Mismatch of Talent

Evidence on Match Quality, Entry Wages, and Job Mobility *

Peter Fredriksson† Lena Hensvik‡ Oskar Nordström Skans§

1 March 2018

Abstract

We examine the impact of mismatch on entry wages, separations, and wage growth using unique data on worker talents. We show that workers are sorted on comparative advantage across jobs within occupations. The starting wages of inexperienced workers are unrelated to mismatch. For experienced workers, on the other hand, mismatch is negatively priced into their starting wages. Separations and wage growth are more strongly related to mismatch among inexperienced workers than among experienced workers. These findings are consistent with models of information updating, where less information is available about the quality of matches involving inexperienced workers.

Keywords: Matching, Job search, Comparative advantage, Employer learning

JEL-codes: J64, J24, J31, J62

*First draft May 2013, first complete version 2 July 2014. We thank Erling Barth, David Card, John Earle, Jan Eeckhout, Pieter Gautier, Georg Graetz, Francis Kramarz, Lisa Laun, Rasmus Lentz, Edwin Leuven, Erik Lindqvist, Andreas Madestam, Espen Moen, Dale Mortensen, Arash Nekoei, Åsa Rosén, Robert Shimer, Uta Schönberg, the editor, and three anonymous referees, as well as participants at the Nordic Data Meeting 2013, SOLE 2014, EALE 2014, Nordic Summer Labor Institute 2015, 3d CReAM Topics in Labour Workshop, LSE, Queen Mary, Royal Holloway, University of Texas at Austin, Warwick, Ecole Polytechnique, Tinbergen, BI in Oslo, DIW, Stockholm University, UCLS, University of Oslo, and VATT for useful comments and suggestions.

†Department of Economics, Uppsala University, Box 513, SE-75120 Uppsala, Sweden (Email: peter.fredriksson@nek.uu.se), UCLS, and IZA. Funding from Marcus and Amalia Wallenberg Foundation and Handelsbanken is gratefully acknowledged.

‡Institute for Evaluation of Labour Market and Education Policy (IFAU), Box 513, SE-751 20 Uppsala, Sweden (Email: lena.hensvik@ifau.uu.se), Uppsala Center for Labor Studies (UCLS), and CESifo. Funding from Ragnar Söderberg Foundation and Handelsbanken is gratefully acknowledged.

§Department of Economics, Uppsala University, Box 513, SE-751 20 Uppsala, Sweden (Email: oskar.nordstrom_skans@nek.uu.se), UCLS, IFAU and IZA. Funding from Ragnar Söderberg Foundation is gratefully acknowledged.

1 Introduction

The allocation of workers across jobs is crucial for wage dispersion, labor productivity and overall efficiency.[1] Idiosyncratic match quality is also a fundamental aspect of several important theoretical contributions on a wide set of topics, including Marimon and Zilibotti (1999) on 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.

This paper asks two key questions: How does the variation in match quality impact entry wages, job separation, and subsequent wage growth? To what extent does the impact of match quality depend on uncertainty about match quality at the hiring stage? In contrast to much of the literature, we base our analysis on a direct measure of match quality (or mismatch), constructed using unique data on worker talents.

Deriving direct and credible evidence on the relationship between uncertainty and job-level match quality has proven difficult. Much of previous work builds on Jovanovic (1979) where match quality is unobserved at the time of hiring, but realized ex post. Thus, revelation of match quality leads to separations when news are bad, and above-average wage growth among stayers when news are good. The traditional empirical approach for evaluating these predictions has been to examine how separations and wages evolve with tenure (e.g., Abraham and Farber 1987, Flinn 1986, and Farber 1999). But a drawback of this strategy is that 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 wages respond to a direct measure of job-level match quality. Our measure of match quality 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) and four traits evaluated by a trained psychologists (social maturity, intensity, psychological energy and emotional stability). Importantly, these variables are measured before labor market entry and hence not endogenous to the outcome of the job-search process. Furthermore, the measures are strongly related to labor market outcomes later in life, but they

[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 (2015) calibrate such a model, and conclude that actual allocations imply very large efficiency losses.

[2]Dustmann et al. (2016) 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.

are not (directly) observed by the employers. Our basic presumption is that each of the particular talents are differentially productive in different jobs, as in Gibbons and Waldman (2004) and Lazear (2009). This should (and do, according to our data) lead to sorting over jobs by comparative advantage among tenured workers, as in Roy (1951). As a consequence, we can use talents among tenured workers as a proxy for the skill requirements of each particular job. In practice, we use this proxy to form a measure of each entrant's distance to optimal match quality, i.e., a measure of *mismatch*. The empirical strategy assumes that potential productive benefits of employing a diverse set of personality types primarily occur across (and not within) different jobs within the same production process. Our results support this assumption.

We base our analysis on economy-wide Swedish administrative data covering both workers and firms. We measure the eight talents for both entrants and incumbents within the same jobs. In line with standard search theory (e.g, Eeckhout and Kircher 2011), we think of jobs as tasks performed within firms. We therefore identify jobs by combining indicators of occupations and establishments. Our analysis data cover more than 300,000 new matches, which makes the job-level analysis possible. Throughout, we control for all observed and unobserved confounders at the job level by including job-year fixed effects. This implies that we analyze the impact of variation in mismatch between different entrants who start the same job, during the same year. Our models also account for the overall market valuation of each talent.

A key objective of the paper is to study the hypothesis that pre-hire uncertainty about match quality affects both the amount of job-level mismatch, and the consequences thereof. We isolate cases where the pre-hire information sets are likely to be limited by drawing on the employer-learning literature pioneered by Farber and Gibbons (1996) and Altonji and Pierret (2001). The basic idea is that labor market experience reduces uncertainty about worker skills and that the market gradually learns about worker abilities. Empirically, these studies use pre-market test scores as proxies for information that is only partially observed by employers who hire inexperienced workers. In our case the focus is slightly different since we study match quality, and not overall market returns. Therefore, we only need to assume that the market is less informed about how well the detailed characteristics of a worker match the detailed skill requirements for each particular job when workers lack prior labor market experience. We validate the results by comparing workers with some prior experience from the same firm to experienced workers without intra-firm experience, and by contrasting experienced job-to-job movers with experienced workers who match from non-employment.

Our analysis is conceptually related to a parallel strand in the literature focusing on occupational mismatch (see, e.g., Gathmann and Schönberg 2010, Groes et al. 2015, Guvenen et al. 2015, and Lise and Postel-Vinay 2016). In line with this emerging literature, our results suggest that the occupational level accounts for a non-trivial component of the overall job-level sorting on talents. But we also demonstrate that a substantial part of sorting (around 50%) is across jobs with the same occupational classifications. Moreover, we correlate estimated job-specific returns to a particular talent with the supply of talent in each job and show that the association is positive, non-trivial in magnitude, and only partly explained by sorting of talented workers in general. The results thus imply substantial sorting on *comparative* advantages across jobs.

Our main results show that match quality at the start of a job matters for subsequent labor market outcomes, and that the impact crucially depends on the available pre-hire information set. Match quality is better among experienced workers, even conditional on job-specific fixed effects and detailed controls for individual skills. Match quality is particularly high among experienced workers who have worked in the firm before, or among experienced workers who are hired directly from another job. Among inexperienced workers, we find that match quality is unrelated to entry wages, but strongly related to separations and wage growth. By contrast, experienced workers receive a wage penalty if they are mismatched, but separation and wage-growth responses are instead more moderate. The results thus suggest more initial uncertainty (more mismatch, no association between mismatch and initial wages) and more learning (a strong impact of mismatch on separations and wage growth) for the inexperienced. Learning is most pronounced for workers with less than 5 years of experience, which is well in line with traditional employer-learning estimates focusing on ability levels (see Lange 2007).[3] For experienced workers, our results are consistent with frictional search, but since initial uncertainty is much lower for this group, there is also less learning. For matches involving experienced workers, the agents accept a smaller set of match qualities (i.e., there is less mismatch) and price the remaining mismatch into wages already from the onset.

To get a sense of the overall costs associated with mismatch, we examine the evolution of subsequent labor earnings. A year after being exposed to mismatch, the earnings loss associated with a standard deviation increase in mismatch amounts to eight percent for the average inexperienced worker.[4] This overall effect reflects a combination of lower wages and an increase in non-employment. The earnings trajectories converge over time and the effects disappear after 4 years.

Our results are robust to a large set of variations of the model. We vary the measurement of mismatch to allow for market-wage weights for each talent and present various other specification checks regarding measurement and functional form. The results are also robust to very flexible controls for worker skills. We further show that the effects are similar for different market segments (high/low skilled). Finally, and perhaps most importantly, we validate the results by using an alternative measure of mismatch built on estimated job-specific wage returns to each of the eight talents; these results are less precise, but well-aligned with those of our main strategy.

Overall, we interpret our set of results as strong evidence for the notion that idiosyncratic job-level match quality has an impact on labor mobility and wages. The mechanisms vary

[3] In the Online Appendix we show that learning within the job is much faster since most of the separation response occurs within a year with a focal point of around six months after the hire.

[4] This estimate is corrected for measurement error.

across groups, however: Ex ante uncertainty and learning are important for inexperienced workers whereas the patterns for experienced workers (in particular job-to-job movers and workers with previous experience from the same firm) are consistent with models of search frictions with little ex ante uncertainty about match quality.

The paper is structured as follows: Section 2 presents a conceptual framework that motivates a set of micro-level predictions. Section 3 describes the data and documents sorting on talents across jobs. Section 4 presents the main empirical results. Section 5 presents a large set of robustness tests and discusses alternative explanations for our results. Section 6 studies earnings trajectories and Section 7 concludes.

2 Conceptual framework

To frame our empirical analysis, this section outlines the basic predictions of a stylized matching model incorporating match-specific productivity and initial uncertainty about this match-specific component. The model can be seen as a simplified version of Jovanovic (1984), albeit with some small differences.[5] We therefore relegate the formal presentation of the model to the Online Appendix (Section A1), where we also derive the predictions outlined here.

2.1 Basic structure

There is undirected search on the part of workers. A (potential) match between a worker and a firm is characterized by some level of match quality. In line with our empirical approach, we think of match quality as the distance from the optimal match – *mismatch* for short. Match productivity is, by definition, falling in mismatch.

At the hiring stage, information is imperfect. Thus, the worker-firm pair initially observes a noisy signal of match-specific productivity. On the basis of the signal, the two parties decide on whether to match, and, if so, agree on an entry wage. Entry wages (and all subsequent wages) are determined by a surplus sharing rule, which allocates a fixed share of the expected surplus of the match to the worker.

As production commences, the worker-firm pair observes productivity, which may yield new information on match quality. Based on this new information, agents decide on (jointly efficient) separations and negotiate new wages for remaining workers.

2.2 Predictions

Here we summarize four predictions that we take to the data. They are all driven by lack of information at the matching stage. When the initial signal is very imprecise, we expect all potential matches to be realized and entry wages will be unrelated to match quality. With a very imprecise initial signal, information updating will also be maximal; thus we expect the

[5] The key differences are that: (i) match quality is bounded by the optimal match; (ii) wages are determined by a sharing rule; and (iii) the outside option is modeled as in Eeckhout and Kircher (2011).

separation response to mismatch, and the impact on wage growth, to be much larger (in the absolute sense) than with a precise initial signal.

1. A less precise initial signal increases initial mismatch When there is little information about match quality, agents will always match (given that the market exists). Conversely, a precise signal truncates the potential distribution of mismatch more than an imprecise one. If the distribution of potential mismatch does not vary with the precision of the initial signal, higher match rates translate into greater exposure to mismatch.

2. A less precise initial signal weakens the negative impact of mismatch on entry wages In the extreme case where the initial signal is very imprecise, the expected surplus of the match at the hiring stage does not depend on mismatch. Since the entry wage is determined by a surplus-sharing rule, the entry wage will be unrelated to mismatch. Conversely, with a more informative initial signal, mismatch is priced into entry wages.

3. A less precise initial signal strengthens the positive impact of mismatch on separations When the initial signal is very imprecise, all potential matches will be realized. But in such a setting, there will also be more scope for information updating when the worker firm pair receives more precise information coming from the fact that they observe productivity. The more information updating there is, the stronger is the separation response.

4. A less precise initial signal strengthens the negative association between mismatch and wage growth within job, among those who do not separate For the set of matches which are not dissolved, learning will eventually be complete and mismatch fully reflected in wages. This holds true no matter the information content of the initial signal; however, there will be more learning if the initial signal was very imprecise. We thus expect wage growth to be more strongly related to initial mismatch among continuing matches in cases where the initial signal was very imprecise.

To examine these predictions, we need a measure of mismatch and proxies for the amount of information available at the time of the match. In the next section we present how we measure mismatch and describe our information proxies. We discuss other interpretations of our results in Section 5.5. We find that these alternative interpretations cannot explain the full range of the results.

3 Data, measurement and descriptive analyses

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 as well as age, education and annual earnings for all Swedish workers during 1985-2008. Information on (full-time equivalent) wages and occupational codes are available for a very large sample of establishments covering almost 50 percent of all private sector workers and all public sector workers from 1997.[6] Information from the draft (detailed below) is available for the period 1969 and 1994. During these years, almost all males went through the draft procedure at age 18 or 19, and enlistment scores are available for 80% of the sample. Thus, military draft scores are available for 25 cohorts of males born between 1951 and 1976. Our sample includes male workers born during 1951-76 who enter new jobs (entrants/new hires) between 1997 and 2008, as well as their tenured male coworkers (incumbents) from the same set of birth cohorts.[7]

3.1 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 sub-tests 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 (mean = 0, standard deviation = 1) within each cohort of draftees.

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.[8] Again, we standardize the scores (mean = 0, standard deviation = 1) within each cohort.

