The effects of educational mismatch on inventor productivity. Evidence from Sweden, 2003-2010

Ioana A. Igna (ioana.igna@outlook.it)

Department of Economics, University of Perugia, Italy

Papers in Innovation Studies

Paper no. 2018/08

Centre for Innovation, Research and Competence in the Learning Economy (CIRCLE)

Lund University

P.O. Box 117, Sölvegatan 16, S-221 00 Lund, SWEDEN

http://www.circle.lu.se/publications

WP 2018/08

The effects of educational mismatch on inventor productivity. Evidence from Sweden, 2003-2010

Ioana A. Igna

Abstract: This paper investigates the influence on inventor productivity of the imperfect occupational match, measured as the number of years of education in excess and in deficit to the required level (educational mismatch). The empirical model draws on a unique database that matches information about individual inventor characteristics, such as age, experience and gender, with patenting performance in Sweden over the period 2003-2010. The results suggest that over-educated (OE) inventors file a number of patents higher than inventors who are appropriately matched (RE), but perform poorly than well-matched inventors who hold a similar level of education. Conversely, under-educated (UE) inventors file a total number of patents lower than inventors who are well-matched (RE), but more than well-matched ones who hold the same level of education. These results conform to the hierarchical pattern of ORU model, well-documented in the literature relating the employees' wages to educational mismatch (i.e. RE>OE>UE). We find that significant differences in returns to education across match and mismatch categories remain even after controlling for individual ability. Our findings are robust to controlling for differences between younger and older inventors, geographical areas and industry of work.

Keywords: Inventor; Productivity; Educational mismatch; Patent

JEL: I26; J24; L6; O3

The effects of educational mismatch on inventor productivity. Evidence from Sweden, 2003-2010

Ioana A. Igna*

Abstract

This paper investigates the influence on inventor productivity of the imperfect occupational match, measured as the number of years of education in excess and in deficit to the required level (educational mismatch). The empirical model draws on a unique database that matches information about individual inventor characteristics, such as age, experience and gender, with patenting performance in Sweden over the period 2003-2010. The results suggest that over-educated (OE) inventors file a number of patents higher than inventors who are appropriately matched (RE), but perform poorly than well-matched inventors who hold a similar level of education. Conversely, under-educated (UE) inventors file a total number of patents lower than inventors who are well-matched (RE), but more than well-matched ones who hold the same level of education. These results conform to the hierarchical pattern of ORU model, well-documented in the literature relating the employees' wages to educational mismatch (i.e. RE>OE>UE). We find that significant differences in returns to education across match and mismatch categories remain even after controlling for individual ability. Our findings are robust to controlling for differences between younger and older inventors, geographical areas and industry of work.

JEL codes: I26, J24, L6, O3

Keywords: Inventor; Productivity; Educational mismatch; Patent

*ioana.igna@outlook.it. Department of Economics, University of Perugia (IT). This paper has been prepared during the author's visiting research experience at Lund University, CIRCLE - Center for Innovation, Research and Competences in the Learning Economy. The author is grateful to Olof Ejermo and Francesco Venturini for their advise and to Michela Vecchi, Alessandro Sterlacchini and participants at SIEPI, XVI Annual Workshop, for very useful comments.

1 Introduction

It is well-established that investment in human capital produces benefits for both individuals and the society taken as a whole. Becker (1962) argues that investment in schooling and on-the-job training increases individuals' knowledge and abilities, which in turn affect positively their job productivity and income. Individuals with higher endowment of human capital have a greater adaptability, especially in turbulent and changing environments, favoring the adoption of new technologies at the aggregate level (Nelson and Phelps, 1966). Human capital directly affects the rate of economic growth (Lucas, 1988) and indirectly being a key input to innovation activities (Romer, 1990; Aghion *et al.*, 1998). In the presence of labour market imperfections, due for instance to discrimination practices (Hsieh *et al.*, 2018) or strict labour market regulation (Griffith and Macartney, 2014), the allocation of human capital between jobs, tasks or firms may be distorted, negatively reverberating on innovation performance and ultimately on the rate of economic growth.

A wide strand of microeconomic literature has studied labour mismatch, investigating the effects on workers' wages, job satisfaction and firm-level productivity (Hersch, 1991; Kampelmann and Rycx, 2012; Levels *et al.*, 2014). Educational mismatch is typically defined as the divergence between worker's level of education and the level demanded by her job or occupation (Desjardins and Rubenson, 2011). However, the literature is still silent about how labour mismatch turns into lower productivity outcomes. In particular, evidence is scarse about the effects of labour mismatch (or misallocation) on the results of innovation activity, which is the key factor for long-lasting productivity growth in knowledge-based economies. This study fills this important gap of the literature by investigating the effects of educational mismatch on inventor productivity.

The labour market mismatch phenomenon is quite relevant in Europe. Recent evidence (OECD, 2016, p. 19) shows that in the countries that took part in the 2012 Survey of Adult Skills 60% of workers are mismatched, either over/under-educated or over/under-skilled. Korpi and Tåhlin (2009) observe that the average level of education has increased significantly in Sweden since 1970s but it has not been followed by an equally increase in the skill requirements by the employers. Consequently, after three decades the incidence rate of over-education is two times higher than that in the early 1970s. This type of inefficiency in the labour market may have detrimental effects for the economic system as a whole, which is likely to influence the rate of innovation and economic growth.

