DISCUSSION PAPER SERIES

IZA DP No. 9585

Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility

Peter Fredriksson
Lena Hensvik
Oskar Nordström Skans

December 2015

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility

Peter Fredriksson
Stockholm University, UCLS, IZA and IFAU

Lena Hensvik
IFAU, UCLS and CESifo

Oskar Nordström Skans
Uppsala University, UCLS, IFAU and IZA

Discussion Paper No. 9585
December 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 9585
December 2015

ABSTRACT

Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility*

We examine the direct impact of idiosyncratic match quality on entry wages and job mobility using unique data on worker talents matched to job-indicators and individual wages. Tenured workers are clustered in jobs with high job-specific returns to their types of talents. We therefore measure mismatch by how well the types of talents of recent hires correspond to the talents of tenured workers performing the same jobs. A stylized model shows that match quality has a smaller impact on entry wages but a larger impact on separations and future wage growth if matches are formed under limited information. Empirically, we find such patterns for inexperienced workers and workers who were hired from non-employment, which are also groups where mismatch is more pronounced on average. Most learning about job-specific mismatch happens within a year. Experienced job-to-job movers appear to match under much less uncertainty. They are better matched on entry and mismatch have a smaller effect on their initial separation rates and later wage growth. Instead, match quality is priced into their starting wages.

JEL Classification: J64, J24, J31, J62

Keywords: matching, job search, comparative advantage, employer learning

Corresponding author:

Peter Fredriksson
Department of Economics
Stockholm University
SE-106 91 Stockholm
Sweden
E-mail: peter.fredriksson@ne.su.se

* First version 2 July 2014. We thank Erling Barth, David Card, John Earle, Jan Eeckhout, Francis Kramarz, Lisa Laun, Rasmus Lentz, Edwin Leuven, Erik Lindqvist, Andreas Madestam, Espen Moen, Dale Mortensen, Arash Nekoei, Åsa Rosén, Robert Shimer and Uta Schönberg as well as seminar and workshop participants at Aarhus University, BI (Oslo), DIW (Berlin), EALE 2014, LSE, Queen Mary University, Royal Holloway, SOLE 2014, Stockholm University, UCLS, University of Oslo, University of Texas at Austin, Nordic Summer Labor Institute, University of Warwick, and VATT for useful comments and suggestions. Peter Fredriksson acknowledges funding from Marcus and Amalia Wallenberg Foundation and Handelsbanken. Lena Hensvik acknowledges funding from Ragnar Söderberg Foundation and Handelsbanken. Oskar Nordström Skans acknowledges funding from Ragnar Söderberg Foundation.

1 Introduction

A longstanding notion within labor economics is that the allocation of workers across jobs is crucial for wage dispersion, labor productivity and overall efficiency.[1] Idiosyncratic match quality is also fundamental to several recent theoretical contributions on a wide set of topics, including the work by Marimon and Zilibotti (1999) on the implications for designing unemployment insurance, Eeckhout and Kircher (2011) on the possibility of identifying sorting from wage data, Gautier et al. (2010) on the interactions between comparative advantage and search frictions, and Helpman et al. (2010) on the impact of trade liberalization on wage inequality and unemployment.

Although match quality is conceptually well-defined, deriving direct, and credible, evidence on the importance of mismatch in the labor market has proven difficult. Much of previous micro-empirical work has been based on the Jovanovic (1979) model where match quality is unobserved at the time of hiring, but realized ex post. The typical approach has been to analyze how exits and wages evolve with tenure (e.g., Abraham and Farber, 1987, Flinn, 1986 and Farber, 1999). A drawback of this strategy is that, e.g., accumulation of firm-specific human capital has the same implications for the associations between wages/separations and tenure as revelations about match quality.[2]

We proceed differently. We use very detailed pre-hire data to assess if separations and entry wages respond to a direct measure of match quality. Empirically, we derive a measure of the distance to optimal match quality, i.e. of the extent of *mismatch*. The calculation of our mismatch index uses data on cognitive abilities and personality traits at age 18. These data include a vector of eight productive "talents": four cognitive skills (inductive, verbal, spatial, and technical ability) as well as four traits evaluated by a trained psychologists (social maturity, intensity, psychological energy and emotional stability). We show that the talents are independently valued at the labor market. Our basic presumption is that each of these particular talents are differentially productive in different jobs, as in Lazear (2009). To corroborate this presumption, we show that (tenured) workers are sorted into jobs with high estimated job-specific returns to their types of skills. Thus, there is sorting into jobs by comparative advantage, as in the Roy

[1]See, e.g., the Roy (1951) model on sorting and wages and the assignment models of Tinbergen (1956) and Sattinger (1975) where the problem of assigning heterogeneous workers to heterogeneous jobs is analyzed. In these (frictionless) models, market prices allocate workers to jobs. A more recent literature combines search frictions and worker/job heterogeneity. Gautier and Teulings (2012) calibrate such a model, and conclude that actual allocations imply very large efficiency losses.

[2]Dustmann et al. (2015) are able to circumvent this problem by contrasting hires through referrals and hires through formal channels. The idea is that there is more information about workers who have obtained their job through a referral. Therefore, their entry wages will be higher and they are less likely to leave the firm. Also, Nagypal (2007) presents an interesting attempt to distinguish explanations based on information about match quality from learning-by-doing, using a detailed structural model. She concludes that most of the variation at longer time horizons is due to learning about match quality. Her identification is based on the assumption that, absent learning-by-doing or learning about match quality, firm-level shocks affect low- and high-tenure workers symmetrically.

(1951)-model.

A key implication of such a selection process is that talents among tenured workers should reflect the skill requirements of each particular job. By combining detailed data on entering workers' talents with equally detailed data on the talents of tenured workers who perform the same job, we are able to infer the quality of new matches from ability indicators that are determined before the match is formed (and, hence, not accumulated on the job).[3]

We study data on recent hires in regressions with job (detailed occupation by establishment and year) fixed effects. This implies that we analyze the impact of variations in match quality between different entrants who start the same job, during the same year. Our models also account for the overall market valuation of entrants' talents and educational attainment. Our main strategy assumes that potential productive benefits of employing a diverse set of personality types primarily occur across (and not within) jobs. The results support this notion. Furthermore, our main results hold also when removing jobs that appear to make use of a more diverse set of personality types.

To frame the empirical analysis, we set up a stylized model where mismatch is differentially observable at the time of hire. If mismatch is partially observed, it should be priced into entry wages, and separations respond only to revelations of mismatch (i.e. mismatch in excess of what was expected at the hiring stage). Entry wages should, on the other hand, be unrelated to mismatch if it is unobserved at the time of recruitment, in which case subsequent separations instead should respond more forcefully. As information is revealed when production commences, wage growth (within jobs) should be more strongly related to match quality if it was unobserved at the time of hire. The amount of initial information available to the matching agents is thus key for both the wage and mobility responses.

Realistically, the available information varies with the characteristics of the match. We use two approaches to implement this idea. The first approach draws on the employer learning literature pioneered by Farber and Gibbons (1996) and Altonji and Pierret (2001). We argue that labor market experience proxies the amount of information available on both sides of the labor market. In particular, for inexperienced workers, it is realistic to assume that the agents fail to observe how well the detailed characteristics of the worker match the detailed skill requirements for each particular job. Note that this assumption is valid even if both sides of the market are able to infer the market value of the opposing

[3]In independent and contemporary work, Guvenen et al. (2015) use the NLSY to examine how (measured) occupational mismatch affects the wage trajectories of workers. Unlike Guvenen et al. (2015) we: (i) focus on job mismatch, and we also show that, for most outcomes, job mismatch is quantitatively more important than occupational mismatch; (ii) focus on new hires (in contrast Guvenen et al. look at existing relationships and, thus, have to deal with the endogeneity of tenure and mismatch), and we also show that for inexperienced new hires mismatch is truly an idiosyncratic shock; (iii) are able to deal with potential confounding factors to a greater extent because of the sheer size of the administrative data (we examine more than 150,000 new hires, whereas Guvenen et al. analyze the careers of 1700 individuals).

agent. The second approach compares workers who are hired from non-employment with workers who are hired from another job. We expect there to be less information available about match quality for those who enter from non-employment.

The results suggest that mismatch matters. In line with the predictions of our stylized model, we find that mismatch is more pronounced among the inexperienced and hires from non-employment, conditional on job specific fixed effects and detailed controls for individual skills. Furthermore, mismatch is unrelated to entry wages among inexperienced workers and workers who have entered from non-employment. In contrast, experienced workers and job-to-job movers receive a wage penalty if they are mismatched. Consequently, we find a pronounced separation response to mismatch among inexperienced workers and entrants from non-employment, whereas the separation response among experienced workers and job-to-job movers is moderate. We also show that wage growth within job is negatively affected by mismatch and that this effect is particularly pronounced among inexperienced workers. We validate the results by using an alternative measure of mismatch built on job-specific estimates of tenured workers' wage returns to each of our eight talents. These results are less precise, but well aligned with those of our main strategy. We also show that most of the separation response appears within the first year, with a peak at six months, suggesting that learning about match quality is a fairly rapid process. In order to estimate the overall consequences of mismatch we study the impact on subsequent annual earnings. The average inexperienced entrant is estimated to lose 13 percent of initial earnings due to mismatch of talents (relative to the optimal match). The effects primarily occur due to an increased frequency of job loss. Earnings trajectories converge over time and the consequences of mismatch disappear after 4-5 years.

Overall, we conclude that the search for match-quality has a substantial impact on the labor market outcomes for both experienced and inexperienced workers. We interpret the evidence as supportive of the notion that workers with little experience and those who search from non-employment form matches with considerable remaining uncertainty, whereas the matches between jobs and experienced workers or job-to-job movers are best characterized by models where information about match quality is revealed before the match is formed.

The paper is structured as follows: Section 2 derives a set of key micro-level predictions from a stylized theoretical model. Section 3 describes the data. Section 4 presents the main empirical results. Section 5 presents a large set of robustness tests. Section 6 studies earnings trajectories and Section 7 concludes.

2 Framework

We set up a stylized matching model incorporating match-specific productivity. The basic structure of the model is the following. Consider a (potential) match between a worker and a firm. This match is characterized by a some level of mismatch (d). The worker-firm pair initially observes a noisy signal of match-specific productivity (d_0). On the basis of the signal, the two parties decide on whether to match, and, if so, agree on an entry wage. As production starts, the worker-firm pair observes productivity. Based on this new (and more precise) information, wages and employment decisions (i.e. separations) are adjusted to reflect current information on match quality.

We use the stylized model to formalize a set of micro-level intuitive predictions regarding the responses of wages and separations to mismatch, under different information scenarios. These predictions are then taken to the data in the remainder of the paper.

Production We assume a constant returns to scale to technology and thus focus on one job. Each worker has a bundle of different skills $s_k(i)$, $k = 1, ..., K$. Productivity depends on how well these skills match with the technology (skill requirement) of the specific job. We measure the relationship between the skills and the technology by the location of the job and the worker in K-dimensional space. Let $d_k(i, j) = |s_k(i) - s_k(j)|$ denote the distance between the location of the worker and the job along the kth dimension and $d = d(i, j)$ the aggregate distance between the worker and the job (we make the empirical measure precise later on).

We take match productivity, $y(d) = y(i, j)$, to be given by

$$y(d) = 1 - \gamma d(i, j) + \theta s(i) + \lambda(j) \tag{1}$$

where $s(i)$ denotes a vector of worker skills, $\lambda(j)$ the quality of the job, and $\gamma > 0$ reflects the substitutability between different skills for a particular job (see Teulings and Gautier 2004). Match productivity is decreasing in the distance between the worker and the job, and thus maximal when $d \to 0$. We let $y^* = 1 + \theta s(i) + \lambda(j)$ denote maximal match productivity. For reasons we make clear below, all outcomes in the model depend on $y(d) - y^* = -\gamma d$. Therefore, we surpress $s(i)$ and $\lambda(j)$ from here on.[4]

Information and learning When the workers and the firms first meet, they observe a (joint) signal, d_0. The signal reveals true match quality with probability α, and a random draw from the distribution of match quality with probability $(1 - \alpha)$. The distribution of match quality is assumed to be uniform on the $(0, 1)$ interval. Using the signal, the

[4]This is in line with our empirical work where we condition on (a polynomial in) individual talent and job fixed effects. Notice, also, that the job quality fixed effect, $\lambda(j)$, subsumes everything about the job, including the skill requirement.

worker-firm pair forms an expectation about match quality. The conditional expectation equals

$$E_0(d|\, d_0) = (1 - \alpha)E(d) + \alpha d_0 \tag{2}$$

and is thus a weighted average of the signal and the unconditional mean $E(d)$; the relative weight attached to the signal is increasing in the probability of an informative signal (α).

The choice on whether to match or not depends on the initial signal (d_0). Once production has commenced, agents learn about match quality by observing production. Conditional on matching, subsequent choices depend on revelations about match quality.

Hiring and wage bargaining We follow Eeckhout and Kircher (2011) when modeling hiring and wage bargaining. We think of three stages: a meeting stage, a revelation stage, and a frictionless stage.[5]

At the meeting stage, each worker is paired randomly with one job. The worker-firm pair observes the initial signal (d_0) and decides on whether to match or to continue searching. Should the agents decide to match, they agree on an entry wage, where workers receive half of the match surplus. Should the agents decide to continue searching, they incur a cost (c) associated with waiting to achieve the frictionless (optimal) stage (see Atakan 2006); we assume that c is shared equally between the two parties.

At the revelation stage, uncertainty about match quality is revealed. The worker-firm pair then decides to continue or to terminate the match. Terminating the match implies waiting until the frictionless stage. The total cost associated with separation is $(c + b)$ – again shared equally; here b denotes the additional cost of separating at the revelation stage. If the parties decide to dissolve the match, they get the pay-offs associated with the optimal allocation.

At the frictionless stage, workers receive the wage associated with the optimal match, w^*, and firms receive profits associated with the optimal match π^*. The assumption that continued search (or dissolution of the match) takes the agents straight to their optimal matches is of course extreme, but Eeckhout and Kircher (2011) show that less extreme assumptions do not alter the substance of the conclusions. The key is that the agents make their decision relative to an outside option that depends on the optimal match (y^*). As our focus is on micro-level predictions, y^* is treated as exogenous.

[5]Eeckhout and Kircher (2011) have no uncertainty and thus have only a meeting stage and a frictionless stage. We add a revelation stage since information may be incomplete at the meeting stage.

2.1 Matching, wages, and separations

The outcomes at the meeting stage depend on the initial signal, see equation (2). At the meeting stage, the expected joint surplus equals[6]

$$E_0(S|\,d_0) = [(1 - p_0(\alpha))E_0(y(d)|\,d_0) + p_0(\alpha)(y^* - (c + b))] - [y^* - c]$$

where $p_0(\alpha)$ denotes the probability of separating at the revelation stage, given the information available at the time of the match. The first term in brackets represents the expected gain from matching; with probability $(1 - p_0(\alpha))$ the match continues to be viable, in which case expected productivity equals $E_0(y(d)|\,d_0) = y^* - \gamma E_0(d|\,d_0)$; with probability $p_0(\alpha)$ the match is destroyed, yielding the joint pay-off $(y^* - (c + b))$. The second term in brackets represents the alternative to matching, i.e., waiting, which yields a pay-off of $(y^* - c)$.

The two parties will match if and only if $E_0(S|\,d_0) = (1 - p_0(\alpha))(c - \gamma E_0(d|\,d_0)) - p_0(\alpha)b > 0$. The matching threshold can thus be written as

$$\gamma E_0(d|\,d_0) + \frac{p_0(\alpha)}{1 - p_0(\alpha)}b < c$$

The left-hand-side represents the (expected) losses associated with matching, and the right-hand-side, the loss associated with waiting. The first term of the left-hand-side is the production loss associated with expected mismatch. The second term on the left-hand-side is the expected additional cost of separating later.

