RATIO

Ratio Working Paper No. 291

The impact of employing mismatched workers on firm productivity, wages and products

Daniel Halvarsson*

Patrik Gustavsson Tingvall**

*daniel.halvarsson@ratio.se, The Ratio Institute, P.O. Box 3203, SE-103 64 Stockholm
**patrik.tingvall@ratio.se, The Ratio Institute, P.O. Box 3203, SE-103 64 Stockholm

The Impact of Employing Mismatched Workers on Firm Productivity, Wages and Profits

Daniel Halvarsson and Patrik Gustavsson Tingvall[1]

Abstract

Educational mismatch in the form of over- and under-educated workers has long been studied in relation to labor market outcomes for individual workers. While its consequences for individual workers and society are dire, we have only anecdotal evidence of its consequences for firms' competitiveness. To bridge this gap, this paper studies the impact of mismatch on firm productivity, wages and profit. The results suggest an asymmetric effect from employing over- and under-educated workers. We find that while employing over-educated workers add to wage cost, there are no matching productivity gains, By contrast, the performance of under-educated workers more than compensates for their wage costs, leading to increased profits at the firm level. The net effect, therefore, in the form of gross operating surplus is significantly negative (positive) when firms employ over- (under-)educated workers. The results suggest that the positive effects primarily stem from under-educated young workers, whereas the losses can be traced to over-educated older workers.

Keywords: Educational mismatch · Productivity · Labor cost · Profits · Proxy variable
JEL: L25 · L60 · J24

[1]Corresponding author Daniel Halvarsson, daniel.halvarsson@ratio.se, The Ratio Institute, and Patrik Gustavsson Tingvall, The Ratio institute, The Swedish Agency for Growth Policy Analysis and Södertörn University. The authors wish to thank seminar participants at the European Economic Association conference 2017, Swedish National Conference in Economics 2016, the Research Institute of Industrial Economics (IFN), and the Ratio Institute for valuable comments and suggestions. Financial support is gratefully acknowledge from Vinnova. The usual caveats apply

1 Introduction

The increasing share of over-educated workers is one of the clearest labor market trends in most high-income countries. Approximately 50 percent of workers in the labor force in OECD countries are employed in jobs that do not match their level of educational attainment. In Sweden alone, the share of over-educated workers doubled from the mid 1970s to 2000 with no sign of a retreat in this trend (Tåhlin 2007). These observations show that education-job mismatch is a subject of concern for individual workers, the employing firms, and society at large (McGuinness 2006). The costs of mismatch for individuals are fairly well understood, especially for over-educated workers. Compared to their well-matched coworkers, overeducated individuals earn a certain rent on their surplus years of education, but their wages tend to be lower than the wages of their well-matched classmates (Rumberger 1987). Furthermore, ample evidence suggests that over-educated workers are more likely than their matched colleges to switch jobs as a result of feeling overqualified, giving rise to additional hiring and firing costs (Verhaest & Omey 2006).

What is less known, however, is whether mismatch impacts the competitiveness of the employing firms (Hartog 2000). According to the efficiency wage hypothesis, workers should earn their marginal product regardless of mismatch. This hypothesis indicates that any cost associated with employing mismatched workers should not affect firm productivity (Duncan & Hoffman 1982). In reality, however, numerous reasons may explain why firms would incur at least some of the costs associated with mismatch; such explanations may involve imperfect information about workers' true skills, job turnover, or business operations in a heavily unionized industry. The previous literature is largely silent about the effects of employing mismatched workers on firm performance. Using a matched employer-employee data set for manufacturing firms, we investigate this question using real data and examine the effects of mismatch on firms' competitiveness.

For the analysis, we depart from a production function approach (Gennaioli et al. 2013) and present a simple model where mismatch, defined by the ORU (over, required, under) education model, is allowed to impact firms' production (Duncan & Hoffman 1982). To gauge the effects of mismatch on firm performance, we draw on the literature on the labor-quality index (Hellerstein & Neumark 1999, Aubert & Crépon 2003, Vandenberghe 2013), which allows for simultaneous links among firm productivity (value added per worker), wages (labor cost per worker), and profits (gross-operating surplus).

In examining the link between firm performance and educational mismatch, we add to the existing literature by (i) adopting a comprehensive approach in analyzing the impact of mismatch on several aspects of competitiveness (wages, productivity, and profits); (ii) using a matched employer-employee data set comprising data from all manufacturing firms in Sweden; and (iii) embedding educational mismatch in a structural framework to yield implications for the links and identification of mismatch parameters.

To test the hypotheses, we use two different estimators. The first relies on a standard SUR (seemingly unrelated regressions) system of productivity and labor costs equations to identify the parameters of a third equation representing profits. To account for the potential endogeneity of the mismatch variables, we adopt a second estimator, namely, the proxy variable approach of Ackerberg et al. (2015) and Vandenberghe 2013. We find that while over- and under-education is insignificant with respect to productivity, it causes average labor costs to rise and fall, respectively, leading to our main finding that over- (under-)education has a negative (positive) effect on profits. That is, the lower wage received by under-educated workers more than compensates for the productivity differences, resulting in a net positive impact on profits. The results are corroborated in a number of robustness tests finding that the profit gains for under-education primarily stem from the productiveness of young workers, whereas the losses in profits for over-education can be traced to the higher wage costs of older workers.

The remainder of this paper is structured as follows. Section 2 presents a brief background regarding previous studies. The following section describes the theoretical framework along with empirical equations for productivity, labor cost, and profits. Section 4 turns to the data, which are described together with the empirical measures of mismatch that enter into the econometric analysis. In Section 5, the econometric method is described, and Section 6 presents the results. The final section discusses the implications and limitations of the results.

2 Literature

The literature contains two dominating approaches for studying the effects of mismatch on firm performance.[1] The first strand of literature is based on Becker (2009) and human capital theory, assuming that wages mirror productivity. If this theory holds, it would be sufficient to examine wages directly to determine the effects of employing mismatched workers on firm productivity. A surplus year of education is simply reflected by a proportional increase in productivity and thereby in wages as well. This line of reasoning is supported by, e.g., Rumberger (1987), who finds that over-educated workers do enjoy a certain rent on their surplus years of education. The study also shows that education increases productivity regardless of potential mismatch. The extent to which a worker can fully utilize his or her education, however, is determined by the limitations imposed on him or her by the job (Rumberger 1987). Although over-educated workers seem to earn a wage premium relative to their well-matched but less educated colleagues, recent studies provide convincing

[1] See McGuinness 2006 and Hartog 2000 and the sources within for a comprehensive survey.

evidence that over-educated workers have lower wages than their well-matched counterparts (Van der Meer 2006, Sicherman 1991, Duncan & Hoffman 1982, Battu et al. 1999, Dolton & Vignoles 2000, Groot 1996, Groot & Van den Brink 2000, Kiker et al. 1997).[2]

Parallel to this literature, researchers have attempted to use the impact of mismatch on job satisfaction to explain the link among mismatch, wages and productivity (Belfield & Harris 2002, Bender & Heywood 2006, Moshavi & Terborg 2002). Notably, satisfied workers are more productive than dissatisfied (mismatched) workers. Studies have shown that the correlation between job satisfaction and firm performance is approximately 15-30 percent (Judge et al. 2001, Iaffaldano & Muchinsky 1985). In fact, several studies demonstrate the effects of over-education on various correlates to workers' productivity, such as higher separation rates and turnover (Allen & Van der Velden 2001, Wolbers 2003), but scarcely any research has specifically examined the impact on firm performance. Hartog 2000, p. 139, for example, remarks that "*[i]t would obviously be highly informative if we knew the effect of over- and under-education on productivity, rather than on wages.*" One exception that does analyze the direct impact of mismatch on firm productivity is the work of Kampelmann & Rycx (2012). Using data on 3000 Belgian firms spanning the period from 1999 to 2006, they examine how the average number of years of over-, required and under-education among employees impacts the labor productivity of firms. To compensate for the increasing level of required and overall educational attainment over the period, they also test whether the lower matching of young workers impacts productivity. In short, Kampelmann & Rycx (2012) find that employing over-educated workers is beneficial for firm productivity, while having younger under-educated workers negatively impacts productivity. These results can be interpreted in line with human capital theory, where higher education translates into higher productivity regardless of mismatch (cf. Rumberger 1987).

Despite the few papers on the link between mismatch and firm performance, a growing body of literature on the demand side of the labor market is studying the effects of heterogeneous labor inputs on firm performance via the so-called index of labor quality (Hellerstein & Neumark 1995). An attractive feature of this methodology is that it facilitates a simultaneous study of the effect on firm productivity (value added per employee), wages (labor cost per employee), and profits as represented by the 'productivity-labor-cost gap' that constitutes the firm's gross operating surplus, henceforth referred to as profit. (Hellerstein & Neumark 1999, Aubert & Crépon 2003, Göbel & Zwick 2009, Van Ours & Stoeldraijer 2011, Vandenberghe 2013).The labor quality index methodology also includes a convenient test of the tenets of human capital theory (Vandenberghe 2013), as the effects from different labor

[2]When human capital theory is discussed as a vehicle for analyzing the impact of educational mismatch, a complication is that the observed wage rent on surplus years of education may be due to unobserved heterogeneity rather than to the extra education (McGuinness & Sloane 2011, McGuinness 2003, Lamo & Messina 2010, Frenette 2004, Bauer 2002). According to this logic, the observed wage rent on surplus education would be due to the workers' inner ability and other unobserved factors that influence both the decision to study and the productivity.

inputs on profits should be insignificant. The labor quality index is commonly applied to other segments of the workforce, e.g., different age categories (Vandenberghe 2013, Aubert & Crépon 2003), occupational categories Kampelmann & Rycx (2012) or workforce diversity (Garnero & Rycx 2012), but it has not yet been applied directly to the human capital function. Thus, while remaining within the human capital framework, the index methodology can directly test whether workers who are over-, required or under-educated are paid 'what they are worth' in terms of their contribution to productivity. More importantly, the method provides a test of the impact on firms' profits and, thus, competitiveness. The next section presents a simple model enabling a test of these linkages.

Before we turn to the model, however, it is important to consider some of the more unrealistic tenets of human capital theory, which relies on clearing labor markets with no rigidity such as unionization and perfect information about workers' productivity. Violation of any of these assumptions can lead to a detachment of workers' productivity from their compensation. In this case, the compensation of workers in the form of wages no longer fully reflects their productivity. On the other hand, market imperfections are not the only explanation for why workers' productivity may be detached from their compensation. One example comes from the worker's own utility function if it includes the compensation given to other workers in addition to his or her own. Having a more compressed wage structure may then foster feelings of cohesiveness among employees (Levine 1991), which can in turn lead to higher productivity. If the wage differential among workers is too large, however, it can result in feelings of unfairness that lead workers to withdraw their effort and, hence, lower their productivity (Akerlof & Yellen 1990). A related concept is the idea of hedonic wages by Lazear & Shaw (2007) in which workers trade off compensation for higher status. According to this theory, high-status jobs generally pay less than their marginal product. However, the reverse can also be predicted for workers with low-status jobs who are compensated beyond their marginal product. In any of these situations, differences in compensation and productivity still reflect an efficient market outcome.