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. 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. 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

[6] 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 of private sector firms is stratified by firm size with the sampling probabilities 3%, 12%, 41%, 70% and 100% for the firm size intervals 1-9, 10-49, 50-199, 200-499 and 500- respectively. Note that the sampling probabilities depend on firm size which implies that small establishments belonging to larger private firms still have a high chance of being included. The wage measure reflects the wage the employee had during the sampling week expressed in full-time monthly equivalents. All wage components, except overtime pay, are included. Thus, the wage measure includes, e.g., piece-rate, performance pay as well as fringe benefits.

[7] Table A1 describes the sample selection process in more detail.

[8] There is also an overall psychological score on a Stanine scale, ranging from 1 to 9, which has been used in Lindqvist and Vestman (2011), Håkansson et al. (2015) and Hensvik and Skans (2016), for example. We use these cruder data in a robustness exercise in Section 5.

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
Adjusted R-square	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. 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. [A]Labels for non-cognitive scores are according to Mood et al. (2012).

attainment constant). The adjusted R2 increases from 0.346 to 0.386 when adding skill measures to a regression with controls for education (and, implicitly holding age and gender constant). As a comparison, adding a second order polynomial in tenure and log firm size raises adjusted R2 by half of this (to 0.364). Most importantly, however, the results in Table 1 show that there is independent, and sufficiently precise, variation in the individual measures of talent.[9]

Table 2 provides a further illustration of the information contained in the skill measures, by showing evidence of the skill diversity across medium-skill occupations.[10] In particular, the table lists the occupation with the highest score along each specific dimension; it also reports the dimension along which workers in a given occupation is least endowed, and the wage rank of the occupation.

Table 2 shows, for instance, that (male) nurses are most heavily endowed with social

[9]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.

[10]Table A4 in the Online Appendix reports analogous information for low- and high-skill occupations.

Table 2: Skill endowments across medium-skill occupations

Skill	Most endowed occupation	Skill endowments			Wage rank
Non-cognitive:		Specific	Average	Least endowed	
Social maturity	Nurses (313)	0.29	0.18	Tech. (=0.03)	0.61
Intensity	Forestry workers (614)	0.33	-0.03	Spat. (=-0.23)	0.26
Psychological energy	Placement officers etc. (342)	0.21	0.15	Tech. (=-0.07)	0.64
Emotional stability	Fire fighters and security guards (515)	0.19	0.05	Spat. (=-0.15)	0.30
Cognitive:					
Inductive	Librarians (243)	0.66	0.15	Int. (=-0.44)	0.56
Verbal	Librarians (243)	0.83	0.15	Int. (=-0.44)	0.56
Spatial	Photographers, image and sound recording (313)	0.29	0.18	Int. (=-0.14)	0.55
Technical	Photographers, image and sound recording (313)	0.38	0.18	Int. (=-0.14)	0.55

Notes: The table focuses on occupations in the middle tercile of the average skill distribution; Table A4 reports analogous information for low- and high-skill occupations. ISCO-codes are reported within parentheses. Data are from 2002 and contain individuals with at least 3 years of tenuree. The wage ranks pertain to the 2002 wage distribution. Labels for non-cognitive scores are according to Mood et al. (2012).

maturity (0.29 standard deviations above average) but score lower on technical abilities; fire fighters score high on emotional stability (i.e., the ability to cope with stress) but they are low on spatial ability; librarians score high on verbal and inductive ability but low on intensity (the ability to take initiative without external pressure). Overall, these associations follow an intuitive pattern; they illustrate the heterogeneity of skill requirements across occupations, and that there is sorting across occupations depending on the talents measured during the draft. We delve into the sorting of skills across jobs within occupations below.

3.2 Jobs and sorting among tenured workers

3.2.1 Jobs

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). This definition of a job allows for the possibility that different job openings use different skill sets also within an occupational category. The reason why we let the job definition be year-specific is in part technical: As explained below, the distinction between new hires and tenured workers is crucial for our analysis; by letting jobs be year-specific we do not use the time dimension for identification and therefore retain an unambiguous separation between incumbents and tenured workers within a job. An additional benefit of this approach is that it allows for technological evolution within cells defined by the combinations of occupations and plants.

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). Throughout, we rely on register data on hires and are thus not able to separate between selection at different stages (e.g. application, interviews, offers, acceptance) of the matching process. 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.[11]

3.2.2 Sorting among tenured workers

Our basic presumption is that different jobs require different sets of skills. To validate this presumption, we first describe the extent of sorting on talents across jobs. As in studies of workplace segregation (e.g., Åslund and Skans 2009), we use the association between own and coworker attributes as a measure of systematic sorting. We focus on workers with at least 3 years of tenure and regress own talents on coworker talents along each of the eight dimensions (i.e., we use 8 observations per worker and year).

Column (1) of Table 3 shows very strong within-job clustering of talents: the correlation between own and coworker talents is 0.89. To isolate sorting on *specific* talents (as opposed to sorting on talent in general), column (2) includes a control for the average ability level among coworkers. Three quarters of sorting remains after adding this control, implying that workers are mainly sorted into jobs on the specific *types* of skills they have.[12]

Our focus on job-level search and matching complements an emerging empirical literature on occupational search (see, e.g., Guvenen et al. 2015). To highlight the value added of using job-level data, we repeat our analysis while controlling for the talent endowment among *all other workers in the same occupation*; see column (3). This exercise shows that half of the job-level sorting remains after accounting for sorting at the occupational level. The final column adds occupation by talent (by year) fixed effects to the specification in column (2). The specification in column (4) thus effectively includes the same covariates at the occupational level as at job level. In comparison to column (2) the correlation is reduced from 0.61 to 0.28, which again suggests that roughly half of the job-level sorting remains after accounting for sorting at the occupational level.

Although accounting for the coworkers skill levels, and all aspects of occupational sorting, reduces the correlation by two thirds (from 0.89 to 0.28), the remaining association is large. We conclude that he labor market is characterized by strong "horizontal" sorting on specific abilities across jobs within occupations.

3.2.3 Sorting of tenured workers according to job-specific wage returns

In this sub-section 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. The question of interest is whether workers sort on these returns.

[11] 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.

[12] Online Table A3 shows 8 regressions, one for each talent ($k = 1, .., 8$). The main takeaway is that the strongest correlation between individual and coworker talent is for the particular talent under consideration.

Table 3: Skill sorting over jobs

	(1)	(2)	(3)	(4)
Mean of skill k, coworkers in job	0.8862***	0.6113***	0.4911***	0.2815***
	(0.0033)	(0.0075)	(0.0064)	(0.0055)
Mean of all skills, coworkers in job		0.3539***		0.3770***
		(0.0065)		(0.0078)
Mean of skill k, occupational peers			0.5281***	
			(0.0069)	
Observations	15,559,712	15,559,712	15,559,712	15,559,712
R-squared	0.1688	0.1712	0.1801	0.1800
Year FE:s	√	√	√	
Year*Occupation*Talent FE:s				√

Notes: Standard errors clustered on the plant by occupation level are reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The unit of observation is a talent by worker combination (i.e., 8 observations per worker). The dependent variable is the standardized individual endowment of talent k among tenured workers. The independent variable of interest is the average endowment of talent k among tenured coworkers in the same job. Column (1) shows the raw association between individual and coworker endowments. Column (2), in addition, accounts for average coworker endowments along all 8 dimensions. Column (3) adds the endowment of other tenured workers in the same occupation to the specification in (1). In column (4), we add occupation*talent*year FE:s (making the year FE:s and all occupational characteristics redundant) to the specification in (2).

As a first step in this analysis, we 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. We therefore run 60,500 separate wage regressions, one for each job-cell where we have at least 10 tenured workers (the regressions also control for age). We then relate these estimates to the intensity of the skill within each job. To this end, we generate a data set with one observation for each combination of skill and job (i.e., 8 observations per job), and run regressions with the estimated returns as the dependent variable (as it is measured with error); Table 4 reports the results.

Column (1) shows the results of a regression where the job-specific return to a particular skill is related to the job-specific intensity of that skill. The association between the skill-specific return and the skill-specific intensity is positive. Column (2) proceeds by controlling for the average skill intensity in order to isolate sorting on *comparative* advantages. Column (3), finally, shows the results of a model with job fixed effects that remove all confounding aspects of the job; the results are identical to the ones in column (2). Workers are (on average) found in jobs where the returns to their talents are higher than average, as suggested by Roy (1951). The magnitudes imply that in jobs where the endowment of a particular skill is a standard deviation higher than in an average job, the return to that skill is twice as large (the average job-level return to a specific skill is estimated to be 0.0062), holding all aspects of the job constant.[13]

[13]Online Figure A1, shows all job-specific skill-return estimates plotted against the skill endowments within the same jobs, separately for each type of talent. In 7 out of 8 cases, the correlation is positive and statistically

Table 4: Job-specific returns and skill levels among tenured workers

	(1)	(2)	(3)
Amount of skill k	0.0082***	0.0059***	0.0059***
	(0.0003)	(0.0008)	(0.0008)
Average amount of all skills		0.0030***	–
		(0.0009)	
Observations	200,952	200,952	200,952
R-squared	0.6968	0.6969	0.6970
Year FE:s	√	√	
Job FE:s			√

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The unit of observation is the plant.The dependent variable is the estimated job-specific return to a skill k among tenured workers. The independent variable of interest is the average amount of skill k among tenured workers in a particular job. Column (1) shows the raw correlation between returns and endowments. In column (2) we account for the average skills among tenured workers. In column (3) we use job (occupation×plant×year) FE:s (which subsume the year FE:s). The mean of the dependent variable is 0.0062.

3.3 Mismatch as entrant-incumbent skill differences

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. In principle, we could use the wage estimates discussed above to build a measure of match-quality, and we also do that in a robustness exercise (see Section 5). However, the estimate job-level returns to talents are very imprecise due to the small number of workers observed within each job.[14] 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 selected (through both entry and exit) on having the right set of talents for the job (see Table 4). Thus, the skill sets of tenured workers is a useful proxy for the skill requirements of each job. Our basic assumption is that a given job requires a certain set of talents; we thus discard the benefits of diversity within jobs (if any) and our estimates will be attenuated if this assumption is not approximately true (we return to this question in a robustness analysis in Section 5). It should be noted, however, that we allow for the possibility that production may benefit from a workforce with diverse talents *across jobs.*

For the purpose of the empirical analysis, we focus on entrants and on tenured workers with at least 3 years of tenure in the current job. All results remain unchanged if we require that the jobs employ at least 10 tenured males with non-missing draft scores (i.e. only using cases where precision is higher), see Section 5. Our generic empirical strategy is to focus on

significant.

[14]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.

the importance of mismatch after removing the direct importance of the vector of individual skills (s_i) through a flexible function $g(s_i)$ as well as the direct importance of coworker skills. In practice we soak up the direct impact of coworker skills by means of job fixed effects (λ_j); the job fixed effects also remove potential associations between match quality and other job-characteristics such as pecuniary or non-pecuniary job amenities. To estimate the impact of mismatch on some outcome Y we thus compare different entrants into the same job, while accounting for the market valuation of their skills, i.e.,[15]

$$Y_{ij} = \beta Mismatch_{ij} + g(s_i) + \lambda_j + \epsilon_{ij} \tag{1}$$

To quantify $Mismatch$, our baseline strategy is to use the distance between the skills of the worker and the skill requirements of the job:

$$Mismatch_{ij} = \sum_{k=1}^{8} |s_{ik} - \bar{s}_{jk}| \tag{2}$$

where s_{ik} denotes the amount 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.

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 and 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 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,

$$\frac{\partial Y_{ij}}{\partial s_{ik}} = \begin{cases} -\beta + g_k'(s_i) \text{ if } s_{ik} < \bar{s}_{jk} \\ \beta + g_k'(s_i) \text{ if } s_{ik} > \bar{s}_{jk} \end{cases} \tag{3}$$

Over-skilled workers are thus not fully rewarded for their talents ($\beta < 0$), but this does not

[15] 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. However, a key difference from the traditional AKM model is that we focus on *entry* wages (in the spirit of Abowd et al., 2006). Since the sample of repeated entrants is small and selected, we prefer to use a parametric function to control for the impact of person characteristics; in our baseline results, $g(s_i)$, is specified as a second order polynomial in each of the eight talents. Section 5.1 and A3.2 show that the results are robust to more flexible polynomial specifications of $g(s_i)$ as well as person fixed effects.

imply that they are paid less than lower-skilled (perhaps perfectly matched) co-workers.[16] Workers who have the right *average* skill level, but the "wrong" composition of talents, receive lower wages if $\beta < 0$.

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. Section 5 also discusses results from using a mismatch index based on firm-specific wage returns to each of the talents.

Given that the measure of mismatch is standardized, the estimates reflect the impact of a standard deviation increase in mismatch. These estimates should be viewed as lower bounds; measurement errors are unavoidable since our proxy for the job skill requirement ($\bar{s}_{jk}$) is an imprecise measure of the actual requirement. In section 6 we provide an exercise where we adjust for these errors.

3.4 Proxies for information

3.4.1 Labor market experience

Our empirical analysis aims to contrast groups where match productivity should be 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 remains largely unobserved among inexperienced workers. For experienced workers, on the other hand, the employer arguably has more information 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 their 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 few years after graduation. Along similar lines, Hensvik and Skans (2016) show, using Swedish data similar to ours, that the wage returns to test scores increase with tenure, and that this pattern is more pronounced among the inexperienced.

Labor market experience is defined as the number of years which the individual is classified as being employed.[17] Since this information is available from 1985 onward, we truncate

[16]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 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.

[17]We use Statistics Sweden's classification system relying on register data on monthly earnings (in Novem-

experience at 13 years of experience for all entrant cohorts. The median entrant in our sample is 36 years old, and has 12 years of experience (see Table 5). 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.[18]

3.4.2 Alternative information proxies

We use two alternative approaches to isolate variation in the pre-hire information set among the group of experienced workers. First, we classify matches on the basis of whether the worker has worked for the same firm, but in another establishment, at some point in the past. Relative to the average experienced worker, there should be more information available about this group of workers; therefore, we expect less exposure to mismatch, more pricing of mismatch into entry wages, and less information updating.