The entry wage is determined by a surplus sharing rule with imperfect information about actual match productivity.

$$w_0(d) = \frac{1}{2}E_0(S|\,d_0) = \frac{1}{2}\left[(1 - p_0(\alpha))(c - \gamma E_0(d|\,d_0)) - p_0(\alpha)b\right] \qquad (3)$$

Notice that entry wages depend on actual mismatch (d) only to the extent that the signal correlates with mismatch.

At the revelation stage, the firm-worker pair revisits the employment relationship and re-negotiates wages. The set of continuing matches is defined by $S(d) = y(d) - (y^* - (c + b)) > 0$. Since $S(d) = (c + b) - \gamma d$, the match continues to be viable if the actual cost of mismatch (γd) is lower than the separation cost $(c + b)$. Separations thus occur if

$$d > \frac{c + b}{\gamma} \equiv d_s \qquad (4)$$

Using the definition of the separation threshold (d_s), we can rewrite the matching thresh-

[6]Throughout we ignore discounting, and thus focus on the expectation of steady state long-run surpluses.

old somewhat. The set of acceptable matches is defined by

$$E_0(d|\, d_0) < d_s - \frac{b/\gamma}{1 - p_0(\alpha)} \equiv d_m(\alpha) \tag{5}$$

and the number of matches is, thus, given by

$$m(\alpha) = \Pr(E_0(d|\, d_0) < d_m) = E(d) + (d_m(\alpha) - E(d))/\alpha \tag{6}$$

From equation (5) it follows that $d_m < d_s$, since matching implies a risk of incurring the additional separation cost (b) in the future.

Agents expect to separate in two distinct scenarios. One is related to the probability of separating if the information obtained at the matching stage was uninformative. The probability that agents receive an uninformative signal is $1-\alpha$. The share of those matches which are destroyed is $1 - d_s$. A second scenario is the probability of separation when the information received was actually informative (which happens with probability α). Despite the fact that information was correct, separations might occur if the information content of the initial signal is sufficiently low. To be specific, separations occur if $\alpha < \bar{\alpha} \equiv (d_m - E(d))/(d_s - E(d)) < 1$. Since $d_m < d_s$ the threshold value is less than unity. In sum, we can write the probability of separating at the revelation stage (p_0) as

$$p_0(\alpha) = (1-\alpha)(1-d_s) + \alpha I(\alpha < \bar{\alpha}) \left(1 - \frac{d_s}{m(\alpha)}\right) \tag{7}$$

where $I()$ denotes the indicator function. $1 - d_s/m$ reflects the probability of separating when the agents received correct information. In the appendix we show that $\partial p_0/\partial\alpha < 0$; that is, if more information is available at the time of the match, agents expect fewer separations at the revelation stage.

To complete the description of the model, we note that the wage, given that the match continues to be viable, is given by

$$w(d) = \frac{1}{2}\left[(c+b) - \gamma d\right] \tag{8}$$

2.2 Predictions

Here we summarize the four predictions that we take to the data. To facilitate interpretation we concentrate on the extreme cases, i.e., $\alpha = 0$ and $\alpha = 1$. We relegate the slightly more complex derivations of how the responses to mismatch varies with marginal changes in α to the appendix.

Predictions 2-4 relate to how the responses to actual mismatch vary with the information content of the signal. The comparative static exercise is thus to change d and then to look at how the magnitude of wage and separation responses varies with the precision of the signal.

1. Exposure to initial mismatch Exposure to initial mismatch depends on how many matches are formed. From (6) we have

$$m(\alpha)_{\alpha\to 1} = d_m(\alpha)_{\alpha\to 1} = c/\gamma < m(\alpha)_{\alpha\to 0} = 1$$

When there is little information about match quality, agents will always match (given that the market exists). In other words, a precise signal truncates the potential distribution of mismatch more than an imprecise one. If the distribution of potential mismatch does not vary with α, higher match rates (induced by lower α) translate into greater exposure to mismatch. The first prediction we take to the data is that less information increases exposure to mismatch.

2. Initial mismatch and entry wages From (3) it follows that

$$\frac{\partial w_0}{\partial d} = -\frac{(1-p_0(\alpha))\gamma\alpha^2}{2} \leq 0$$

and so

$$\left.\frac{\partial w_0}{\partial d}\right|_{\alpha\to 1} - \left.\frac{\partial w_0}{\partial d}\right|_{\alpha\to 0} = -\frac{\gamma}{2} < 0$$

This is the second prediction we take to the data. With greater information content of the initial signal, the extent to which mismatch is priced into entry wages increases.

3. Initial mismatch and the separation rate At the revelation stage, separations are deterministic and determined by (4). Let us instead look at the separation rate: $s = p_0 = (1-\alpha)(1-d_s) + \alpha I(\alpha < \bar{\alpha})(1 - \frac{d_s}{m})$. For a marginal match (i.e. a match where $d \to d_s$), the effect of a marginal increase in d equals

$$\frac{\partial s}{\partial d} = (1-\alpha) + \frac{\alpha I(\alpha < \bar{\alpha})}{m(\alpha)} \geq 0$$

At the extremes, only the first term is relevant. And therefore

$$\left.\frac{\partial s}{\partial d}\right|_{\alpha\to 1} - \left.\frac{\partial s}{\partial d}\right|_{\alpha\to 0} = -1 \leq 0$$

Thus a more precise initial signal lowers the impact on separations. This is the third prediction we take to the data.[7]

[7]Even though we can sign the difference in the separation response to mismatch at the extremes, the separation response is not monotonous in α over the entire range of α; see appendix.

4. Initial mismatch and wage growth within jobs Define $\Delta w = w(d) - w_0(d)$, where $w(d)$ is given by (8) and $w_0(d)$ by (3). We have

$$\frac{\partial \Delta w}{\partial d} = -\frac{\gamma}{2}\left[1 - (1 - p_0)\alpha\right] \leq 0$$

and so

$$\left.\frac{\partial \Delta w}{\partial d}\right|_{\alpha \to 1} - \left.\frac{\partial \Delta w}{\partial d}\right|_{\alpha \to 0} = \frac{\gamma}{2} \geq 0$$

This is the fourth prediction we take to the data. With greater information content of the initial signal, the effect of mismatch on wage growth within job falls in absolute value.

3 Data and measurement

We use data from administrative employment registers collected by Statistics Sweden and test scores from the Swedish War Archives. The complete data contain annual employer-employee records for the universe of the Swedish workforce during 1985-2008, with unique person and establishment identifiers. The basis of our analysis is all male workers who enter new jobs (entrants) between 1997 and 2008, and their tenured male coworkers (incumbents).[8] To these data we add socioeconomic background characteristics and military enlistment scores for both entrants and incumbents. Information from the draft is available for all males who did the draft between between 1969 and 1994. During these years, almost all males went through the draft procedure at age 18 or 19, which means that our sample consists of 25 cohorts of male entrants born between 1951 and 1976.

We also add information on wages (adjusted for working hours) and occupational codes to the data. This information is available for a very large sample of private sector establishments covering almost 50 percent of all private sector workers and all public sector workers.[9]

3.1 Measuring talents

The data from the draft procedure include four different measures of cognitive skills and four measures of non-cognitive skills. These test results are not publicly displayed (although available for research purposes). The cognitive measures are based on four subtests measuring: (i) inductive skill (or reasoning), (ii) verbal comprehension, (iii) spatial ability, and (iv) technical understanding. The tests are graded on a scale from 0 to 40 for some cohorts and from 0 to 25 for others. To achieve comparability across cohorts, we standardize the test scores within each cohort of draftees.

[8] We focus on this period since 1997 is the first year that we have occupation information in our data.

[9] Wage and occupation information is collected during a measurement week (in September-November) each year, conditional on being employed for at least one hour during the sampling week. The sampling is stratified by firm size and industry; small firms in the private sector are underrepresented.

The non-cognitive measures are based on behavioral questions in a 20-minute interview with a trained psychologist. On the basis of the interview, the draftee is scored along four separate dimensions. According to Mood et al. (2012), who provide a detailed discussion of the tests, the four scores should be interpreted as capturing (i) social maturity, (ii) psychological energy (e.g., focus and perseverance), (iii) intensity (e.g., activation without external pressure) and (iv) emotional stability (e.g., tolerance to stress). The non-cognitive dimensions are graded from 1 to 5.[10] We standardize these test scores within each cohort of draftees.

To show that each of the measured talents have some independent information content, we relate them to prime-age (age 35) wages within our sample.[11] Table 1 shows the results. Column (1) does not control for education, while column (2) controls for level-of-education fixed effects. The results imply that all skill measures have precisely determined returns, even conditional on educational attainment.[12] On average, a standard deviation increase in a talent is associated with an increase of wages by 2.5 percent (1.5 percent, holding educational attainment constant). Most importantly, however, the results in Table 1 show that there is independent, and sufficiently precise, variation in the individual measures of talent.[13]

3.2 Measuring match quality

Our general approach to measuring match quality is one of job-specific skill weights, as in Lazear (2009). To this end, we use data on individual skills of recent hires and relate them to a proxy for the productivity of these skills in the specific job for which they are hired.

3.2.1 Definition of a job

We define a *job* as an occupation×plant×(entry year) combination. We use (the Swedish version of) the ISCO-88 (International Standard Classification of Occupations 1988) standard at the 3-digit level. Occupations are reported by the employer and the 3-digit level allows us to distinguish between 113 occupations (for instance accountants/lawyers or mining/construction workers). The definition of a job allows for the possibility that technologies differ across plants within an occupational category, and that there is tech-

[10]There is also an overall psychological score on a Stanine scale, which ranges from 1 to 9. These data have been used in some previous studies such as Lindqvist and Vestman (2011) and Håkansson et al. (2015). We use these cruder data in a robustness exercise in Section 5.

[11]The results in Böhlmark and Lindquist (2006) and Bhuller et al. (2015) suggest that earnings at roughly age 35 gives a good approximation of life-time earnings.

[12]This is fairly remarkable, in particular since Grönqvist et al. (2010) estimate the reliability ratio of overall cognitive ability to 73% and the reliability ratio of overall non-cognitive ability to 50%.

[13]The average correlation between two cognitive (non-cognitive) components is 0.59 (0.54). The average correlation between one cognitive and one non-cognitive component is 0.25.

Table 1: Wage returns to skill

	(1)	(2)
Cognitive skills:		
Inductive skill	0.0373***	0.0216***
	(0.0008)	(0.0007)
Verbal skill	0.0253***	0.0031***
	(0.0007)	(0.0007)
Spatial skill	0.0095***	0.0028***
	(0.0006)	(0.0006)
Technical skill	0.0350***	0.0209***
	(0.0007)	(0.0006)
Non-cognitive skills:^{A}		
Social maturity	0.0308***	0.0242***
	(0.0007)	(0.0007)
Intensity	0.0046***	0.0049***
	(0.0006)	(0.0006)
Psychological energy	0.0277***	0.0182***
	(0.0007)	(0.0006)
Emotional stability	0.0260***	0.0205***
	(0.0007)	(0.0006)
Observations	343,440	343,440
R-squared	0.3185	0.3862
Year FE:s	√	√
Educational attainment FE:s		√

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. ^{A}labels for non-cognitive scores are according to Mood et al. (2012). The sample includes all males aged 35 during 1997-2001 who have non-missing information on wages and test-scores. Regressions are weighted by sampling weights to adjust for underrepresentation of small firms in the private sector.

nological evolution within cells defined by occupation and plant.[14]

We define *entrants* (new hires) as workers who enter a new establishment without ever having worked there before (at least since 1985, thus not in the last 12+ years). We define a *separation* (after entry) as a case when a worker is not observed at the entry establishment during any of the two years following the year of entry.[15]

3.2.2 Sorting on job-specific wage returns

As a first step towards turning our talent measures into a measure of match quality, we analyze how *job-specific wage returns* to a particular talent is related to worker endowments of this particular skill. The basic idea is that such job-specific returns are informative about the usefulness of the particular talent in the production process at the job.

In practice, we first estimate the returns to each of the eight standardized test scores within each job (plant×occupation×year) for workers with at least three years of tenure within our sample. To this end, we run 60,500 separate wage regressions, one for each job-cell where we have at least 10 tenured workers. These regressions also control for age.

Figure 1 relates the estimated returns to a specific skill within a job (x-axis) to the average endowments of the same skill among the tenured employees within the same job (y-axis). For seven out of the eight talents, there is a positive (and statistically significant) relationship between the two. Thus, workers are (on average) found in jobs where the returns to their talents are higher than average, as suggested by Roy (1951).

3.2.3 Mismatch as entrant-incumbent skill differences

In principle, one can use the wage estimates discussed above to build a measure of match-quality, and we also do that in a robustness exercise (Section 5). However, the returns to talents estimated at the job-level are very imprecise due to the small numbers of workers observed within each job.[16] Instead we design an explicit measure of mismatch based on comparing the talent-endowments of entrants to the talent-endowments of tenured workers.

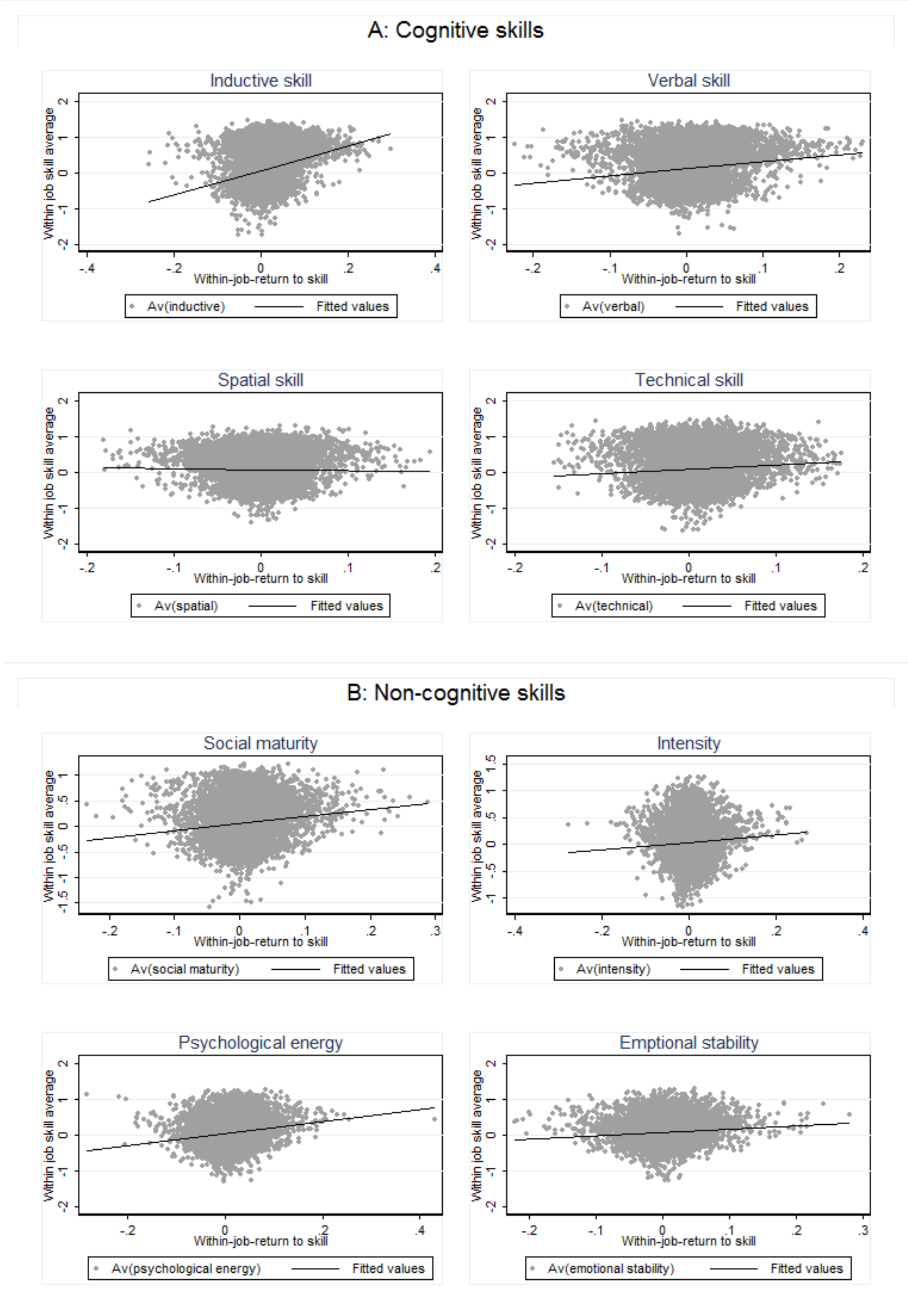
We do this by contrasting the eight cognitive and non-cognitive talents among new hires with those of tenured workers in the same jobs. The rationale for doing so is that tenured workers are likely to be selected (through both entry and exit) on having the right set of talents for the job if match quality matters (as suggested by the figures presented above). Thus, the skill sets of tenured workers can identify the skill requirements of

[14]In Section 5, we verify that matching towards the specific employer (within occupations) carries an important dimension of job-level sorting.