To study the effects of educational mismatch on inventor productivity, we conduct a cross-sectional analysis on a sample of inventors residing in Sweden, covering the period 2003-2010. We estimate the so-called ORU model which measures the effect of Over-education, Required education and Under-education on inventor productivity, defined by both the number of patent applications and the number of forward citations.

Our main finding is that inventors holding a number of years of education in excess (over-education) or in deficit (under-education) to the modal educational level (required education) in a given occupation are associated with a lower inventive productivity compared to inventors having a job position requiring the exact level of competencies (well-matched). We check the effects of ORU model by including a set of control variables related to both demographic and employment characteristics. Moreover, we are able to control for individual idiosyncratic characteristics. In particular, we investigate whether the inventive performance of matched and mismatched inventors is affected by school proficiency, as measured by the high-school grade. This variable controls for individual ability and is included to check to some extent the simultaneity feedbacks between patent productivity and educational mismatches (ORU). The inclusion of these controls does not affect our main results, which remain highly significant. Furthermore, we assess the robustness of our results by investigating the pattern of ORU model along some crucial inventor characteristics. Our results show that formal education matters more for younger inventors as younger well-matched inventors are more productive than their older counterparts. It implies that the opportunity cost of younger over-educated is higher than that of older over-educated, while the opposite is true for under-educated inventors. We also find that the pattern and significance of ORU variables are robust distinguishing inventors on the basis of geographical areas and the R&D intensity of industrial sectors, where over-educated inventors have a higher opportunity cost in low-tech industries, while under-educated are mostly penalized in high-tech industries.

Our study makes some important contributions to the literature. This work is related to the microeconomic literature on the determinants of inventor productivity (Zwick *et al.*, 2017; Frosch, 2011), and in particular to the studies that analyze the link between inventor productivity and job mobility (Nakajima *et al.*, 2010; Hoisl and de Rassenfosse, 2017). This study is also related to the literature on the effects of educational mismatch on workers' earnings and firm productivity. A large strand of literature investigates the impact of educational mismatch on workers' wages and demonstrate the existence of substantial differences in wage returns across matched and mismatched categories (Korpi and Tåhlin, 2009; Levels *et al.*, 2014). Kampelmann and Rycx (2012) document that a higher level of required education positively influence the firm productivity, whereas additional years of over-education (under-education) are beneficial (detrimental) for firm productivity. Finally, our results contribute to the macroeconomic literature on resource misallocation by measuring the microeconomic effects of labour mismatch on inventive productivity. Jones (2011) suggests that the allocation of resources across industries and firms plays an important role in explaining the total factor productivity differentials across countries. Jovanovic (2014) argues that a better assignment of workers to the right task is likely to reduce job turnover and accelerate the rate of economic growth.

The paper is structured as follows. Section 2 surveys the literature. Section 3 discusses the econometric approach and the factors that influence the inventive performance. Section 4 presents the data and the measurement of educational mismatch. Section 5 shows the descriptive statistics and the empirical results. Finally, Section 6 concludes.

2 Related literature

A large body of the literature has paid attention to the determinants of inventor productivity, focusing on both individual and employer characteristics. Existing evidence shows that the level of education plays a key role in explaining the research productivity of scientists and engineering. Giuri *et al.* (2007) show that the majority of inventors have a university degree and only a small share of inventors hold a PhD degree. Jung and Ejermo (2014) focus on Swedish inventors and document that over the period 1985-2007 an increasingly larger proportion of inventors hold a PhD degree. Onishi and Nagaoka (2012) analyze the productivity of a sample of Japanese inventors and find that inventors with PhD degrees more productive in terms of quantity and quality of patents than their counterparts with lower educational levels. Similarly, Mariani and Romanelli (2007) find that European inventors with a high level of education and employed in large firms are more likely to apply for a large number of patent applications over their career.

One of the studies that more convincingly address the issue of the causal link between schooling and innovation is Toivanen and Väänänen (2016). These authors examine the causal effect of engineering graduation on individuals' propensity to patent inventions. Using data on Finnish inventors and the university proximity as an instrument, the authors document that the political decision to establish new engineering universities between 1950 and 1981 increased the probability that individuals from the nearby regions were enrolled in engineering graduation programs and this was associated with a higher patenting productivity.

Zwick *et al.* (2017) is the first paper to investigate the effects of several dimensions of human capital on patent history data over the entire career of the universe of German inventors. The authors find that divergent thinking, problem-solving skills, risk attitude and the ability to benefit from networking have a positive correlation with inventor productivity. Still, high schooling attainment maintains a prominent role in explaining inventive capacity, even after controlling for other dimensions of human capital. Some studies, by contrast, do not find a significant association between the level of education and patent productivity (Schettino *et al.*, 2013). Hoisl (2007) shows that the level of education has no influence on inventor productivity, whereas the composition of the research team together with the use of external sources of knowledge affect in a large extent the inventive output. Giuri and Mariani (2013) provide evidence that inventors with a high level of education and working on projects with highly scientific content are capable to create wider research networks. The best-qualified inventors are more likely to detect and absorb external knowledge, regardless of the geographical location of their sources.

A bunch of studies observe that inventors' returns to education varies among technology classes. Formal education appears to be more significant in science-based sectors such as electrical engineering, chemicals and pharmaceuticals, whereas informal accumulation of experience plays a major role in mechanical engineering (Giuri *et al.*, 2007; Jung and Ejermo, 2014).