3 Theoretical framework

We begin the presentation of the theoretical model by providing the ORU model for individual workers, followed by the framework that includes a modified version of the said model at the firm level. This setup naturally leads to three linked equations, with one each for productivity, a mirror equation for labor cost, and one for profit—all including the relevant firm-level ORU variables—which together constitute six competing hypotheses regarding the impacts of ORU on the firm.

3.1 Decomposition of attained education

Duncan & Hoffman (1982) decompose individuals' attained education (E) and job match into over-, required- and under-education, also known as the ORU model, where

$$E = Over + Required - Under. \tag{1}$$

Before restating the model at the firm level, we need some notation regarding individual workers' ORU. In the ORU framework, worker j's education attainment, given by E_{jt}, can thus be written as

$$E_{jt} = R_{kt} + \overbrace{(E_{jt} - R_{kt})}^{Mismatch}, \tag{2}$$

where mismatch stands for the years of under- or over-education depending on whether $E_{jt} - R_{kt} \lesseqgtr 0$.[3] In the literature that studies wage effects from mismatch, most studies depart from the Mincerian earnings function (Harthog, 2000). At the firm level, however, there is no such trail to follow. Adopting standard human capital theory hence seems the most promising strategy. Worker j's human capital h_{jt} is therefore approximated using the return to educational attainment $\mu_j E_{jt}$ of the form $h_{jt} \cong e^{\mu_j E_{jt}}$ with μ_j as the (Mincerian) return to education. Inserting the ORU expression from eq. (2) for educational attainment, we obtain

$$h_{jt} \cong e^{\mu_j E_{jt}} = e^{\mu_j R_{kt} + \mu_j (E_{jt} - R_{kt})}. \tag{3}$$

We divide mismatch into two parts. For over-education, we can define $O_{jt} \equiv \mathbf{1}_{E>R}\,(E_{jt} - R_{kt})$, where $\mathbf{1}_{E>R}$ is the indicator function taking the value 1 for $E_{jt} > R_{kt}$ and 0 otherwise. Similarly for under-education, $U_{jt} \equiv \mathbf{1}_{E<R}\,(R_{kt} - E_{jt})$, where $\mathbf{1}_{E<R}$ takes the value 1 for $E_{jt} < R_{kt}$ and 0 otherwise. Relaxing the assumption of symmetric effects of over- and under-education, a general ORU model in terms of individual worker's human capital can be formulated as

$$h_{jt} \cong e^{\mu_j (E_{jt})} = e^{\mu_j R_{kt} + \omega_j O_{it} - \upsilon_j U_{it}}. \tag{4}$$

letting μ_j be the return from required education associated with occupation k. However, to examine the impact from ORU on the firm, we need a suitable model of firm production.

[3] A unique partition can be made only at a fixed point in time because individuals within a given profession may enroll in additional education or change to a profession with a different educational requirement. See, e.g., Mavromaras & McGuinness (2012), who studies the dynamic aspect of over-education.

3.2 Labor productivity

We depart from a version of the Cobb-Douglas production function used in Gennaioli et al. (2013).[4] There are $i = 1, ..., N$ number of firms that are, in each period t ($t = 1, ..., T$), in the business of producing a single good Y_{it} by using K_{it} amounts of physical capital and H_{it} amounts of human capital. We focus on the value added per employee, henceforth referred to as productivity (c.f. Vandenberghe 2013), given by

$$Y_{it}/L_{it} = K_{it}^{\alpha} H_{it}^{\beta} L_{it}^{-1} e^{\nu_{it}}. \tag{5}$$

Firms face an uncertain environment that is captured by the stochastic term ν_{it}, which can include elements such as short-term shocks to productivity and unobserved heterogeneity (see Section 5 on the method for a detailed discussion). At any point in time, the firm employs a total of L_{it} number of workers, each with an amount of human capital that corresponds to h_j, which together constitute the aggregated human capital H_{it} of the firm. To form H_{it}, we first need to decide on a principle by which the set of workers' human capital h_j can be aggregated within the i firm. Since 'workers of a firm' corresponds to 'citizens of a country', we can adapt an approach from macroeconomics (Gennaioli et al. 2013) by taking the expected value of a first-order Taylor expansion of $h(\mu_j, E_{jt})$ evaluated at the average Mincerian return $1/L_{it} \sum_j \mu_j = \bar{\mu}_i$ and the average attained education $1/L_{it} \sum_j E_j = \bar{E}_{it}$ of the L_{it} workers in the firm. Letting μ_j and E_{jt} have a bivariate distribution given by $S_t(\mu_j, E_{jt})$, we define firm-level human capital as

$$H_{it} := L_{it} \mathbb{E}_{it}(h_{it}) \cong L_{it} e^{\bar{\mu}_i \bar{E}_{it}} = L_t \bar{h}_{it}, \tag{6}$$

where L_{it} is the number of workers in firm i at time t. To accommodate for subsets of workers having over-, required, or under-education in H_{it}, we rely extensively on this definition. More specifically, we derive an ORU version of H_{it} in Appendix A that yields

$$\widetilde{H}_{it} \cong L_{it} e^{\bar{\mu}_i \bar{R}_{it}} \left[1 + \bar{\omega}_i \bar{O}_{it} + \bar{\upsilon}_i \bar{U}_{it}\right]. \tag{7}$$

The level of human capital in a firm is expressed as a function of the average level of required education $\bar{R}_{it}$, over-education $\bar{O}_{it}$, and under-education $\bar{U}_{it}$. Averages are computed over firm i's total number of L_{it} workers at time t. The return to each educational category is given here by $\bar{\mu}_i$, $\bar{\omega}_i$ and $\bar{\upsilon}_i$. Observe that we have reversed the sign of $\bar{U}_{it}$ meaning that the coefficient for under-educated workers $\bar{\upsilon}$ should be negative.[5] In deriving eq. (7), the

[4] The human capital production function of Gennaioli et al. (2013) is embedded in an occupational choice general equilibrium setting. We merely draw on their inclusion of human capital and their general specification.

[5] By multiplying U_{it} with (-1), interpretation becomes easier since a negative estimated coefficient translates into a negative contribution to productivity.

return coefficients $\bar{\mu}_i$, $\bar{\omega}_i$, $\bar{\upsilon}_i$ are firm specific. Here, we assume that the returns are symmetric across firms, allowing us to drop the indexation, or alternatively, we could estimate the equations using a random-coefficient model. We choose the former because of its comparative simplicity and because the random coefficient assumption may not be empirically justified. Henceforth, we also dispense with the bar notation above the ORU variables and their associated coefficients to simplify the notation further. Since $\widetilde{H}_{it}$ is log-linear, it can be directly inserted into the production function 5. After taking the logarithm, we arrive at the version of labor productivity in eq. (5) suitable for empirical estimation, as follows:

$$lp_{it} = \alpha k_{it} + (\beta - 1)\, l_i + \beta\mu R_{it} + \beta\omega O_{it} + \beta\upsilon U_{it} + \nu_{it}. \tag{8}$$

with lowercase letters for the natural logarithm and where $lp_{it} = \log Y_{it}/L_{it}$. This specification comprises a convenient model for analyzing the impact of mismatch on productivity. Note that this expression is similar to the empirical model in Kampelmann & Rycx (2012). One important difference, however, is the β term for the share of human capital adjoined to the ORU variables. Without this term, the impact of workers is not correctly identified. The presence of β also invalidates a direct identification of the ORU effect on the profit equation (see section 5)

According to human capital theory (e.g., Becker 2009), the coefficient for under-educated workers υ should be negative.[6] The reason is that these workers have less than the required level of attained education and should therefore contribute less to labor productivity. The reverse reasoning applies to over-educated workers giving a positive ω. As argued by Rumberger (1987), more education should necessary lead to higher productivity, regardless of mismatch. The coefficient μ can be fitted with different interpretations. Consider, for example, firms for which the workers' education closely corresponds to the required educational level; then, μ reflects general human capital H_{it} as defined in (6), which is expected to contribute positively to productivity. The interpretation becomes more difficult when considering firms with multiple occupations and a large share of mismatched workers (i.e., when $\bar{O}_{it} \gg 0$ or $\bar{U}_{it} \gg 0$, or both). Then, $\bar{R}_{it}$ more reflects the general occupational structure of a firm, against which workers' educational attainment is compared. Many firms have multiple layers of occupations (Tåg 2013) that can require high levels of human capital at the top and a broad base of workers employed in occupations that are less human capital intensive. To capture some of the variation embedded in firms' occupational structure, we therefore include a set of occupational controls (see Section 4 for a variable description).

Thus, with regard to the empirical equation, we form two hypotheses regarding the

[6] Disregarding other factors that could also potentially affect productivity.

expected contribution from over- and under-educated workers to lp_{it}.

$$\text{(i):} \qquad \omega > 0, \tag{9}$$

$$\text{(ii):} \qquad \upsilon < 0. \tag{10}$$

3.3 Labor cost

To estimate the effect of heterogeneous labor inputs on average labor cost, the literature on the labor quality index relies on the fact that a firm's total labor cost can be written as a sum of the labor cost for different groups of a firm's workforce, which can in turn be expressed as the labor cost per employee, analogous to the specification of productivity (value added per employee) (see, e.g., Aubert & Crépon 2003). By using $\widetilde{H}_{it}$, we can apply the same methodology to the ORU framework. To arrive at an expression for the average labor cost that mirrors the labor quality methodology, it is convenient to depart from the first-order condition for profit maximizing. Using the Cobb-Douglas production function in eq. (5) with a linear cost function, we obtain the following first-order condition:

$$\beta K_{it}^{\alpha} \left(H_{it}'\right)^{\beta-1} = \breve{w}. \tag{11}$$

Total labor cost is here given by $\breve{w}\widetilde{H}_{it} \equiv W_{it}$. Eq. (11) can hence be restated using $W_{it}/\widetilde{H}_{it}$ for $\breve{w}$. Collecting terms and taking the logarithm results in the following expression:

$$\begin{aligned} lc_{it} = & \log \beta^{w} + \alpha^{w} k_{it} + \left(\beta^{w} - 1\right) l_{i} + \beta^{w} \mu^{w} R_{it} + \beta^{w} \omega^{w} O_{it} \\ & - \beta^{w} \upsilon^{w} U_{it} + \nu_{it}^{w}, \end{aligned} \tag{12}$$

which describes the average labor cost from employing ORU workers with $lc_{it} = \log W_{it}/L_{it}$. This expression represents the complementary side of firm productivity and is the mirror image of labor productivity in the sense that it contains the same variables and parameter configurations up the constant $\log \beta$. This difference is emphasized in the labor quality literature (see, e.g., Vandenberghe 2013) in that it allows for a direct comparison of the productivity effect from a particular labor input to the effect it levies on the cost side of the firm. Given that the standard conditions in human capital theory are fulfilled, the empirical version of the two equations should yield the same hypotheses regarding the coefficients ω^{w} and υ^{w} that

$$\text{(iii):} \qquad \omega^{w} > 0, \tag{13}$$

$$\text{(iv):} \qquad \upsilon^{w} < 0. \tag{14}$$

A further implication is that the same parameter estimates should apply across eq. (8) and (12), which means that we can form two additional hypotheses of equality between the parameters

$$\text{(v):} \qquad \omega - \omega^w = 0, \tag{15}$$

$$\text{(vi):} \qquad \upsilon - \upsilon^w = 0. \tag{16}$$

In the presence of imperfect competition, e.g., labor market rigidity, heavily unionized industries, or uncertainty about workers' productivity, these hypotheses may be rejected. However, as mentioned in the literature section, interdependent utility configurations among workers or hedonic wage-setting could also result in distributional differences between productivity and compensation. Although these theories do not relate to educational mismatch per se, a common trait is that they predict a compressed compensation structure relative to that of productivity, which might in turn invalidate the equality hypotheses on efficiency grounds. Hence, regardless of the reason, a rejection of the equality hypotheses can be seen as an indirect test of the basic form of human capital theory and a test of the net effects on firm performance from employing heterogeneous ORU workers. To test the two latter hypotheses, we construct a third equation that constitutes firms' profit in the form of its gross operating surplus, i.e., the productivity-labor cost gap (Hellerstein & Neumark 1999, Aubert & Crépon 2003, Göbel & Zwick 2009, Van Ours & Stoeldraijer 2011, Vandenberghe 2013).[7]