[15]We impose the two year requirement to avoid defining recalls as separations. To avoid including lay-offs due to plant closures, we only include entrants into establishments that remain in the following year.

[16]Note also that we need within-job variation in skill sets to identify the job-level returns to skills. Since job-level sorting reduces the skill variation within jobs, precision is reduced even further.

Figure 1: Correlation between skills and skill returns among tenured workers



Notes: The figure illustrates the relationship between the average job-specific skill endowments among tenured workers and the estimated job-level returns to skills holding age constant. Slope (standard error) of the regression lines, from top left to bottom right: 1.18 (0.05); 0.71 (0.06); -0.11 (0.06); 0.41 (0.07); 0.23 (0.04); 0.12 (0.04); 0.18 (0.04); 0.08 (0.03).

each job. Our basic assumption is that a given job requires a certain set of talents; it is still possible that a production process (e.g. an establishment) benefits from having a workforce with diverse talents *across jobs* (we return to the possible benefits from skill diversity within job in a robustness analysis in Section 5).

For the purpose of the empirical analysis, we focus on entrants and tenured workers with at least 3 years of tenure in the current job. To measure the talents of incumbent workers with reasonable precision, we require that the job employs at least 10 tenured males with non-missing draft scores (in Section 5 we show that results are similar even if we only require one tenured worker). Our generic empirical strategy is to focus on the importance of mismatch after removing the direct importance of all other job-characteristics through job fixed effects (λ_j) and the direct importance of the vector of individual skills (s_i) through a flexible function $g(s_i)$. Hence, when studying the impact of mismatch on some outcome Y we compare different entrants into the same job, while accounting for the market valuation of their skills, i.e.,[17]

$$Y_{ij} = \beta Mismatch_{ij} + g(s_i) + \lambda_j + \epsilon_{ij} \tag{9}$$

To quantify $Mismatch$, our baseline strategy is to use the distance between the skills of the worker and the skill requirements of the job ($d(i, j)$ in terms of Section 2) as:

$$\hat{d}(i, j) = Mismatch_{ij} = \sum_{k=1}^{8} |s_{ik} - \bar{s}_{jk}| \tag{10}$$

where s_{ik} denotes the level of talent k for worker i, and $\bar{s}_{jk}$ denotes the average talent along same dimension among incumbent (tenured) workers. We aggregate the deviations of each of the eight talents to an overall mismatch index, and then standardize the overall index to mean zero, with a unit standard deviation for ease of interpretation.

Obviously, the mismatch index captures mismatch along the *horizontal* dimension ("the worker has a different set of talents than incumbent workers"). The *vertical* dimension ("the worker is over- or under-skilled relative to the skill requirement") can also affect the measure, but only net of the market valuation of the skills. To see this, consider a case where the outcome is log wages. Mismatch reduces wages (i.e., $\beta < 0$), but the overall impact of skills is positive (i.e., $g'() > 0$). Then, increasing s_{ik} from a starting point of $Mismatch = 0$, holding everything else constant, would lower wages through the introduction of mismatch, but also increase wages through the market value of the talent. In contrast, if s_{ik} was reduced, both effects would be negative. Thus, the wage return

[17]The model is related to the AKM-model of Abowd et al. (1999) with the two extensions that we include the mismatch term and that we replace the firm effects of the AKM model by job-effects. On the other hand, we need to rely on a parametric function to control for the impact of person characteristics since we are focusing on entry wages and separation responses and the sample repeated entrants is small and less representative. In appendix section A1.5 we present results from models that also include person fixed effects.

from a marginal increase in s_{ik}, *kinks* at the point where $s_{ik} = \bar{s}_{jk}$, but this does not imply that the marginal return to additional skills within the job turns negative. Formally,

$$\frac{\partial Y_{ij}}{\partial s_{ik}} = \begin{cases} -\beta + g_k'(s_i) \text{ if } s_{ik} < \bar{s}_{jk} \\ \beta + g_k'(s_i) \text{ if } s_{ik} > \bar{s}_{jk} \end{cases} \tag{11}$$

Over-skilled workers are thus not fully rewarded for their talents ($\beta < 0$), but this does not imply that they are paid less than lower-skilled (perhaps perfectly matched) co-workers.[18] Workers who have the right *average* skill level, but the "wrong" composition of talents, receive lower wages if $\beta < 0$.

Notably, the formulation captures the key aspects of the theoretical model. The mismatch measure isolates the extent to which an individual worker is found in a job where his skills deviates from the match with the highest idiosyncratic returns.

We present a large number of variations and robustness checks with respect to the measurement of mismatch in Section 5, including strategies accounting for the fact that the wage returns are different for the different skills, models that account for the possibility that some jobs require a diverse set of personality types, and models estimated for different market segments. The section also discusses results from using a mismatch index based on firm-specific wage returns to each of the talents.

3.3 Proxies for information

Our empirical analysis aims to contrast groups where match productivity is likely difficult to observe at the hiring stage to groups where match productivity is likely easier observed. Our main approach is to classify matches on the basis of the workers' previous labor market experience. In particular, we conjecture that match productivity remain largely unobserved among inexperienced workers. For experienced workers, on the other hand, the employer arguably have more information about about how suitable a given worker is for a specific job (see Farber and Gibbons 1996 and Altonji and Pierret 2001). Such information can come from work histories, previous wages or references related to jobs that are similar to the job under consideration. Analogously, experienced employees should have more information regarding where his/her bundle of talents can be put to most productive use. Thus, match productivity is likely to be, at least partially, observed ex ante for experienced workers but not necessarily for the inexperienced. In line with this view, Lange (2007) shows that most of the market learning takes place within the first

[18] To be precise, with a sufficiently strong market valuation, over-skilled workers are still remunerated for their additional skills. Having more skills than required for a job, reduce the within-job wage returns of additional skills, but not necessarily the overall wage; readers familiar with "ORU" extensions of the Mincer wage regression (Duncan and Hoffman, 1981) will recognize this implication. Note also that our specification also allows for incentives to move from j into a worse match j' ($Mismatch_{ij'} > Mismatch_{ij}$), if ($\lambda_{j'} - \lambda_j > 0$) is large enough to compensate for the drop in match quality.

few years after graduation. Along similar lines, Hensvik and Skans (2014) show, using Swedish data similar to ours that the wage returns to test scores increases with tenure, and that this pattern is more pronounced among the inexperienced.

As an alternative approach we classify matches on the basis of whether the worker entered from non-employment or from another job. Here, we expect there to be less information available about match-specific productivity for those who are hired directly from non-employment.

Labor market *experience* is defined as the number of years which the individual is classified as being employed according to Statistics Sweden's classification system.[19] Since this information is available from 1985 onwards, we truncate experience at 13 years of experience for all entrant cohorts. The median entrant in our sample is 35 years old, and has (at least) 13 years of experience. As mentioned above, we primarily focus on the contrast between inexperienced and experienced workers; for the purpose of the main analysis, inexperienced workers are those with less than 5 years of experience (in line with the speed of market learning estimated by Lange, 2007) while experienced workers have at least 5 years of experience.[20] We define *job-to-job* movers as workers who were employed in the previous year and treat all others as entrants from non-employment.

3.4 Descriptive statistics

Table 2 shows the characteristics of the new hires (entrants) in our sample as well as some basic information about the occupations they enter. Since our main analysis focuses on entrants with at least 10 tenured coworkers within the same job, our sample consists of larger establishments (655 employees on average) than an overall sample of entrants (144 employees) during the same time period.[21] The separation rate, defined as the probability of leaving the establishment within the first year after being hired, equals 21 percent in our sample (29 percent in the overall sample); see the first row of Table 2. As expected, inexperienced workers have a higher separation rate and a lower share of inexperienced recent hires were employed during the previous year (job-to-job movers). Figure A1 shows how separation probabilities and wages evolve with tenure within our sample. Consistent with the earlier literature, these cross-sectional data show a robust negative relationship between tenure and separation and a robust positive relationship between wages and tenure.

For illustrative purposes, the lower half of Table 2 categorizes the occupations at the fairly crude 1-digit ISCO level. Most of the jobs in our data fall in the categories "professionals", "technicians", and "machine operators", which taken together comprise

[19]The classification relies on register data on monthly earnings (in November). It uses employment thresholds calibrated to mimic the employment definition of the Labor Force Surveys.

[20]The key results are also shown for a wide range of experiences and the results corroborate this split of the data.

[21]Table A1 contains information which is analogous to Table 2 for all male entrants during 1997-2008.

Table 2: Entrants 1997-2008

	All		Inexp. 0-4 yrs	Exp. 5+ yrs
	(1)	(2)	(3)	(4)
	mean (SD)	median	mean (SD)	mean (SD)
Separation rate	.21		.24	.20
ln(Entry wage)	10.06 (.37)	10.00	9.82 (.25)	10.11 (.37)
Age	36.2 (7.9)	35.0	27.1 (4.1)	37.9 (7.3)
Experience at entry	12.5 (5.1)	13.0	2.2	13.0
Job-to-job mobility	.82		.46	.88
Entry establishment size	655 (1,180)	243	710 (1,109)	645 (1,193)
Education:				
Primary school less than 7 years	0.07		0.05	0.08
High school short (less than 2 years)	0.41		0.34	0.43
College short (less than 2 years)	0.48		0.60	0.46
PhD long (Doctoral)	0.03		0.01	0.03
Entry occupation:				
Legislators, senior officials and managers	.05		.04	.05
Professionals	.29		.34	.29
Technicians and associate professionals	.24		.20	.25
Clerks	.04		.05	.04
Service workers and shop sales workers	.06		.07	.06
Skilled agricultural and fishery workers	.00		.00	.00
Craft and related trades workers	.09		.06	.09
Plant machine operators and assemblers	.18		.18	.18
Elementary occupations	.05		.05	.05
Mismatch	.00 (1.00)	-.17	.04 (1.02)	.00 (1.00)
Observations	154,681		24,383	130,298

Notes: The table shows the characteristics of the entrants in the year of entry. Medians and standard deviations are provided in brackets where these are relevant.

71 percent of our sample. We explore the extent to which our key results vary between different occupational levels in Section 5.

The final row of Table 2 shows the values of the (standardized) mismatch index in in our sample. Inexperienced workers are mismatched to a greater extent than experienced workers. The difference between the two groups corresponds to 0.04 of a standard deviation, we return to this issue in the next section.[22]

4 Results

In this section we provide empirical tests of the predictions of Section 2. Our main approach is to use experience groups as a proxy for the amount of information about match quality. But we also demonstrate that all the results hold true if we classify

[22]The average mismatch (before subtracting the mean) is 2.58 standard deviations for inexperienced workers.

workers on the basis of employment status (non-employed vs another job) prior to entry.

4.1 Prediction 1: Information and exposure to initial mismatch

As argued in section 3.3, the observability of match quality is likely to differ across experience groups. In particular, initial match quality is more likely unobserved for inexperienced workers. We should thus observe more mismatch among inexperienced workers than among experienced workers. To test this prediction, we analyze how mismatch varies with experience (and/or previous employment) using:

$$Mismatch_{ij} = \beta_x + g(s_i) + \gamma Z_i + \lambda_j + \epsilon_{ij} \tag{12}$$

where i refers to individuals, j to "jobs" (j = *occupation* $\times$ *plant* $\times$ *entry year*), and x are experience groups defined as x=0-4, 1-5,..., 8-12 (+13 yrs. is the reference category); $g(s_i)$ is a flexible control function (vector) in all individual talents; Z_i includes age and 11 education dummies[23]; and λ_j are job fixed effects.

Figure 2 shows average exposure to mismatch by experience group (i.e. $\hat{\beta}_x$), relative to the most experienced. It is clear that initial mismatch decreases with experience. The results are thus consistent with the notion that there is less information, and therefore more mismatch, when inexperienced workers are hired.

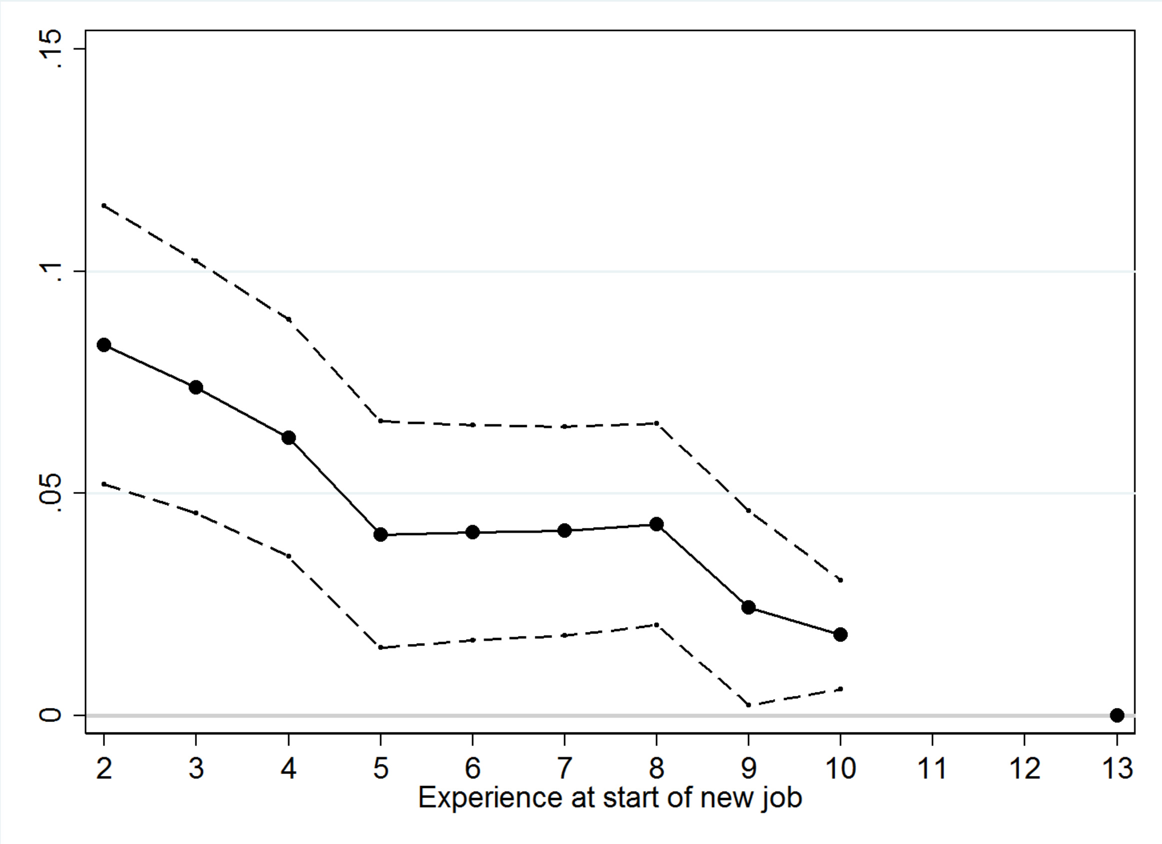
Column (1) of Table 3 compares the inexperienced (0-4 years of experience) to all other experience groups (while restricting the impact of covariates to be the same across groups). Mismatch is 0.031 standard deviations higher among the inexperienced than among experienced workers.

