# Online Appendix for Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility

*By* Peter Fredriksson, Lena Hensvik, and Oskar Nordström Skans*

This online Appendix contains a formal presentation of our conceptual framework and a number of additional results for the paper "Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility"

## 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.[1]. To save on notation, we write match productivity as $y(d)$ from here on.

* Fredriksson: Department of Economics, Uppsala University, Box 513, SE-75120 Uppsala, Sweden, UCLS, and IZA (email: peter.fredriksson@nek.uu.se); Hensvik: 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; Skans: Department of Economics, Uppsala University, Box 513, SE-751 20 Uppsala, Sweden, UCLS, IFAU, and IZA (email: oskar.nordstrom_skans@nek.uu.se).

[1] 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.

### 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 $\alpha$, 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 $(\alpha)$.

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

### Hiring and wage bargaining

We follow Eeckhout and Kircher (2011) when modeling hiring and wage bargaining. We think of three stages: a meeting stage, a revelation stage, and a frictionless stage.[2]

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 $\pi^*$. 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

[2]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.

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.

### A. *Matching, wages, and separations*

At the meeting stage, the expected joint surplus equals[3]

$$\begin{aligned} E_0(S|\, d_0) &= [(1-p_0)E_0(y(d)|\, d_0) + p_0(y^* - (c+b))] - [y^* - c] \\ &= (1-p_0)(c - \gamma E_0(d|\, d_0)) - p_0 b \end{aligned}$$

where $p_0$ denotes the probability of separating at the revelation stage (which given our distributional assumption about $d_0$, only depends on $\alpha$). 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

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

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 re-negotiates 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 ($\gamma d$) is lower than the separation cost ($c+b$). Separations occur if

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

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

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 $\alpha$ since $p_0$ depends on $\alpha$. 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 $\alpha$). 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

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

### B. *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 $\alpha$ 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 $\alpha$ is bounded from below. In particular, the upper bound on $d_0$ implies that $\alpha$ 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 $\alpha$; notice also that $d_m(\alpha)$ is a positive function of $\alpha$ 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 $\alpha$. 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 $\alpha$ 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 $\eta(\underline{\alpha}) = (d_m(\underline{\alpha}) - d_s)/(1 + d_s - d_m(\underline{\alpha})) < 1$).

### Prediction 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.[4] We thus focus on how the match rate varies 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 $\mathit{\Delta}(\alpha) = (d_s - d_m(\alpha))/(d_m(\alpha) - E(d)) > 0$. Increasing $\alpha$ has a direct

[4]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 $c$ 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 $c$.

negative effect and an indirect (positive) effect, via the dependence of $d_m$ on $p_0$ (with an increase in $\alpha$, $p_0$ 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 $m(\underline{\alpha}) = 1 > m(1) = d_m(1)$.

PREDICTION 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 $\left.\frac{\partial w_0}{\partial d}\right|_{\alpha=\underline{\alpha}} = -\frac{(1-p_0)\gamma\underline{\alpha}^2}{2} > \left.\frac{\partial w_0}{\partial d}\right|_{\alpha=1} = -\frac{(1-p_0)\gamma}{2}$.

PREDICTION 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

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

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 $\alpha$ 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

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

PREDICTION 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

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

PREDICTION 5

**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.A).

The change in the variance of the observed mismatch distribution ($\Delta$var) 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 of the paper, 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 of the main text. 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 in the main text. 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.

### A3. Additional results

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

#### A. *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 ($\tau$) 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 $\lambda_j$ 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.

#### B. *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.

### *C. 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 of the paper, 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.[5] 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.[6] 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.

### D. *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.

### E. *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).

[5]The annual separation response for the entire sample is 0.0075, see Table **??**, 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.

[6]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.

## REFERENCES

**Atakan, Alp E.** 2006. "Assortative Matching with Explicit Search Costs." *Econometrica*, 74(3): 667–680.

**Eeckhout, Jan, and Philipp Kircher.** 2011. "Identifying Sorting – In Theory." *Review of Economic Studies*, 78(3): 872–906.

**Mood, Carina, Jan O. Jonsson, and Erik Bihagen.** 2012. "Socioeconomic persistence across generations: The role of cognitive and non-cognitive processes." In *From Parents to Children: The Intergenerational Transmission of Advantage.* , ed. John Ermisch, Markus Jäntti and Timothy M Smeeding, Chapter 3. Russell Sage Foundation.