Another key factor of creative performance widely debated in the literature concerns inventor's age. Existing evidence suggests that the relationship between age and innovation performance follows an inversely U-shaped form, with most inventions produced by individuals aged between 35 and 50 years (see Frosch, 2011, for a review). Moreover, the inventor's average age varies across sectors: inventors are significantly younger in knowledge intensive sectors than in most experience-based disciplines (Mariani and Romanelli, 2007).

Research productivity has been largely investigated in relation to job mobility. Mobile R&D workers are found to contribute more to firm's patenting activity than immobile inventors (Kaiser *et al.*, 2015), and among movers, academic inventors with more valuable patents are more likely to be hired away by private organizations (Crespi *et al.*, 2007). Job mobility may also affect the firm's performance, but existing evidence goes in different directions. Rahko (2017) finds that outbound mobility of inventors weakens the source firm's patenting performance, in particular when mobile inventors worked in the firm's core technology or have moved to technologically close firms. By contrast, Liu *et al.* (2016) document that inventors in firms characterized by a high inventor turnover appear to contribute significantly to firm's inventive performance than inventors in firms with low job mobility.

At individual level, Hoisl (2007) studies the mobility of a large sample of German inventors and finds the existence of a simultaneous relationship between inventors' mobility and their productivity level. Mobile inventors appear to be more productive than non-mobile ones, but highly-productive inventors turn out to be less likely to move. Miguélez and Moreno (2013) examine the role of different characteristics of local labour market for inventors on regional patenting at European level. The authors find that labour mobility and research networks are positively associated to the regional innovation performance.

Trajtenberg and Shalem (2009) investigate the incentives to mobility, and the resulting performance, of a sample of software inventors in the US. The authors document that inventors who are more likely to move have more cited patents and their innovative content is quite general. By contrast, inventors who are less likely to move have more claims and more original patents. The authors argue that this different propensity to move is related to asymmetric information between the inventor and her employee, since inventors have more information to evaluate their invention than their employer. In this context, when a firm fails to detect the value of a patent, the inventor has the incentive to move to firms with a greater capacity to recognize the true quality of her innovation (and hence to appropriately rewarding her). According to Hoisl (2009), inventor mobility depends on the imperfect matching process between the worker and the job (or the firm), namely inventors who find themselves in poor matches are willing to move in order to improve the quality of their match with the new employer.

Nakajima *et al.* (2010) investigate the factors enabling a better job matching for inventors. The authors consider that referrals through job networks can be used by firms in order to evaluate the unobserved ability of prospective employees and thus to assess the quality of the match prior to the hiring. Focusing on the US patents, the authors document that networked inventors exhibit a higher number of granted patents and have longer job tenure than non-networked inventors: networking seems to facilitate the match between inventor and the new hiring firm and, as a result, mobile inventors are less likely to quit the new employment. Alternatively, the authors interpret their findings as a result of a learning mechanism, namely the inventor improves his competences, skills and know-how through the interaction with her past collaborators at the new firm. Similarly, Ahlin and Ejermo (2017) study the mobility of early-career Swedish inventors, finding a positive effect of job mobility on the quality and quantity of patents. This effect can be explained by a learning mechanism of the inventor in the new job position.

Whereas Nakajima *et al.* (2010) offer an indirect evidence of the matching effects, Hoisl and de Rassenfosse (2017) is the first study to investigate the direct effects of the job matching on the inventor productivity following a move. The authors measure the match quality between the inventor and the hiring firm with the portion of inventor's earlier skills that could be useful to the new firm, so-called "knowledge fit"[1]. Their findings show a negative and significant effect of the knowledge fit on the productivity gains following a move. In other words, when the inventor owns knowledge and skills which are not useful to the new firm, the match quality is low and thus are the productivity gains after a move.

A strand of the literature on inventor productivity has also focused on the relationship between migration and invention. Kerr and Lincoln (2010) and Hunt and Gauthier-Loiselle (2010) find a positive and significant effect of skilled migrants on the production of patents in the US market. By contrast, Zheng and Ejermo (2015) document that migrants in Sweden have a lower probability to become inventors compared to the natives. Moreover, conditional on being an inventor, migrants also show a lower inventive performance, even when they moved to Sweden as a child and obtained their education degree in Sweden. Montobbio *et al.* (2015) study the contribution of migrant workers to innovation at sectoral level over the period 1994-2005. Focusing on three countries, Germany, UK and France, the authors provide evidence that highly-educated migrants have a positive but smaller effect on innovation compared to highly-educated natives. Moreover, this effect appears to be stronger in high-tech manufacturing industries than in services. In a similar vein, Zheng (2017) investigate the role of migrant inventors in Sweden over the period 1985-2007. The author looks at the interaction between

[1]The variable of knowledge fit is measured using a self-reported method, namely asking to a sample of 869 inventors whether "a significant part of [their] previous inventive experience was no longer applicable to the new organization's inventive activities". The possible answers are defined on a 6-point Likert scale: "fully agree, ..., fully disagree, does not apply" (Hoisl and de Rassenfosse, 2017, pag. 39).

the level of education of the inventor, her type of occupation and her sector of employment. It is found that in high-tech knowledge-intensive services sector migrants are more likely to patent than natives, and this may be explained by their higher educational level. Furthermore, in high-tech manufacturing, low-tech and medium-low tech knowledge intensive services sector migrants are less likely to patent because they are over-represented in low-skill occupations such as clerks, machine operators and assemblers.