Column (2) adds an indicator for entering from non-employment (the omitted category is job-to-job movers). Adding the non-employment indicator reduces the coefficient on the inexperienced, but only marginally. In column (3) we replace the "inexperienced"-dummy with a full set of years-of-experience fixed effects. Entrants from non-employment are exposed to more mismatch than job-to-job movers. The difference across the two groups equals 1.2 percent of a standard deviation.

Considering that our models account for the direct impact of talents, the level of education, age and job fixed effects (i.e. that the analysis is conducted between workers entering the same jobs), we view the evidence in Table 3 as strongly suggesting that inexperienced workers and workers entering from non-employment, on average, are less well matched at the start of a new job than experienced job-to-job movers. In light of our theoretical framework, this result is consistent with the prediction that match-quality signals contain more noise for the inexperienced and previously non-employed.

[23]From now on, we divide the education levels in table 2 into the following 11 categories: Primary school <7 yrs; Primary school 7-9 yrs.; High school short, <2 yrs.; High school medium, 2 yrs.; High school long, 3 yrs.; College short, <2 yrs.; College medium short, 2 yrs.; College medium long, 3 yrs.;

Figure 2: Initial mismatch by experience



Notes: The figure plots the estimated coefficients on the experience dummies in equation 12. The experience groups are 5-year moving averages (+/- 2 years). Dashed lines are 95% confidence bands.

Table 3: Determinants of mismatch

	(1)	(2)	(3)
Inexperienced (0-4 yrs.)	0.0308***	0.0252***	
	(0.0069)	(0.0073)	
Non-employment to employment		0.0192***	0.0118**
		(0.0066)	(0.0070)
Observations	153,481	153,481	153,481
R-squared	0.6968	0.6969	0.6970
Education dummies	√	√	√
Entrant test scores	√	√	√
(Entry occupation×Entry Year×Plant) FE:s	√	√	√
Experience FE:s			√

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. Mismatch is measured at the time of hiring and experience is measured at the start of the new job. "Entrant test scores" include 2nd order polynomials in each of the eight talents.

4.2 Prediction 2: Initial mismatch and entry wages

To examine the prediction that lower match quality reduces wages, as long as the signal is sufficiently informative, we run wage regressions separately by experience group (x). The regressions relate entry wages to mismatch at the time of the hire:

$$\ln(Entry\,Wage_{ij}^{x}) = \beta^{x} Mismatch_{ij} + g_{w}^{x}(s_i) + \gamma^{x} Z_i + \lambda_j^{x} + \epsilon_{ij}^{x}. \tag{13}$$

As above, i refers to individuals, j to "jobs" (j = *occupation* $\times$ *plant* $\times$ *entry year*), and x to experience groups; Z_i controls for age and education, while $g_w^x(s_i)$ is a second order polynomial in each of the eight talents (we provide robustness checks with even more flexible functional forms in Section 5). As discussed above, we include the flexible skill controls to hold outside opportunities for the worker constant. The job fixed effects control for everything that is specific about plants and occupations and their interactions (by year), including the direct impact of the skill requirements of the job, job amenities and all potential direct effects from the skill levels of the tenured workers.

Figure 3 presents the first set of results. It plots estimates of the coefficient on the mismatch index (i.e. $\hat{\beta}^x$) by detailed experience group (the experience groups are 5-year moving averages). We expect mismatch along partly observed dimensions to be priced. If mismatch is unobserved at the time of hire, the entry wage should be unrelated to mismatch. The figure shows that entry wages are unrelated to mismatch for workers with up to 5 years of experience. For more experienced workers, we find a negative effect on entry wages.[24] For workers with at least 13 years of experience at the start of the new job, a standard deviation increase in mismatch lowers the entry wage by 1.7 percent. Note that this wage penalty is conditional on job-fixed effects, hence workers may still receive a net wage premium when entering a job with low match quality if the average job-level wage premium is sufficiently high (and vice versa for employers who may trade off match quality and worker skill levels).

Table 4 presents more detailed regression results. Panel A splits the sample into workers with more or less than 5 years of experience. Column (1) displays the results for inexperienced workers showing that the entry wage is unrelated to initial mismatch in this group. The coefficient estimate is very small (-0.17 percent) and precisely determined. Column (2) instead displays results for all workers with at least 5 years of prior experience at the start of the new job.[25] Among these workers, a standard deviation increase in mismatch reduces wages by 1.4 percent. Column (3) shows that the difference across experience groups is statistically significant.

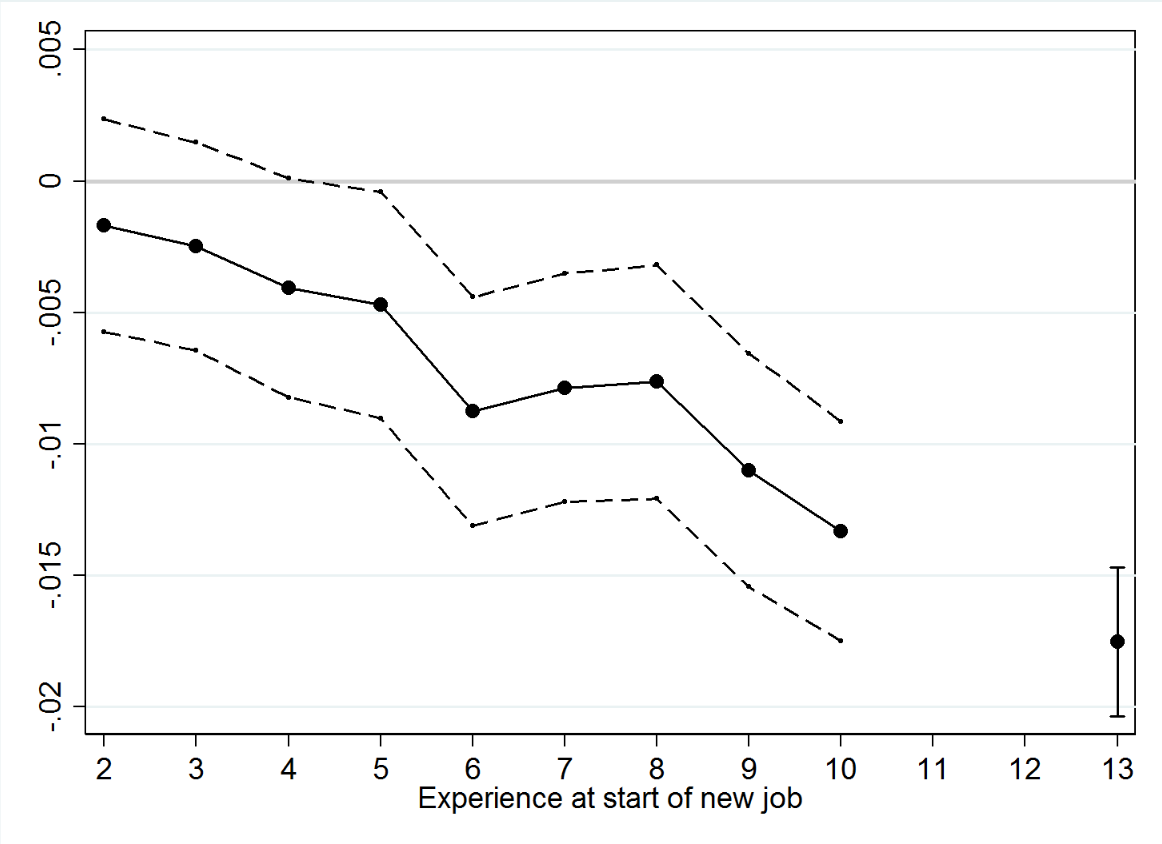
Panel B of Table 4 instead splits the sample according to whether workers have entered

College long, 4 yrs.; PhD short (Licentiate); PhD long (Doctoral).

[24] The pattern is consistent with the employer learning patterns estimated by Lange (2007).

[25] In the pooled regression we include a dummy for each level of experience.

Figure 3: The entry wage response to mismatch, by experience group



Notes: Each dot is an estimate of the wage response to initial mismatch within 5-year experience bins (+/- 2 years). The sample consists of entrants in 1997-2008. Experience can be traced back to 1985; it is truncated at 13 years for workers with 13 years experience or longer. Dashed lines are 95% confidence bands.

Table 4: Entry wage responses to mismatch

	(1)	(2)	(3)
Panel A:	Inexp. 0-4 yrs.	Exp. 5+ yrs.	P-val. for differences
Mismatch	-0.0017	-0.0139***	0.0000
	(0.0021)	(0.0009)	
Observations	24,383	130,298	
R-squared	0.8613	0.8386	
Panel B:	From non-emp.	From job	
Mismatch	-0.0013	-0.0118***	0.0000
	(0.0019)	(0.0010)	
Observations	28,247	125,234	
R-squared	0.8766	0.8378	
Education dummies	✓	✓	
Entrant test scores	✓	✓	
(Entry occupation×Entry Year×Plant) FE:s	✓	✓	

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The sample consists of entrants in 1997-2008. All regressions include a full set of birth cohort and experience fixed effects. The test score controls are 2nd order polynomials in each of the eight test score domains. Column (3) displays the p-value of the difference between the estimates in columns (1) and (2).

from non-employment (Column 1) or from another job (Column 2). We expect mismatch to be priced to a lesser extent for individuals entering from non-employment if there is more remaining uncertainty at the time of recruitment for this group. The results are very much in line with those presented in Panel A. For those entering from non-employment, entry wages are unrelated to mismatch; for job-to-job movers, on the other hand, a standard deviation increase in mismatch reduces the entry wage by 1.2 percent and the difference across groups is statistically significant (Column 3).

The results imply that the inexperienced and entrants from non-employment (i.e. the groups where mismatch is more prevalent, see Table 3) are the groups where mismatch has the lowest impact on entry wages, which is consistent with the notion of more ex ante uncertainty among these two groups.[26]

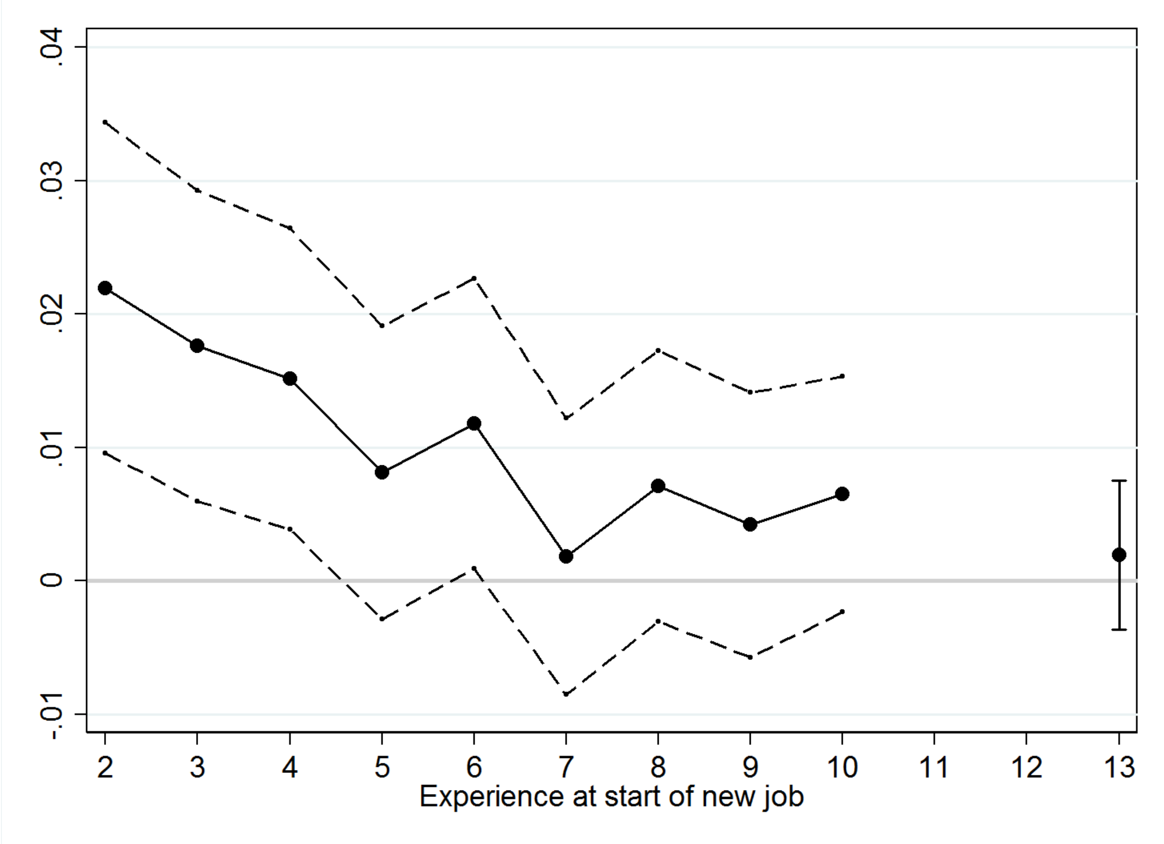
4.3 Prediction 3: Initial mismatch and separations

4.3.1 Baseline results

The relationship between separations and mismatch should be the flip side of the entry wage response. Section 2 (and intuition) suggests that if mismatch is unobserved at the time of hiring, higher mismatch leads to separation (if the price of mismatch is higher than any separation cost). To examine the validity of this prediction, we run regression

[26] In a robustness analysis in Section 5 we show that mismatch, in general, has a roughly linear relationship to wages within experience groups.

Figure 4: The separations response to mismatch, by experience group



Notes: Each dot is an estimate of the separation response to initial mismatch within 5-year experience bins (+/- 2 years). The sample consists of entrants in 1997-2008. Experience can be traced back to 1985; it is truncated at 13 years for workers with 13 years experience or longer. Dashed lines are 95% confidence bands.

models which are identical to (13), but with the first year separation rate as the outcome of interest:

$$1^{st}\ year\ Separation_{ij}^{x} = \beta^{x} Mismatch_{ij} + g^{x}(s_i) + \gamma_s^{x} Z_i + \lambda_j^{x} + \epsilon_{ij}^{x} \tag{14}$$

Figure 4 plots the estimates of the coefficient of interest (β^x). For inexperienced workers, we find that a standard deviation increase in mismatch raises separations by 2.2 percentage points. This corresponds to almost a tenth of the average separation probability for this group. The impact is considerably smaller for experienced workers; beyond 7 years of experience the relationship between separations and mismatch is not statistically significant.