Second, we contrast job-to-job movers with hires from non-employment. Here, we expect there to be less information available about match-specific productivity for those who are hired directly from non-employment, e.g., because their job references date further back. We define *job-to-job* movers as workers who were employed in the previous year and treat all others as entrants from non-employment.

We only define these proxies for experienced workers. The reason for this is that the sample of experienced workers is sufficiently large to permit a further division into subgroups (see Table 5).[19] Notice also that doing the analysis within the experienced group implies that we ask whether the alternative information proxies adds information relative to the amount of information available in matches involving experienced workers in general. We analyze how these alternative information proxies are related to the exposure to mismatch in Section 4.1 and leave the analysis of how they change the impact of mismatch on wages and separations to Section 5.4.1.

3.5 Analysis data

Table 5 shows the characteristics of the new hires (entrants) in our sample. Since our analysis requires at least one male tenured coworker (from the draft cohorts) within the same job as the entrants, our sample consists of larger establishments (387 employees on average) than an overall sample of entrants (144 employees) during the same time period (see Table A2). The separation rate, defined as the probability of leaving the establishment within the first year after being hired, equals 23 percent in our sample (29 percent in the overall sample);

ber). This register data definition of employment is calibrated to correspond to employment according to the Labor Force Surveys.

[18] We also show results for a finer classification of experience groups, and these results corroborate our main findings.

[19] The Online Appendix present separate results for the inexperienced group. The estimates are less precise than for the experienced group, but the signs of the estimates are consistent with our underlying information story; see Table A7.

Table 5: Entrants 1997-2008

	All		Inexp. 0-4 yrs	Exp. 5+ yrs
	(1)	(2)	(3)	(4)
	mean (SD)	median	mean (SD)	mean (SD)
Separation rate	.23		.28	.22
ln(Entry wage)	10.04(.36)	9.97	9.78 (.24)	10.08 (.36)
Age	36.8 (8.0)	36.0	27.3 (4.2)	38.4 (7.3)
Experience at entry	11.6 (5.8)	12.0	2.3	13.2
Job-to-job mobility	.82		.46	.88
Prior within firm experience	.29		.11	.32
Entry establishment size	387 (909)	92	424 (890)	379 (911)
Education:				
Compulsory or less	0.08		0.05	0.08
High school	0.42		0.33	0.44
College	0.50		0.62	0.48
Entry occupation:				
Legislators, senior officials and managers	.07		.03	.07
Professionals	.29		.36	.28
Technicians and associate professionals	.25		.20	.26
Clerks	.04		.06	.04
Service workers and shop sales workers	.09		.12	.08
Skilled agricultural and fishery workers	.01		.00	.01
Craft and related trades workers	.09		.05	.10
Plant machine operators and assemblers	.12		.12	.12
Elementary occupations	.04		.05	.04
Skills:				
Cognitive	.20 (.98)		.31 (1.01)	.18 (.97)
Non-cognitive	.15 (.98)		.14 (1.02)	.15 (.97)
Mismatch	.00 (1.00)	-.19	.02 (1.02)	.00 (1.00)
Observations	328,651		47,360	281,291

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

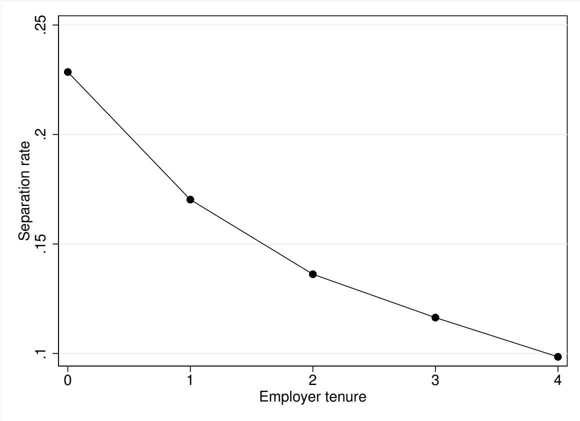
see the first row of Table 5. Inexperienced workers have a higher separation rate, a lower share of job-to-job movers, and fewer workers with some prior experience with the firm. For illustration, the table categorizes occupations at a crude 1-digit ISCO level. Two thirds of the sample are "professionals", "technicians", or "machine operators".[20] Table 5 also shows that the average new hire is more skilled than the average member of a cohort; moreover, there appears to be stronger selection on cognitive skill among the inexperienced than the experienced.[21] The final row of the table shows the values of the (standardized) mismatch index. Inexperienced workers are mismatched to a greater extent than experienced workers. We return to this issue in the Section 4.1.

[20]We explore the extent to which our key results vary between different occupational levels in Section 5.

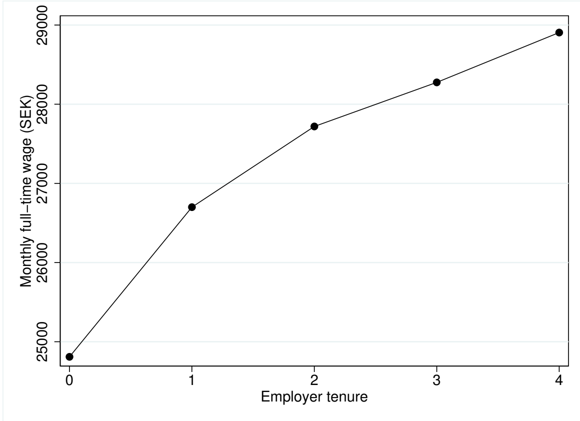
[21]This is due to the inexperienced having more education than the experienced. Notice that our regressions include detailed educational attaiment FE:s and flexible controls for ability scores.

Figure 1: Separations, wages and within job variance in talents by tenure

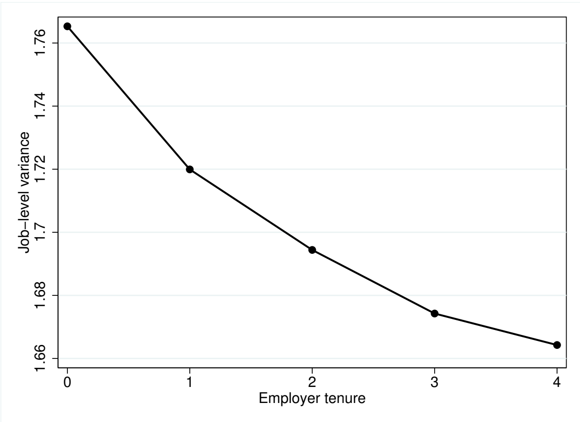
(a) Separations



(b) Wages



(c) Within job variance in talents



Notes: Panels (a) and (b) show the separation rate and mean wage by tenure. Panel (c) shows the variance in the eight talents among workers within the same job and tenure category.

Figures 1 (a) and (b) show how separation probabilities and wages evolve with tenure in the cross-section. Consistent with the earlier literature, the relationship between tenure and separation is negative whereas the relationship between wages and tenure is positive. Figure 1 (c) shows that the variance in our (pre-determined) talents within jobs falls with tenure. This suggests that job-level sorting is partly due to higher separation rates among workers whose talents differ from their coworkers.

Online Figure A2 graphs tenure-profiles of separations and wages separately for experienced and inexperienced workers. Experience is measured at the start of the new job (i.e., when tenure = 0). The wage and separation profiles are steeper for inexperienced workers, in particular during the first year, consistent with more learning in matches involving inexperienced workers. Along the same lines, the within-job variance in talents is falling much faster for the inexperienced than for the experienced. Finally, Figure A2 (d) shows the variance of wages and how it evolves with job tenure. The evolution of the wage variance within jobs reflects two opposing forcing. If match quality is more or less known from the outset, selection (i.e., that the worst matches are destroyed) implies that the wage variance falls with tenure. If initial match quality is uncertain, learning will imply that wages are increasingly tied to match quality. Learning per se implies that the wage variance gradually increases with tenure. The evolution of the wage variance across the two groups (experienced and inexperience) line up strikingly well with these conjectures. Among the experienced, selection is the predominant force since the wage variance falls with tenure. Among the inexperienced, learning appears to be the dominant force since the wage variance is increasing in tenure.

4 Results

4.1 Information and exposure to initial mismatch

As argued in section 3.4, initial match quality is more likely to be unobserved for inexperienced workers. We should thus observe more mismatch among inexperienced workers than among experienced workers. To test this prediction, we run regressions of the following kind:

$$Mismatch_{ij} = \sum_x \alpha_x x_i + g(s_i) + \gamma Z_i + \lambda_j + \epsilon_{ij} \tag{4}$$

where i refers to individuals, j to "jobs" ($j = occupation \times plant \times entry\,year$), and x_i denotes experience group indicators (the omitted reference category has 13+ years of experience); $g(s_i)$ is a 2nd order polynomial in each of the eight individual talents; Z_i includes age and 11 educational attainment fixed effects;[22] and λ_j are job fixed effects.

Figure 2 shows the estimates on the experience group indicators (i.e. $\hat{\alpha}_x$). It is clear that mismatch decreases with the amount of previous labor market experience at job entry. The results are thus consistent with there being less information, and therefore more mismatch, in matches involving inexperienced workers.

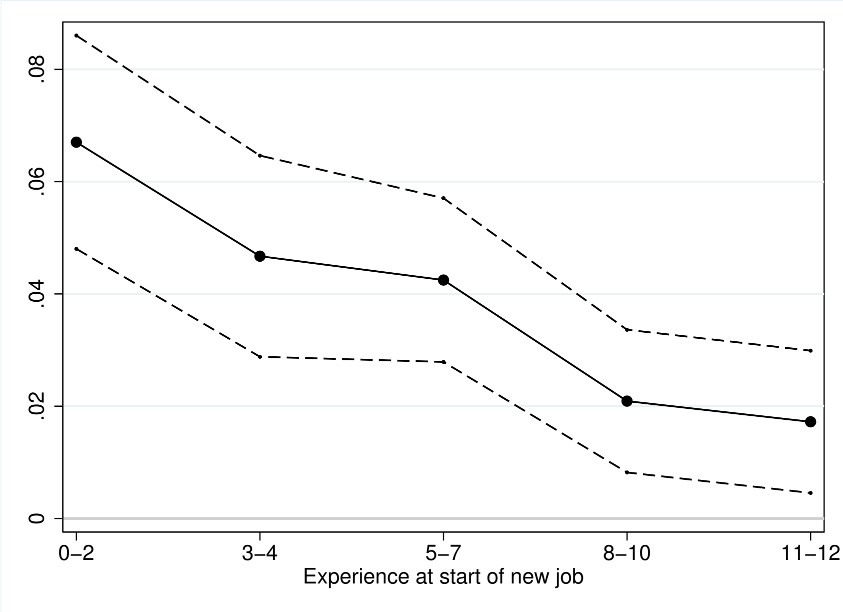
Column (1) of Table 6 compares the experienced (5+ years of experience) with the inexperienced (0-4 years of experience). Mismatch is 0.028 standard deviations lower among the experienced than among inexperienced workers.

Column (2) instead focuses on heterogeneity among the group of experienced workers. It adds an indicator for previous employment at another plant within the same firm, as well as dummies for each year of labor market experience in the $[5, 12]$ range. As expected, workers who have worked at another plant within the same firm in the past are exposed to better match quality.

Column (3) instead includes a dummy for job-to-job movers (relative to hires from non-employment), again controlling for experience non-parametrically. The results show that mismatch is lower among experienced job-to-job movers. This is consistent with hires from unemployment having a wider acceptance set, due to greater uncertainty or other factors.

[22] From now on, we divide the education attainment 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.; College long, 4 yrs.; PhD short (Licentiate); PhD long (Doctoral).

Figure 2: Initial mismatch by experience



Notes: The figure plots estimates of α_x; see equation (4). Workers with 13+ years of experience is the reference category. Mismatch and experience is measured at the start of the new job. Dashed lines are 95% confidence bands.

Considering that our models account for the direct impact of talents, the level of education, age and job fixed effects (i.e., the analysis compares workers entering the same jobs), we view the evidence in Table 6 as strongly suggesting that inexperienced workers (and outside hires) are less well-matched at the start of a new job than experienced hires. The results are thus consistent with the prediction that match-quality-signals are more noisy for inexperienced workers. However, without exploring the impact of mismatch on wage trajectories and separations, we cannot rule out that match quality differs because of frictions rather than information uncertainty. In what follows, we explore exactly these dimensions, focusing on labor market experience as the information proxy; we return to the alternative information proxies (prior firm experience and job-to-job mobility) in Section 5.4.

4.2 Initial mismatch and entry wages

To examine the prediction that lower match quality reduces entry wages when the signal is sufficiently informative, we analyze how entry wages are related to mismatch (at the time of the hire) across experience groups:

$$\ln(Entry\,Wage_{ij}) = \sum_x \beta_x^w x_i Mismatch_{ij} + g^w(s_i) + \sum_x \alpha_x^w x_i + \gamma^w Z_i + \lambda_j^w + \epsilon_{ij}^w. \tag{5}$$

Notation is analogous to equation (4); $g^w(s_i)$ is thus a second order polynomial in each of the eight talents (we provide robustness checks with more flexible functional forms in Section 5) and λ_j^w denote job fixed effects. As discussed above, we include the flexible skill controls to

Table 6: Determinants of mismatch

	(1)	(2)	(3)
	Baseline measure	Alternative measures for the experienced	
	Labor market experience	*Within-firm experience*	*Job-to-job transition*
Experienced (>4 yrs.)	-0.0275***		
	(0.0060)		
Firm experience		-0.0228***	
		(0.0065)	
Job-to-job			-0.0154**
			(0.0076)
Observations	328,651	281,291	281,291
R-squared	0.7886	0.8008	0.8008
Education FE:s	√	√	√
Entrant test scores	√	√	√
Job 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. Within-firm experience is an indicator for having some prior experience from another plant within the same firm. Job FE:s are (Entry occupation×Entry Year×Plant) FE:s. Experience FE:s control for each year of experience.

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.