**Teulings, Coen N., and Pieter A. Gautier.** 2004. "The Right Man for the Job." *Review of Economic Studies*, 71(2): 553–580.

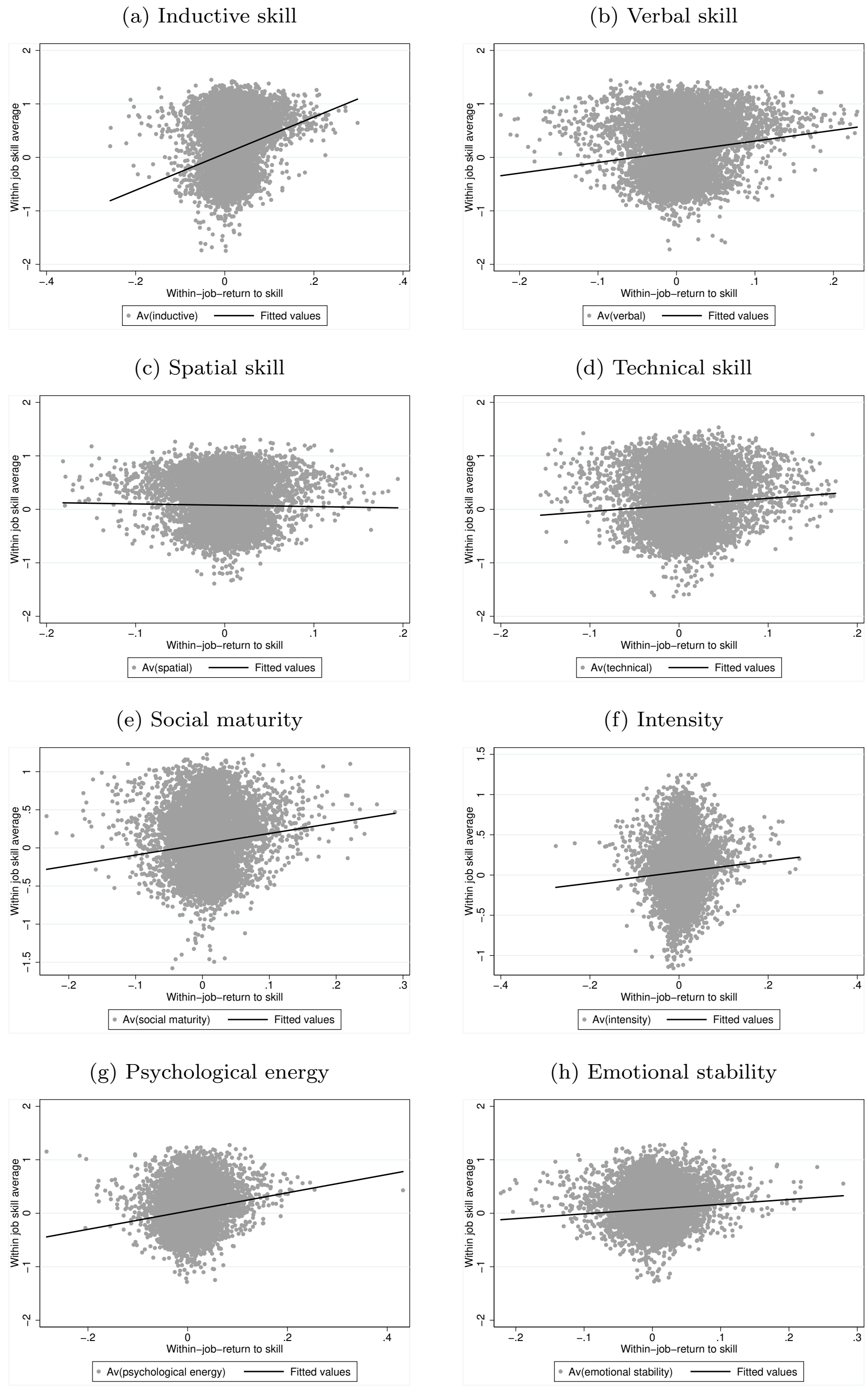


Figure A1. Correlation between skills and skill returns among tenured workers

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

Table A4—Skill endowments across occupations

**(a) Low-skilled occupations (bottom third of average skills)**

| Skill | Most endowed occupation | Skill endowments | | | Wage rank |
|---|---|---|---|---|---|
| *Non-cognitive:* | | Specific | Average | Least | |
| 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 (middle third of average skills)**

| Skill | Most endowed occupation | Skill endowments | | | Wage rank |
|---|---|---|---|---|---|
| *Non-cognitive:* | | Specific | Average | Least | |
| 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/sound recording (313) | 0.29 | 0.18 | Int (-0.14) | 0.55 |
| Technical | Photographers, image/sound recording (313) | 0.38 | 0.18 | Int (-0.14) | 0.55 |