3 Estimation strategy and econometric issues

In order to study the effect of educational mismatch on inventor productivity, we rely on the empirical model mostly used in the analyses of the effects of labour mismatch on wages and firm productivity (Hartog, 2000; Mahy *et al.*, 2015) In particular, we employ the so-called ORU model which explains the effect of Over-education, Required education and Under-education on inventor productivity. We estimate the following model:

$$y_i = \beta_0 + \beta_1 req_i + \beta_2 over_i + \beta_3 under_i + \beta_4 X_i + \varepsilon_i \tag{1}$$

where i denotes the inventor. The dependent variable y_i is expressed in terms of the row number or quality adjusted number of inventors' patent applications between 2003 and 2010. X_i represents a set of control variables.

The patent quality depends on the impact that the innovation has on the course of the related technology field. The patent value is measured in the literature with an array of indicators such as the number of claims, family size, patent scope, backward citations and forward citations (Tong and Frame, 1994; Hall *et al.*, 2001; Harhoff *et al.*, 2003; Lanjouw and Schankerman, 2004). In this study, we use the number of forward citations (NFC) to measure the quality of patents.

We estimate the relationship between the ORU variables and the inventor productivity by applying a negative binomial model with robust standard errors. Given the discrete nature of our dependent variables, the application of a linear regression model can result into inefficient and inconsistent estimates. The most basic model to treat with count variables is the Poisson regression which assumes equality between the conditional mean of the dependent variable and the conditional variance. If over-dispersion is present, the Poisson regression model produces consistent but inefficient estimates as the standard errors are downward biased. By contrast, the negative binomial regression model accounts for over-dispersion, allowing the conditional variance of the outcome variable to exceed the conditional mean. Looking at our data, we observe that the distribution of inventor's patent applications is right-skewed: nearly 60% of inventors have just one patent application from 2003 to 2010 and the variance (10.78) is larger than the mean (2.39). With respect to the number of forward citations, almost 53% of inventors have not received any citations and just 21% of them received only one citation over the entire period of analysis. Accordingly, the variance (11.9) is by far larger than the mean (2.59). To test for over-dispersion, we also perform the Likelihood-ratio test in order to check for the significance of the dispersion parameter. We reject the null hypothesis that the dispersion parameter is statistically equal to zero, and therefore the negative binomial model results the most appropriate model for our analysis. We apply this regression model in a cross-sectional analysis given the very low variability of the educational mismatch status of each inventor over the time horizon of our analysis (despite our data allow us to explore a within-individual variation of inventor productivity).

Our analysis may be potentially affected by econometric issues. The relationship that we are interested to identify is whether an increase in the years of required education, over- (under-) education, is followed by an increase (decrease) in the inventor patenting productivity. However, there are three main sources of bias that may hamper the identification of this relationship. First, our ORU variables may be correlated with the error term, namely with some omitted variables that may explain part of the variation of inventive productivity. Since we conduct a cross-sectional analysis, we are not able to control for individual fixed effects (i.e. for time-invariant individual heterogeneity). However, due to the rich data available, we control for a wide set of inventor characteristics and hence the bias associated with the lack of account of idiosyncratic heterogeneity is likely to be small.

The second issue is related to measurement errors in education variable. Statistics Sweden takes into account only the educational degrees that have been acquired in Sweden. In other words, those individuals that have acquired their highest degree in a foreign country are registered with the lower degree acquired in Sweden (Joona *et al.*, 2014, p. 6). For this reason, the required years of education may be affected by measurement errors that lead to biased coefficients for both over-education and under-education.

Finally, our results may be affected by problems of reverse causation. Inventors may be better allocated to the job because they have a higher patenting productivity. For instance, the most talented inventors may find more easily a better job and hence be less involved in imperfect job matching. In other words, the mismatch is a reflection of inventor's ability and causation would run from inventive productivity to the status of (mis)match. Reverse causation may be addressed by instrumental variables (IV) that are related to educational mismatches (i.e. the endogenous variables) and unrelated to the patenting activity. Formally, one should exploit information on some individual characteristics that is related to the educational attainment of the inventors but not to their inventive productivity. This is not an easy task due to the nature of the data source. As an alternative, one may exploit information on factors that hamper inventor mobility between jobs, and hence indirectly affect their job matches. Indeed, most inventors do not change their status of mismatch or match during the period of analysis. This may suggest that most of them face different types of constraints (e.g. family ties) that prevent them from relocating geographically in order to improve the quality of their job matching.

However, since the better job matching of more productive inventors is likely to be driven by their inner inventive ability, as long as this attitude is correlated to their educational performance, one may add an indicator of graduation performance and see whether this affects the magnitude and significance of the impact of ORU variables. In other words, adding a control for individual (education-related) ability such as the high-school grade, we may be able in some extent to neutralize the effect of reverse causality. Clearly, this check does not fully address estimation bias associated with simultaneity feedbacks, but can somehow provide some insights on how severe is this issue in our regression analysis.

Control variables

Our rich data source allows us to use a set of standard covariates that have been found in the literature to be related to inventor productivity and may plague estimation of educational mismatch effect. The first group of controls contains information on the inventors' age, gender, nationality (i.e. foreign-born status) and, albeit in a subsample, the high-school grade. The second group includes the firm size, the sector of activities and geographical area of work. We have also information on the technology field on which a patent relies.