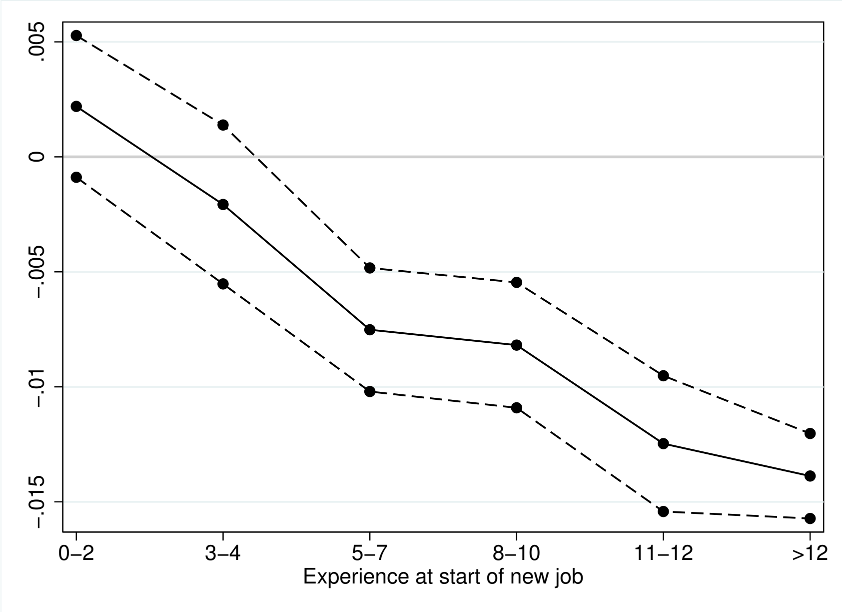
Equation (5) also includes a complete set of interactions between the experience indicators (x_i) and mismatch.[23] We are primarily interested in how estimates of β_x^w vary across experience groups. Figure 3 thus plots estimates of the coefficient on the mismatch index by experience group.[24] We expect 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 the least experienced workers. For more experienced workers, on the other hand, we find a significant negative effect on entry wages. For workers with more than 12 years of experience at the start of the new job, a standard deviation increase in mismatch lowers the entry wage by 1.4 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).

Panel A of Table 7 presents the results of separate regressions for the inexperienced (0-4 years of experience) and the experienced (5+ years of experience) (column (1) is for the full

[23]The complete set of interactions thus subsumes the main effect of mismatch.

[24]The fact that the the estimates are most precise for the group with more than 12 years of experience has to do with the fact that most of the sample belongs to this group.

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



Notes: Each dot is an estimate of the entry wage response to initial mismatch by experience group. The sample consists of entrants in 1997-2008. Dashed lines are 95% confidence bands.

sample).[25] Column (2) displays the results for inexperienced workers, confirming that the entry wage is unrelated to initial mismatch in this group. The coefficient estimate is very small (-0.0022) and precisely determined. Column (3) instead shows results for workers with at least 5 years of prior experience at the start of the new job. Among these workers, a standard deviation increase in mismatch reduces wages by 1.2 percent. The difference across experience groups is statistically significant.[26]

In Section 5 below, we discuss a number of exercises that validate the robustness of these results and discuss the role of our two alternative information proxies (prior within-firm experience and job-to-job mobility).

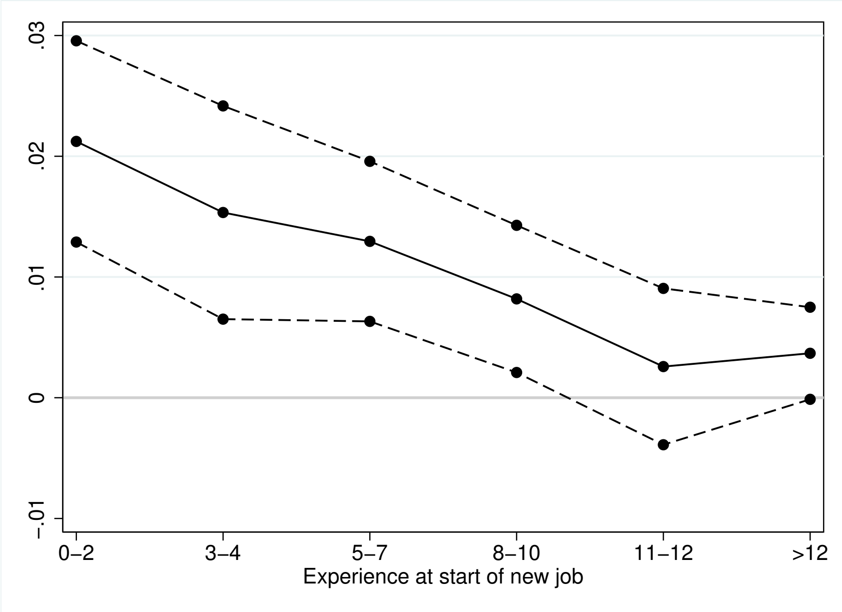
4.3 Initial mismatch and separations

The relationship between separations and mismatch should be the opposite of the entry-wage response. If mismatch is unobserved at the time of hiring, higher mismatch leads to separations as long as the production loss associated with revealed mismatch is higher than the separation cost (see Online Section A1). To examine the validity of this prediction, we run regression models identical to (5), but with the first-year separation rate as the outcome:

[25] By estimating separate models we allow the job fixed effects and all controls to be different for inexperienced/experienced. In this respect, these models are more flexible than the pooled regression defined by (5). From a practical point of view, the results do not depend on how flexibly we enter the job fixed effects and the control variables.

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

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



Notes: Each dot is an estimate of the separation response to initial mismatch by experience group. The sample consists of entrants in 1997-2008. Dashed lines are 95% confidence bands.

$$1^{st}\ year\ Separation_{ij} = \sum_{x} \beta^s_x x_i Mismatch_{ij} + g^s(s_i) + \sum_{x} \alpha^s_x x_i + \gamma^s Z_i + \lambda^s_j + \epsilon^s_{ij} \tag{6}$$

Figure 4 plots the estimates of the coefficients of interest ($\hat{\beta}^s_x$). For the least experienced workers, we find that a standard deviation increase in mismatch raises separations by 2.1 percentage points. The impact is considerably smaller for experienced workers: beyond 10 years of experience, the relationship between separations and mismatch is statistically insignificant.

Panel B of Table 7 reports the results for a split sample analysis corresponding to the wage analysis above. A standard deviation increase in mismatch increases separations among inexperienced workers by 2.1 percentage points. This corresponds to 8 percent of the average separation probability for this group. The impact is considerably lower among experienced workers: a standard deviation increase in mismatch raises separations by 0.6 percentage points. Again, the difference across groups is statistically significant; we analyze the robustness of the results, as well as the role of prior within-firm experience and job-to-job mobility, in Section 5.

We have also studied the timing of the separation response using higher frequency (monthly) data. The results, presented in the Online Appendix (Section A3.3), suggest that most of the separations occur during the first six months.[27]

[27] The results thus suggest that employing firms learn faster than the market, consistent with the notion of asymmetric learning discussed in, e.g., Schönberg (2007).

4.4 Initial mismatch and wage growth within jobs

The impact of mismatch on wage growth among remaining workers 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 match surplus. We examine the validity of this prediction by estimating the relationship between mismatch and wage growth by experience group, using a model which is completely analogous to the starting wage/separation models explained above:

$$\Delta \ln Wage_{ij} = \sum_x \beta_x^d x_i Mismatch_{ij} + g^d(s_i) + \sum \alpha_x^d x_i + \gamma^d Z_i + \lambda_j^d + \epsilon_{ij}^d \tag{7}$$

We calculate wage growth as the 3-year difference in log wages for individuals who have stayed in the same plant. Unfortunately, the sample is reduced to around a quarter of the original size for two reasons. First, the wage data are collected via sampling and we lose plants that randomly exit the sampling frame. Second, wage growth within job is only observed for the selected sub-sample that stay on in the same job.[28]

The results shown in Figure 5 suggest that there is more information updating among inexperienced workers, as evidenced by the fact that wage changes are negatively related to (revelations of) mismatch in matches involving this group of workers. Panel C of Table 7 shows results from split samples. Inexperienced workers (column 2) who are subjected to a standard deviation increase in initial mismatch have 5.7 percent lower wage growth. For experienced workers, the estimate in column (3) corresponds to a much lower reduction in wage prospects (-1.5 percent). The difference across the two groups is substantial (-4.1 percent) and in line with our prediction; however, it is not statistically significant at conventional levels (see column 4). If we instead estimate a pooled model where mismatch is interacted with an inexperienced dummy (akin to Figure 5), we find similar estimates with a difference (-3.2 percent) between groups which is statistically significant at the 5 percent level.

Overall, the results in Table 7 are well in line 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; for the same reason, the separation and wage growth responses are lower among experienced workers than among inexperienced workers. All of this suggests that mismatch to a large extent is factored in already at the time of hiring for experienced workers.

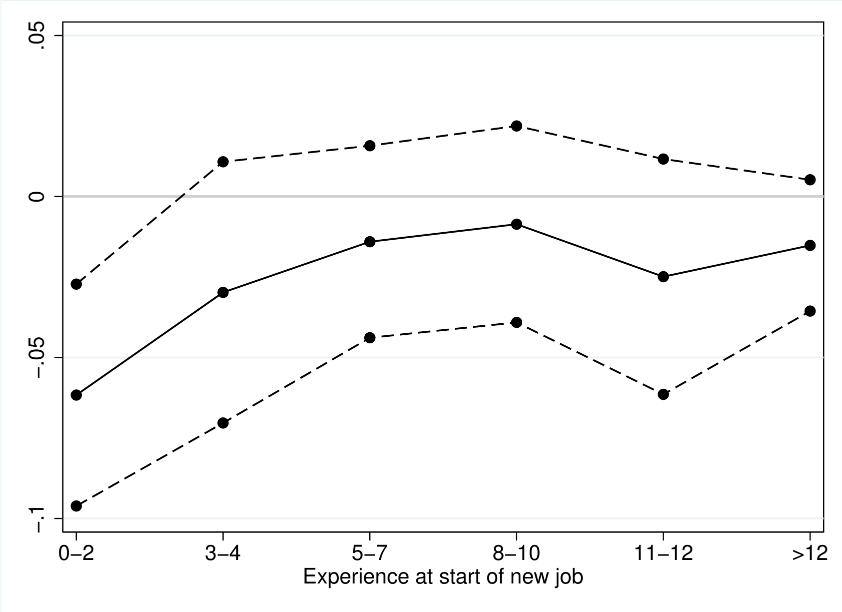
[28] Sample selection will imply that average wage growth is positive, but it does not necessarily affect the marginal impact of an increase in mismatch on wage growth.

Table 7: Responses to mismatch

	(1)	(2)	(3)	(4)
	All	Inexp. 0-4 yrs.	Exp. 5+ yrs.	P-val. for differences
		A. ENTRY WAGES		
Mismatch	-0.0097***	-0.0022	-0.0119***	0.0000
	(0.0007)	(0.0021)	(0.0008)	
Mean dep. var.	10.04	9.78	10.08	
Observations	328,651	47,360	281,291	
R-squared	0.8848	0.9050	0.8874	
		B. SEPARATIONS		
Mismatch	0.0075***	0.0214***	0.0058***	0.0114
	(0.0016)	(0.0068)	(0.0018)	
Mean dep. var.	0.23	0.28	0.22	
Observations	328,651	47,360	281,291	
R-squared	0.6083	0.7371	0.6379	
		C. WAGE GROWTH		
Mismatch	-0.0203**	-0.0568*	-0.0152	0.1282
	(0.0083)	(0.0323)	(0.0098)	
Mean dep. var.	6.97	7.13	6.95	
Observations	72,107	11,275	61,253	
R-squared	0.7207	0.8476	0.7306	
Education FE:s	√	√	√	
Entrant test scores	√	√	√	
Job 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 years-of-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). Job FE:s are (Entry occupation×Entry Year×Plant) FE:s.

Figure 5: The wage growth response to mismatch, by experience group



Notes: Each dot is an estimate of the wage growth response to initial mismatch by experience group. The sample consists of entrants in 1997-2008. Dashed lines are 95% confidence bands.

5 Robustness and alternative explanations

This section documents the robustness of our main results and provide alternative estimates that shed further light on the interpretation of these results. The focus is on split-sample regressions (experienced vs. non-experienced) on entry wages and separations. In Section 5.1 we present basic specification checks regarding measurement and functional form; Sections 5.2 and 5.3 study the role of occupations and labor market segments. Section 5.4 reports results from the alternative information proxies and alternative ways of measuring match quality. Section 5.5 discusses possible alternative explanations for our main results.

5.1 Basic robustness checks

Table 8 reports the results of a set of basic robustness checks; for ease of reference, Panel A shows the baseline estimates.

Panel B shows that the results are robust to more flexible skill controls. More specifically, we include all (8×7) interactions between the eight test scores; this has a very limited impact on the estimates. In Online Section A3.2 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; they thus take the direct effect of individual skill into account, but also capture other unobserved dimensions of worker ability (and outside options) potentially not captured by the test scores. However, the identifying data set is much smaller, and less representative, as workers must be recorded as new hires at sampled workplaces twice within the sample period. To gain precision, we thus pool across experience groups. The results, although statistically imprecise, show a very similar pattern to those of the main model for the same sample.

Panel C of Table 8 examines whether the effects of mismatch are non-linear. We pursue this robustness check for three reasons. First, there may be ranges of inaction, either because of measurement errors or because mobility/separation costs are substantial. 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 varies 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 show that although the impact of mismatch on entry wages is somewhat non-linear (the absolute size of the effect tends to be larger for high values of mismatch), the estimates on the second order terms are small.[29] Further, we find that separations are literally linear in mismatch. This indicates that allowing for non-linearities in mismatch is not crucial.

Panel D of Table 8 only includes jobs with at least 10 tenured workers. This zooms in on a smaller sample with a more precise measure of skill requirements. Naturally, the sample size drops substantially. Nevertheless, our results are remarkably stable. Overall, the absolute sizes of the estimates are somewhat higher which is consistent with the view that we get a more precise measure of skill requirements when we only include jobs with at least 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 the final column of Table 1. The results do not change.