3.4 Productivity-labor cost gap

To form an expression for this profit measure, the labor cost equation (lc_{it}) is subtracted from that of productivity (lp_{it}). This step is the final piece of the equation system, and together with productivity and labor cost, it provides insight on how ORU workers affect not only productivity, wages but also profits. It is formally written as:

$$\pi_{it} = \eta_k k_{it} + \eta_l l_i + \eta_3 R_{it} + \eta_O O_{it} + \eta_U U_{it} + \nu'_{it}. \tag{17}$$

where $\pi_{it} = \log Y_{it}/L_{it} - \log W_{it}/L_{it} = \log Y_{it}/W_{it}$ and $\eta_k = \alpha - \alpha^w$; $\eta_l = \beta - \beta^w$; $\eta_R = \beta\mu - \beta^w \mu^w$; $\eta_O = \beta\omega - \beta^w \omega^w$; $\eta_U = \beta\upsilon - \beta^w \upsilon^w$; $\nu'_{it} = \nu_{it} - \nu^w_{it}$. Unlike for the equations of productivity and labor cost, the parameters of eq. (17) cannot be identified directly by regressing the ORU variables on π_{it}, which means that the two latter hypotheses cannot be evaluated based on running a separate regression on eq. (17) alone. This problem is

[7] Gross operating surplus refers to the surplus value added that remains after expenditures on labor inputs have been subtracted. From this surplus, firms can pay investors and pay taxes. It also provides a source of financing for investments (definition from the OECD).

discussed in Section 5. Next, we turn to the data and the construction of the empirical measures that correspond to the variables entered into the models in the current section.

4 Data and variable creation

To empirically test hypotheses (i) to (vi), we need to use information on worker and firm characteristics. Our firm- and individual-level data are from Statistics Sweden. Information on workers is obtained from the LISA database, spanning the period 1990-2013, which contains full population data on the labor force (individuals aged 16-65 years). The LISA database supplies us with information on earnings, age, gender, educational attainment, occupation, and location. Individual-level data are tagged with the identity of the plant and firm tied to the worker. Firm-level data cover all firms and supplies us with information on firms' inputs and outputs, such as investments, capital stock, expenditures, profits, raw material inputs, number of employees, and value added. The data also contain regional information and NACE codes for industry classification. Due to missing data for some variables, the analysis covers the period 2001-2010.

Since our aim is to examine the effects from employing mismatched workers, we exclude students and individuals outside the labor force. We also exclude individuals younger than 19 or older than 64, as they are categorized as not being part of the labor force. Finally, we restrict the sample to the private sector, since the public sector is primarily in the market for the provision of public goods.

To determine the number of years of education attained, we have information on educational attainment based on the international standard classification of education (ISED 97). Since we are interested in knowing the distance (years) to the required level of education, we apply the following enumeration of 'years of education': For pre-primary school (level 0), 0 years of schooling is assigned. For primary education (level 1 and level 2), an additional 9 years is assigned. For individuals whose highest attained education consists of secondary schooling (level 3), 1, 2 or 3 years is added to the 9 years depending on the length of their education. For individuals who attained less than 2 years of higher education or vocational training (level 4), 4 years of schooling is added to the 9 years of primary schooling. For individuals who completed 2, 3, 4 or 5 years of higher education (level 5), the corresponding number of years is added to the 3 years of secondary education; hence, these individuals are assigned 14, 15, 16, or 17 years, respectively. Individuals with 2 or 4 years of postgraduate-level education (level 6) are assigned an additional 2 or 4 years following their 4-year higher education, which makes their total years of education either 18 or 20 years. Hence, 20 years of education is the highest possible amount of education attained.

To define over- and under-education, we follow the consensus in the ORU literature and define required education in relation to the typical educational attainment of all workers in a given occupation. To sort workers by occupation, we use information on 115 3-digit (SSYK) occupational categories based on the ISCO-88 classification. We see, for instance, that workers with occupational code 211 representing 'Physicists, chemists and related professionals' had 16 years of required education (mode) in 2001. To take another example, 'Metal- and mineral-products machine operators' (code 821) employed $44,045$ workers in 2001, but that number declined to 38,805 in 2010. While the required education remained at 11 years throughout the period, average over-education rose from 0.355 years to 0.514 years, and under-education increased from -0.595 to -0.437. This trend is not specific to this occupation but represents a general trend observed for a substantial number of industries.[8]

We also collect basic information on workers' age, gender, and country of birth (outside the EU25, USA, Great Britain and Australia) and the local labor market where they work. Since the analysis is conducted at the firm level, individual characteristics and ORU variables are aggregated. Guided by the approximation of human capital in eq. (7), we calculate firm-year averages of years of over-, required- and under-education. We also calculate the proportion of immigrants, females, and average worker age each year within firms. We define 75 local labor markets according to Korpi & Clark (2015). For each local labor market, we calculate the average level of educational attainment among workers in the region and the total number of employed workers to account for regional differences in the supply of workers.

Firm-level data are restricted to the manufacturing sector, and firms with at least 10 employees with all variables measured in monetary terms are to be no less than 1000 SEK (approx. 100 Euro). To reduce the impact of outliers, we drop the bottom and top 1st percentiles of value added per worker and labor costs per worker, leaving us with 55,795 firm-year observations. All monetary variables are deflated using industry deflators. To find appropriate industry-level deflators for firms in the different industries, we rely on NACE rev. 2 to the largest extent possible. For firms that still exist in the data in 2007, the rev. 2 code is rolled back for earlier years. For firms that existed prior to 2007 that have only information regarding their 2002 classification (rev. 1 and or rev 1.1), we use deflators based on the NACE rev. 1 regime. Instead of explicitly translating industry affiliation from 2002 to 2007 or backward, we can avoid the mapping problem. Rather than using industry dummy variables in the empirical models, we resort to demeaning all variables based on their industry-year to ensure that we use a correct industry classification.

The main dependent variables we consider are value added per employee (lp_{it}) *(VA per empl)* and labor costs per employee (lc_{it}) *(LC per empl)*, which combine to constitute a direct measure of profit π_{it} by way of the gross operating surplus. Value added and labor

[8]Supplementary description for all occupations can be provided by request from the authors.

Table 1: Descriptive statistics

Variable	Obs	Mean	Std.dev	Min	Max
VA per empl (log)	55795	6.172	0.446	-0.640	10.640
LC per empl (log)	55795	5.940	0.268	-1.016	8.118
Capital (log)	55795	8.339	1.841	0.644	16.853
Labor (log)	55795	3.556	1.045	2.303	9.982
Materials (log)	55795	9.748	1.660	0.000	17.785
Req-ed	55795	10.839	0.882	9.000	20.000
Over-ed	55795	0.920	0.494	0.000	3.563
Under-ed	55795	0.512	0.325	0.000	4.200
Age	55795	42.428	4.520	23.462	58.714
Immigrants	55795	0.029	0.059	0.000	0.952
Females	55795	0.227	0.182	0.000	1.000
LLM avg ed	53142	11.928	0.374	10.923	12.799
LMM size (log)	53142	11.791	1.475	7.273	14.226
Nr occupations	55470	10.634	7.217	1.000	84.000
Occ high/low	55795	0.863	2.526	0.000	166.000
Req-sd	55795	1.395	0.569	0.000	4.634

Summary statistics for variables are presented as they are enter into the analysis. All monetary variables are deflated using appropriate 2-digit industry deflators from OECD statistics.

costs are deflated using the industry deflator for value added (VALP). From Table 2, we note that average productivity is higher than average labor cost (which should be equal from the Cobb-Douglas function). In addition, the standard deviation of labor cost is approximately half the size of productivity, which suggests a compressed wage structure. This can also be seen in Figure (1), which plots the kernel density of the respective dependent variable. *Capital* is defined as the deflated book value of machines and buildings. *Labor* corresponds to the average number of workers employed by the firm in the current year. It is recalculated to full-time equivalents using industry information on total hours worked and total employment from OECD statistics. Intermediate inputs, used as a proxy variable, captures expenditures on *Materials* deflated by the industry deflator (INTP) for intermediate goods.

To measure educational mismatch, we rely on eq. (7) of the extended ORU model. Here, we follow Kampelmann & Rycx (2012), who use the mode of education, which represents the level of education attained by most workers employed in one of 115 occupational categories in each given year. To obtain the aggregated measure of required education ($req\text{-}ed_{it}$), the firm-year average is computed for all $j = 1, ..., L_{it}$ employees using the required education $R_{kt,j\in k}$ linked to each employee at the 3-digit occupational level ($k = 1, ..., 115$),

$$Req\text{-}ed_{it} = \frac{1}{L_{it}} \sum_{j=1}^{L_{it}} R_{kt,j\in k} \tag{18}$$

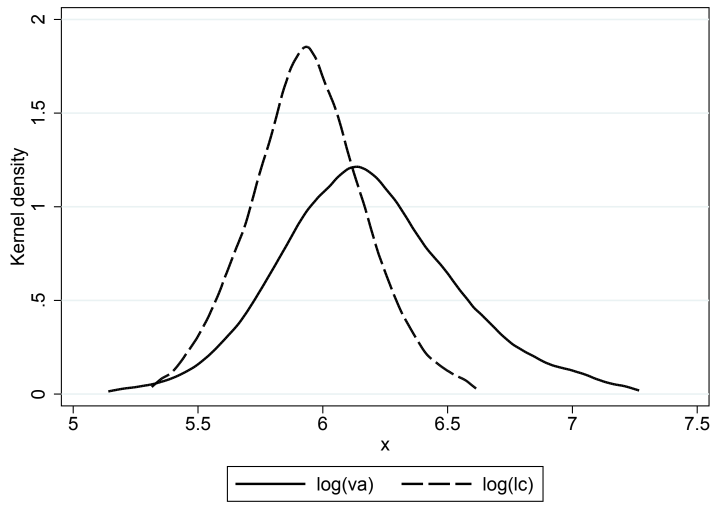


Figure 1: Kernel density of log productivity and log labor cost for firms in manufacturing represented in our sample over the period 2001-2010.