The log of the inventor's age, both in linear and quadratic form, is included in order to account for the inverse U-shaped relationship between innovation productivity and inventor's age (Frosch, 2011, pag. 415). Younger workers are more likely to be over-educated in their occupation. During their career, they acquire work experience and more skills than those provided by their educational attainment, and most likely their job matching improves. By contrast, older workers are most likely to be under-educated as educational attainments in Sweden (Jung and Ejermo, 2014), as well as across OECD countries, has grown over time (McGowan and Andrews, 2015b).

As for the gender, empirical research suggests the presence of gender differences in inventive productivity (Azoulay *et al.*, 2007). Schettino *et al.* (2013) show that patents filed by men are more valuable than those filed by women, while the number of patent applications is not significantly affected by gender differences. By contrast, Whittington and Smith-Doerr (2005) suggest that the quality of patents filed by women is the same or better than that filed by men. In general, women are also more likely to be overeducated than men in the labour market (McGoldrick and Robst, 1996; Karakaya *et al.*, 2007).

The literature on the impact of migration on labour market outcomes or inventive activity is quite wide. Foreign-born inventors are found to perform poorly both in terms of patent quantity and quality than natives in Sweden (Zheng and Ejermo, 2015). Moreover, foreign-born workers are more likely to be over-educated than natives across OECD countries (Joona *et al.*, 2014; Levels *et al.*, 2014).

High-school grades are included as a control for individual ability and are used to check the robustness of our results on the effects of educational mismatch.

As regard firm size, large firms are more likely to patent and to produce high valuable patents as they usually operate in international markets and thus have a higher strategic propensity toward patent protection (Ahlin and Ejermo, 2017). By contrast, small firms have more financial constraints and thus patent more valuable inventions (Leiponen and Byma, 2009). Large firms are also more likely to employ over-educated workers than small firms as they can anticipate future skills needed for their activities. By contrast, large firms are less likely to employ under-educated workers as they have a better managerial quality (McGowan and Andrews, 2015b). Furthermore, patenting activity varies across industries. It is higher in high-tech manufacturing industries and knowledge-intensive services than in other manufacturing industries, agriculture and services (Zheng and Ejermo, 2015). To account for these differences we have included twenty-three industry dummies.

In order to take into account the knowledge spillovers among inventors we control for the number of inventors working in the same geographical area, their average productivity across both geographical areas and industries of work. In particular, we expect that in the areas and industries characterized by a higher concentration of inventors, there are high levels of knowledge transfers. As a result, the inventor productivity is likely to increase (Orlando and Verba, 2005).

Finally, we include five dummies to cover the technology fields of the patents. These dummies take into account the fact that the technology fields with a higher number of patent applications are more likely to have a higher number of citations than fields with a lower number of patents applications (Ejermo and Kander, 2011).

4 Data

In this study we conduct a cross-sectional analysis on a sample of inventors residing in Sweden, covering the period 2003-2010. The final database uses several data sets from Statistics Sweden. The first data set

(Patent Database) contains information on inventors and their patent applications at the European Patent Office (EPO), extracted from the Worldwide Patent Statistics (PATSTAT) database. The inventors included in the EPO records with Swedish addresses have been identified through the Swedish security number (SSN) assigned to each resident in Sweden[2]. The identification was made following two steps (see Jung and Ejermo, 2014, for a detailed description of data construction)[3].

The Patent Database has been merged with the Longitudinal integration database for health insurance and labor market studies (LISA) data sets on individual and employer characteristics using the individual and employer key identifiers, respectively. We use the Patent Database to construct our two dependent variables and the patent technological classes. This database is available from 1978 to 2010. The dependent variables are defined both in terms of quantity and quality of the patents. For the former, we calculate the total counts of inventors' patent applications between 2003 and 2010, not corrected for the contribution of more inventors to the same patent (i.e. counts are not fractional). Also, for each patent, we calculate the amount of forward citations received by each inventor up to 5 years following the publication date. We therefore compute the cites-weighted number for each patent. Most of studies on patent value show that forward cites appear to be correlated with the value of inventions (Trajtenberg, 1990). However, this measure of quality has some drawbacks. On one hand, patents do not offer a comprehensive overview on new inventions in the economy as not all new inventions are patented (Griliches, 1998). On the other hand, the value of inventions varies to a large extent in their technological impact and are strongly skewed (Hall *et al.*, 2001). Despite these limitations, the number of forward citations is widely used in the literature and is considered a good predictor of patent quality (Lanjouw and Schankerman, 2004).

Following Squicciarini *et al.* (2013), inventors' patents are grouped by their main technology field as follows: electrical engineering, instruments, chemistry, mechanical engineering and other fields (e.g. furniture, civil engineering and other consumer goods).

Our analysis is restricted to inventors for which information on the occupational status and the level of education was available. The educational variable is classified according to the Swedish Educational Terminology (SUN2000s) which is linked to the ISCED 1997. In order to compute the mismatch indicators, we have converted the educational categories into years of education (see Table 9 in the Appendix 1). The analysis is focused on a specific occupational category of inventors, namely those belonging to 21 and 31 occupations at two digits, as defined by the Swedish standard classification of occupations[4]. These occupations are related to professionals and technicians in physical, engineering, mathematical and natural sciences which represent the majority of employees who apply for a patent, as reflecting the main profiles employed in R&D tasks (Kaiser *et al.*, 2015, pag. 96)[5]. Besides, as standard in this literature (see Bauer, 2002 and Allen and Van der Velden, 2001), only

[2]A SSN is assigned to any resident living in Sweden for more than one year.