**(c) High-skilled occupations (top third of average skills)**

| Skill | Most endowed occupation | Skill endowments | | | Wage rank |
|---|---|---|---|---|---|
| *Non-cognitive:* | | Specific | Average | Least | |
| 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/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, Jonsson and Bihagen (2012).

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 | √ | √ |

*Note:* 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.

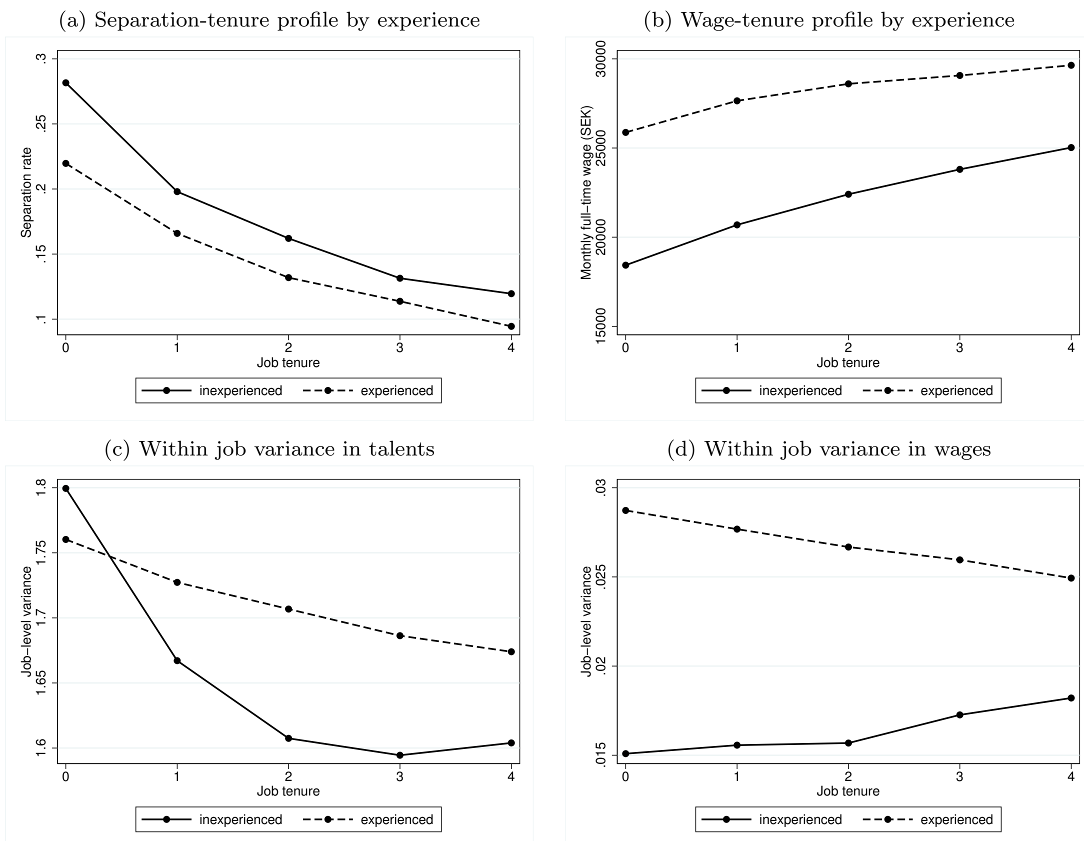


Figure A2. Tenure profiles by experience groups

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

Table A6—Responses to mismatch with worker fixed effects

| | Baseline All new hires (1) | Baseline Repeat new hires (2) | Worker FE:s Repeat new hires (3) |
|---|---|---|---|
| Panel A. 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 |
| Panel B. 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 FE:s | √ | √ | |
| Entrant test scores | √ | √ | |
| Job FE:s | √ | √ | |
| Entry occupation FE:s | | | √ |
| Entry Year FE:s | | | √ |
| Job skill requirements | | | √ |
| Worker FE:s | | | √ |

*Note:* Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. Job FE:s are (Entry occupation×Entry Year×Plant) FE:s. 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).