The measures of over- and under-education are computed analogously by using E_{jt} to correspond to a worker's educational attainment at time t.

$$\textit{Over-ed}_{it} = \frac{1}{L_{it}} \sum_{j=1}^{L_{it}} \mathbf{1}_{E>R} \left(E_{jt} - R_{kt,j\in k}\right) \tag{19}$$

$$\textit{Under-ed}_{it} = \frac{1}{L_{it}} \sum_{j=1}^{L_{it}} \mathbf{1}_{R>E} \left(R_{kt,j\in k} - E_{jt}\right) \tag{20}$$

where $\mathbf{1}_{E>R}$ corresponds to the indicator function taking the value 1 for over-educated workers and 0 otherwise. For $\textit{Under-ed}_{it}$, $\mathbf{1}_{R>E}$ is defined similarly but as 1 for under-educated workers and 0 otherwise. Table 1 shows that average over-education is more pronounced in the sample with an average year of 0.92 year compared to 0.512 years for under-education. As for the dependent variables, the corresponding density plots for educational mismatch are plotted in Figure 2. In general, higher levels of over-education are observed among employees in the manufacturing sector.

To control for age structure, we include the average age of the workforce. Recognizing that age is related to ORU (Kampelmann & Rycx 2012), in the estimations, we endogenize average age along with the structural variables of capital, labor and ORU (see the next section on the method for details). In addition to age, we include three sets of additional controls. First, to control for worker heterogeneity, we include *Immigrants* as the proportion of the workforce born outside EU 25, North America, Australia, and Great Britain and

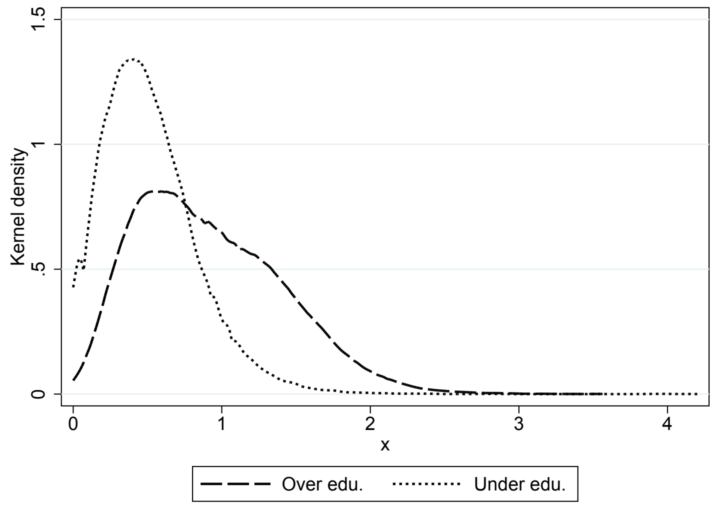


Figure 2: Kernel density of avg. over- and under-education for employees in manufacturing firms represented in our sample over the period 2001-2010.

Gender for the proportion of females in the workforce. Second, following Gennaioli et al. (2013), we include regional variables that potentially affect firm productivity. These are *LLM avg edu* and *LLM size*, and they capture human capital abundance and the size of local labor markets, respectively.

To control for occupational structure, we include a third set of controls. Here, we consider the total number of occupation in a firm in any given year (*Nr occ*). The average firm has 10 occupations, but the number ranges from 1 to 84. Since the ORU parameters are defined at the occupational level, not controlling for this heterogeneity risks having bias in the coefficients on educational mismatch. To capture the type of multi-layered occupational structures, e.g., a broad base of workers occupied in jobs with a low level of required education followed by fewer layers of workers in occupations that require a higher education, we also compute the number of workers in each firm for any given year who are employed in the occupation with the highest and lowest required education, respectively, and form the ratio of high/low (*Occ high/low*). This variable takes the value of 1 for firms with an equal number of workers in the top as in the bottom of the multi-layered organizational structure order by the level of required education. The average firm has a high/low ratio of 0.86, reflecting an average wider base of workers employed in occupations with a low required educational level compared to the number of workers in occupations that require relatively higher levels of education. However, neither the number of occupations nor the high/low ratio necessarily reflects the variation in required education between occupations in firms. A firm can still have a large number of occupations with a distinct occupational hierarchy without necessarily having the occupations be far apart in terms of educational requirements. To capture this heterogeneity, we additionally calculate the standard deviation of required education (*Req-sd*) for all occupations represented by the workers in the firm. These controls may capture spurious correlations transmitted via the ORU variables to firm productivity and labor cost.

5 Empirical method and identification

There are several challenges in estimating the empirical versions of equations (8), (12), and (17). The first problem arises when trying to estimate the structural parameters in the profit equation. If we run separate regressions on (17), there is no way of identifying the relevant parameters to effectively test the validity of hypotheses (v) that $\omega - \omega^w = 0$ and (vi) that $\upsilon - \upsilon^w = 0$. Consider, e.g., the coefficient for over-education (O_{it}); in the profit equation, it is given by $\eta_O = \beta\omega - \beta^w\omega^w$. From the coefficient of l_{it} given by $\eta_l = \beta - \beta^w$, there is no way of recovering $\omega - \omega^w$. Simply dividing by η_l would not provide a valid identification strategy, since

$$\frac{\eta_O}{\eta_l} = \frac{\beta\omega - \beta^w\omega^w}{\beta - \beta^w}. \tag{21}$$

For example, if $\beta = \beta^w$, the above expression is undefined. While it is possible to recover the estimates $\hat{\omega}$ and $\hat{\omega}^w$ from running separate regressions for labor productivity and labor cost, we cannot test the significance in hypothesis (v) unless the error terms are independent. However, since both equations are derived from the same production function, the error terms are likely dependent. To resolve this issue, we can stack the labor productivity and labor cost equation to form a system of equations that can be estimated jointly with the seemingly unrelated regression (SUR) estimator. The estimates of $\hat{\beta}$ and $\hat{\beta}^w$ can then be recovered to solve for $\hat{\omega} - \hat{\omega}^w$ as a non-linear combination of $\hat{\beta}$ and $\hat{\beta}^w$. Importantly, this approach completely dispenses with the need to run a separate regression for the profit equation.

The remaining econometric challenges all come from various components of the empirical error term ν_{it}. Here, we consider a standard Hicks-neutral error term of the following form:

$$\nu_{it} = \delta_i + \varphi_{it} + \varepsilon_{it}, \tag{22}$$

where δ_i reflects unobserved characteristics specific to each firm, φ_{it} represents any short-term productivity shocks that are partly known to the firm, and ε_{it} is a stochastic term. Beginning with δ_i, it is well known that a large part of firm heterogeneity is firm specific. This includes management skills, organizational form and other aspects that change very little from year to year, including any time-invariant industry and regional characteristics that may impact a subset of firms. Not controlling for the variance in δ_i across firms introduces omitted variable bias in the SUR estimator. To account for any such heterogeneity, we choose to estimate a first-differenced version of the system, which effectively expunges the δ_i term. This estimator is henceforth referred to as FDSUR. To our knowledge, we are not aware of any attempts in the previous literature to resolve this problem.

Turning to the short-term productivity term φ_{it}, we need to account for the well-known fact that factor inputs $(k_{it}, l_{it}, \tilde{h}_{it})$ can be correlated with productivity shock φ_{it}, making factor inputs endogenous and thus violating the assumption that $\mathrm{E}[\phi_{it}|k_{it}, l_{it}, \tilde{h}_{it}] = 0$. This problem is sometimes referred to as a "transmission or simultaneity bias" (Gandhi et al. 2013), and it arises if firms (or decision makers in firms) choose to adjust the level of factor inputs based on information unavailable to the econometrician merely observing the *ex post* result of such choices. There are numerous ways to address this problem, and the associated body of literature is large (see, e.g., Olley & Pakes 1996, Levinsohn & Petrin 2003, Ackerberg et al. 2006, 2015, Gandhi et al. 2013). A popular strategy is to use some observable variable that responds quickly to changes in productivity. With such a proxy for productivity, the endogenous component φ_{it} can ideally be inverted out. In their groundbreaking paper, Levinsohn & Petrin (2003) suggest that intermediate inputs, such as materials or electricity expenses, are suitable proxy variables for φ_{it}. While they mainly focused on the endogenous aspect of capital, Ackerberg et al. (2006) later noted in a working paper version of Ackerberg et al. (2015) that the same argument could equally well be applied to the input of labor. They also provided an appropriate extension of the Levinsohn & Petrin (2003) model to account for the endogeneity of both capital and labor inputs.

For our purposes, the ORU variables are part of firms' human capital inputs, which makes the Levinsohn & Petrin (2003) model more attractive. However, a limitation in both Ackerberg et al. (2015) and Levinsohn & Petrin (2003) is the inattention to the fixed-effect term of δ_i, which, in addition to the considerations mentioned above, could also be a source of endogeneity. We therefore choose to proceed following Vandenberghe (2013), who extends the approach of (Ackerberg et al. 2006, 2015) to account for δ_i together with a vector of control variables (henceforth referred to as FDPV).[9] Importantly, Vandenberghe (2013) also allows for a more generous labor-input vector that is consistent with our ambition to include several ORU variables in the model. With this approach, we can account for the endogeneity of not only l_{it} and k_{it} but also the ORU variables that are embedded in the firm-level human capital term $\widetilde{H}_{it}$ in eq. (7). While the FDSUR estimator is used for its simplicity as a baseline reference, we focus on the results from using the FDPV estimator when discussing the main results in the next section.

Let m_{it} be the log of material inputs of firm i at time t. If m_{it} is fully flexible, firms will respond to short-term productivity shocks (φ_{it}) by quickly adjusting their purchases

[9] The abbreviation FDPV stands for first-difference proxy variable. As we shall see, this approach also makes use of first differences to handle the fixed-effect term.

of intermediate inputs. The original idea comes from Levinsohn & Petrin (2003) regarding how fluctuations in m_{it} can be used to recover information about φ_{it}. In addition to φ_{it}, m_{it} also depends on the level of other quasi-fixed input factors, such as capital (Levinsohn & Petrin 2003), labor inputs (Ackerberg et al. 2015) or *labor quality*, as in Vandenberghe (2013). In our case, m_{it} is also a function of R_{it}, O_{it} and U_{it}, and we write

$$m_{it} = f\left(\varphi_{it}, k_{it}, l_{it}, R_{it}, O_{it}, U_{it}\right), \tag{23}$$

or to simplify the notation, we can write the right-hand side in matrix form as $f\left(\varphi_{it}, \boldsymbol{X}_{it}\right)$.[10] Under the assumption that f is monotonic in all its arguments (Ackerberg et al. (2015)), φ_{it} can be inverted out and expressed as a non-parametric function of intermediate inputs, capital, and ORU variables to yield

$$\varphi_{it} = f^{-1}\left(m_{it}, \boldsymbol{X}_{it}\right) \tag{24}$$

Substituting for this expression in (22) results in an error term in which the productivity terms have been replaced with a function of directly observable variables

$$\nu_{it} = \delta_i + f^{-1}\left(m_{it}, \boldsymbol{X}_{it}\right) + \varepsilon_{it}. \tag{25}$$

In place of the endogenous φ_{it} term, the new error term now contains an unknown function f. However, since f can be approximated, this manipulation provide a possible way around the endogeneity problem.