Panel F separates between mismatch in the cognitive and non-cognitive dimensions by introducing two separate indexes, one for each dimension. The coefficients are of very similar magnitudes, and not significantly different from one another.

Panel G instead collapses the skill-vector to a cognitive and non-cognitive aggregate score to calculate mismatch across these two (cruder) dimensions. The results are smaller than the baseline, but the qualitative picture remains unchanged. This suggests that there is independent information regarding match quality in the full vector of (eight) skill components; information which is lost when using the cognitive and non-cognitive aggregates.

5.2 Jobs and occupations

Our main specification controls for the characteristics of the job through a job fixed effect. An alternative approach is to control for a flexible function of average incumbent skills along

[29]For experienced workers, the impact of mismatch, when evaluated at a standard deviation above the mean, is -1.2%; 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: 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	-0.0022	-0.0119***	0.0214***	0.0058***
	(0.0021)	(0.0008)	(0.0068)	(0.0018)
Observations	47,360	281,291	47,360	281,291
R-squared	0.9050	0.8874	0.7371	0.6379
B. All skill interactions				
Mismatch	-0.0023	-0.0107***	0.0225***	0.0060***
	(0.0022)	(0.0009)	(0.0071)	(0.0019)
Observations	47,360	281,291	47,360	281,291
R-squared	0.9052	0.8875	0.7376	0.6379
C. Non-linearities in mismatch				
Mismatch	-0.0013	-0.0110***	0.0200***	0.0057***
	(0.0022)	(0.0009)	(0.0070)	(0.0019)
$Mismatch^2$	-0.0014*	-0.0012***	0.0021	0.0001
	(0.0008)	(0.0003)	(0.0029)	(0.0007)
Observations	47,360	281,291	47,360	281,291
R-squared	0.9050	0.8387	0.5968	0.4807
D. Restricting the sample to jobs with at least 10 tenured coworkers				
Mismatch	-0.0012	-0.0141***	0.0233**	0.0051***
	(0.0020)	(0.0009)	(0.0065)	(0.0020)
Observations	24,360	129,415	24,360	129,415
R-squared	0.8600	0.8388	0.5980	0.4819
E. Weighted mismatch index				
Mismatch	-0.0022	-0.0118***	0.0202***	0.0051***
	(0.0020)	(0.0009)	(0.0067)	(0.0018)
Observations	47,360	281,291	47,360	281,291
R-squared	0.9050	0.8874	0.7371	0.6379
F. Mismatch in cognitive and non-cognitive ability				
$Mismatch_{cognitive}$	-0.0028	-0.0079***	0.0152**	0.0043**
	(0.0019)	(0.0008)	(0.0062)	(0.0018)
$Mismatch_{non-cognitive}$	0.0003	-0.0081***	0.0133**	0.0035*
	(0.0023)	(0.0009)	(0.0072)	(0.0019)
Observations	47,360	281,291	47,360	281,291
R-squared	0.9050	0.8874	0.7371	0.6379
G. Mismatch (based on overall cognitive and non-cognitive scores)				
Mismatch	-0.0014	-0.0076***	0.0128**	0.0034*
	(0.0017)	(0.0007)	(0.0054)	(0.0014)
	47,360	281,291	47,360	281,291
	0.9062	0.8885	0.7402	0.6422
Education FE:s	√	√	√	√
Entrant test scores	√	√	√	√
Job FE:s	√	√	√	√

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The specification is the same as in Table 7. Job FE:s are (Entry occupation×Entry Year×Plant) FE:s.

each of the 8 dimensions through a flexible function corresponding to $g(s)$. Table 9 shows (compare panels A and B) that we gain precision (in particular for the inexperienced, where the sample is much smaller) and that the point estimates are similar.

Another aspect of cross-job heterogeneity, that has been highlighted by Guvenen et al. (2015), is that the returns to specific skills may vary across occupations. As shown in Section 3.2.2, there is ample sorting on talents both across occupations and across jobs within the same occupation (in roughly equal amounts). In order to take out the role of sorting on occupations-specific returns, Panel C shows results from specifications that control for the interactions between each of the eight individual skills and the average intensity of that skill within the occupation (as a rudimentary way of allowing the returns to skill to vary across occupations). More than half of the effects of interest remain even after controlling for the intersection of skill-specific returns and skill endowments at the occupational level, suggesting that occupations do matter but that misallocation of talents across jobs within occupations has a remaining impact on outcomes. As an alternative exercise we have analyzed how the impact of mismatch in the actual job differs from the impact of mismatch measured relative to another (random) job within the same occupation within our sample.[30] The results displayed in Online Figure A4 confirm the picture that both the occupational component and the job-specific component matter for the outcomes of interest.

5.3 Labor market segments, diversity and horizontal mismatch

Table 10 investigates the role of heterogeneity across occupations and other indicators of labor market segments.

In panel A we redefine a "job" as the interaction between 3-digit occupation, plant, entry year (as before) and *education level*. Thus, we classify, for instance, lawyers entering a certain establishment as having different jobs if they have 4-year diplomas rather than 3-year diplomas (both are possible). Results remain robust.

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 for the high-end jobs, which could imply that initial mismatch would be priced to a greater extent in these jobs. However, this argument 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 two panels is that the impact of mismatch is similar across

[30] The reason for randomly drawing a comparison job within the same occupation is that job mismatch unavoidably contains some measurement error. By randomly drawing a comparison job we ensure that the measures of job mismatch and occupational mismatch are plagued by the same amount of measurement error.

Table 9: Jobs and occupations

	(1)	(2)	(3)	(4)
	ENTRY WAGES		SEPARATIONS	
	Inexp.	Exp.	Inexp.	Exp.
	0-4 yrs.	5+ yrs.	0-4 yrs.	5+ yrs.
A. Baseline				
Mismatch	-0.0022	-0.0119***	0.0214***	0.0058***
	(0.0021)	(0.0008)	(0.0068)	(0.0018)
Observations	47,360	281,291	47,360	281,291
R-squared	0.9050	0.8874	0.7371	0.6379
B. Job skill requirements in place of Job FE:s				
Mismatch	-0.0037***	-0.0154***	0.0206***	0.0054*
	(0.0010)	(0.0006)	(0.0029)	(0.0011)
Observations	47,360	281,291	47,360	281,291
R-squared	0.5438	0.6556	0.0466	0.0275
C. Own skills interacted with occupation skills				
Mismatch	-0.0032***	-0.0083***	0.0151***	0.0034
	(0.0012)	(0.0007)	(0.0035)	(0.0013)
Observations	47,360	281,291	47,360	281,291
R-squared	0.5459	0.6575	0.0474	0.0277
Education dummies	√	√	√	√
Entrant test scores	√	√	√	√
Job FE:s (A). Skill requirements (B-C)	√	√	√	√

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 years-of-experience fixed effects. The test score controls are 2nd order polynomials in each of the eight test score domains. Job (Entry occupation×Entry Year×Plant) FE:s in panel A. In panels B-C we replace the job FE:s with FE:s for entry year, entry occupation, and a 2nd order polynomial in job skill requirements (average skill along each of the eight dimensions among tenured workers)

Table 10: 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	-0.0004	-0.0110***	0.0219**	0.0075**
	(0.0033)	(0.0014)	(0.0111)	(0.0030)
Observations	47,360	281,291	47,360	281,291
R-squared	0.9473	0.9374	0.8525	0.7928
B. High vs. low-skill jobs				
Mismatch	-0.0000	-0.0136***	0.0222***	0.0046*
	(0.0026)	(0.0009)	(0.0091)	(0.0022)
Mismatch×High-skilled job	-0.0040	0.0031***	-0.0013	0.002
	(0.0029)	(0.0012)	(0.0098)	(0.0026)
Observations	47,360	281,291	47,360	281,291
R-squared	0.9050	0.8874	0.7371	0.6379
C. High vs. low-skill workers				
Mismatch	-0.0004	-0.0114***	0.0200***	0.0049**
	(0.0028)	(0.0009)	(0.0093)	(0.0021)
Mismatch×High education	-0.0004	-0.0006	0.0076	0.0017
	(0.0034)	(0.0012)	(0.0117)	(0.0024)
Observations	47,360	281,291	47,360	281,291
R-squared	0.9050	0.8641	0.7105	0.5740
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. Zooming in on horizontal mismatch				
Mismatch	-0.0032	-0.0078***	0.0258***	0.0046
	(0.0020)	(0.0011)	(0.0058)	(0.0022)
Observations	23,654	140,485	23,654	140,485
R-squared	0.5369	0.6255	0.0509	0.0273
Education dummies	√	√	√	√
Entrant test scores	√	√	√	√
Job FE:s (A-D). Skill requirements (E)	√	√	√	√

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. Job (Entry occupation×Entry Year×Plant) FE:s in panels A-D. In panel E we replace the job FE:s with FE:s for entry year, entry occupation, and a 2nd order polynomial in job skill requirements (average skill along each of the eight dimensions among tenured workers) types of positions. The results do not 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 compute the (eight) skill variances among tenured workers in each job with at least 10 tenured workers. Then, we remove the quarter of jobs with the highest skill dispersion (using the job-level mean of the eight variances). The idea is that jobs with a high variance in skills among tenured workers are more likely to use production processes with cross-worker complementarities within the same job-classification.[31] As shown in the table, the results are robust.

Panel E, focuses on the part of the sample where the *average skill level* of the entrant matches that of the incumbents. The intention is to fully isolate the horizontal aspect of mismatch, i.e., when workers only differ in the extent to which their *composition* of skills match the required composition within the job. We implement this idea by only using the half of the sample where the deviation between the levels of entrant and incumbent skill averages is the smallest. Since this reduces the effective sample, in particular the number of cases where there are multiple entrants into the same job, we use the parametric control for skill requirements instead of the job fixed effects for this exercise. The results are similar to the main results, suggesting that horizontal mismatch is mainly driving our baseline estimates.

5.4 Alternative information proxies and mismatch measures

5.4.1 Alternative information proxies

In section 4.1, we showed that experienced workers with prior within-firm experience and workers who are hired from another job are better matched. But how do these characteristics change the relationship between mismatch and economic outcomes? To measure these effects, we interact the impact of mismatch with these alternative information proxies for the group of experienced workers.[32] Table 11 reports the results.[33] Columns (2) and (3) show that the alternative information proxies push the results in the same direction as labor market experience; mismatch produces larger wage penalties and smaller separation responses when there is more information available. However, it is noteworthy that the variable capturing within-firm experience (a priori a very strong proxy for information) reduces the wage responses less than the experience dummy does. This suggests that a lot of uncertainty is removed by labor market experience, a result which is broadly consistent with Schönberg (2007) who shows (in the context of ability levels) that market learning is more important

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

[32] To improve precision, we opt for the parametric skill-requirements specification which has 2nd order polynomials in incumbent test scores in place of the job fixed effects. This buys us precision, without changing the point estimates.

[33] For completeness, Table A7 reports analogous results for the inexperienced.

than asymmetric, firm-specific, learning.[34]

5.4.2 Measuring mismatch from wage returns

Here we generate an alternative measure of mismatch based on the idea that job-level wage *returns* to specific talents can inform us about the usefulness of these talents in the production process. We use the estimates from the 60,500 job-cell regressions, which formed the basis of the analysis in Table 4 in Section 3. These estimated job level returns are used to calculate an alternative mismatch measure, defined as follows:

$$Mismatch_{ij}^{returns} = \sum_{k=1}^{K}(\hat{\beta}_k - \hat{\beta}_{jk})s_{ik} \tag{8}$$

where $\hat{\beta}_{jk}$ is the estimated return to 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. 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.[35]

Table 12 shows the results when we estimate the key equations using the mismatch measure defined in equation (8). 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 of Table 7. For the inexperienced, the wage impact is -0.0045 (compared to -0.0022 in the baseline); for the experienced, the wage impact is -0.0166 (-0.0119 in the baseline). The wage-based mismatch metric gives slightly weaker separation responses (0.0149 instead of 0.0214 for the inexperienced and 0.0034 instead of 0.0058 for the experienced), but qualitatively the results line up remarkably well with our baseline results.

5.5 Alternative explanations for the results

Here we raise the issue of whether there are alternative possible 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.

Our baseline analysis has a flavor of peer effects models in the sense that the measurement of mismatch is based on the correspondence between the talents of entrants and the talents

[34] If we interact the within-firm experience dummy with years of labor market experience we find that the role of within-firm experience is reduced as labor market experience grows which is consistent with the notion that there is less remaining uncertainty for all workers with longer labor market experience.

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

Table 11: Responses to mismatch with alternative information measures

	(1)	(2)	(3)
	Baseline	Alt. info. measures for the experienced	
	Labor market experience	*Firm experience*	*Job-to-job mobility*
		A: ENTRY WAGES	
Mismatch (*MM*)	-0.0001	-0.0142***	-0.0100***
	(0.0009)	(0.0006)	(0.0011)
*MM**Labor mkt exp.>4 yrs.	-0.0149***		
	(0.0009)		
*MM**Any firm exp.		-0.0036***	
		(0.0008)	
*MM**Job-to-job			-0.0059***
			(0.0011)
Mean dep. var.	10.04	9.78	10.08
Observations	328,651	281,291	281,291
R-squared	0.6658	0.6556	0.6600
		B: SEPARATIONS	
Mismatch	0.0233***	0.0087***	0.0148***
	(0.0022)	(0.0012)	(0.0025)
*MM**Labor mkt exp.>4 yrs.	-0.0183***		
	(0.0022)		
*MM**Any firm exp.		-0.0109***	
		(0.0017)	
*MM**Job-to-job			-0.0110***
			(0.0025)
Mean dep. var.	0.23	0.28	0.22
Observations	328,651	281,291	281,291
R-squared	0.0311	0.0284	0.0275
Education dummies	√	√	√
Entrant test scores	√	√	√
Incumbent test scores	√	√	√
Year FE:s	√	√	√
Occupation 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.