[3]In the first step, the individual's SSN was added by a commercial firm according to her name and address, and in the second step, Statistics Sweden has identified each individual's SSN with that contained in the Longitudinal integration database for health insurance and labor market studies (LISA), available from 1990 to 2012. The LISA database contains demographic and employment information on all individuals aged 16 and over who were registered in Sweden on December 31 of each year. Finally, each SSN has been replaced with an individual key identifier.

[4]The Swedish Occupational Classification (SSYK) is linked to ISCO-88 and is organized into 10 major categories: 1. legislators, senior official and managers; 2. professionals; 3. technicians and associate professionals; 4. clerks; 5. service workers and shop sales; 6. skilled agricultural and fishery workers; 7. craft and related trade workers; 8. plant and machine operators and assemblers; 9. elementary occupations; 10. armed forces.

[5]In order to identify the main occupations in which R&D workers are employed, OECD (2015, pag. 162) states that: *ISCO-08 is the relevant reference document: researchers are classified in ISCO-08 Major Group 2, "Professionals", and in "Research and Development Managers" (ISCO-08, 1223); technicians and equivalent staff are classified in ISCO-08 Major Group 3, "Technicians and associate*

inventors who are "purely employed", and thus are not self-employed or hold more than one job, are considered in this analysis. Information on employment status is available as of 2003, thus our analysis covers the period 2003-2010.

LISA data sets and Patent Database are used to construct the following control variables. Inventor's age, both in linear and quadratic form, is expressed in logarithm. The gender is a dummy variable (equal to 1 if female). The inventor's industry of work is classified according to the Swedish Standard Industrial Classification (SNI 1992 and SNI 2007, built on ISIC Rev. 3), and it is grouped into 23 categories. The firm size is a categorical variable grouped into three classes: small firms (1-99 employees), medium firms (100-499) and large firms (500 employees and more). We also include the firm's patent portfolio, as measured by the log of the total number of patent applications that the employer filed to the EPO from 2003 to 2010, the number of inventors working in the same geographical area (i.e. metropolitan, urban and countryside areas), and the average productivity of inventors computed across geographical areas and industries of work, respectively. The latter variable is measured as the ratio of the total number of patent counts to the number of inventors in the same geographical area or industry of work.

The Swedish and Foreign Background data set is used to construct the foreign-born variable. It is measured as a dummy which assumes the value zero if the inventor has Swedish background (the inventor is born in Sweden and at least one of her parent is a Swedish citizen), and the value one if the inventor has foreign background (the inventor is born abroad and at least one of her parent is a foreigner).

Finally, the inventor's high-school grade is taken from the Secondary Schooling Registry database. This variable refers only to inventors who graduated from high-school between 1973 and 1996, the period in which the national grading system in Sweden was unified, and thus allow the comparability of grades among inventors (Björklund *et al.*, 2003).

The final sample consists of a cross-section of 3,127 inventors and it is reduced to 2,150 when the high school grade is included. We take the first occurrence of the control variables that are time-variant during the period examined (such as firm size, age, age^2 and the firm's patent portfolio).

Measuring Educational Mismatch

Educational mismatch is based on the comparison between the individual's educational attainment and that demanded by her occupation. There are in the literature three main approaches to measure the demanded level of education: the job analysis, the self-reported and the realized matches.

In the job analysis approach, job experts determine the correspondence between the codes assigned to each occupation and those allocated to each level of education. Examples include the correspondence between the International Standard Classification of Occupations (ISCO) code and the International Standard Classification of Education (ISCED). This method presents some drawbacks (Hartog, 2000). First, it is based on the information available to the expert in defining the actual requirements to perform the job (Verhaest and Omey, 2006; Dahlstedt, 2011). Second, although the job requirements are likely to be quite stable over time, jobs within each occupation are assumed to be homogeneous in terms of skill content. Finally, the results of the match are comparable only in countries that use the same educational encoding (McGuinness, 2006, Quintini, 2011).

professionals"; and other R&D supporting staff are essentially found in ISCO-08 Major Groups 4, "Clerks"; 6, "Skilled agricultural and fishery workers"; and 8, "Plant and machine operators, and assemblers". By convention, R&D personnel working in defence are classified in ISCO-08 Major Group 0, "Armed forces occupations".

The self-reporting approach consists in asking workers about the educational requirements of their job, or alternatively, whether they think to be over-educated, under-educated or correctly matched with respect to the qualification required by their job (Duncan and Hoffman, 1981, Verhaest and Omey, 2006, Levels *et al.*, 2014). This method presents some limitations. First, it depends on the way the question has been formulated. As observed by Leuven and Oosterbeek (2011, p. 10), questions about required education are formulated in different ways across studies, some of them focusing on the recruitment standards to *get* a job, and others by highlighting the requirements to *perform* a job. As a result, measures of required education are not directly comparable. Second, the respondents may overstate (or understate) the requirements of their job (Hartog, 2000) or may reflect instead job satisfaction (Boualam *et al.*, 2014). Conversely, this method has the advantage of being job-specific rather than to refer to an occupation which collects a group of heterogeneous jobs (McGowan and Andrews, 2015a).