We can now turn to the productivity equation in (8) and insert the expression for (25) to obtain the following empirical version of the equation:

$$\log Y_{it}/L_{it} = \alpha k_{it} + \beta \tilde{h}_{it} + f^{-1}\left(m_{it}, \boldsymbol{X}_{it}\right) + \boldsymbol{\Gamma}' \boldsymbol{Z}_{it} + \delta_i + \varepsilon_{it}, \tag{26}$$

where $\boldsymbol{Z}_{it}$ is a vector containing the set of additional control variables described in Section 4. The situation for the labor cost equation is analogous and is therefore not illustrated here. Because $f^{-1}\left(m_{it}, \boldsymbol{X}_{it}\right)$ is non-parametrically defined Ackerberg et al. 2015,[11] the empirical equation can be restated in terms of another second non-parametric function ϕ_t that encompass all of $\alpha k_{it}, \beta \tilde{h}_{it}$, $f^{-1}\left(m_{it}, \boldsymbol{X}_{it}\right)$, and $\boldsymbol{\Gamma}' \boldsymbol{Z}_{it}$ as follows:

$$\log Y_{it}/L_{it} = \phi_t(m_{it}, \boldsymbol{X}_{it}, \boldsymbol{Z}_{it}) + \delta_i + \varepsilon_{it}. \tag{27}$$

[10] It is possible to allow the function f to vary over time, but following Vandenberghe (2013), we consider f to be common to all firms and fixed throughout the period.

[11] In the paper by Gandhi et al. (2013), the identification using a value added production function is criticized. However, Ackerberg et al. (2015) shows that as long as the gross production function is Leontief in materials (intermediate inputs), identification is accomplished.

To eliminate the fixed-effect term δ_i, Vandenberghe (2013), we resort to first differencing with the following expression:

$$\Delta \log Y_{it}/L_{it} = \Delta\phi_t \left(m_{it}, \boldsymbol{X}_{it}, \boldsymbol{Z}_{it}\right) + \Delta\varepsilon_{it}, \tag{28}$$

where Δ is the usual first-difference operator. This expression can then be estimated non-parametrically via OLS to yield an unbiased estimate of the composite ϕ_{it} function.[12] We approximate ϕ_t by way of first differencing a second-degree polynomial expansion of $k_{it}, l_{it}, R_{it}, O_{it}, U_{it}$, and the average employee age Age_{it} in the firm. While the age variable belongs to the set of control variables, it is closely related to both l_{it} and the ORU variables. We therefore choose to include it in the polynomial instead of the $\boldsymbol{Z}_{it}$ term. With regard to the rest of the control variables, they are simply added linearly to the polynomial (see Vandenberghe (2013) for more details). The resulting first-difference estimates $\hat{\boldsymbol{\tau}}^{FD}$ can then be adjoined to the polynomial terms in *levels* to form the unbiased estimate of ϕ_t given by $\hat{\phi}_t^{FD}$. For given values of the parameter vector $\boldsymbol{\beta} = [\alpha, \beta, \mu, \omega, \upsilon, \gamma_{Age},]$, we can subtract the production function variables from $\hat{\phi}_{it}^{FD}$ to partial out the productivity term φ_{it} as a function of the $\boldsymbol{\beta}$ vector, giving

$$\varphi_{it}\left(\boldsymbol{\beta}\right) = \hat{\phi}_{it}^{FD} - \boldsymbol{X}_{it}\boldsymbol{\beta} - \boldsymbol{Z}_{it}\hat{\boldsymbol{\Gamma}}^{FD}, \tag{29}$$

where $\boldsymbol{Z}_{it}$ is treated as free variables. Thus, the control variables are entered into eq. (29) in levels with their first difference estimate recovered from $\hat{\boldsymbol{\tau}}^{FD}$. To identify the $\boldsymbol{\beta}$ vector, we want to form appropriate moment conditions. Following Ackerberg et al. (2015) and Vandenberghe (2013), we make the additional assumption that φ_{it} follows a first-order Markow process as described by

$$\begin{aligned} \varphi_{it}\left(\boldsymbol{\beta}\right) =& E\left[\varphi_{it}\left(\boldsymbol{\beta}\right)|\varphi_{it-1}\left(\boldsymbol{\beta}\right)\right] + \xi_{it}\left(\boldsymbol{\beta}\right), & (30)\\ =& g\left(\varphi_{it-1}\left(\boldsymbol{\beta}\right)\right) + \xi_{it}\left(\boldsymbol{\beta}\right), & (31) \end{aligned}$$

where g is again some non-parametric function.[13] To evaluate (31), we approximate g with a 3rd-degree polynomial expansion of φ_{it-1} with $g\left(\varphi_{it-1}\right) \approx \gamma_1\varphi_{it-1} + \gamma_2\varphi_{it-1}^2 + \gamma_3\varphi_{it-1}^3$ that we fit using OLS. The predicted residuals are then used to form the following moment

[12] With respect to eq. (27), Vandenberghe (2013) differs from Ackerberg et al. (2015) in two regards. First, Ackerberg et al. (2015) do not include the $\boldsymbol{Z}_it$ vector of controls in the ϕ_t function. Second, Vandenberghe (2013) accounts for the fixed-effect term δ_i, while AFC assumes that it is contained in ε_{it}.

[13] In this regard, we differ from Vandenberghe (2013), who treat all variables as endogenous, and we thereby include their coefficient vector of controls $\boldsymbol{\Gamma}^{FD}$ in the $\boldsymbol{\beta}$ vector. In this aspect, our our approach is more similar to that of Levinsohn & Petrin (2003), who allows for freely varying variables when estimating the vector for endogenous variables.

conditions of the structural parameters:

$$\mathbb{E}\left[\xi_{it}\left(\boldsymbol{\beta}\right)|\boldsymbol{X}_{it}\right]=0. \tag{32}$$

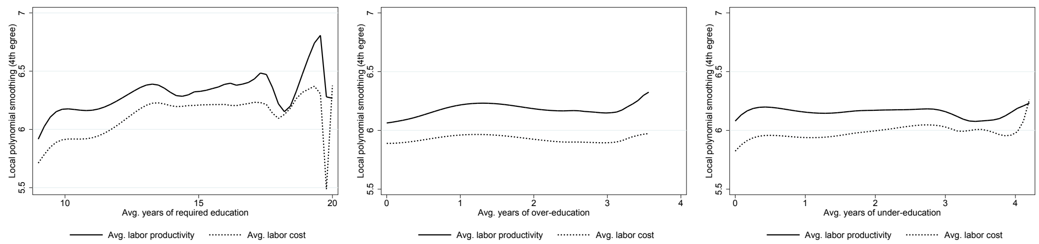
To find the $\boldsymbol{\beta}$ vector that satisfies the moment condition, we use generalized method of moments. Note that the above moment condition relies on the timing of firm inputs. It requires that both capital and human capital parameters are chosen at $t-1$ or earlier. There is a discussion in the literature regarding the timing of labor. If labor is chosen later, the moment condition should be modified by using lagged labor variables in $\hat{\boldsymbol{X}}_{it}$ to ensure identification. However, using lags, we find that it significantly reduces the power of the estimator (Ackerberg et al. 2015). We often found that lagging labor and ORU variables in the moment conditions led to non-convergence of the GMM estimator. We therefore choose to use contemporaneous variables in $\boldsymbol{X}_{it}$. Non-parametric standard errors are computed using 250 firm-block bootstraps.

Finally, as discussed earlier, we cannot recover the ORU parameters from directly estimating the profit equation. To test hypotheses (v) and (vi) of whether $\omega-\omega^{w}=0$ and $\upsilon-\upsilon^{w}=0$, we can employ the bootstrapping procedure. Since we have 250 point estimates for each variable in both the productivity and labor cost equations, we can directly observe potential cross-equation correlation, allowing us to form estimates and test the significance of the parameters in the profit equation without estimating it directly as we did for the FDSUR estimates. While we follow the broad strokes of the approach outlined in Vandenberghe (2013), we offer a new and novel way of identifying the structural parameters of human capital in a system with more than one equation—in the FDSUR model without concern for endogeneity and in the FDPV model when we allow for endogenous human capital variables.

6 Results

Figure 3 provides a first glance at the relation between labor cost and labor productivity for different levels of education and mismatch. In the figure, productivity and labor cost have been fitted with a 4th-degree polynomial over the degree of required education and educational mismatch. For all ORU measures, labor productivity is larger than labor cost, indicating a positive profit for both over- and under-education. Specifically, we note that productivity is steadily increasing up to 14 years of required education, after which the function flattens out.[14]. For over- and under-education, two opposite trends are suggested. For additional years of over-education, productivity increases while labor costs remain at

[14]The large variation for higher values of the ORU variables reflects the small number of firms with ORU of that magnitude



(a) Average required education (b) Average over-education (c) Average under-education

Figure 3: Local polynomial smoothing (4th degree) of productivity and labor cost for degrees of ORU among workers

the same level, while for under-education, the reverse seems to be true. For each additional year of under-education, productivity remains at the same level while labor cost increases. Hence, in terms of hypotheses (i) to (iv), only (i) is supported by ocular inspection, i.e., indicating that over-education is positively associated with higher productivity. Regarding hypothesis (v) and (vi), the figures lend little support in favor of the null hypotheses that profit is unaffected by over- and under-education. However, these figures do not tell us much about the causality between the ORU variables and firm productivity, labor cost or profit; therefore, the rest of this section presents the econometric results from the FDSUR and FDPV estimators.

The main results are presented in Table 2. The table contains two sets of estimates: columns (1) to (3) present the FDSUR estimates for productivity, labor cost, and profit, while columns (4) to (6) present the corresponding estimates from the FDPV estimation. The latter are also our preferred results when discussing the findings. The structural ORU coefficients cannot be observed directly from the estimates but must be backed out. These results are presented in the table under *Implied* parameters, and when discussing the estimates for the ORU parameters, we focus on the implied parameter estimates of μ, ω and υ of required, over- and under-educated employees instead of the raw estimates for *Req-ed*, *Over-ed*, and *Under-ed* that constitute $\hat{\beta\mu}$, $\hat{\beta\omega}$ and $\hat{\beta\upsilon}$. For the profit equation, the implied ORU parameters refer to the difference of the implied estimates; thus, $\hat{\mu} - \hat{\mu}^w$, $\hat{\omega} - \omega^w$ and $\hat{\upsilon} - \hat{\upsilon}^w$.[15]

Turning to the results, we see that the estimate of β is close to the unconditional value of β at 0.79 inferred from Figure 1. For both α and β, the estimates from FDPV are consistently higher than those for FDSUR, potentially reflecting a downward endogeneity bias in the FDSUR models. Regarding the ORU estimates, our first observation is that the

[15]To control for spurious correlation, all variables are demeaned by their corresponding industry year average as in Vandenberghe (2013). We prefer this approach over dummy variables as it ensures convergence of the GMM model.