Table 12: Responses to mismatch with alternative mismatch measure

	(1)	(2)	**(3)**
	Inexp.	**Exp.**	**P-val. for**
	0-4 yrs.	5+ yrs.	**differences**
		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	$\surd$	$\surd$	
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. 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 $\omega = \left(\sqrt{\sum_k \text{var}(\hat{\beta}_{jk})}\right)^{-1}$ to adjust for small sample error in the estimates of the job level returns.

of tenured workers. However, our analysis compares entrants into the same job, and these entrants will by definition be exposed to the same set of peers. Thus, the first-order effect of peer quality (i.e. the "standard" peer effect) is accounted for by the job fixed effects. In the "horizontal" dimension, we measure mismatch for people with a similar level (but a different composition) of talents, so any potential peer effects must arise from benefits of being similar to other workers. In the "vertical" dimension, we measure absolute differences. Hence, to explain these 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, the wages of well-matched workers would 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 could explain why we find entry wages to be unrelated to mismatch among inexperienced workers.[36] However, we do not believe that this is a credible explanation 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 10 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

This section illustrates the long-run consequences of mismatch. To this end, we document the long-run effects of initial mismatch on future wages, employment, and earnings. We focus on the inexperienced workers since our previous evidence shows that entry wages are unrelated to mismatch, which suggests that mismatch can be treated as a purely idiosyncratic shock for this group of workers.[37]

We use the first entry cohorts in our data (1997-99), allowing us to follow the same sample over 10 years. We estimate the effects for all individuals in these entry cohorts, independently of whether individuals remained in the same job or not. Figure 6 shows the long-run responses of wages, non-employment, and annual earnings to initial mismatch. The first observation in each panel is for the year after the match is formed. Non-employment initially increases, wages decline, and earnings decrease. This initial response is consistent with the view that (some) poor matches are destroyed when match quality is revealed. The non-employment (and earnings) responses fade away and the point estimates are close to zero after 4 years. Over the longer haul, wages, non-employment, and earnings are completely unrelated to initial mismatch.

The point estimate for the log earnings change during the year after the match was formed is -0.027 (SE: 0.009). 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 based on wage returns discussed in section 5.4. Regressing this alternative mismatch measure on the baseline mismatch measure gives an estimate of 0.346 (SE: 0.027). If the measurement errors

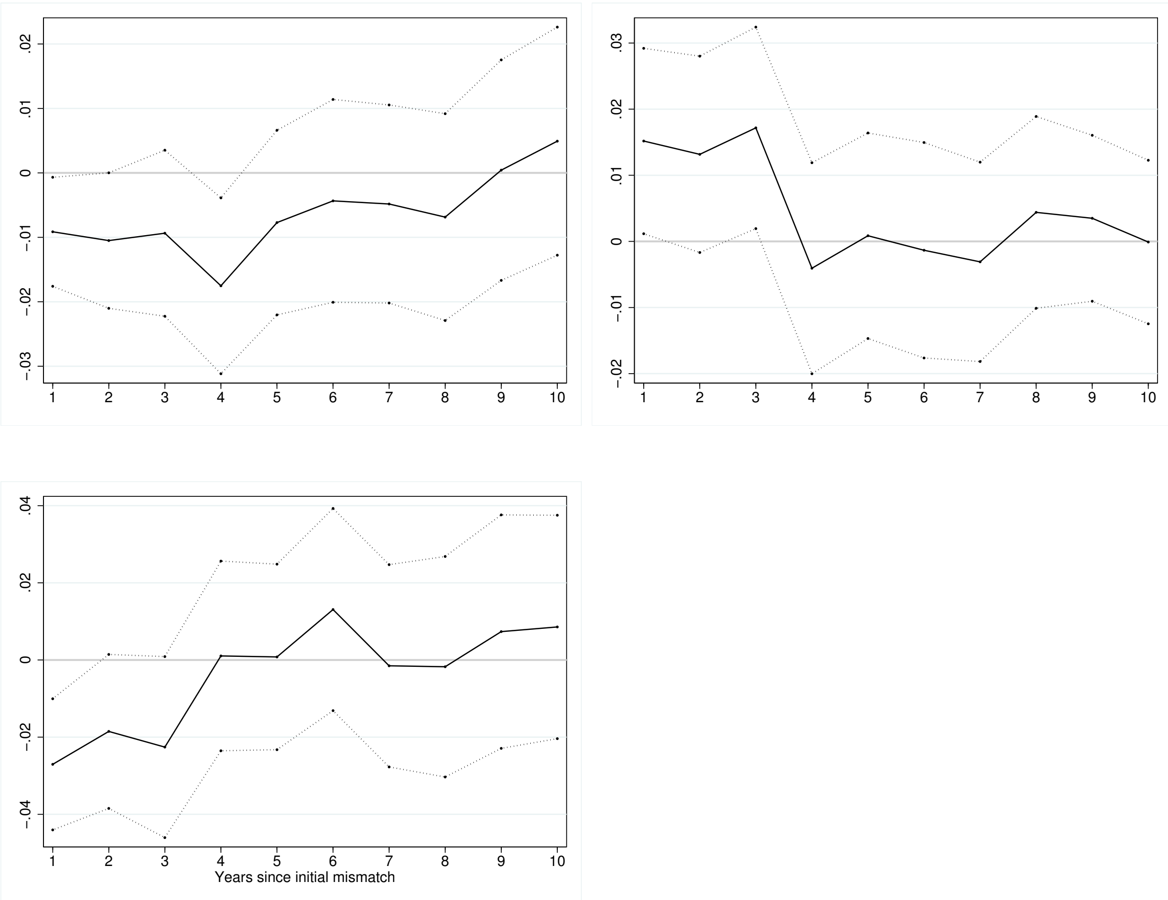
[36] A related concern is that unionization rates and exposure to collective bargaining differ across the various groups that we use to proxy uncertainty. Collective bargaining exposure differs mainly at the industry level, so any differences across groups are handled by the job fixed effects. Unionization is, in general, associated with less wage dispersion and fewer separations. If the experienced (or workers with prior experience in the firm) are unionized to a greater extent, we would expect to see a smaller separation response and a smaller (absolute) wage response; we don't see the latter in our results.

[37] The fact that entry wages are unrelated to mismatch, suggests there was no information available about mismatch at the time of the match. If mismatch was unknown, workers cannot self-select on the basis of mismatch. Therefore, exposure to mismatch should not be correlated with unobserved characteristics of workers, which implies that we can estimate the effects of initial mismatch on subsequent labor market outcomes.

Figure 6: Initial mismatch and subsequent outcomes for inexperienced workers

(a) Wages

(b) Non-employment



(c) Earnings

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. (1). The sample is restricted to entrants who we can follow the entire follow-up period (entrants in 1997-1999). Dashed lines are 95% confidence bands.

are uncorrelated, this suggests that a one standard deviation increase in true mismatch leads to an 8 percent drop in earnings one year after matching. Over time the large initial drop fades away through onward mobility to better matches.

7 Conclusions

We have examined the direct impact of mismatch on wages and job mobility using unique Swedish data on a multitude of talents, detailed occupational information, wages, and 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 main 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 further document that workers are sorted into jobs where their coworkers have similar talents. We also show that much of this sorting is *talent-specific*, and that about half of the sorting is across jobs with the same occupational classification. Moreover, we show that workers in jobs with high returns to a specific skill have higher than average levels of this skill.

The main part of the paper documents several novel facts about job mismatch: First, mismatch is higher among inexperienced workers. Second, starting wages are unrelated to mismatch for inexperienced workers whereas experienced workers receive a wage penalty if they are mismatched. Third, we find a non-trivial separation response to mismatch among inexperienced workers; the impact for experienced workers is small, however. Fourth, wage growth within jobs is negatively related to initial mismatch, and this effect is more pronounced among inexperienced workers. Fifth, within the group of experienced workers, proxies for additional information (prior within-firm experience, being hired from non-employment) are associated with higher initial wage penalties and smaller separation responses. Finally, we show that a standard deviation increase in mismatch reduces annual earnings by 8 percent for the average inexperienced entrant in the year after the match was formed. The earnings losses persist beyond the immediate impact, but disappear after four years.

We interpret the results as suggesting that systematic sorting on skills across jobs is a fundamentally important aspect of the labor market. But the impact of this sorting process on key labor market outcomes is a function of the information available at the time of hiring. For inexperienced workers, agents do not appear to fully observe how well the detailed characteristics of the worker match the skill-requirements of the job before production starts; these workers thus match under considerable uncertainty as in Jovanovic (1979). As a result, we see a post-match learning process with match quality determining separations and within-job wage growth for this group of workers. On the other hand, market learning seems to wash out most of the uncertainty over a few years time since prior within-firm experience adds less to initial wage losses than general labor market experience does. All in all, the matching process for inexperienced workers contains important elements of information uncertainty and learning, whereas the process for experienced workers appears best described by frictional models where (most of) the information about match quality is available already when matches are formed.

References

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

Abowd, John, Francis Kramarz and Sebastien Roux (2006), 'Wages, mobility and firm performance: Advantages and insights from using matched worker-firm data', *Economic Journal* **116**, F245–F285.

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.

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 (2016), 'Referral-based job search networks', *Review of Economic Studies* **83**(2), 514–546.

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.

Gathmann, Christina and Uta Schönberg (2010), ‘How general is human capital? a task-based approach’, *Journal of Labor Economics* **28**(1), 1–50.

Gautier, Pieter A. and Coen N. Teulings (2015), ‘Sorting and the output loss due to search frictions’, *Journal of the European Economic Association* **13**(6), 1136–1166.

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.

Gibbons, Robert and Michael Waldman (2004), ‘Task-specific human capital’, *The American Economic Review* **94**(2), 203–207.

Groes, Fane Naja, Philipp Kircher and Iourii Manovskii (2015), ‘The u-shapes of occupational mobility’, *Review of Economic Studies* **82**(2), 659–692.

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 (2016), ‘Social networks, employee selection and labor market outcomes’, *Journal of Labor Economics* **34**(4), 433–454.

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

Jovanovic, Boyan (1984), ‘Matching, turnover, and unemployment’, *Journal of Political Economy* **92**(1), 108–122.

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.

Lise, Jeremy and Fabien Postel-Vinay (2016), Multidimensional skills, sorting, and human capital accumulation, Manuscript, University of Minnesota.

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.

Schönberg, Uta (2007), ‘Testing for asymmetric employer learning’, *Journal of Labor Economics* **25**, 651–692.

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.

Åslund, Olof and Oskar Nordström Skans (2009), ‘How to measure segregation conditional on the distribution of covariates’, *Journal of Population Economics* **22**(4), 971–981.

Online Appendix

A1 A formal presentation of the model

Production The technology is constant returns to scale; thus we 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. Mismatch between the skills of the worker and skill requirement of the job along the kth dimension is measured by $d_k(i,j) = |s_k(i) - s_k(j)|$ and we denote the aggregate distance between the worker and the job by $d = d(i,j)$.

Match productivity, $y(i,j)$, is assumed to be given by

$$y(i,j) = 1 - \gamma d(i,j) + \theta s(i) + \lambda(j) \tag{A1}$$

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 $y(i,j) - y^* = -\gamma d(i,j)$. Therefore, we suppress $s(i)$ and $\lambda(j)$ below.[38]. To save on notation, we write match productivity as $y(d)$ from here on.

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

$$E_0(d|\, d_0) = (1 - \alpha)E(d) + \alpha d_0 \tag{A2}$$

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,

[38] 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.

and a frictionless stage.[39]

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^*); with an optimally determined reservation wage rule individuals will be climbing the job-ladder towards the optimal match. As our focus is on micro-level predictions, y^* is treated as exogenous.

A1.1 Matching, wages, and separations

At the meeting stage, the expected joint surplus equals[40]

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

where p_0 denotes the probability of separating at the revelation stage (which given our distributional assumption about d_0, only depends on α). The first term in brackets represents the expected gain from matching; with probability $(1 - p_0)$ 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 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 match if and only if $E_0(S|\, d_0) > 0$. The matching threshold can thus be written as

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

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

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

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.

$$w_0(d) = \frac{1}{2}E_0(S|\, d_0) = \frac{1}{2}\left[(1-p_0)(c - \gamma E_0(d|\, d_0)) - p_0 b\right] \tag{A3}$$

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 renegotiates wages. The set of continuing matches is defined by $S(d) = y(d) - (y^* - (c+b)) = (c+b) - \gamma d > 0$. The match thus continues to be viable if the actual cost of mismatch (γd) is lower than the separation cost ($c+b$). Separations occur if

$$d > \frac{c+b}{\gamma} \equiv d_s \tag{A4}$$

Using the definition of the separation threshold (d_s), we can rewrite the matching threshold somewhat. The set of acceptable matches is defined by

$$E_0(d|\, d_0) < d_s - \frac{b/\gamma}{1-p_0} \equiv d_m \tag{A5}$$

Notice that d_m depends on α since p_0 depends on α. The number of matches is, m, is given by

$$m = \Pr(E_0(d|\, d_0) < d_m) = E(d) + (d_m - E(d))/\alpha \tag{A6}$$

From equation (A5) 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

$$p_0 = (1-\alpha)(1-d_s) + \alpha I(\alpha < \bar{\alpha})\left(1 - \frac{d_s}{m}\right) \tag{A7}$$

where $I()$ denotes the indicator function. $1-d_s/m$ reflects the probability of separating when the agents received correct information. 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.

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

$$w(d) = \frac{1}{2}\left[(c+b) - \gamma d\right] \tag{A8}$$

A1.2 Predictions we take to the data

Let us begin by establishing some notation and some restrictions which must hold true for the market to exist. At this stage we make explicit that the match threshold (d_m), the number of matches (m), and separation expectations (p_0), all depend on α by equations (A5), (A6), and (A7). We thus write $d_m(\alpha)$, $m(\alpha)$, and $p_0(\alpha)$.