In the realized matches approach the required amount of education is computed taking into account the statistical distribution of educational attainments within each occupation. Specifically, the required level is defined as the modal value or, alternatively, the number of years of education in the range of one standard deviation around the sample mean (Verdugo and Verdugo, 1989, Kiker *et al.*, 1997, Joona *et al.*, 2014). The realized matches approach is generally applied when the other two methods are not available. It presents a number of inconveniences (Verhaest and Omey, 2006). First, like the job analysis method, it ignores the heterogeneity of jobs within the same occupational category, thus allowing only one educational level to be appropriate for each occupation (Flisi *et al.*, 2017). Second, it compares workers graduated in different periods and hence it is highly sensitive to cohort effects (OECD, 2013). As a result, it does not take into account the changes in the technology requirement of the job (Hartog, 2000). Finally, the required level of education is endogenously inferred by the characteristics of the workers included in each occupation (Hartog, 2000).

In this study we use the third method and our choice is dictated essentially by the data availability. In particular, the required years of education (RE) are computed as the mode years of education within each occupation (i.e. sub-groups 21 and 31 of the Swedish Occupational Classification (SSYK)) and for each year from 2003 to 2010. In order to identify the inventors who are over-educated or under-educated we follow the method adopted in Levels *et al.* (2014, p. 967). We calculate the years of over-education as the difference between the inventors years of education and the required education within a given occupation. We assign zero to the over-education variable in case the years of education and the required education are the same (i.e. well-matched inventors) or the required education is higher. By contrast, we calculate the years of under-education as the difference between the required education and the inventors years of education within a given occupation. We assign zero to the under-education variable in case the years of education and the required education are the same (i.e. well-matched inventors) or the years of education are higher. In this way under-education assumes positive values. Formally, the years of over-education and under-education are computed as follows:

$$OE = EA - RE \quad \text{if} \quad EA > RE$$
$$OE = 0 \quad \text{if} \quad EA \leq RE$$

$$UE = RE - EA \quad \text{if} \quad RE > EA$$
$$UE = 0 \quad \text{if} \quad RE \leq EA$$

where OE indicates over-education, RE is required years of education, UE represents under-education and EA is the educational attainment converted into years of education.

5 Empirical results

Descriptive statistics

Table 1 shows that the total number of patent applications per inventor is 2.73 over the period 2003 to 2010, with a range between 1 and 88 patents per inventor. The total number of forward citations over the period is about 3.31, and up to a maximum of 104 citations. Figure 1 shows that the distribution of inventors' productivity is strongly right-skewed for both the quantitative and the qualitative patenting indicators. The majority of inventors have just one patent over the period 2003-2010. Likewise, most of inventors have received one citation up to 5 years after the publication. The sample is composed by 88% by men and 11% of inventors are foreign-born. The inventors' average years of education is equal to 15.86, which corresponds to a Master's degree. Similarly, the mean value of the required years of education is 15.69. Over-educated inventors have one year of education in excess, and up to a maximum of 8 years. Under-educated inventors have less than one year of education in deficit, and up to a maximum of 9 years. Over the period 2003-2010 inventors have an average age of 40, with a range between 22 and 68. The firms in which inventors are employed are of large dimensions, namely with more than 500 employees, and hold an average of 57.85 patents. Finally, 66% of inventors works in metropolitan areas (i.e. Stockholm, Gothenburg and Malmö) whilst a tiny percentage works in countryside areas.

Table 2 displays that 62% of inventors are highly-qualified, holding a Master's degree or a PhD degree, whereas 15% of inventors have a Bachelor's degree. Nearly 84% of inventors have a degree in Engineering and 10% in Science and Maths; only a small percentage is distributed to the remaining fields.

Table 3 illustrates that the sample of inventors is mainly concentrated in the manufacturing sector, in which around 21% are in electrical and optical equipment industry, 11% in machinery and equipment n.e.c. and about 18% in transport equipment (i.e. motor vehicles, trailer and semi-trailer). Around 23% of inventors work in firms offering business services related to R&D, science and engineering[6].

[6]The Tables 1 - 3 show only inventors with at least a secondary degree because in Statistic Sweden the data for field of education and sectoral classification are only available for workers with a level of education equal or higher than secondary school.

Table 1: **Descriptive statistics (2003-2010)**

Variables	Mean	Standard Deviation	Minimum	Maximum
Total number of patents applications	2.73	4.08	1	88
Total number of forward citations	3.31	5.65	1	104
Years of education	15.86	2.81	7	21
Required education (years)	15.69	0.92	13	16
Over-education (years)	1.01	1.89	0	8
Under-education (years)	0.83	1.46	0	9
Female (dummy)	0.11	0.32	0	1
Age	40	1.26	22	68
Foreign-born (dummy)	0.11	0.32	0	1
Firm size (nr of workers)				
Small size	0.10	0.30	0	1
Medium size	0.14	0.35	0	1
Large size	0.75	0.43	0	1
Patent portfolio	57.85	7.73	1	1070.62
Areas of work				
Metro areas	0.66	0.47	0	1
Urban areas	0.31	0.46	0	1
Countryside areas	0.03	0.18	0	1
Observations	3,127			

Source: Statistics Sweden and CIRCLE data on inventors. Only inventors employed in 21 and 31 occupations are included. Since the majority of inventors do not change their level of mismatch over time, we take the first occurrence of age, firm's patent portfolio and firm size (time-variant variables). Over-education is set to zero if the inventor is well-matched or under-educated, whereas under-education is set to zero if the inventor is well-matched or over-educated. Years of education and areas of work are not included in the regressions.