Table 2: Estimation results of the main models

	Prod.	Lab. cost FDSUR	Profit	Prod.	Lab. cost FDPV	Profit
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Labor	-0.223***	-0.169***	-0.053***	-0.195***	-0.147***	-0.048***
	(0.010)	(0.007)	(0.009)	(0.012)	(0.008)	(0.011)
Capital	0.035***	0.019***	0.016***	0.045***	0.024***	0.021***
	(0.003)	(0.001)	(0.003)	(0.003)	(0.002)	(0.013)
Req-ed	-0.019**	0.019***	-0.038***	-0.032***	0.018***	-0.050***
	(0.008)	(0.004)	(0.007)	(0.009)	(0.005)	(0.007)
Over-ed	-0.001	0.019***	-0.020**	-0.005	0.020***	-0.025***
	(0.009)	(0.005)	(0.008)	(0.009)	(0.005)	(0.009)
Under-ed	0.013	-0.016***	0.029***	0.023**	-0.017**	0.040***
	(0.011)	(0.006)	(0.010)	(0.011)	(0.006)	(0.010)
Age	-0.007***	-0.001**	-0.006***	-0.008***	0.000	-0.007***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Immigrants	0.058	0.103***	-0.045	0.014	0.088***	-0.074
	(0.061)	(0.035)	(0.054)	(0.061)	(0.031)	(0.055)
Females	-0.059*	-0.026	-0.032	-0.055*	-0.025	-0.031
	(0.032)	(0.018)	(0.028)	(0.031)	(0.016)	(0.027)
LLM ed	-0.029	-0.007	-0.022	-0.025	-0.006	-0.018
	(0.022)	(0.011)	(0.020)	(0.022)	(0.012)	(0.021)
LLM size	0.004	0.000	0.004	0.004	0.001	0.004
	(0.006)	(0.003)	(0.005)	(0.006)	(0.003)	(0.006)
Nr Occ	0.004***	0.003***	0.002***	0.003***	0.002***	0.001***
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Occ high/low	-0.000	0.000	-0.001	0.000	0.000	0.000
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
Req-sd	-0.007	0.002	-0.009**	-0.006	0.003	-0.008*
	(0.005)	(0.002)	(0.004)	(0.005)	(0.002)	(0.005)
Constant	0.007***	0.015***	-0.008***			
	(0.001)	(0.000)	(0.001)			
Implied						
β (*Labor*)	0.777***	0.831***	-0.044***	0.805***	0.853***	-0.048***
	(0.010)	(0.007)	(0.008)	(0.012)	(0.008)	(0.011)
μ (*Req-ed*)	-0.025**	0.023***	-0.048***	-0.040***	0.021***	-0.061***
	(0.010)	(0.005)	(0.009)	(0.005)	(0.006)	(0.004)
ω (*Over-ed*)	-0.001	0.023***	-0.027***	-0.006	0.024***	-0.030***
	(0.011)	(0.006)	(0.010)	(0.005)	(0.006)	(0.005)
υ (*Under-ed*)	0.017	-0.019***	0.034***	0.029**	-0.019**	0.049***
	(0.014)	(0.007)	(0.013)	(0.006)	(0.007)	(0.006)

Standard errors are in parentheses, with significance levels *** ($p < 0.01$), ** ($p < 0.05$), and * ($p < 0.1$). Standard errors are robust for FDSUR estimates, whereas all FDPV estimates are computed from 250 block-bootstrap samples. Estimates from the profit equation are computed from the difference in estimates between productivity and labor cost. The endogenous variables in FDPV are *Labor*, *Capital*, *Req-ed*, *Over-ed*, *Under-ed*, and *Age*, while the rest of the variables are treated as "free" and thus exogenous. The model is estimated for 39,713 observations with a total of 7,317 spanning the period 2001-2010.

labor cost and productivity equation depict a somewhat different situation than the one observed in Figure 3.

We begin with the coefficient for required education μ. The results for firm productivity are consistently negative and significant for both estimators. Therefore, the positive trend observed in Figure 3 (a) does not hold when controlling for confounding factors. The estimate in column (4) suggests that average required education causes productivity to decline by 4 percent for one additional year in required education, whereas labor cost increases by approximately two percent per additional year of required education (col. 5). We also note that for labor cost, the FDSUR estimate is close to the FDPV estimate. In column 6, we note that one additional year of required education is associated with a decrease in profit of approximately six percent. This result is strikingly different from that obtained by Kampelmann & Rycx (2012) in two regards. First, they present evidence that one additional year of required education contributes positively to value added per worker by approximately two percent. Second, in their estimates of a profit equation, no significant effect could be discerned for required education or for any other ORU variables.

Turning to the results for over-education ω, the impact of having over-educated workers on productivity is not different from zero in either col. (1) or (4). Hence, these results reject hypothesis (i) that over-education leads to productivity gains. This result is also surprising since it contrasts with one of the key findings in Kampelmann & Rycx (2012). Turning to labor cost, the results in columns (2) and (5) suggest that one additional year of over-education increases labor cost by 2.4 percent. This result is in line with hypothesis (ii) that over-education positively impacts labor cost. Combining the negative and insignificant productivity effect with an upward pressure of wages, the results in columns (3) and (6) suggest that over-education is associated with decreased profits. The non-neutrality of the effect of over-education on profit rejects the first of the null equality hypothesis (v) that employing over-educated workers has a symmetric effect (hence, no effect) on profit.

The final piece for understanding how ORU affects the demand side is the impact of under-education υ. Starting with the effect on productivity, educational mismatch in the form of under-education is not significantly different from zero for the FDSUR estimator, but when we account for endogeneity, the FDPV estimator registers a positive and significant productivity effect of having under-educated workers. Since the latter estimator is preferred, the estimate can be interpreted as indicating that the average productivity increases by 2.9 percent for every additional year of under-education among workers. This result goes against hypothesis (ii), which posits a decrease in productivity following under-education. In addition to the positive productivity effect, under-education reduces labor cost by 1.9 percent per additional year of under-education. Thus, firms with under-educated workers experience a positive productivity effect and reduced labor cost, which in combination leads to increased profits seen in both cols. (3) and (6). For one additional year of under-education, profit increases by 4.9 percent (FDPV), which rejects the equality hypothesis (vi) that profit is unaffected by employing either over- or under-educated workers.

The results above for over- and under-education constitute our key findings. While firm performance in terms of profit is negatively affected by having over-educated employees, the reverse seems to be the case for employing under-educated workers. Before turning to the robustness analysis, we can make a number of observations regarding the estimates of the control variables (age, proportion of immigrants and females) used in the analysis. Here, we see that only the age variable affects productivity and labor costs in a way that affects the profit. Slightly puzzling is that the effect is negative across all columns. To the extent that average age controls for other work-related skills, the expected sign of the effect is positive. We can, however, speculate that the negative effect comes from non-linearities in the effect across the age distribution. If the effect from higher average age reflects that of an inverted U-shape, it is plausible that a negative effect from having an aging workforce dominates the positive effect from having younger employees. The effects from immigrants and gender is less conclusive. While the proportion of immigrants is estimated to positively contribute to both productivity and labor cost, it is only significant in the latter. For the proportion of females among the employees, the estimated effect is the reverse, with a negative contribution to productivity and labor cost, but only in the former is it significantly different from zero. For both immigrants and gender, however, neither effect causes a significant change in firms' profits. Turning to the regional variables, we find no evidence of significant externalities related to human capital or the size of local labor markets. However, having a heterogeneous local labor market, reflected by the number of occupations, contributes positively to productivity, wages and profit. A possible explanation for this result is that a diversified job market increases the probability of (heterogeneous) workers finding jobs that match their unique competence. Finally, we may observe a weakly significant negative estimate of the standard deviation of required education on profit.

6.1 Robustness analysis

In this section, we investigate a number of scenarios that could potentially affect our main result. Specifically, we analyze if demography of workers, firm size, asymmetric firm exit and entry, and the removal of marginally mismatched workers alter the results. The positive productivity effect from having under-educated workers is in contrast to Kampelmann & Rycx (2012), who found negative productivity effects from having under-educated workers, especially under-educated young workers. To address this problem, we split the labor force into young and old workers (cutoff point: 35 years old). Hence, instead of three ORU variables as in the previous analysis, we now run the regressions using the two sets of ORU variables; one for old and one for young simultaneously. The results are presented in Panel A in Table

3. Here, we see that the positive impact of under-educated workers on profit mainly stems from young under-educated workers. In addition, we see that the negative effect from over-education can completely be attributed to the over-education of older workers. In addition to controlling the impact of young and old workers, we also divide workers into age categories with 10 years per category because there is an upward trend in educational attainment and an over-representation of older workers in the under-educated category. Thus, if older workers are systematically categorized as 'more' under-educated and if they compensate for their relatively low education by acquiring more experience and on-the-job-skills, this effect could very well be dominating any negative effect. We therefore construct a new metric of 'required education' that is conditional on the age distribution by coarsening workers' age into the following categories: 25 years or younger, 26 to 35 years, 36 to 45 years, 46 to 56 years, and 56 years or older. We then compute the mode educational level and mismatch for each occupation-year-age category. The results from running the FDPV estimator suggest that the negative (positive) effect found for over-(under-)education are preserved for profit. Specifically, firms still experience net gains from having under-employed workers, but they do not seem to affect total labor costs once age is accounted for in the ORU variables. These results are available upon request.

Since the required educational level is reduced to a scalar value associated with each occupation and year, it is unlikely that this value perfectly captures the true required educational level. To increase the likelihood that we measure true under- and over-education, we proceed and condition individuals to have an educational level above or below one standard deviation of educational attainment in the respective occupation and year. By invoking this restriction, we increase the likelihood that R at least belongs to the closed interval $[R - \sigma(E), R + \sigma(E)]$, with the result that only severely over-educated or under-educated workers are classified as mismatched. The results presented in Panel B are qualitatively similar to those in the main model. If entrant and/or exiting firms differ systematically from other firms with regard to the average educational mismatch, this could yield different results when estimating the models using a full vs. balanced panel. Hence, to examine whether the results are robust with respect to entry and exit, we proceed by estimating the model on a balanced panel. The results are presented in Panel C. The greatest difference from the main results in Table 2 is that over-education no longer has a negative and significant effect on profit using FDSUR. However, after controlling for the endogeneity of structural parameters, the negative result found earlier for profit still persists, but with reduced significance. Otherwise the parameter estimates for the profit equation is qualitatively similar to those in the main model, suggesting that the potential bias from entry and exit does not qualitatively alter the main results.