Table 5 reports the results in more detail. Column (1) of Panel A contains the results for inexperienced workers and column (2) shows the results of a pooled regression for all workers with at least 5 years of experience. The separation response to mismatch is considerably larger among inexperienced workers than among experienced workers. A standard deviation increase in mismatch increases separations among inexperienced workers by 2.2 percentage points and by 0.5 percentage points among experienced workers;

Table 5: Separation responses to mismatch

	(1)	(2)	(3)
	Inexp. 0-4 yrs.	Exp. 5+ yrs.	P-val. for differences
Mismatch	0.0220***	0.0050***	0.0042
	(0.0063)	(0.0019)	
Observations	24,383	130,298	
R-squared	0.5968	0.4807	
	From non-emp.	From job	
Mismatch	0.0114***	0.0062***	0.3269
	(0.0057)	(0.0020)	
Observations	28,247	125,234	
R-squared	0.6189	0.4879	
Education dummies	√	√	
Entrant test scores	√	√	
(Entry occupation×Entry Year×Plant) FE:s	√	√	

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The sample consists of entrants in 1997-2008. All regressions include a full set of birth cohort and experience fixed effects. The test score controls are 2nd order polynomials in each of the eight test score domains. Column (3) displays the p-value of the difference between the estimates in columns (1) and (2).

the difference across groups is statistically significant.

Panel B of Table 5 shows the results for workers entering from non-employment (Column 1) compared to job-to-job movers (Column 2). Among entrants from non-employment, the effect on separations is positive and amounts to an increase by 1.1 percentage points for a standard deviation increase in mismatch. For job-to-job movers, on the other hand, there is a smaller separation response, amounting to 0.6 percentage points. Again, the pattern of the results are very much in line with the results for different experience groups.

Overall, the results in Tables 4 and 5 agree well with the interpretative framework in Section 2. Because there is more information about experienced workers, and experienced workers are likely to have more information on where their skills are most apt, entry wages are negatively related to mismatch. The separation response is lower among experienced workers than among inexperienced workers. All of this suggests that mismatch is factored in already at the time of hiring for this group of workers. Analogously, there is likely to be more information about match quality for job-to-job movers than for entrants from non-employment. Therefore, mismatch is priced into the entry wages of job-to-job movers, and we observe a smaller separation response among job-to-job movers than among entrants from non-employment.

Since the results based on the categorization of workers according to their prior employment status and their prior labor market experience are so similar, we drop the division based on prior employment status from here on. (The interested reader can find them in

Appendix A1.3, and they continue to be very much in line with the results for experience groups.)

4.3.2 Timing of the separation response

The timing of the separation response provides information about the speed of learning about mismatch. In order to shed light on the speed of learning we need to use higher-frequency data than the annual information we have used so far. We therefore tap monthly separation-indicators. These are unfortunately of somewhat lower quality; they are based on the last month of wage payments from each employer within each calender year, and this creates some uncertainty regarding the exact month of separation since wage payments sometimes are delayed. The data are described in greater detail in appendix section A1.4.

Figure 5 shows the separation response by months since the start of the new job. To gain precision, we pool all experience groups.[27] As before we use moving averages to increase precision. The first point in the figure represents (quarterly) separations within the first 1-3 months after the start of the new job, the second represents the response after 2-4 months, and so forth. The results show that the peak of the separation occurs approximately 6 months since the start of the new job. In general, the speed of adjustment is thus fairly rapid. We find no evidence of separation responses after 1 year. Swedish employment protection may contribute to the peak at 6 months, since employment protection legislation allows for a 6 months probation period (during which both agents can terminate the contract at will) at the start of a new permanent contract.[28] This implies that 6 months could be a focal point, and incentives, from both the employer and the employee side, are to some extent geared towards terminating the contract at 6 months if the match quality is poor.

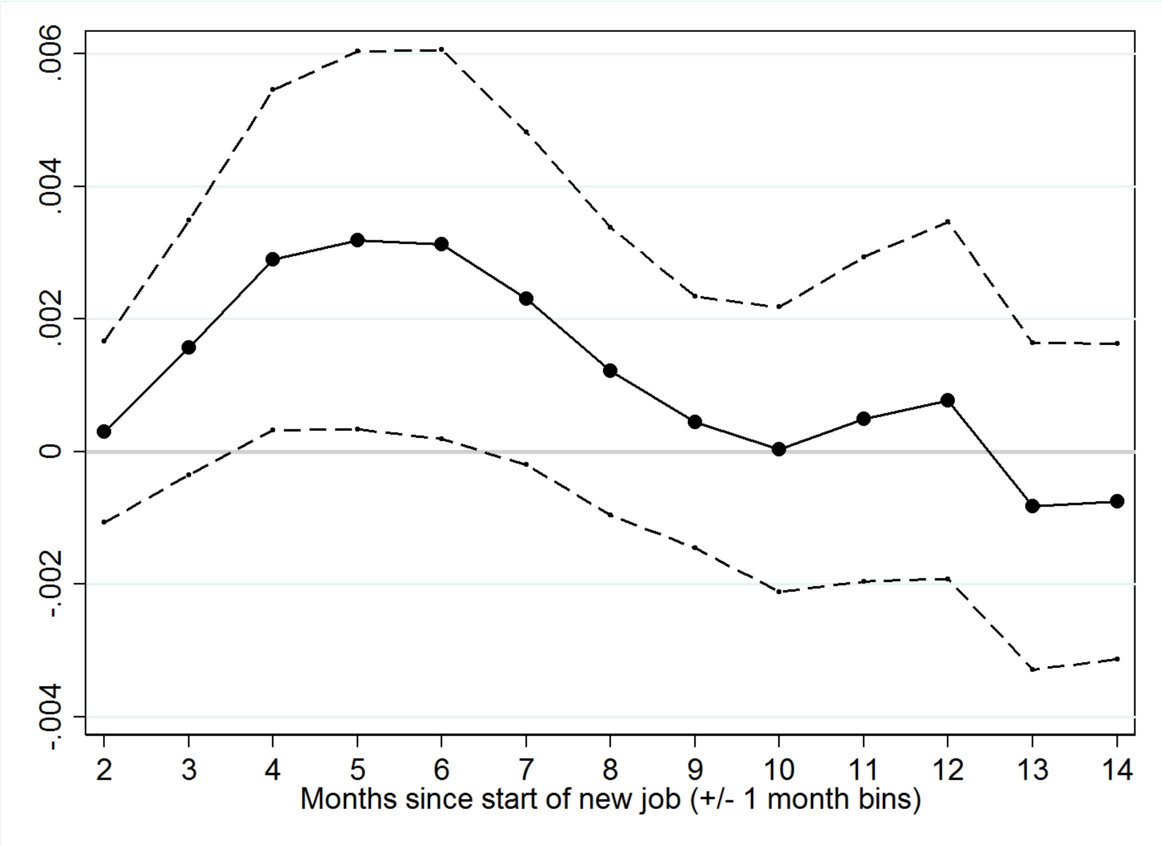
4.4 Prediction 4: Initial mismatch and wage growth within jobs

Section 2 showed that the impact of mismatch on wage growth should be more negative for groups where there is more initial uncertainty about mismatch. This prediction comes from the fact that, over the longer run, uncertainty about initial mismatch is revealed and this should be reflected in wages as long as wages are positively related to the size of the matching surplus. Table 6 examines the validity of this prediction by estimating wage growth equations separately for inexperienced and experienced workers. We calculate

[27] The annual separation response for the sample is 0.007, see appendix Table A4, column (1). The separation response among those with less than 5 years of experience is larger with an almost identical time profile, but the responses are less precisely estimated.

[28] OECD characterizes Swedish Employment Protection Legislation as being around average in terms of overall strictness. The rules concerning the use of temporary contracts are however very flexible, whereas the rules pertaining to layoffs (in particular for cause) among workers on permanent contracts are rather stringent.

Figure 5: Timing of the separation response



Notes: The figure displays the response to initial mismatch within 3 month-bins (+/- 1 month). We calculate the monthly duration of employment using an indicator for the first and the last month of remuneration from each employer. Since entry occupations are measured in September or October (depending on sampling month), we focus on workers that entered their new job in the period August-October of each year. Dashed lines are confidence bands.

wage growth as the 3-year difference in log wages for individuals who have stayed in the same plant, and estimate the regressions separately by experience group. Notice that the samples are reduced to around a quarter of the original size. The sample reduction has two origins. First, the wage data are collected via sampling; thus we lose a substantial fraction of observations because plants randomly exit the sampling frame. Second, wage growth within job is (of course) only observed for the selected sub-sample that stay on in the same job.[29]

Table 6: The impact of mismatch on wage growth within job

	(1)	(2)	(3)
	Inexp. (0-4 yrs.)	Exp. (5+ yrs.)	P-val for differences
Mismatch	-0.0503*	-0.0145	0.1557
	(0.0265)	(0.0093)	
Observations	6,827	31,357	
R-squared	0.7881	0.6170	
Education dummies	√	√	
Entrant test scores	√	√	
(Entry occupation×Entry Year×Plant) FE:s	√	√	

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. Wage growth refers to the 3-year difference in log wages for individuals who have stayed on in the same plant. All regressions include a full set of birth cohort and experience fixed effects. The test score controls are 2nd order polynomials in each of the eight test score domains. Column (3) displays the p-value of the difference between the estimates in columns (1) and (2)

The results in column (1) suggest that inexperienced workers who are subjected to a standard deviation increase in initial mismatch see their wages grow at a 5 percent lower rate than the average worker in the same group. For experienced workers, the estimate in column (2) corresponds to a reduction in wage growth by 1.4 percent. The difference across the two groups is thus -3.6 percent. These estimates are in line with our prediction, but the difference across groups is not statistically significant at conventional levels (see column 3).[30]

The lack of statistical significance comes from the fact that we estimate the regressions very flexibly, allowing all coefficients to vary across the two groups. If we pool the two regressions, only allowing mean wage growth and the impact of mismatch to differ across the two groups, we find that mismatch is associated with 4.3 percent lower wage growth for inexperienced workers and with 1.5 percent lower wage growth for experienced workers. The difference across the two groups is -2.8 percent, which is statistically significant at the 5 percent level.

[29]Sample selection will imply that average wage growth is positive, but does not necessarily affect the marginal impact of an increase in mismatch on wage growth.

[30]The results are similar, but less precise, if we compare entrants from non-employment versus another job, see Table A2 in Appendix. Mismatch is associated with 2.8 percent (se 0.0225) lower wage growth for entrants from non-employment compared to 1.7 percent (se 0.0097) for entrants from another job.

5 Robustness

This section is primarily devoted to the robustness of our main results. We address several issues, e.g., the measurement of the mismatch index, selection on unobserved characteristics, and the sample restrictions. In Section 5.1 we present basic specification checks regarding measurement and functional form; Section 5.2 studies the role of occupations and market segments, and Section 5.3 discusses results from alternative ways of measuring match quality. We end this section with a discussion (see Section 5.4) of how the results relate to explanations other than the interpretative framework we presented in Section 2.

5.1 Basic specification checks

Table 7 presents a battery of robustness checks; for easy reference we reiterate the baseline estimates in Panel A.[31]

Panel B shows that the results are robust to more flexible controls for skills. More specifically, we include all (8×7) interactions between the eight test score domains, which has a very limited impact on the estimates. In appendix Section A1.5 we push the idea of additional skill controls even further by estimating models with individual fixed effects. These fixed effects obviously hold all time-invariant characteristics of the individual constant, and thus take the direct effect of individual skill into account but they also capture other unobserved dimensions of worker ability (and outside options), potentially not captured by the test scores. However, in order to estimate the model we need to make some compromises. The identifying data set is much smaller, and less representative, as workers must to be recorded as a new hire at sampled workplaces at least twice. This forces us to pool across experience groups to gain precision. The results, although statistically imprecise, are quite similar to those of the main model.

Panel C examines whether the effects of mismatch are non-linear. We pursue this robustness check for three reasons. First, one may suspect that there are ranges of inaction, either because there is some measurement error in skills or because mobility/separation costs are substantial. If so, there should be an initial range of inaction until mismatch surpasses a certain threshold when separations should increase. Second, there is unavoidably some arbitrariness in specifying the mismatch index. The correct functional form of mismatch depends on the (unknown) production technology. Finally, the fact that the extent of mismatch vary with experience and previous employment status could potentially explain differences in responses if the impact of mismatch is highly non-linear. The results presented in Panel C shows, however, that although the impact of mismatch on entry wages is somewhat non-linear (the absolute size of the effect tends to be larger for

[31] Results for job-to-job movers vs. entrants from non-employment is presented in the appendix (Table A3).

Table 7: Basic specification checks

	ENTRY WAGES		SEPARATIONS	
	Inexp.	Exp.	Inexp.	Exp.
	0-4 yrs.	5+ yrs.	0-4 yrs.	5+ yrs.
	(1)	(2)	(3)	(4)
A. Baseline				
$Mismatch$	-0.0017	-0.0139***	0.0220***	0.0050***
	(0.0021)	(0.0009)	(0.0063)	(0.0019)
Observations	24,383	130,298	24,383	130,298
R-squared	0.8613	0.8386	0.5968	0.4807
B. All skill interactions				
$Mismatch$	-0.0022	-0.0128***	0.0220***	0.0054***
	(0.0022)	(0.0010)	(0.0067)	(0.0020)
Observations	24,383	130,298	24,383	130,298
R-squared	0.8615	0.8387	0.5975	0.4808
C. Non-linearities in mismatch				
$Mismatch$	-0.0005	-0.0126***	0.0206***	0.0055***
	(0.0021)	(0.0010)	(0.0066)	(0.0020)
$Mismatch^2$	-0.0012*	-0.0015***	0.0014	-0.0006
	(0.0007)	(0.0003)	(0.0021)	(0.0007)
Observations	24,383	130,298	24,383	130,298
R-squared	0.8613	0.8387	0.5968	0.4807
D. No restriction on # tenured workers				
$Mismatch$	-0.0020	-0.0115***	0.0174**	0.0060***
	(0.0025)	(0.0008)	(0.0080)	(0.0017)
Observations	36,194	298,619	36,194	298,619
R-squared	0.9099	0.8840	0.7583	0.6260
E. Weighted mismatch index				
$Mismatch$	-0.0015	-0.0138***	0.0218***	0.0053***
	(0.0020)	(0.0009)	(0.0063)	(0.0019)
Observations	24,383	130,298	24,383	130,298
R-squared	0.8613	0.8386	0.5968	0.4807
F. Mismatch in cognitive and non-cognitive ability				
$Mismatch_{cognitive}$	-0.0027	-0.0091***	0.0137**	0.0037**
	(0.0018)	(0.0009)	(0.0056)	(0.0018)
$Mismatch_{non-cognitive}$	0.0012	-0.0101***	0.0171**	0.0031
	(0.0025)	(0.0011)	(0.0073)	(0.0022)
Observations	24,383	130,298	24,383	130,298
R-squared	0.8613	0.8386	0.5968	0.4807
G. Mismatch (based on overall cognitive and non-cognitive scores)				
$Mismatch$	-0.0014	-0.0078***	0.0117**	0.0028*
	(0.0016)	(0.0007)	(0.0048)	(0.0014)
	24,383	130,298	24,383	130,298
	0.8613	0.8384	0.5965	0.4807
Education dummies	√	√	√	√
Entrant test scores	√	√	√	√
(Entry occupation×Entry Year×Plant) FE:s	√	√	√	√

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The specification is the same as in Tables 4 and 5.

high values of mismatch), the estimates on the second order terms are small.[32] Further, we find that separations are literally linear in mismatch. We take this to indicate that allowing for non-linearities in mismatch is not crucial.

Panel D of Table 7 relaxes the restriction on the number of tenured workers. Our main strategy has been to restrict the analysis to jobs with at least 10 tenured workers, in order to have a reasonably precise measure of the skill requirements of the job. Panel D drops this restriction and thus includes all jobs with at least one tenured worker. Without the restriction, sample size increases substantially. Nevertheless, our results are remarkably stable. Overall the absolute sizes of the estimates are somewhat lower which is consistent with the view that we get a less precise measure of skill requirements when we include jobs with less than 10 tenured workers.

