0.079**	
		(0.025)	
Max $\hat{\alpha}$			-0.081**
			(0.025)
Mean Educ	-0.009	-0.009	-0.009
	(0.005)	(0.005)	(0.005)
Firm size	-0.056***	-0.047**	-0.045**
	(0.015)	(0.015)	(0.015)
Observations	4045	4045	4045

Notes: Standard errors are clustered at the firm level. ***p<0.01, **p<0.5, *p<0.1.