Finally, we restrict the sample to firms with at least 50 employees, reducing the sample to 10,809 observations and 2,175 firms. Larger firm size means that our ORU measures

Table 3: Results for robustness tests

	Prod.	Lab. cost	Profit	Prod.	Lab. cost	Profit
Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A					
	Young			Old		
μ (*Req-ed*)	-0.045***	0.015*	-0.059***	-0.01	0.030***	-0.040***
	(0.004)	(0.007)	(0.004)	(0.004)	(0.006)	(0.004)
ω (*Over-ed*)	0.025*	0.032***	-0.008	-0.002	0.026***	-0.028**
	(0.006)	(0.008)	(0.006)	(0.005)	(0.008)	(0.005)
υ (*Under-ed*)	0.027	-0.009	0.035*	-0.005	-0.019**	0.013
	(0.010)	(0.014)	(0.009)	(0.005)	(0.008)	(0.005)
	Panel B			Panel C		
	ORU (plus minus 1 sd)			Balanced		
μ (Req-ed)	-0.036***	0.022***	-0.058***	-0.031**	0.028***	-0.059***
	(0.005)	(0.006)	(0.004)	(0.006)	(0.007)	(0.006)
ω (*Over-ed*)	-0.006	0.022***	-0.028***	0.007	0.033***	-0.026**
	(0.005)	(0.005)	(0.004)	(0.006)	(0.007)	(0.006)
υ (*Under-ed*)	0.026*	-0.021***	0.047***	0.020	-0.022*	0.042***
	(0.006)	(0.007)	(0.005)	(0.008)	(0.011)	(0.008)
	Panel D					
	Large firms (+50 empl.)					
μ (*Req-ed*)	-0.065*	0.048***	-0.113***			
	(0.015)	(0.015)	(0.015)			
ω (*Over-ed*)	-0.029	0.065***	-0.094**			
	(0.016)	(0.017)	(0.016)			
υ (*Under-ed*)	0.044	-0.043*	0.087*			
	(0.024)	(0.022)	(0.024)			

Standard errors are in parentheses, with significance levels *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$). Standard errors are robust for FDSUR estimates, whereas FDPV estimates are computed from 250 block-bootstrap samples. Profit estimates are computed from the difference in estimates between productivity and labor cost. The endogenous variables in FDPV are *Labor*, *Capital*, *Req-ed*, *Over-ed.*, *Under-ed*, and *Age*, while the rest of the variables are treated as "free", thus exogenous.

become less sensitive to changes in ORU for one or two (a few) workers. From the results in Panel D, we note that when excluding small firms, the effects on productivity from under-education are no longer significant. Second, the effects on labor cost become larger than those found in the main sample. Third, the profit for large firms is still negatively affected by having employees with required and over-education and positively affected by having under-educated workers. The positive effect, however, is only weakly significant. For over-(under-)education, the negative (positive) estimate translates into a decrease (increase) in the profit by 9.4 percent (8.7 percent). In short, while the sample of larger firms did not reproduce all the individual estimates of the ORU variables, the overall effects on profit are similar, further suggesting that the results found using the main sample also apply to larger firms.

7 Conclusions

The increased levels of over-education workers (in many Western countries) are among the most frequently discussed labor market issues among both policy makers and researchers. Without excess demand to absorb the greater proportion of highly educated students, educational mismatch risks becoming a more persistent phenomenon that is costly both to society and to the individuals themselves. While the effect on workers have been thoroughly analyzed, less is known about how mismatch impacts firm performance and competitiveness. The purpose of this paper has been to deepen our understanding of how employing mismatched workers, in the form of over-, required and over-education (ORU), impacts productivity, labor cost and profits. To accomplish this goal, we present a simple model where we embed the ORU model in the human capital production function at the firm level. Drawing on the methodology of a labor quality index that studies the impact from heterogeneous labor inputs (Vandenberghe 2013), we provide a system of three equations that is suitable for empirical testing that links mismatch with productivity, wage cost and profits.

According to human capital theory (Becker 2009), workers are paid according to their productivity. Taking the efficiency wage hypothesis as the null hypothesis, our model presents a simple way of testing this theory in a setting of heterogeneous labor inputs in the form of workers ORU. Using a matched employer-employee dataset of Swedish manufacturing firms observed in the period 2001-2010, we take a structural approach using a first-differenced version (Vandenberghe 2013) of the proxy variable method of Ackerberg et al. (2015) to estimate the model.

Our key result is that ORU among employees seems to have asymmetric effects on firms productivity, labor cost per employee and profitability. Contrary to as suggested by human capital theory, we find no evidence that having over-educated employees leads to productivity gains; rather, we observe an increase in costs. This is not the case when firms employ under-educated workers, where we find a positive impact on productivity and, at the same time, a reduction in labor costs. Hence, the net effect as measured by the gross operating surplus suggests that having over-(under-)educated workers has a negative (positive) effect of approximately -3 (+4.9) percent for each additional year of over-(under-)education among employed workers. Disaggregating the analysis based on worker age, we find that the negative effects of over-education can be ascribed to older employees, whereas the positive effect can be traced to young under-educated workers.

This study, however, is not without limitations. First, we do not explore possible dynamic effects of educational mismatch. It is conceivable that employing mismatch workers, particularly younger over-educated workers, leads to increased productivity gains over time, whereas labor costs may increase even faster. One explanation behind the missing positive effect on productivity from employing over-educated workers could relate to the contribution of young or inexperienced but well-educated workers. Although they immediately impose a higher cost on firms, the expected productivity gains may not materialize until later years. Finally, based on the trend observed for ORU workers in the manufacturing industry, we also caution against generalizing the interpretation to other sectors of the economy. One concern when examining other sectors using our structural approach is the use of materials as a proxy variable to handle the endogeneity of the ORU variables, as these are not reported for many service-oriented firms. Nevertheless, using alternative instruments to address the endogeneity problem and comparing the results from manufacturing to those for service firms seem to be a natural direction in which to proceed.

References

Ackerberg, D. A., Caves, K. & Frazer, G. (2015), 'Identification properties of recent production function estimators', *Econometrica* **83**(6), 2411–2451.

Ackerberg, D., Caves, K. & Frazer, G. (2006), Structural identification of production functions. working paper, https://ideas.repec.org/p/pra/mprapa/38349.html.

Akerlof, G. A. & Yellen, J. L. (1990), 'The fair wage-effort hypothesis and unemployment', *The Quarterly Journal of Economics* pp. 255–283.

Allen, J. & Van der Velden, R. (2001), 'Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search', *Oxford economic papers* pp. 434–452.

Aubert, P. & Crépon, B. (2003), 'Age, wage and productivity: Firm-level evidence', *Economie et Statistique* (363), 95–119.

Battu, H., Belfield, C. R. & Sloane, P. J. (1999), 'Overeducation among graduates: a cohort view', *Education economics* **7**(1), 21–38.

Bauer, T. K. (2002), 'Educational mismatch and wages: a panel analysis', *Economics of Education Review* **21**(3), 221–229.

Becker, G. S. (2009), *Human capital: A theoretical and empirical analysis, with special reference to education*, University of Chicago Press.

Belfield, C. R. & Harris, R. D. (2002), 'How well do theories of job matching explain variations in job satisfaction across education levels? evidence for uk graduates', *Applied Economics* **34**(5), 535–548.

Bender, K. A. & Heywood, J. S. (2006), 'Job satisfaction of the highly educated: The role of gender, academic tenure, and earnings', *Scottish Journal of Political Economy* **53**(2), 253–279.

Dolton, P. & Vignoles, A. (2000), 'The incidence and effects of overeducation in the uk graduate labour market', *Economics of education review* **19**(2), 179–198.

Duncan, G. J. & Hoffman, S. D. (1982), 'The incidence and wage effects of overeducation', *Economics of Education Review* **1**(1), 75–86.

Frenette, M. (2004), 'The overqualified canadian graduate: the role of the academic program in the incidence, persistence, and economic returns to overqualification', *Economics of Education Review* **23**(1), 29–45.

Gandhi, A., Navarro, S. & Rivers, D. (2013), On the identification of production functions: How heterogeneous is productivity? Working paper, University of Wisconsin-Madison and University of Western Ontario.

Garnero, A. & Rycx, F. (2012), The heterogeneous effects of workplace diversity on wages and productivity. Working paper, http://www.aiel.it/Old/bacheca/Capua/papers/.

Gennaioli, N., La Porta, R., Lopez-de Silanes, F. & Shleifer, A. (2013), 'Human capital and regional development*', *The Quarterly journal of economics* **128**(1), 105–164.

Göbel, C. & Zwick, T. (2009), Age and productivity-evidence from linked employer employee data. ZEW-Centre for European Economic Research Discussion Paper, number 09-020.

Groot, W. (1996), 'The incidence of, and returns to overeducation in the uk', *Applied Economics* **28**(10), 1345–1350.

Groot, W. & Van den Brink, H. M. (2000), 'Overeducation in the labor market: a meta-analysis', *Economics of education review* **19**(2), 149–158.

Hartog, J. (2000), 'Over-education and earnings: where are we, where should we go?', *Economics of education review* **19**(2), 131–147.

Hellerstein, J. K. & Neumark, D. (1995), 'Are earnings profiles steeper than productivity profiles? evidence from israeli firm-level data', *Journal of Human Resources* pp. 89–112.

Hellerstein, J. K. & Neumark, D. (1999), 'Sex, wages, and productivity: An empirical analysis of israeli firm-level data', *International Economic Review* **40**(1), 95–123.

Iaffaldano, M. T. & Muchinsky, P. M. (1985), 'Job satisfaction and job performance: A meta-analysis.', *Psychological bulletin* **97**(2), 251.

Judge, T. A., Thoresen, C. J., Bono, J. E. & Patton, G. K. (2001), 'The job satisfaction–job performance relationship: A qualitative and quantitative review.', *Psychological bulletin* **127**(3), 376.

Kampelmann, S. & Rycx, F. (2012), 'The impact of educational mismatch on firm productivity: Evidence from linked panel data', *Economics of Education Review* **31**(6), 918–931.

Kiker, B. F., Santos, M. C. & De Oliveira, M. M. (1997), 'Overeducation and undereducation: Evidence for portugal', *Economics of Education Review* **16**(2), 111–125.

Korpi, M. & Clark, A. W. (2015), 'Internal migration and human capital theory: To what extent is it selective?', *Economic Letters* **136**(November 2015), 31–34.

Lamo, A. & Messina, J. (2010), 'Formal education, mismatch and wages after transition: Assessing the impact of unobserved heterogeneity using matching estimators', *Economics of Education Review* **29**(6), 1086–1099.

Lazear, E. P. & Shaw, K. L. (2007), 'Personnel economics: The economist's view of human resources', *The Journal of Economic Perspectives* **21**(4), 91.

Levine, D. I. (1991), 'Cohesiveness, productivity, and wage dispersion', *Journal of Economic Behavior & Organization* **15**(2), 237–255.

Levinsohn, J. & Petrin, A. (2003), 'Estimating production functions using inputs to control for unobservables', *The Review of Economic Studies* **70**(2), 317–341.

Mavromaras, K. & McGuinness, S. (2012), 'Overskilling dynamics and education pathways', *Economics of Education Review* **31**(5), 619–628.

McGuinness, S. (2003), 'Graduate overeducation as a sheepskin effect: evidence from northern ireland', *Applied Economics* **35**(5), 597–608.