In Panel E we explore whether mismatch in terms of the skills that are highly rewarded in the labor market is more important than mismatch in other dimensions. To examine this issue we weight the components of the mismatch index with the estimated wage returns to the particular components; as weights we use the returns reported in Table 1. Using this weighted mismatch index does not change the results.

Panel F instead inquires whether it is mismatch in the cognitive or non-cognitive dimension that primarily matters. We thus introduce two separate indexes, one for the cognitive and the other for the non-cognitive dimension. The coefficients on mismatch along the cognitive and non-cognitive dimensions are not significantly different from one another.

Panel G finally focuses on mismatch in these two (cognitive vs non-cognitive) dimensions rather than the full vector of eight skills. We do this by calculating a mismatch index based only on cognitive and non-cognitive skill aggregates. The results are about half the size relative to the baseline, but the overall picture remains unchanged. The attenuation of the point estimates implies that there is independent information regarding match quality in the full vector of (eight) skill components; this information is lost when relying on the cruder cognitive and non-cognitive aggregates.

5.2 Occupations and labor market segments

Next, we present a number of variations of the analysis focused on investigating the role of heterogeneity across occupations and other indicators of labor market segments. The results are displayed in Table 8.

In panel A we first make the definition of a "job" even more precise by defining it as the interaction between 3-digit occupation, plant, entry year (as before) and *education*

[32] For experienced workers, we find that the impact of mismatch, when evaluated at a standard deviation above the mean, is -1.6%; evaluated at a standard deviation below the mean, the impact is -1.0%. We have also percentile ranked the mismatch index and allowed the effect of mismatch to vary across percentiles. Again, the effect appears to be broadly linear across the mismatch distribution.

Table 8: Occupation and labor market segments

	(1)	(2)	(3)	(4)
	ENTRY WAGES		SEPARATIONS	
	Inexp.	Exp.	Inexp.	Exp.
	0-4 yrs.	5+ yrs.	0-4 yrs.	5+ yrs.
A. Job*Education fixed effects				
Mismatch	0.0008	-0.0119***	0.0217**	0.0059**
	(0.0030)	(0.0014)	(0.0098)	(0.0029)
Observations	24,383	130,298	24,383	130,298
R-squared	0.9230	0.9023	0.7596	0.6730
B. High vs. low-skill jobs				
Mismatch	-0.0008	-0.0186***	0.0221***	0.0045*
	(0.0025)	(0.0010)	(0.0082)	(0.0023)
Mismatch×High-skilled job	-0.0034	0.0050***	-0.0010	0.0013
	(0.0027)	(0.0012)	(0.0083)	(0.0025)
Observations	24,383	130,298	24,383	130,298
R-squared	0.8554	0.8278	0.5959	0.4806
C. High vs. low-skill workers				
Mismatch	-0.0032	-0.0163***	0.0227***	0.0033
	(0.0029)	(0.0010)	(0.0086)	(0.0022)
Mismatch×High education	0.0009	0.0005	-0.0029	0.0040
	(0.0028)	(0.0013)	(0.0083)	(0.0025)
Observations	24,383	130,298	24,383	130,298
R-squared	0.8570	0.8313	0.5960	0.4807
D. Excluding diverse jobs				
Mismatch	-0.0026	-0.0155***	0.0205***	0.0068***
	(0.0025)	(0.0012)	(0.0074)	(0.0024)
Observations	17,869	98,142	17,869	98,142
R-squared	0.8606	0.8283	0.5889	0.4761
E. Job vs. occupation mismatch				
Mismatch (actual job)	-0.0025	-0.0095***	0.0156*	0.0055**
	(0.0028)	(0.0013)	(0.0087)	(0.0026)
Mismatch (random job w/in occupation)	0.0012	-0.0070***	0.0102	-0.0007
	(0.0028)	(0.0013)	(0.0087)	(0.0026)
Observations	24,383	130,298	24,383	130,298
R-squared	0.8613	0.8387	0.5968	0.4807
Education dummies	√	√	√	√
Entrant test scores	√	√	√	√
(Entry occupation×Entry Year×Plant) FE:s	√	√	√	√

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. Sample consists of entrants in 1997-2008. All regressions include a full set of birth cohort and experience fixed effects. The test score controls are 2nd order polynomials in each of the eight test score domains.

level. In practice, we are thus defining (e.g.) lawyers entering a certain establishment as having different jobs if they have 4-year diplomas rather than 3-year diplomas (both are possible). This extension leaves the baseline results unaffected, however.

Next we examine whether mismatch has different implications depending on whether the position is high-skilled or low-skilled. One reason for this extension is that the losses associated with mismatch may be larger at the higher end of the job-complexity scale. If so, firms may invest more resources in screening, which could imply that initial mismatch would be priced to a greater extent in these jobs. However, this relies on the presumption that it is equally hard to observe the relevant skills for high-level and low-level positions. We present two ways of categorizing jobs into high- and low-skilled positions. In panel B we classify the job depending on the skill class of the occupation and in panel C we classify the job depending on whether the individual entrant has high or low education.

The basic message emerging from these panels is that the impact of mismatch is similar across the distribution of positions. In panel B, the only significant difference is the extent to which initial mismatch is priced among experienced workers. In panel C, the only significant difference is that high-skill experienced workers are more likely to separate. Neither of these results suggest that there is more information about workers and jobs at the higher end of the labor market.

Panel D aims to address the potential concern that some jobs are defined by the need for *diversity* in personality types, rather the need for a specific type of worker. To this end, we first compute the job-level variance of each skill across tenured workers and calculate the averages of these (eight) variances by job. Then, we remove the quarter of jobs with the highest variance. The basic assumption is that jobs with a high variance in skills across tenured workers are more likely to be using production processes with cross-worker complementarities within the same job-classification (i.e. a form of measurement error in our definition of a job).[33] As shown in the table, the results remain very robust.

Panel E deals with the role of the establishment in the definition of a job. A priori, it is conceivable that the patterns we observe are driven by a search for the right occupation, and not the right job. If this was true, we would observe equally strong patterns if we defined our measure of mismatch relative to *any* job in the same occupation. We therefore constructed a separate measure of mismatch towards a random job within the same occupation, using the same sample restrictions for these jobs as for the main analysis (thus, keeping issues related to statistical precision equal across the two terms). We then rerun the main analysis including both the actual (baseline) job-level mismatch score and the corresponding score for mismatch towards a random job within the same occupation. The results displayed in Panel E show that "pure" occupational mismatch do matter to

[33] One form of such complementarity is teamwork, where team members can complement each others skills. Ideal data would then define the skill needs according to the missing skills, and not the skills of the incumbent workers.

some extent, in particular for the wages of experienced workers, but (importantly) the main patterns are primarily driven by actual job-level mismatch.

5.3 Alternative approaches

Our baseline mismatch index is based on the distance between the talents of new entrants and tenured workers. To validate our strategy we have conducted two additional exercises. The first exercise, shown in the appendix (Table A5), analyzes how the variance of talents across different workers is related to the tenure of these workers. The results show that the variance is lower among workers with longer tenure and that this declining pattern is stronger among workers with less pre-hire experience. This result is consistent with the notion that workers, over time, are selected on a job-specific set of talents, and that this process is particularly pronounced among inexperienced workers.

The second exercise is based on the idea that job-level wage *returns* to specific talents of can inform us about the usefulness of these talents in the production process. Here we use the estimates from the 60,500 job-cell regressions shown initially in Figure 1. We then use these estimated job level returns to calculate an alternative mismatch measure. It is defined as follows:

$$Mismatch_{ij}^{returns} = \sum_{k=1}^{K} (\hat{\beta}_k - \hat{\beta}_{jk}) s_{ik} \tag{15}$$

where $\hat{\beta}_{jk}$ is the estimated return skill k in job j; $\hat{\beta}_k$ is the estimated market return to skill k and s_{ik} denotes individual skill. According to this measure, an individual entrant is mismatched if the returns to his particular skill set (within his job) is low relative to the average market returns to the same skill set.

An advantage of this measure is that it directly relates the pay-offs of staying within the match to the outside option. A disadvantage is that we must rely on the very noisy estimated returns, sometimes from very small cells. Naturally, a non-trivial part of the variance illustrated in Figure 1 is related to sample noise in small job-cells. To prevent attenuation bias stemming from poor precision of $\hat{\beta}_{jk}$ in these small cells, we weight the regression by the inverse of the sampling error.[34]

Table 9 shows the results when we estimate the key equations using the mismatch measure defined in (15). To make the results comparable to the baseline estimates, we standardize the mismatch index as before. The results provide a very similar picture as the baseline estimates: With the baseline mismatch metric we estimated the impact of mismatch for the inexperienced to -0.17 percent; here it is estimated to -0.45 percent. The impact of mismatch for experienced workers was -1.39 percent in the baseline; here

[34]The weights are constructed as $\omega = \left(\sqrt{\sum_k \mathrm{var}(\hat{\beta}_{jk})}\right)^{-1}$.

Table 9: Responses to mismatch with alternative mismatch measure

	(1)	(2)	
	Inexp.	**Exp.**	
	0-4 yrs.	5+ yrs.	
		ENTRY WAGES	
$Mismatch_{ij}^{returns}$	-0.0045	-0.0166***	0.016
	(0.0035)	(0.0014)	
Observations	19,621	99,901	
R-squared	0.8314	0.8362	
		SEPARATIONS	
$Mismatch_{ij}^{returns}$	0.0149	0.0034	0.384
	(0.0094)	(0.0030)	
Observations	19,621	99,901	
R-squared	0.5153	0.4271	
Education dummies	√	√	
Entrant test scores	√	√	
(Entry occupation×Entry Year×Plant) FE:s	√	√	

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. All regressions include a full set of birth cohort and experience fixed effects. The test score controls are 2nd order polynomials in each of the eight test score domains. The regressions are weighted by $\omega = \left(\sqrt{\sum_k \mathrm{var}(\hat{\beta}_{jk})}\right)^{-1}$ to adjust for small sample errror in the estimates of the job level returns..

it is estimated to -1.66 percent. With the wage based mismatch metric we obtain slightly weaker separation responses but the pattern across the inexperienced and experienced groups is identical. Overall, we interpret these results (despite the lack of precision) as providing a very strong validation of the results based on our baseline measure.

5.4 Alternative explanations: Peer effects, preferences and wage compression

Here we raise the issue of whether there are alternative explanations for our results. We address three alternative explanations: (i) peer effects; (ii) preferences; and (iii) differential wage dispersion in different segments of the labor market.

Let us first address the question of whether our results are consistent with a standard peer-effects model. Our baseline analysis has a flavor of peer effects models, since the measurement of mismatch is based on the correspondence between the talents of entrants and tenured workers. However, there are several aspects of our analysis that differentiates it from all peer effects models we are aware of. The first, and main, difference is that we compare entrants into the same job and these entrants will all be exposed to the same set of peers, which implies that the first-order effect (i.e. the "standard" peer effect) of peer quality is accounted for by the job fixed effects. The second difference is that we rely on a vector of talents, which implies that mismatch arises also for people with a similar level (but a different composition) of talents, so any potential peer effects must arise from benefits of being similar. The third difference is that we sum all absolute deviations between the talents of entering workers relative to the talents of tenured workers, regardless of whether these deviations are positive or negative. To explain our results by a peer effects model, it would have to be better for a low-skilled workers to work in a low-skilled environment, whereas the reverse needs to be true for the high skilled. Clearly, this is not what a standard peer-effects model would predict. All in all, we do not believe that our results should be interpreted within the framework of a (standard) peer-effects model.

Another alternative explanation is related to preferences of the workers. Workers may have a preference for working with people who have similar traits and talents as themselves. While this could explain why mismatched workers separate to a greater extent, it cannot explain the wage patterns we observe. Indeed, if preference for similarity would be the main driving force, we expect that well-matched workers would pay a compensating wage differential for similarity, i.e., their wages should be lower than the wages of mismatched workers. This is clearly different from what we observe in the data.

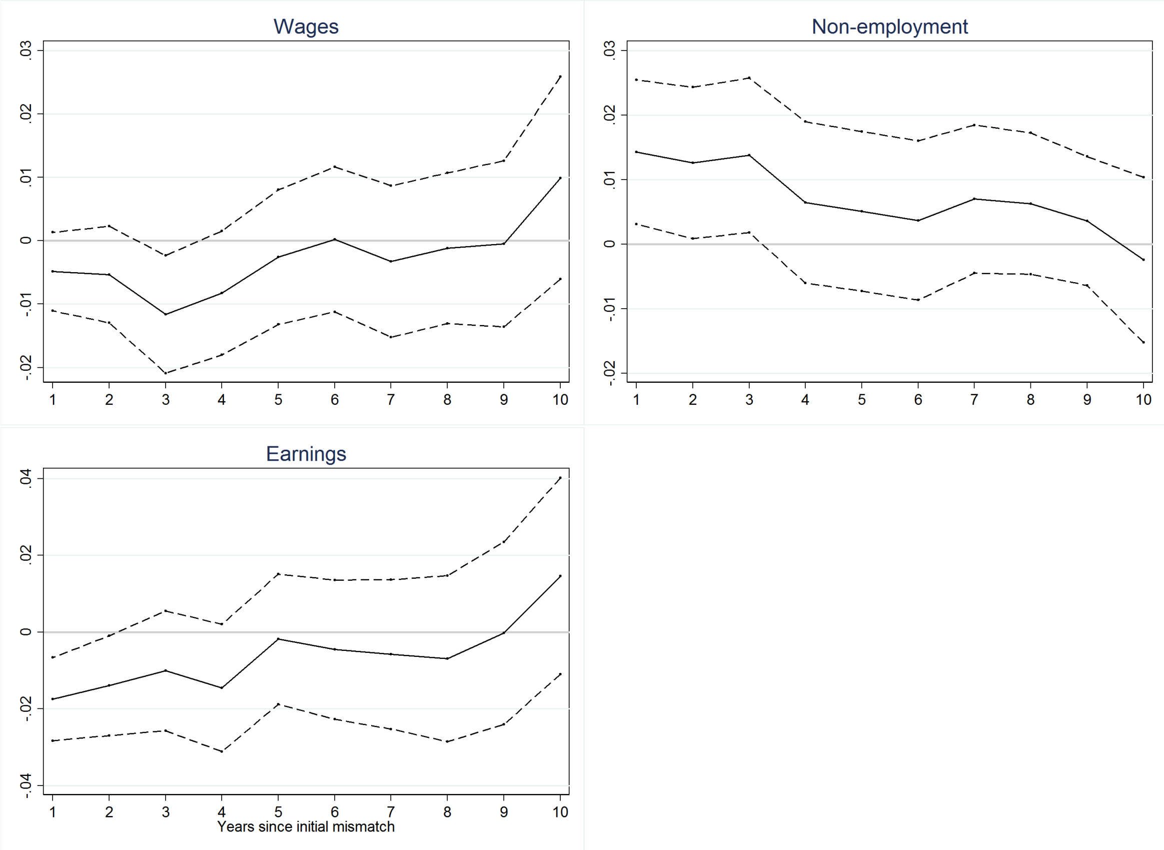
Finally, a potential concern is that experienced and inexperienced workers are found in different segments of the labor market. Now, if inexperienced workers are in the lower segments of the market, and there is wage compression from below, this may be one reason we observe that entry wages are unrelated to mismatch for inexperienced workers. However, we do not believe that this should be a major concern since we fail to find any systematic differences in the impact of mismatch when we stratify the analysis by job-level or education in Table 8 above. Furthermore, we find lower (log) wage dispersion among experienced workers in the low-skilled segments (0.205) than among inexperienced workers in high-skilled segments of the labor market (0.254). Hence, the fact that we find wage responses for low-skilled experienced workers, but not for high-skilled inexperienced workers, cannot be a function of differences in wage structures between these two groups.