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

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

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. A less precise initial signal increases initial mismatch If the distribution of potential mismatch does not vary with the precision of the initial signal, higher match rates translate into greater exposure to mismatch.[41] We thus focus on how the match rate varies

[41] We assume that the cost of delay (c) is unrelated to uncertainty. In our empirical work we treat matches involving, e.g., inexperienced workers as matches where there is more uncertainty about mismatch. If c

with the precision of the initial signal.

From (A6), 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.

It may also be instructive to focus on the extreme cases, $\alpha = \underline{\alpha}$ and $\alpha = 1$. We have $m(\underline{\alpha}) = 1 > m(1) = d_m(1)$.

2. A less precise initial signal weakens the negative impact of mismatch on entry wages From (A3) 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

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

In the extreme cases, we have $\frac{\partial w_0}{\partial d}\Big|_{\alpha=\underline{\alpha}} = -\frac{(1-p_0)\gamma\underline{\alpha}^2}{2} > \frac{\partial w_0}{\partial d}\Big|_{\alpha=1} = -\frac{(1-p_0)\gamma}{2}$.

3. A less precise initial signal strengthens the positive impact of mismatch on separations 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

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

varies by experience group, exposure to mismatch reflects uncertainty *and* search friction (c). Prediction 1 is thus not robust to allowing variation in the cost of delay. However, as a first approximation, the responses to variation in mismatch, Predictions 2-4 below, are robust. In the extreme cases, $\alpha = 0$, $\alpha = 1$, it is straightforward to verify that the response magnitudes never involve c.

It is straightforward to verify that

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

Thus in the extreme cases, the separation is falling in the precision of the initial signal. For marginal changes in α matters are slightly more complex. 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. A less precise initial signal strengthens the negative impact of mismatch on wage growth (within job) Define $\Delta w = w(d) - w_0(d)$, where $w(d)$ is given by (A8) and $w_0(d)$ by (A3). 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 (see section A3.1).

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

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

A2 Additional descriptives

Table A1 shows the various stages in the sampling selection process and Table A2 reports some basic descriptive statistics for all male entrants born between 1951-1976.

Table A3 presents results that parallel Table 3, but in this instance we show correlations between individual skills and coworker skills for each of the eight skills considered. Table A3 shows that the strongest correlation is for the particular talent under consideration (see main diagonal).

Figure A1 relates to Table 4. It shows all estimated job-specific skill-returns plotted against the skill endowments within the same jobs, separately for each type of talent. In 7 out of 8 cases, the correlation between job-specific returns and job-specific skill endowments is positive and statistically significant.

Table A4 relates to Table 2. Relative to Table 2, it expands the set of occupations to include the low-end and the top-end of the skill distribution. Table A4 thus shows the occupation with the highest score along a particular dimension, separately for the low-, the middle-, and the high-end of the skill distribution. We also report the dimension in which employed workers in a given occupation is least endowed, and the wage rank of the occupation. The table shows, for instance, that among low-skill occupations: miners score high on emotional stability (but low on inductive ability); furniture carpenters score relatively high on spatial ability (but are low on verbal ability). Among high-skill occupations, medical doctors score high on a variety of skill measures; this includes both cognitive and non-cognitive traits. Pilots seem to be emotionally stable (but are relatively low on verbal ability). Notice that all of the measurements are made before the individuals self-select into the various occupations.

Table A1: Sample selection

	Worker-year observations
All male entrants 1997-2008	5,385,589
... born between 1951-1976	2,784,253
... in sampled firms	707,337
... with at least one male tenured coworker born between 1951-1976	328,651

Table A2: All male entrants 1997-2008

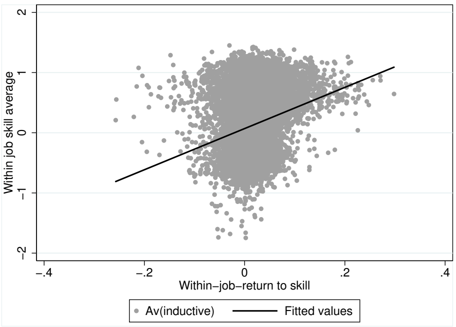
	mean (SD)	median
Separation rate	.29	
Age	36.4 (8.0)	36
Experience at entry	11.3 (5.5)	12
Job-to-job mobility	.73	
Prior within firm experience	.12	
Entry establishment size	144 (498)	22
Education:		
Compulsory or less	.13	
High school	.50	
College	.38	
Observations	2,784,253	

Table A3: Skill sorting over jobs by talent

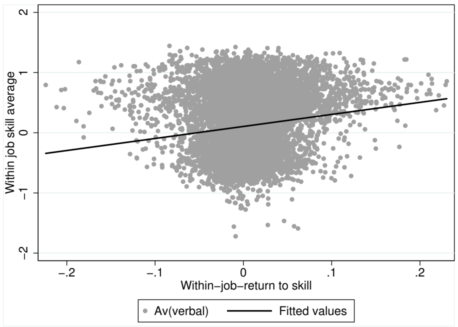
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Cognitive skills				*Non-cognitive skills*			
	Inductive	Verbal	Spatial	Technical	Social maturity	Intensity	Psychological energy	Emotional stability
Cognitive skills:								
Inductive	**0.4212*****	0.2977***	0.1503***	0.1259***	0.0869***	0.0435***	0.1199***	0.0521***
	(0.0084)	(0.0068)	(0.0082)	(0.0067)	(0.0080)	(0.0083)	(0.0082)	(0.0084)
Verbal	0.2714***	**0.4374*****	0.0880***	-0.0233***	0.1682***	-0.0536***	0.1602***	0.0834***
	(0.0063)	**(0.0075)**	(0.0075)	(0.0061)	(0.0076)	(0.0080)	(0.0080)	(0.0086)
Spatial	0.0583***	0.0417***	**0.2820*****	0.1307***	-0.0193***	-0.0069	-0.0150**	0.0172**
	(0.0055)	(0.0054)	**(0.0094)**	(0.0059)	(0.0064)	(0.0072)	(0.0070)	(0.0071)
Technical	0.1117***	0.0286***	0.2486***	**0.6422*****	0.0205***	0.0577***	0.0009	0.0549***
	(0.0050)	(0.0049)	(0.0064)	**(0.0065)**	(0.0060)	(0.0063)	(0.0058)	(0.0061)
Non-cognitive skills:								
Social maturity	0.0698***	0.1107***	0.0039	0.0172**	**0.3295*****	0.0972***	0.1752***	0.1476***
	(0.0062)	(0.0063)	(0.0080)	(0.0069)	**(0.0095)**	(0.0084)	(0.0084)	(0.0084)
Intensity	-0.0759***	-0.1351***	-0.0355***	-0.0061	0.1117***	**0.5137*****	0.1152***	0.1931***
	(0.0048)	(0.0050)	(0.0065)	(0.0051)	(0.0073)	**(0.0108)**	(0.0088)	(0.0080)
Psychological energy	0.1530***	0.1755***	0.0656***	0.0437***	0.1246***	-0.0196**	**0.2285*****	0.0951***
	(0.0062)	(0.0062)	(0.0078)	(0.0062)	(0.0086)	(0.0098)	**(0.0118)**	(0.0088)
Emotional stability	0.0155***	0.0167***	0.0411***	0.0468***	0.1158***	0.1734***	0.1316***	**0.2329*****
	(0.0060)	(0.0061)	(0.0079)	(0.0062)	(0.0081)	(0.0093)	(0.0086)	**(0.0115)**
Observations	1,944,964	1,944,964	1,944,964	1,944,964	1,944,964	1,944,964	1,944,964	1,944,964
Adjusted R-squared	0.2740	0.2608	0.1656	0.2235	0.1445	0.0678	0.1340	0.1076
Year FE:s	√	√	√	√	√	√	√	√

Notes: Robust standard errors clustered on the establishment*occupation level in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The estimates are based on the same model as in column (1) of Table 3, for each of the 8 talents.

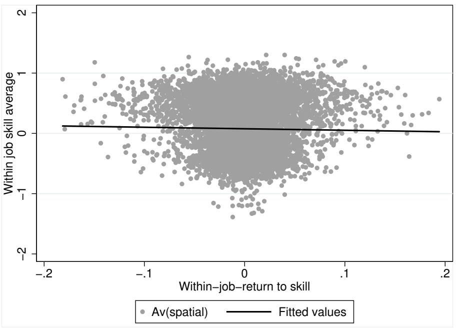
Figure A1: Correlation between skills and skill returns among tenured workers



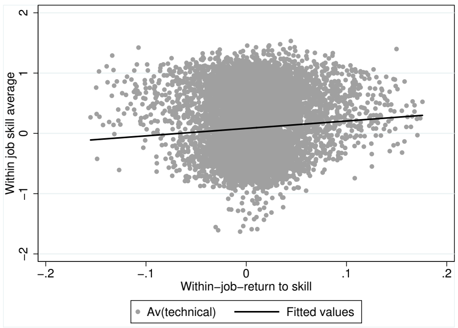
(a) Inductive skill



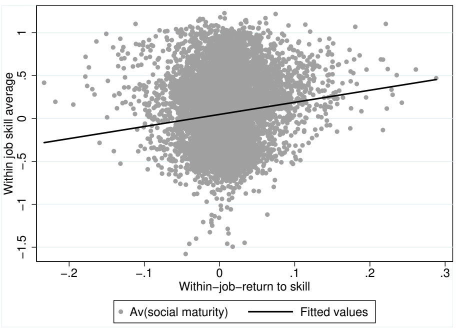
(b) Verbal skill



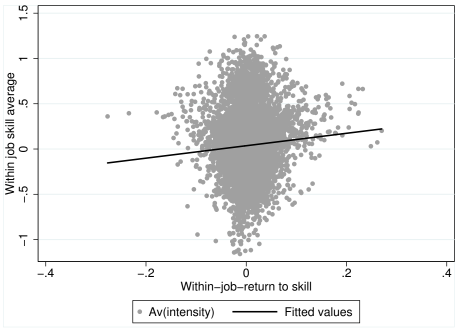
(c) Spatial skill



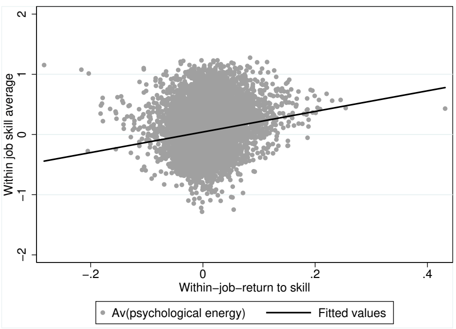
(d) Technical skill



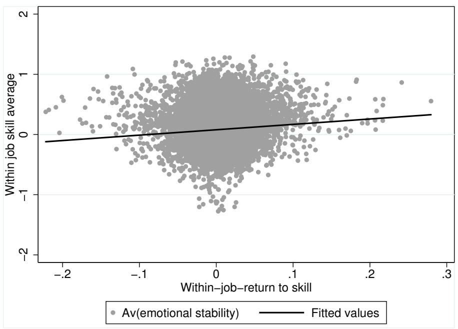
(e) Social maturity



(f) Intensity



(g) Psychological energy



(h) Emotional stability

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).

Table A4: Skill endowments across occupations

(a) Low-skilled occupations

Skill	Most endowed occupation in bottom third of average skills	Skill endowments			Wage rank
Non-cognitive:		Specific	Average	Least endowed	
Social maturity	Restaurant workers (e.g., cooks) (512)	-0.11	-0.35	Tech (=-0.46)	0.09
Intensity	Miners (711)	0.01	-0.23	Ind (=-0.59)	0.70
Psychological energy	Dairy and poultry producers (612)	-0.13	-0.33	Ind (=-0.54)	0.10
Emotional stability	Miners (711)	-0.08	-0.26	Ind (=-0.59)	0.70
Cognitive:					
Inductive	Storage workers (413)	-0.28	-0.48	Tech (=-0.40)	0.29
Verbal	Storage workers (413)	-0.29	-0.47	Tech (=-0.40)	0.29
Spatial	Furniture carpenters (742)	-0.20	-0.39	Verb (=-0.47)	0.14
Technical	Wood and paper processing (814)	-0.27	-0.44	Verb (=-0.41)	0.59

(b) Medium-skilled occupations

Skill	Most endowed occupation in middle third of average skills	Skill endowments			Wage rank
Non-cognitive:		Specific	Average	Least endowed	
Social maturity	Nurses (313)	0.29	0.18	Tech (=0.03)	0.61
Intensity	Forestry workers (614)	0.33	-0.03	Spat (=-0.23)	0.26
Psychological energy	Placement officers etc. (342)	0.21	0.15	Tech (=-0.07)	0.64
Emotional stability	Fire fighters and security guards (515)	0.19	0.05	Spat (=-0.15)	0.30
Cognitive:					
Inductive	Librarians (243)	0.66	0.15	Int (=-0.44)	0.56
Verbal	Librarians (243)	0.83	0.15	Int (=-0.44)	0.56
Spatial	Photographers, image and sound recording (313)	0.29	0.18	Int (=-0.14)	0.55
Technical	Photographers, image and sound recording (313)	0.38	0.18	Int (=-0.14)	0.55

(c) High-skilled occupations

Skill	Most endowed occupation in top third of average skills	Skill endowments			Wage rank
Non-cognitive:		Specific	Average	Least endowed	
Social maturity	Medical doctors (222)	0.81	0.42	Int (=0.26)	0.99
Intensity	Police officers (345)	0.69	0.21	Tech (=0.11)	0.74
Psychological energy	Medical doctors (222)	0.84	0.40	Int (=0.26)	0.99
Emotional stability	Pilots and naval officers (314)	0.66	0.34	Verb (=0.32)	0.98
Cognitive:					
Inductive	Medical doctors (222)	1.10	0.22	Int (=0.26)	0.99
Verbal	Medical doctors (222)	1.11	0.23	Int (=0.26)	0.99
Spatial	University research and teaching (213)	0.73	0.19	Int (=0.04)	0.83
Technical	Architects and engineers (214)	0.90	0.28	Int (=0.23)	0.90

Notes: ISCO-codes are reported within parentheses. Data are from 2002 and contain individuals with at least 3 years of tenure. The wage ranks pertain to the 2002 wage distribution. Labels for non-cognitive scores are according to Mood et al. (2012).