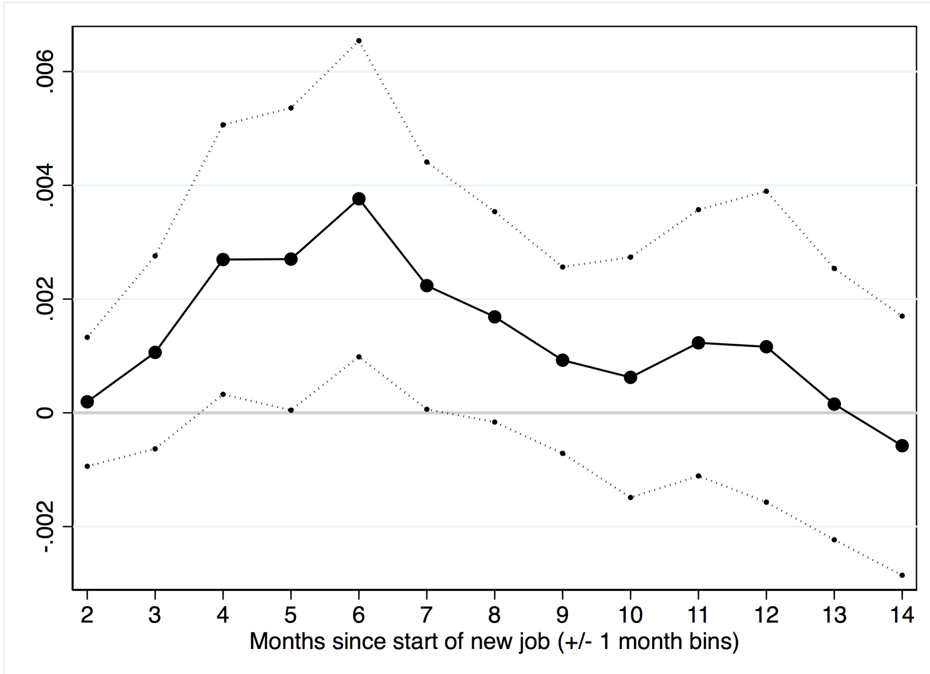


FIGURE A3. TIMING OF THE SEPARATION RESPONSE

*Notes:* The figure displays the response to initial mismatch within 3 month-bins (+/- 1 month). Dashed lines are 95% confidence bands.

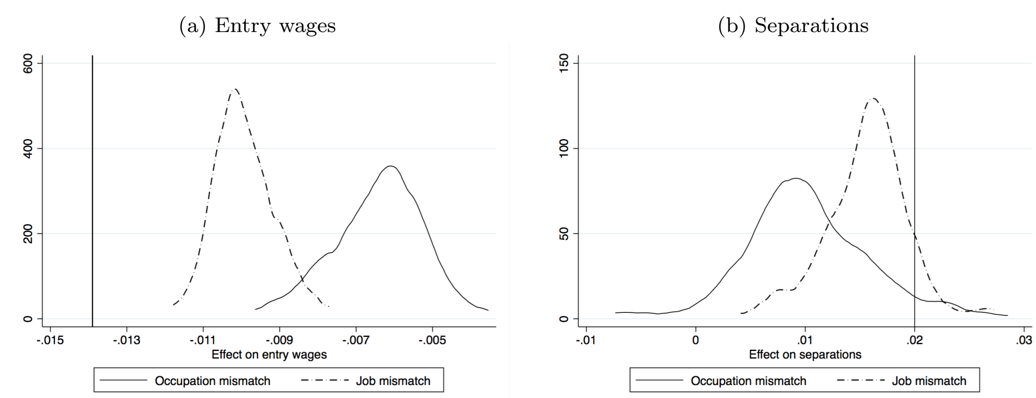


FIGURE A4. JOB AND OCCUPATION MISMATCH

*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.

Table A7—Responses to mismatch with alternative information measures

| | Alt. info. measures for the inexperienced | |
|---|---|---|
| | *Firm experience* | *Job-to-job mobility* |
| | (1) | (2) |
| *Panel A. 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 |
| *Panel B. 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 FE:s | √ | √ |
| 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.