6 Initial mismatch and overall earnings losses

Our stylized theoretical model (by assumption) does not deliver any predictions about the long-run consequences of mismatch. However, in order to provide insight into the labor market impact of mismatch it is useful to document the long run patterns and overall earnings responses. This section therefore provides some basic results regarding the long-run effects of mismatch on future wages, employment, and earnings. We focus on the inexperienced workers.

We use the first entry cohorts in our data (1997-99), which allow us to follow the same sample over 10 years. The long-run responses of wages (independently of whether individuals remained in the same job or changed jobs), non-employment and annual earnings to initial mismatch are shown in Figure 6. The first observation in each panel is for the

Figure 6: Initial mismatch and subsequent outcomes



Notes: The figure displays the response to initial mismatch on wages, the probability of being non-employed and annual earnings in year t+1 to t+10. The panels are based on the regression specification in eq. 9. The sample is restricted to entrants who we can follow the entire follow-up period (entrants in 1997-1999). Dashed lines are confidence bands.

year after matching. Non-employment initially increases (and earnings decrease). This initial response is consistent with the view that poor matches are destroyed when match quality is revealed. The non-employment (and earnings) responses fade away and point estimates are close to zero after 4-5 years. Over the longer haul, wages, non-employment, and earnings are completely unrelated to initial mismatch.

The point estimate for log earnings losses during the year after matching is -0.017 (se 0.006). However, the magnitude of this effect is likely to be deflated due to measurement errors in our measure of mismatch. In order to correct for measurement errors we can use the relationship between our main measure and the alternative measure discussed in section 5.3. Regressing this alternative mismatch measure on the baseline mismatch measure gives an estimate of 0.346 (se 0.027). If the measurement errors are uncorrelated between these measures, this suggests that a one standard deviation increase in true mismatch of talents leads to a 5 percent drop in log earnings one year after matching. Since the average inexperienced entrant has a mismatch score of 2.58 standard deviations from the optimal match, this suggest that mismatch reduces initial annual earnings by 13 percent for the average inexperienced entrant.

7 Conclusions

We have examined the direct impact of mismatch on wages and job mobility using unique Swedish data containing information on a multitude of talents, detailed occupational information, wages, and the indicators for the identity of the employer. Our empirical approach builds on the idea that any sorting model will imply that tenured workers are selected on having the right skills for the job. To measure mismatch we thus compare how well the talents of recently hired workers correspond to the talents of incumbent workers performing the same job.

As a prelude to our empirical analysis we show that each component of our vector of talents (inductive-, verbal-, spatial, and technical ability as well as social maturity, intensity, psychological energy, and emotional stability) is independently valued on the labor market, even conditional on educational attainment. We also show that workers in jobs with high returns to a specific skill, have higher than average levels of this skill. In addition, we document several novel facts about mismatch: First, mismatch is higher among inexperienced workers and workers who have entered from non-employment. Second, starting wages are unrelated to mismatch for inexperienced workers and workers hired from non-employment whereas experienced workers and job-to-job movers receive a wage penalty if they are mismatched. Third, we find a non-trivial separation response to mismatch among inexperienced workers and entrants from non-employment, whereas the impact for experienced workers and job-to-job movers is small. Fourth, the adjustment to mismatch is relatively fast: mismatch of talents predicts separation during the first year after recruitment but not thereafter. Fifth, wage growth within jobs is negatively related to initial mismatch, and this effect is more pronounced among inexperienced workers. Finally, we show that mismatch reduces annual earnings by 13 percent for the average inexperienced entrant in the year after the match was formed, primarily through an increased risk of non-employment. The earnings losses are fairly persistent, but disappear within five years.

We interpret the differential outcomes across groups as being a function of the information available at the time of hiring. For inexperienced workers, and entrants from non-employment, it is realistic to assume that both the prospective employee and the prospective employer fail to observe how well the detailed characteristics of the worker match the skill-requirements of the job. We therefore conclude that inexperienced workers, and those who search from non-employment, appear to match under uncertainty as in Jovanovic (1979). In contrast, the matching process for experienced job-to-job movers appears to be best described by a model where information about match quality is available already at the initial matching stage. Overall, the results support the notion that the misallocation of workers, and the inability to observe match quality for marginal applicants, is a fundamental problem in the labor market.

References

Abowd, John, Francis Kramarz and David Margolis (1999), 'High wage workers and high wage firms', *Econometrica* **67**(2), 251–333.

Abraham, Katherine G. and Henry S. Farber (1987), 'Job duration, seniority, and earnings', *American Economic Review* **77**, 278–297.

Altonji, Joseph G. and Charles R. Pierret (2001), 'Employer learning and statistical discrimination', *Quarterly Journal of Economics* **116**(1), 313–350.

Atakan, Alp E. (2006), 'Assortative matching with explicit search costs', *Econometrica* **74**(3), 667–680.

Bhuller, Manudeep, Magne Mogstad and Kjell G. Salvanes (2015), 'Life-cycle bias and the returns to schooling in current and lifetime earnings', *Journal of Labor Economics* (forthcoming).

Böhlmark, Anders and Matthew J. Lindquist (2006), 'Life-cycle variations in the association between current and lifetime income: Replication and extension for Sweden', *Journal of Labor Economics* **24**, 879–896.

Duncan, Greg J. and Saul D. Hoffman (1981), 'The incidence and wage effects of overeducation', *Economics of Education Review* **1**(1), 75–86.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg and Herbert Brucker (2015), 'Referral-based job search networks', *Review of Economic Studies* (forthcoming).

Eeckhout, Jan and Philipp Kircher (2011), 'Identifying sorting – in theory', *Review of Economic Studies* **78**, 872–906.

Farber, Henry S. (1999), Mobility and stability: The dynamics of job change in labor markets, *in* O.Ashenfelter and D.Card, eds, 'Handbook of Labor Economics', Vol. 3, North-Holland, pp. 2439–2483.

Farber, Henry S. and Robert Gibbons (1996), 'Learning and wage dynamics', *Quarterly Journal of Economics* **111**(4), 1007–1047.

Flinn, Christopher J. (1986), 'Wages and job mobility of young workers', *Journal of Political Economy* **94**(3), S88–S110.

Gautier, Pieter A. and Coen N. Teulings (2012), Sorting and the output loss due to search frictions. Manuscript, VU University Amsterdam.

Gautier, Pieter A., Coen N. Teulings and Aico van Vuuren (2010), 'On-the-job search, mismatch and efficiency', *Review of Economic Studies* **77**, 245–272.

Grönqvist, Erik, Björn Öckert and Jonas Vlachos (2010), The intergenerational transmission of cognitive and non-cognitive abilities, Discussion Paper 7908, CEPR.

Guvenen, Fatih, Burhanettin Kuruscu, Satoshi Tanaka and David Wiczer (2015), Multidimensional skill mismatch, Working Paper 21376, NBER.

Håkansson, Christina, Erik Lindqvist and Jonas Vlachos (2015), Firms and skills: The evolution of worker sorting, Working Paper 2015:9, IFAU.

Helpman, Elhanan, Oleg Itskhoki and Stephen Redding (2010), 'Inequality and unemployment in a global economy', *Econometrica* **78**, 1239–1283.

Hensvik, Lena and Oskar Nordström Skans (2014), 'Social networks, employee selection and labor market outcomes', *Journal of Labor Economics* (forthcoming).

Jovanovic, Boyan (1979), 'Job matching, and the theory of turnover', *Journal of Political Economy* **87**, 972–990.

Lange, Fabian (2007), 'The speed of employer learning', *Journal of Labor Economics* **25**(1), 1–35.

Lazear, Edward P. (2009), 'Firm-specific human captial: A skill-weights approach', *Journal of Political Economy* **117**(5), 914–940.

Lindqvist, Erik and Roine Vestman (2011), 'The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment', *American Economic Journal: Applied Economics* **3**(1), 101–128.

Marimon, Ramon and Fabrizio Zilibotti (1999), 'Unemployment vs. mismatch of talents: Reconsidering unemployment benefits', *Economic Journal* **109**(455), 266–291.

Mood, Carina, Jan O. Jonsson and Erik Bihagen (2012), Socioeconomic persistence across generations: The role of cognitive and non-cognitive processes, *in* J.Ermisch, M.Jäntti and T. M.Smeeding, eds, 'From Parents to Children: The Intergenerational Transmission of Advantage', Russell Sage Foundation, chapter 3.

Nagypal, Eva (2007), 'Learning by doing vs. learning about match quality: Can we tell them apart?', *Review of Economic Studies* **74**, 537–566.

Roy, A.D. (1951), 'Some thoughts on the distribution of earnings', *Oxford Economic Papers* **3**(2), 135–146.

Sattinger, Michael (1975), 'Comparative advantage and the distribution of earnings and abilities', *Econometrica* **43**, 455–468.

Teulings, Coen N. and Pieter A. Gautier (2004), 'The right man for the job', *Review of Economic Studies* **71**, 553–580.

Tinbergen, Jan (1956), 'On the theory of income distribution', *Weltwirtschaftliches Archiv* **77**, 156–173.

A1 Appendix

A1.1 Details of the model

Here we consider the model predictions in closer detail. First we derive p_0, the probability of separation. Let $C = 1$ indicate the event that the signal is correct. Then $p_0 = \Pr(C = 0)\Pr(d > d_s | E_0(d) < d_m, C = 0) + \Pr(C = 1)\Pr(d > d_s | E_0(d) < d_m, C = 1)$. The first term in this expression is simply equal to $(1-\alpha)(1-d_s)$. $\Pr(d > d_s | E_0(d) < d_m, C = 1) > 0$ if and only if $m \equiv E(d) + (d_m - E(d))/\alpha > d_s$. Since $d_m < d_s$, this cannot happen for all values of α. In particular, $\Pr(d > d_s | E_0(d) < d_m, C = 1) > 0$ if and only if $\alpha < \bar{\alpha} \equiv (d_m - E(d)/(d_s - E(d)) < 1$. Thus

$$\Pr(d > d_s | E_0(d) < d_m, C = 1) = \begin{matrix} \frac{d_m - E(d) - \alpha(d_s - E(d))}{\alpha E(d) + d_m - E(d)} = 1 - \frac{d_s}{m} \text{ if } \alpha < \bar{\alpha} \\ 0 \text{ if } \alpha \geq \bar{\alpha} \end{matrix}$$

where m denotes the number of matches (given that the signal was informative). In sum we can write the probability of separation as

$$p_0(\alpha) = (1-\alpha)(1-d_s) + \alpha I(\alpha < \bar{\alpha})\left(1 - \frac{d_s}{m(\alpha)}\right)$$

where $I(\cdot)$ denotes the indicator function. If $\alpha < \bar{\alpha}$, this is an implicit function in p_0, since the number of matches depends on p_0 via the matching threshold d_m.

The fact that d_0 is bounded by the (0,1) interval also implies that α is bounded from below. In particular, the upper bound on d_0 implies that α must be greater than

$$\underline{\alpha} = \frac{d_m(\underline{\alpha}) - E(d)}{1 - E(d)}$$

Now $0 \leq \underline{\alpha} < \bar{\alpha} = [(d_m(\bar{\alpha}) - E(d))/(d_s - E(d))]$. Notice that $d_m > E(d)$ is a requirement for the market to exist for all values of α; notice also that $d_m(\alpha)$ is a positive function of α via its dependence on $p_0(\alpha)$. Thus, if we require that $d_m(\alpha) \to E(d)$ (from above) when $\alpha \to 0$, then $\underline{\alpha} \to 0$. So, if we assume that the agents are basically indifferent between matching and waiting when the signal is very imprecise, the extreme case $\alpha \to 0$ is part of the solution. For future reference it is useful to note that $m(\underline{\alpha}) = 1$ and $m(\bar{\alpha}) = d_s$.

We begin by showing that p_0 is decreasing in α. Intuitively, this should be the case. And it is straightforward to verify that $p_0(\underline{\alpha}) = 1 - d_s$ (since $m(\underline{\alpha}) = 1$), $p_0(\bar{\alpha}) = (1-\bar{\alpha})(1-d_s)$ (since $m(\bar{\alpha}) = d_s$), and $p_0(1) = 0$. The elasticity of the non-separation margin with respect to α is given by

$$\eta(\alpha) \equiv -\frac{\partial p_0}{\partial \alpha}\frac{\alpha}{1-p_0} = \begin{matrix} \frac{\alpha d_s(1-m+\psi)}{(1-p_0)m\Omega} > 0 \text{ if } \alpha < \bar{\alpha} \\ \frac{\alpha(1-d_s)}{(1-p_0)} > 0 \text{ if } \alpha \geq \bar{\alpha} \end{matrix}$$

where $\psi \equiv \frac{d_m - E(d)}{\alpha m} < 1$ and $\Omega \equiv 1 + \frac{d_s}{m}\frac{d_s - d_m}{(1-p_0)m} > 0$. Now $\eta(\alpha) < 1$. (Suffice it to note that $\eta(1) = (1-d_s) < 1$; $\eta(\bar{\alpha}) = \bar{\alpha}(1-d_s)/(d_s + \bar{\alpha}(1-d_s)) < 1$; and $\eta(\underline{\alpha}) = (d_m(\underline{\alpha}) - d_s)/(1 + d_s - d_m(\underline{\alpha})) < 1$).

1. Exposure to initial mismatch From (6), it follows that

$$\frac{\partial m}{\partial \alpha}\frac{\alpha}{m} = -\psi\left[1 - \Delta(\alpha)\eta(\alpha)\right]$$

where $\Delta(\alpha) = (d_s - d_m(\alpha))/(d_m(\alpha) - E(d)) > 0$. Increasing α has a direct negative effect and an indirect (positive) effect, via the dependence of d_m on p_0 (with an increase in α, p_0 declines, and therefore d_m increases). Since $\eta(\alpha) < 1$, a sufficient condition for the direct effect to be larger than the indirect effect is that $\Delta(\alpha) < 1$. Since $\Delta'(\alpha) < 0$, it suffices to find a condition that guarantees that $\Delta(1) < 1$. If $c > b + \gamma E(d)$, then

$$\frac{\partial m}{\partial \alpha}\frac{\alpha}{m} = -\psi\left[1 - \Delta(\alpha)\eta(\alpha)\right] < 0$$

The meaning of the condition $c > b + \gamma E(d)$ is that the net cost associated with waiting $(c - b)$ is greater than the production loss associated with the mean of the potential mismatch distribution.