A3 Additional results

This section reports a set of additional results which are referred to in the main text.

A3.1 Variance of skills and tenure

One implication of the theory outlined in Section A1 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 (x) and tenure (τ) as:

$$\sigma^2_{jx\tau} = \frac{1}{K}\sum_{k=1}^{K}\sigma^2_{kjx\tau}$$

We then examine how the dispersion of skills varies with experience and tenure using the following equation:

$$\sigma^2_{jx\tau} = \delta_1 Inexp. + \delta_2 Tenure + \delta_3 Inexp. \times Tenure + \lambda_j + \epsilon_{jx\tau} \tag{A9}$$

where λ_j denotes "job" (Occupation×Year×Plant) fixed effects. Column (1) of Table A5 shows the results. There is somewhat higher variability of skills among inexperienced entrants compared to entrants who accumulated more pre-hire experience. However, as expected the difference with respect to experience falls with tenure, suggesting that remaining inexperienced workers become more like remaining experienced workers.

Column (2) conducts an analogous exercise for the within-job variance in wages. Here the interaction between the inexperienced dummy and tenure is positive, reflecting that there is more learning among the inexperienced than among the experienced and, therefore, variation in match quality gets translated into variation in wages for this group over the course of the match.

Figure A2 illustrates the results condensed in Table A5.

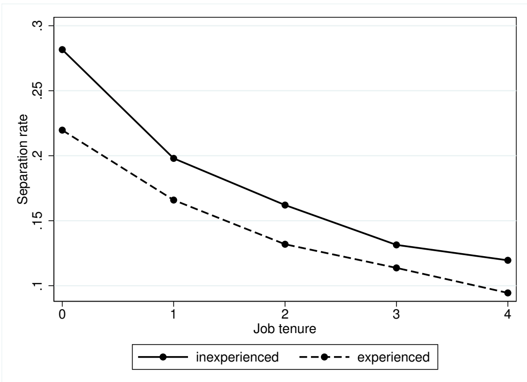
Table A5: Skill dispersion and tenure

	(1)	(2)
	Within-job variance in:	
	Talents	*Wages*
Inexperienced	0.0087	-0.0196***
	(0.0303)	(0.0008)
Tenure	-0.0209***	-0.0010***
	(0.0063)	(0.0002)
Inexperienced×Tenure	-0.0332**	0.0013***
	(0.0144)	(0.0004)
Observations	290,415	290,415
Adj. R-squared	0.169	0.167
Job FE:s	√	√

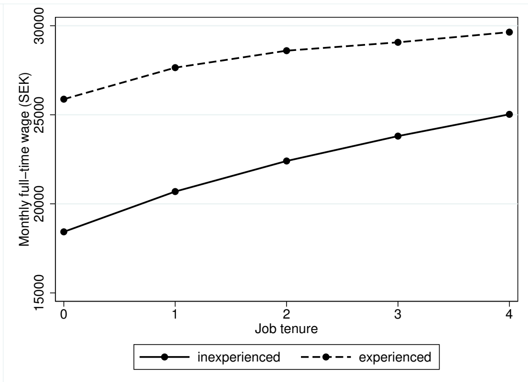
Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the mean variance in wages/skills within the job-experience group-tenure cell. Job FE:s = (Occupation×Year×Plant) FE:s.

Figure A2: Tenure profiles by experience groups

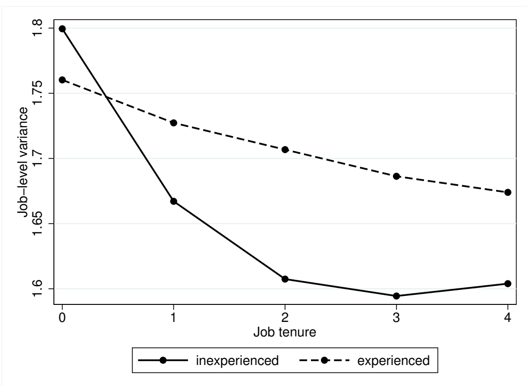
(a) Separation-tenure profile by experience



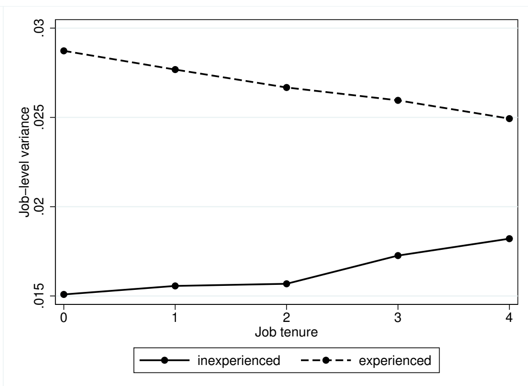
(b) Wage-tenure profile by experience



(c) Within job variance in talents



(d) Within job variance in wages



Notes: Figures (a) and (b) show how the separation rate and mean wage evolve with by employer tenure, separately for experienced and inexperienced workers. Figure (c) and (d) show the variance in the eight talents/wages among workers within the same job and tenure category.

A3.2 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 A6 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.

Table A6: Responses to mismatch with worker fixed effects

	(1)	(2)	(3)
	Baseline	Baseline	Worker FE:s
	All new hires	Repeat new hires	Repeat new hires
		ENTRY WAGES	
Mismatch	-0.0097***	-0.0091***	-0.0036**
	(0.0007)	(0.0014)	(0.0009)
Observations	328,651	135,325	135,325
R-squared	0.8848	0.9080	0.9308
		SEPARATIONS	
Mismatch	0.0075***	0.0070**	0.0086***
	(0.0016)	(0.0033)	(0.0030)
Observations	328,651	135,325	135,325
R-squared	0.6083	0.6977	0.4700
Education dummies	√	√	
Entrant test scores	√	√	
(Entry occupation×Entry Year×Plant) FE:s	√	√	
Entry occupation			√
Entry Year			√
Job skill requirements			√
Worker FE:s			√

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. Column (3) include worker fixed effects and replaces the job FE:s with FE:s for entry occupation, entry year, and a 2nd order polynomial in job skill requirements (average skill along each of the eight dimensions among tenured workers).

A3.3 The timing of the separation response

Here we probe deeper into the timing of the separation response. The exact timing of the response conveys information on how fast the worker-firm pair learns about mismatch. To examine this issue, we need higher-frequency data than the annual information we use in the main text. We therefore tap monthly separation-indicators.

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.

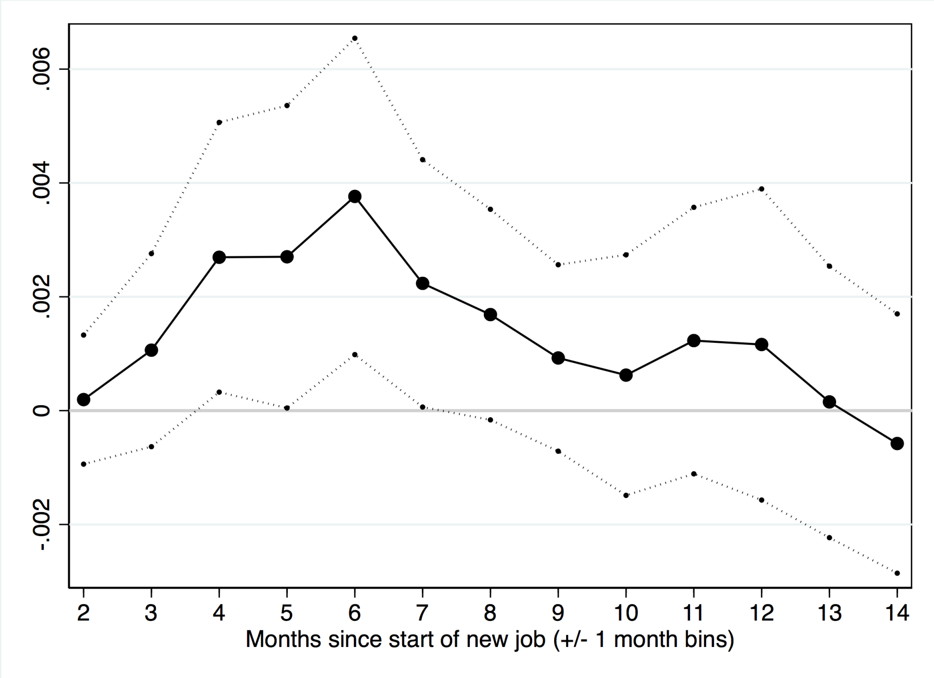
One potential concern with the monthly indicators is that the first and last month of compensation are self-reported by the employers, which increases the risk of measurement error. In our sample, 35 percent of the separations occur in December (conditioning on entry in August-October), which seems high 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.

Figure A3 shows the separation response by months since the start of the new job. To gain precision, we pool all experience groups.[42] We use moving (quarterly) averages to increase precision in the figure (i.e. 1-3 months, 2-4 months,... after the start of the new job). The results show that the peak of the separation occurs after approximately 6 months. In general, the speed of adjustment is thus fairly rapid. We find no evidence of separation responses after 1 year. Employment protection in Sweden may contribute to the peak at 6 months, since employment protection legislation allows for a 6 months initial probation period during which both agents can terminate the contract at will.[43] This implies that 6 months could be a focal point and incentives from both the employer and the employee side can be geared towards terminating bad matches after 6 months.

[42] The annual separation response for the entire sample is 0.0075, see Table 7, column (1). The monthly 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.

[43] 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 A3: Timing of the separation response



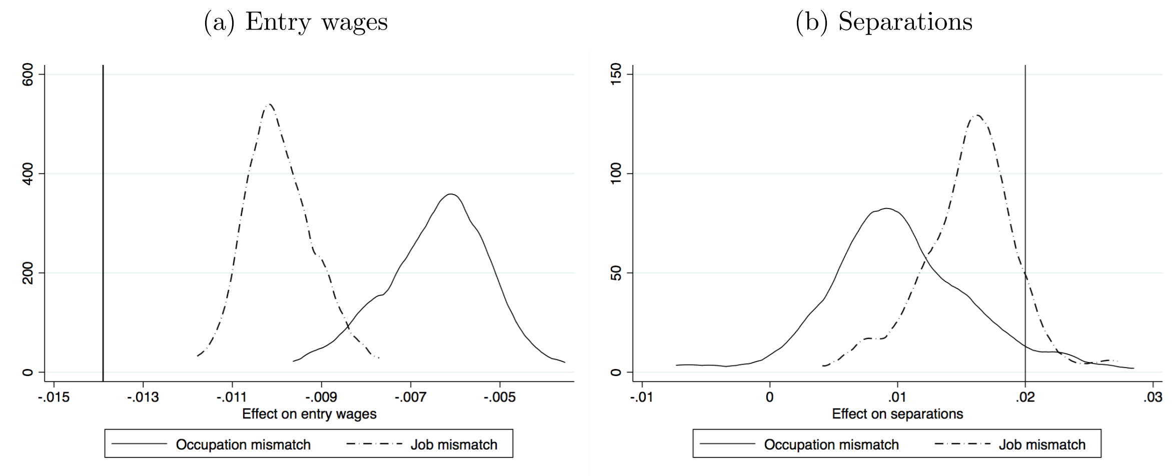
Notes: The figure displays the response to initial mismatch within 3 month-bins (+/- 1 month). Dashed lines are 95% confidence bands. See Online Section A3.3 for a more precise description of the data.

A3.4 Jobs and occupations

Figure A4 reports the results of a simulation exercise where we randomly draw another job within occupation, calculate mismatch for this randomly drawn job, and then estimate models including job and occupational mismatch simultaneously. The reason for randomly drawing a job is that we want the mismatch at the two levels to be equally mismeasured.

Figure A4 shows that mismatch both has an occupational component and a job component. Of these two components, job mismatch is somewhat more relevant for the outcomes.

Figure A4: Job and occupation mismatch



Notes: Figure (a) shows the entry-wage impact of mismatch among the experienced and Figure (b) the separation impact among the inexperienced. The estimates for occupational mismatch were generated by randomly drawing another job within occupation and then calculating the mismatch measure with respect to this randomly drawn job. Then we estimated models including job and occupational mismatch simultaneously. Vertical lines in the figure indicate our baseline estimates.

A3.5 Alternative information proxies for the inexperienced

Table A7 relates to Table 11 in the main text. It examines the effects of the alternative information proxies for the inexperienced group. For this group, we have fewer observations and, therefore, we grapple a bit with precision. Nevertheless, the signs of all estimates are consistent with our underlying information story. The absolute size of the separation response is larger for individuals hired from non-employment (which is consistent with there being less information available about mismatch for this group at the time of the match). The estimates are also consistent with there being more information available about workers with some prior experience in the firm. For these workers, mismatch is negatively priced into their entry wages to a greater extent and there is a smaller separation response to mismatch (in the absolute sense).

Table A7: Responses to mismatch with alternative information measures

	(2)	(3)
	Alt. info. measures for the inexperienced	
	Firm experience	*Job-to-job mobility*
	ENTRY WAGES	
$Mismatch$ (MM)	-0.0030***	-0.0034***
	(0.0011)	(0.0012)
MM*Any firm exp.	-0.0070***	
	(0.0021)	
MM*Job-to-job		-0.0002
		(0.0014)
Observations	47,360	47,360
R-squared	0.5451	0.5515
	SEPARATIONS	
$Mismatch$	0.0212***	0.0249***
	(0.0030)	(0.0035)
MM*Any firm exp.	-0.0063	
	(0.0066)	
MM*Job-to-job		-0.0102**
		(0.0041)
Observations	47,360	47,360
R-squared	0.0469	0.0471
Education dummies	√	√
Entrant test scores	√	√
Incumbent test scores	√	√
Year FE:s	√	√
Occupation 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.