McGuinness, S. (2006), 'Overeducation in the labour market', *Journal of economic surveys* **20**(3), 387–418.

McGuinness, S. & Sloane, P. J. (2011), 'Labour market mismatch among uk graduates: An analysis using reflex data', *Economics of Education Review* **30**(1), 130–145.

Moshavi, D. & Terborg, J. R. (2002), 'The job satisfaction and performance of contingent and regular customer service representatives: A human capital perspective', *International Journal of Service Industry Management* **13**(4), 333–347.

Olley, G. S. & Pakes, A. (1996), 'The dynamics of productivity in the telecommunications equipment', *Econometrica* **64**(6), 1263–1297.

Rumberger, R. W. (1987), 'High school dropouts: A review of issues and evidence', *Review of educational research* **57**(2), 101–121.

Sicherman, N. (1991), '"overeducation" in the labor market', *Journal of labor Economics* pp. 101–122.

Tåg, J. (2013), 'Production hierarchies in sweden', *Economics Letters* **121**(2), 210–213.

Tåhlin, M. (2007), *Överutbildning I Sverige: utveckling och konsekvenser*, SNS förlag, Stockholm.

Van der Meer, P. H. (2006), 'The validity of two education requirement measures', *Economics of Education Review* **25**(2), 211–219.

Van Ours, J. C. & Stoeldraijer, L. (2011), 'Age, wage and productivity in dutch manufacturing', *De Economist* **159**(2), 113–137.

Vandenberghe, V. (2013), 'Are firms willing to employ a greying and feminizing workforce?', *Labour Economics* **22**, 30–46.

Verhaest, D. & Omey, E. (2006), 'The impact of overeducation and its measurement', *Social Indicators Research* **77**(3), 419–448.

Wolbers, M. H. (2003), 'Job mismatches and their labour-market effects among school-leavers in europe', *European Sociological Review* **19**(3), 249–266.

Appendix

A Firm-level human capital with ORU decomposition of workers' education

To integrate the ORU definition with firm-level human capital, we draw extensively on the literature on heterogeneous labor inputs (Hellerstein & Neumark 1999, Aubert & Crépon 2003, Vandenberghe 2013). We simplify the notation by dropping firm-year indexation *it*. The total number of workers in a given firm at a point in time is thus denoted by L. Instead of using a single measure of labor inputs, Hellerstein & Neumark (1999) and later Aubert & Crépon (2003) introduce a more flexible input that can accommodate workers of different kinds. This measure is referred to as *labor quality* (henceforth LQ). Consider $k = 1, ..., n$ mutually exclusive subsets of workers in a firm, with a total of L_k workers for group k. Workers in a group have the same marginal productivity given by θ_k; then, the LQ of the firm is defined by summing over the n productivity weighted labor units $\theta_k L_k$, and thus,

$$LQ = \sum_{k=1}^{n} \theta_k L_k. \tag{33}$$

To estimate θ_k in the setting of a log-linear production function (e.g. a Cobb-Douglas function), LQ is first rewritten in terms of relative productivities for some reference group k^*. This can be accomplished by factoring out the productivity parameter of the k^* group θ^* together with the total number of workers L from 33, resulting in

$$LQ = \theta_{k^*} L \left(\frac{L_{k^*}}{L} + \sum_{k-\{k^*\}}^{n} \frac{\theta_k}{\theta_{k^*}} \frac{L_k}{L} \right).$$

In this expression, the number of workers in a group is expressed as a share of the total L. For group k, this is given by L_k/L, with a productivity relative to the reference group given by θ_k/θ_{k^*}. Next, using the fact that $L_{k^*}/L = 1 - \sum_{k-\{k^*\}}^{n} L_k/L$ and substituting for L_{k^*}/L, we obtain the final expression. After collecting terms, LQ can be expressed as follows:

$$LQ = \theta_{k^*} L \left(1 + \sum_{k-\{k^*\}}^{n} \left(\frac{\theta_k}{\theta_{k^*}} - 1 \right) \frac{L_k}{L} \right). \tag{34}$$

Taking the logarithms and using the Taylor expansion of $\log(1 + x) \approx x$ for small x, the relative productivity of group k, $\theta_k/\theta_{k^*} - 1$, can be directly estimated from L_k/L added linearly to the log production function.[16]

Since we aim to study the productivity effects from employees with ORU education, we need a suitable expression of firm-level human capital H_{it} expressed as a function of the degree of ORU education among workers. To accomplish this goal, we adopt the above approach and form three distinct subsets of workers, indexed by o, r, and u, respectively. Additionally, we add more structure to (33) by modeling the productivity of worker type k, θ_k, as a function of the ORU variables. Let L^o, L^r and L^u be the number of workers with over-, required and under-education. By construction, each worker can be part of exactly one group at each point in time, which means that $L = L^o + L^r + L^u$. In addition, let $\bar{h}$ denote the average human capital of the L workers, and let $\bar{h}^o$, $\bar{h}^r$, and $\bar{h}^u$ be the average human capital of workers in the corresponding group. Thus, we can write firm-level human capital (on LQ form) as a sum of its constituent productive group components

$$H = \bar{h}L = \bar{h}^o L^o + \bar{h}^r L^r + \bar{h}^u L^u. \tag{35}$$

This setting is analogous to the one in (33) with $\theta_k = \bar{h}_k$. To construct the group-specific human capital functions, we proceed by defining three separate functions of human capital for the individual worker j depending on which of the groups r, u or o he or she belongs to. These are in turn given by

$$h_j^r = e^{\mu R} \qquad \forall j \in \{R\}, \tag{36}$$
$$h_j^u = e^{\mu R + \omega_j O_j} \qquad \forall j \in \{O\}, \tag{37}$$
$$h_j^o = e^{\mu R + \upsilon_j U_j} \qquad \forall j \in \{U\}. \tag{38}$$

Starting with workers in the r group, their educational level is the same level as the required R. These individuals are defined as *matched* with returns to education given by μ. For workers who are over-educated, each has the same benchmarked component of μR but also a worker-specific return ω_j to the years of over-education $O_j = E_j - R > 0$. The human capital function is defined analogously for under-educated workers, with the same fixed benchmark capital μR together with a worker-specific return υ_j to the years of under-education $U_j = E_j - R < 0$. In this setup, we consider only one occupation with a single level of required education with a given rate of return. In practice, we can allow for many occupations within the firm with different required education together with an individual specific return to required education μ_j, but for simplicity, we abstract from these features here.

To aggregate the human capital function, we take the expected value of the first-order multivariate Taylor expansion of (36-38) evaluated at the group means of each parameter,

[16] See, e.g., Vandenberghe (2013) for a recent application of the LQ method to study the relative productivity of females in different age groups.

resulting in

$$\bar{h}^r = \mathbb{E}_r\left(h_j^r\right) \cong e^{\bar{\mu}\bar{R}}, \tag{39}$$

$$\bar{h}^o = \mathbb{E}_o\left(h_j^o\right) \cong e^{\bar{\mu}\bar{R}+\bar{\omega}\bar{O}}, \tag{40}$$

$$\bar{h}^u = \mathbb{E}_u\left(h_j^u\right) \cong e^{\bar{\mu}\bar{R}+\bar{\upsilon}\bar{U}}. \tag{41}$$

Inserting (39-41) back into the expression for H in (35) gives our first result:

$$H = L^r e^{\bar{\mu}\bar{R}} + L^o e^{\bar{\mu}\bar{R}+\bar{\omega}\bar{O}} + L^u e^{\bar{\mu}\bar{R}+\bar{\upsilon}\bar{U}}. \tag{42}$$

With firm-level aggregates of all ORU parameters embedded in the H_{it} function it remains to be shown how the expression in (42) can be fit into the production function described in (5).[17] To accomplish this goal, we parallel the LQ method and proceed by defining variables for the relative size of each of the ORU groups denoted by $\lambda^o := L^o/L$, $\lambda^r := L^r/L$, and $\lambda^u := L_u/L$. An important aspect of the above functions is that the human capital of mismatched workers is expressed relative to $e^{\mu R}$, which allows us to express the aggregate $\bar{h}^o$ and $\bar{h}^u$ in terms of $\bar{h}^r$,

$$\bar{h}^o = \bar{h}^r e^{\bar{\omega}\bar{O}},$$
$$\bar{h}^u = \bar{h}^r e^{\bar{\upsilon}\bar{U}}.$$

This allows us to factor $Le^{\bar{\mu}\bar{R}}$ from (42) and write the expression in terms of λ^o, λ^r, and λ^u,

$$H = Le^{\bar{\mu}\bar{R}}\left[\lambda^o e^{\bar{\omega}\bar{O}} + \lambda^r + \lambda^u e^{\bar{\upsilon}\bar{U}}\right]. \tag{43}$$

Using the fact that $\lambda^r = 1 - \lambda^o - \lambda^u$, inserting it into (43) and collecting terms, we write H in the form

$$H = Le^{\bar{\mu}\bar{R}}\left[1 + \lambda^o\left(e^{\bar{\omega}\bar{O}} - 1\right) + \lambda^u\left(e^{\bar{\upsilon}\bar{U}} - 1\right)\right]. \tag{44}$$

Observe that for small $\bar{\omega}\bar{O}$ and $\bar{\upsilon}\bar{U}$, the exponential functions can be approximated using a linear expansion around 0 by $e^x \cong 1 + x$, which lets us write H as

$$H \cong Le^{\bar{\mu}\bar{R}}\left[1 + \bar{\omega}\bar{O}\lambda^o + \bar{\upsilon}\bar{U}\lambda^u\right]. \tag{45}$$

Since (45) is log-linear for small values of $\bar{\omega}\bar{O}\lambda^o + \bar{\upsilon}\bar{U}\lambda^u$, we use the inverse of the exponential approximation, which states that $\log(1+x) \cong x$. Finally, since the averages $\bar{O}$ and $\bar{U}$ are taken with respect to the number of workers in the respective groups, multiplying by the groups, the relative shares λ^o and λ^u results in $\bar{O}\lambda^o = (L^o/L)(1/L^o)\sum_{j=1}^{L^o} O_j =$

[17] While μR are constants, making the average notation redundant in (42), we leave it here for clarity to show how we use averaging to go from the individual level to the group level.

$1/L \sum_{j=1}^{L^o} O_j := \overline{O}_j$, which is simply the average years of over-education with the average calculated with respect to all L_{it} workers in the firm. The same applies to the $\bar{U}\lambda_u$ term. As a result, we can define our ORU version of firm-level human capital as

$$H \cong Le^{\bar{\mu}\bar{R}} \left[1 + \bar{\omega}\overline{O} + \bar{\upsilon}\overline{U}\right]. \tag{46}$$

This expression satisfies both of the desirable properties stated at the beginning of this section: (i) it allows for a direct identification of the productivity effects $\bar{\mu}$, $\bar{\omega}$, and $\bar{\upsilon}$ from having certain levels of required, over-, and under-education among employees at the firm level, and (ii) the expression is fully compatible with the log-linear form of the Cobb-Douglas production function and can thus be estimated once the usual concerns regarding the potential endogeneity of input factors in the production function are considered.