2. Initial mismatch and entry wages From (3) it follows that entry wages are falling in d:

$$\frac{\partial w_0}{\partial d} = -\frac{(1-p_0)\gamma\alpha^2}{2} \leq 0$$

And so

$$\frac{\partial^2 w_0}{\partial d \partial \alpha} = -\gamma\alpha(1-p_0)\left[1 + \frac{\eta}{2}\right] \leq 0$$

3. Initial mismatch and the separation rate The separation rate is given by: $s = p_0 = (1-\alpha)(1-d_s) + \alpha I(\alpha < \bar{\alpha})(1 - \frac{d_s}{m})$. For a marginal match (i.e. a match where $d \to d_s$), we have $\partial s/\partial d = -\partial s/\partial d_s$, and therefore

$$\frac{\partial s}{\partial d} = (1-\alpha) + \frac{\alpha I(\alpha < \bar{\alpha})}{m} \geq 0$$

It is straightforward to verify that

$$\left.\frac{\partial s}{\partial d}\right|_{\alpha=\underline{\alpha}} = 1 > \left.\frac{\partial s}{\partial d}\right|_{\alpha=\bar{\alpha}} = 1 - \bar{\alpha} > \left.\frac{\partial s}{\partial d}\right|_{\alpha=1} = 0$$

On the interval $\alpha \in [\bar{\alpha}, 1]$, $\partial^2 s/\partial d \partial \alpha < 0$; on the interval $\alpha \in [\underline{\alpha}, \bar{\alpha})$, $\partial^2 s/\partial d \partial \alpha > 0$, however. In particular

$$\frac{\partial^2 s}{\partial d \partial \alpha} = \begin{array}{c} \frac{1}{m}\left[1 - m - \frac{\partial m}{\partial \alpha}\frac{\alpha}{m}\right] > 0 \text{ if } \alpha < \bar{\alpha} \\ -1 < 0 \text{ if } \alpha \geq \bar{\alpha} \end{array}$$

4. Initial mismatch and wage growth within jobs Define $\Delta w = w(d) - w_0(d)$, where $w(d)$ is given by (8) and $w_0(d)$ by (3). We have

$$\frac{\partial \Delta w}{\partial d} = -\frac{\gamma}{2}\left[1 - (1 - p_0)\alpha\right] \leq 0$$

and

$$\frac{\partial^2 \Delta w}{\partial d \partial \alpha} = -\frac{\partial^2 w_0}{\partial d \partial \alpha} = \gamma\alpha(1 - p_0)(1 + \frac{\eta}{2}) \geq 0$$

5. Variance of mismatch by tenure Finally we show that the variance of the observed mismatch distribution declines with tenure. This relates to the point that we should observe a decline in the variance of talents with tenure if mismatch is relevant.

The change in the variance of the observed mismatch distribution (Δvar) is given by

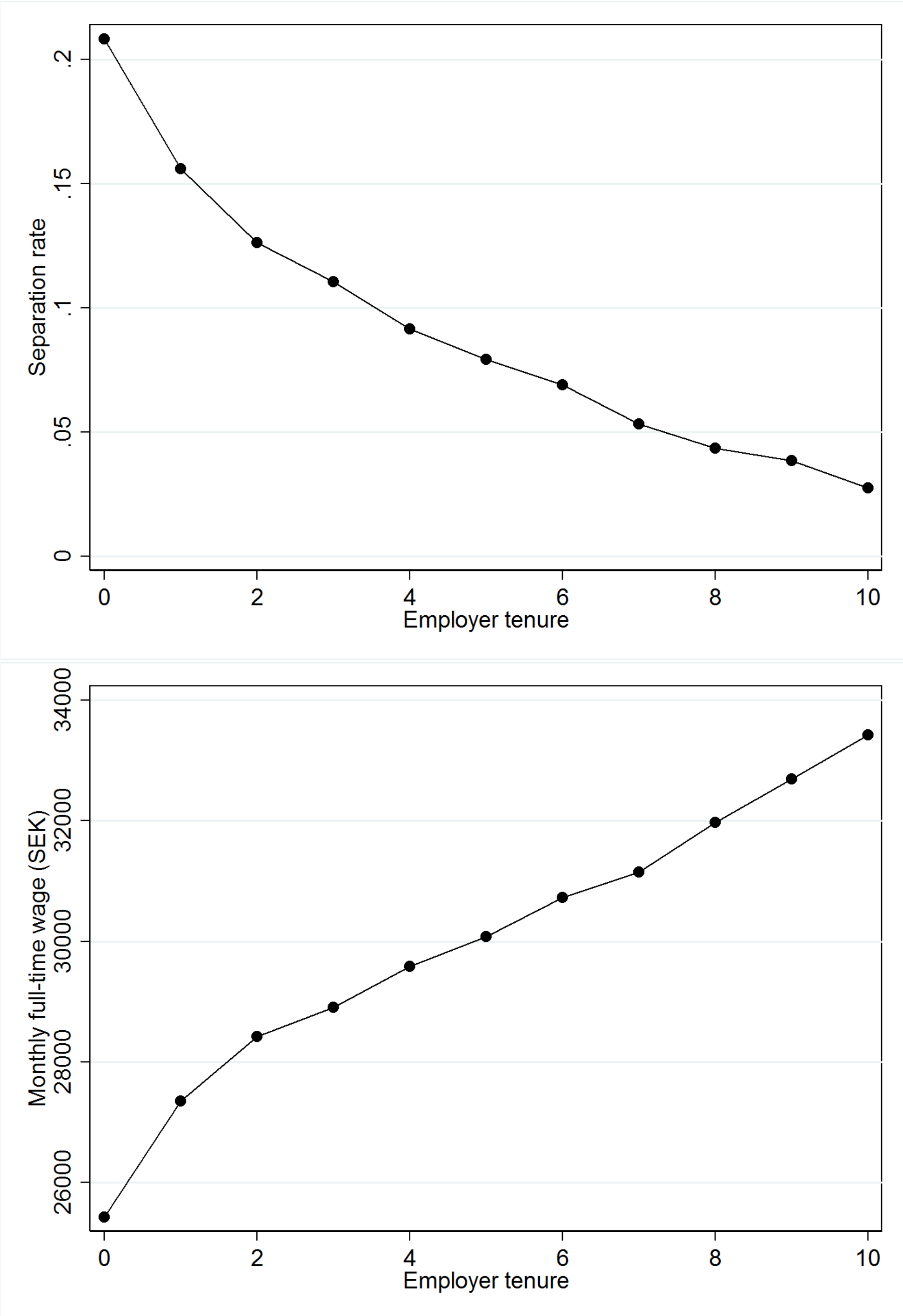
$$\Delta \text{var} = -\left[(1 - \alpha)(1 - d_s^2) + \alpha I(\alpha < \bar{\alpha})(m^2 - d_s^2)\right] \text{var}(d) \leq 0$$

It follows that $\Delta\text{var}(\underline{\alpha}) = -(1 - d_s^2)\text{var}(d) < \Delta\text{var}(\bar{\alpha}) = -(1 - d_s^2)(1 - \bar{\alpha})\text{var}(d) < \Delta\text{var}_{\alpha \to 1} = 0$. In general

$$\frac{\partial \Delta \text{var}}{\partial \alpha} = \begin{array}{c} \text{var}(d)\left[(1 - m^2) + 2m^2 \frac{\partial m}{\partial \alpha}\frac{\alpha}{m}\right] > 0 \text{ if } \alpha < \bar{\alpha} \\ \text{var}(d)(1 - d_s^2) > 0 \text{ if } \alpha \geq \bar{\alpha} \end{array}$$

A1.2 Additional descriptives

Figure A1: Separations and wages by tenure



Notes: Figure A shows the share of workers separating from the job for each level of worker tenure and Figure B shows the average monthly full-time equivalent wage for each level of worker tenure.

Table A1: All male entrants 1997-2008

	mean	SD	median
Separation rate	.29	.46	0
Age	36.4	8.0	36
Experience at entry	11.3	5.5	12
Entry from employment	.73	.44	1
Entry establishment size	144	498	22
Education:			
Primary school less than 7 years	.02	.13	0
Primary 7-9 years	.11	.31	0
High school short (less than 2 years)	.03	.18	0
High school short (2 years)	.28	.45	0
High school long (3 years)	.18	.38	0
College short (less than 2 years)	.11	.32	0
College short (2 years)	.07	.26	0
College long (3 years)	.11	.31	0
College long (4 years)	.08	.27	0
PhD short (Licentiate)	.11	.04	0
PhD long (Doctoral)	.01	.08	0
Observations	2,784,253		

A1.3 Estimates by prior employment status

Table A2: The impact of mismatch on wage growth within job

	(1)	(2)	(3)
	From non-emp.	**From job**	**P-val for diff.**
Mismatch	-0.0278	-0.0172*	0.6136
	(0.0225)	(0.0097)	
Observations	7,472	30,429	
R-squared	0.8049	0.6252	
Education FE:s	$\surd$	$\surd$	
Entrant test scores	$\surd$	$\surd$	
(Entry occupation×Entry Year×Plant) FE:s	$\surd$	$\surd$	

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The specification is the same as in Table 6.

Table A3: Robustness: Estimates by previous employment status

	(1)	(2)	(3)	(4)
	Entry wages		Separations	
	From non-empl.	**Job-to-job**	**From non-empl.**	**Job-to-job**
A. Baseline				
Mismatch	-0.0013	-0.0118***	0.0114***	0.0062***
	(0.0019)	(0.0010)	(0.0057)	(0.0020)
Observations	28,247	125,234	28,247	125,234
R-squared	0.8766	0.8378	0.6189	0.4879
B. All skill interactions				
Mismatch	-0.0018	-0.0110***	0.0119*	0.0068***
	(0.0020)	(0.0010)	(0.0061)	(0.0021)
Observations	28,247	125,234	28,247	125,234
R-squared	0.8769	0.8379	0.6195	0.4880
C. Non-linearities in mismatch				
Mismatch	-0.0008	-0.0105***	0.0087	0.0067***
	(0.0020)	(0.0010)	(0.0063)	(0.0021)
$Mismatch^2$	-0.0004	-0.0017***	0.0023	-0.0006
	(0.0006)	(0.0004)	(0.0022)	(0.0007)
Observations	28,247	125,234	28,247	125,234
R-squared	0.8766	0.8378	0.6189	0.4880
D. No restriction on # tenured workers				
Mismatch	-0.0018	-0.0106***	0.0129**	0.0052***
	(0.0019)	(0.0008)	(0.0065)	(0.0018)
Observations	58,175	274,588	58,175	274,588
R-squared	0.9204	0.8839	0.7617	0.6350
E. Weighted mismatch index				
Mismatch	-0.0012	-0.0119***	0.0112*	0.0069***
	(0.0019)	(0.0010)	(0.0058)	(0.0020)
Observations	28,247	125,234	28,247	125,234
R-squared	0.8766	0.8378	0.6189	0.4880
F. Cognitive vs. non-cognitive ability				
$Mismatch_{cognitive}$	-0.0004	-0.0076***	0.0082	0.0042**
	(0.0018)	(0.0009)	(0.0052)	(0.0018)
$Mismatch_{non-cognitive}$	-0.0015	-0.0087***	0.0073	0.0044*
	(0.0022)	(0.0011)	(0.0066)	(0.0023)
Observations	28,247	125,234	28,247	125,234
R-squared	0.8766	0.8378	0.6189	0.4879
G. Mismatch (based on overall cognitive and non-cognitive scores)				
Mismatch	-0.0007	-0.0065***	0.0094**	0.0034**
	(0.0015)	(0.0007)	(0.0046)	(0.0015)
Observations	28,247	125,234	28,247	125,234
R-squared	0.8765	0.8377	0.6190	0.4879
Education FE:s	√	√	√	√
Entrant test scores	√	√	√	√
(Entry occupation×Entry Year×Plant) FE:s	√	√	√	√

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The specification is the same as in Tables 4 and 5.

A1.4 Monthly data

In addition to the annual employment records used for the main analysis, we have information on the first and last month of remuneration from each employer. We use this information to measure the length (in months) of each employment spell. As described in Section 3, our wage and occupation data are collected during a measurement week once every year (in September-November depending on the employer). Therefore, we calculate the monthly employment duration for entrants who started a new job in August-October, in order to obtain a reliable mapping between the starting month and the entry wage/occupation. The average job spell lasts for 35 months, almost three years.

One potential concern is that the first and last month of compensation are self-reported by the employers, which increases the risk of measurement error. In our sample, 35 percent of the separations occur in December (conditioning on entry in August-October), which seems high even if we consider that a disproportionate number of employment relationships are likely to terminate in December for natural reasons. For the sake of our analysis it is however important to remember that such measurement error will only be a problem if the probability of misreporting is correlated with the degree of initial mismatch, which seems highly unlikely.

A1.5 Worker fixed effects

As an additional robustness check we introduce individual fixed effects. These fixed effects obviously hold all time-invariant characteristics of the individual constant, and thus take the direct effect of individual skill into account. The advantage is that any unobserved dimensions of worker ability (and outside options), potentially not captured by the test scores, are accounted for.

There are two disadvantages, however. Introducing worker fixed effects is extremely taxing on the data, since it requires repeated observations per worker. Thus a given worker must be recorded as a new hire, at least twice. Apart from the obvious sample reduction caused by the elimination of those that were recorded as new entrants once, there is a further reduction caused by the sampling of establishments in the wage data. Second, workers who are repeat new hires may be non-representative for the population of new hires; along the observed dimension they are slightly less experienced (10.9 yrs. compared to 11.3 yrs.) and (by construction) tend to be job-to-job movers to a somewhat greater extent (85 compared to 82 percent).

To deal with the first problem we are forced to pool all experience groups. To provide a comparison, column (1) of Table A4 shows the estimates from the baseline specification for all new hires, when the inexperienced and experienced are pooled together. Column (2) shows the baseline specification for repeat new hires (notice that the sample is reduced to 27 percent of the original sample); despite our concerns the estimates are comparable to column (1). Column (3) finally shows the results when we introduce worker fixed effects. We think the estimates are reassuringly stable across specifications. The entry wage response to mismatch is lower than in the baseline specification, while the separation response is somewhat higher. Of course, the worker fixed effects estimates are substantially less precise, and the differences in the estimates across columns (2) and (3) are not statistically significant.

Table A4: Responses to mismatch with worker fixed effects

	(1)	(2)	(3)
	Baseline	Baseline	Worker FE:s
	All new hires	Repeat new hires	Repeat new hires
$Mismatch$	-0.0109***	-0.0118***	-0.0054
	(0.0008)	(0.0020)	(0.0049)
Observations	154,681	41,309	41,309
R-squared	0.7818	0.8771	0.8730
$Mismatch$	0.0072***	0.0086*	0.0110
	(0.0017)	(0.0045)	(0.0077)
Observations	154,681	41,309	41,309
R-squared	0.4472	0.5990	0.4777
Education dummies	$\checkmark$	$\checkmark$	
Entrant test scores	$\checkmark$	$\checkmark$	
(Entry occupation×Entry Year×Plant) FE:s	$\checkmark$	$\checkmark$	
Entry occupation			$\checkmark$
Entry Year			$\checkmark$
Worker FE:s			$\checkmark$

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors (reported in parentheses) are robust to heteroscedasticity. Column (1) shows the baseline estimates when we pool both experience groups. Column (2) restricts the sample to workers who we observe entering a new job at least twice and column (3) include worker fixed effects.

A1.6 Variance of skills and tenure

One implication of the theory outlined in Section 2 is that pre-hire differences between inexperienced and experienced workers should be smaller among those that remain within jobs since the worst matches are destroyed. We test this prediction by calculating the average skill dispersion within each job (j), experience (at entry) group (e) and tenure (τ) as:

$$\sigma^2_{je\tau} = \frac{1}{K}\sum_{k=1}^{K}\sigma^2_{kje\tau}$$

Then, we examine how the dispersion of skills vary with experience and tenure using the following equation:

$$\sigma^2_{je\tau} = \delta_1 Inexp. + \delta_2 Inexp. \times Tenure + \delta_3 Tenure + \lambda_j + \epsilon_{je\tau} \tag{A1}$$

where λ_j denotes "job" (Occupation×Year×Plant) fixed effects. Table A5 suggests that there is higher variability of skills among inexperienced entrants compared to entrants who accumulated more pre-hire experience (column 1) or who come from non-employment (column 2). However, as expected the difference with respect to experience falls with tenure, suggesting that remaining inexperienced workers become more like remaining experienced workers.

Table A5: Skill dispersion and tenure

	(1)	(2)
Inexperienced	0.0097	
	(0.0062)	
Inexperienced×Tenure	-0.0116***	
	(0.0034)	
From non-employment		0.0321***
		(0.0060)
From non-employment×Tenure		-0.0065*
		(0.0035)
Tenure	-0.0087***	-0.0094***
	(0.0015)	(0.0000)
Observations	531,633	531,633
Adj. R-squared	0.7165	0.7166
(Occupation×Year×Plant) FE:s	√	√

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the mean variance in skills within the job-experience group-tenure cell.