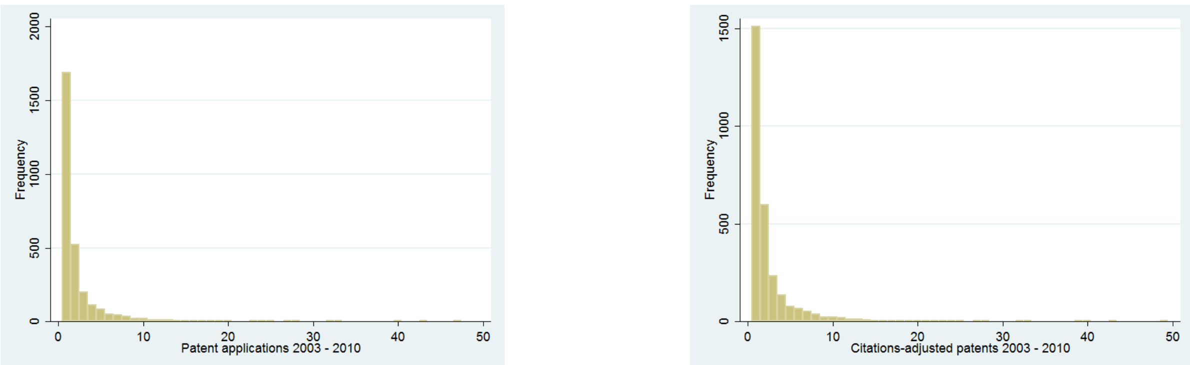

Figure 1: **Histograms for the number of patent applications and forward citations (2003-2010)**

Table 2: **Levels and Fields of Education**

Variables	Share	Standard Deviation
Level of education		
Lower secondary level	0.01	0.09
Upper secondary level	0.10	0.30
Post secondary level (less than 2 years)	0.12	0.32
Post secondary level (up to three years)	0.15	0.36
Post secondary level (up to five years)	0.42	0.49
Postgraduate	0.20	0.40
Field of education		
General education	0.02	0.13
Teaching	0.00	0.04
Humanities	0.01	0.08
Social Sciences	0.01	0.11
Science and Maths	0.10	0.31
Engineering	0.84	0.37
Agriculture	0.00	0.03
Health	0.01	0.12
Services	0.00	0.05
Observations	3,127	

Source: Statistics Sweden and CIRCLE data on inventors. The levels of education are classified according to Swedish Standard Classification of Education (SUN) and aggregated as follows (see Table 9 in Appendix 1): Lower secondary school (categories 10 and 20); Upper secondary school (31, 32, 33); Post secondary level (less than 2 years) (41); Post secondary level (up to three years) which corresponds approximately to a Bachelor's degree (52, 53); Post secondary level (up to five years) which corresponds approximately to a Master's degree (54, 55); Postgraduate (60, 62, 64).

Table 3: **Sectoral classification for the sample of inventors (2003-2010)**

Variables	Share	Standard Deviation
Manufacturing		
Food, beverages and tobacco	0.004	0.064
Textile and leather products	0.003	0.051
Wood and wood products	0.003	0.054
Pulp, Paper and paper products, publishing and printing	0.053	0.224
Coke, refined petroleum products and nuclear products	0.004	0.067
Chemicals and chemicals products	0.067	0.250
Rubber and plastic products	0.012	0.108
Other non-metallic mineral products	0.002	0.047
Basic metal and fabricated metal products	0.061	0.239
Machinery and equipment N.E.C.	0.112	0.315
Electrical and optical equipment	0.214	0.410
Transport equipment	0.182	0.386
Manufacturing N.E.C.	0.001	0.036
Other sectors		
Construction	0.004	0.064
Wholesale and Retail Trade, Repair and household goods	0.032	0.176
Renting and business activities	0.235	0.424
Education	0.006	0.076
Health and social work	0.002	0.044
Observations	3,127	

Source: Statistics Sweden and CIRCLE data on inventors. Only inventors employed in 21 and 31 occupations are included. The Swedish sector classification (SNI2002) is linked to ISIC Rev. 3. The remaining sectors (hotels and restaurants; financial intermediation; activities of household; extra territorial organizations and bodies) do not report observations of inventors employed in occupation 21 and 31 (SSYK, linked to ISCO-88) or in some cases (agriculture, hunting and forestry; mining and quarrying; electricity, gas and water supply; transport, storage and communication; public administration and defence; other community, social and personal service activities) contain a small fraction of inventors.

Baseline estimates

Table 4 and Table 5 present our baseline results of the relationship between the educational mismatch variables (ORU) and inventor productivity, measured in terms of quantity and quality of patents, respectively[7].

Estimates yielded by the negative binomial regression are based on standard errors robust to heteroskedasticity. In Table 4, Column 1 reports the basic specification including only the ORU variables. The estimated coefficients show that there exists a statistically significant association between educational mismatches and the number of inventors' patent applications. In particular, inventors holding a number of years of education in excess (over-education) or in deficit (under-education) to the modal value of education (i.e. required education) in a given occupation are associated with a lower inventive productivity relative to inventors with a similar educational background, but rightly matched to occupation (i.e. $\beta_3 < \beta_1 > \beta_2$). These results confirm the empirical evidence found in the literature studying the relationship between educational mismatch and workers' earnings (Korpi and Tåhlin, 2009; Quintini, 2011). In particular, over-educated inventors filed a number of patents higher than inventors who are well-matched in the same occupation (0.09), but perform poorly than well-matched inventors who hold a similar amount of education ($0.21 > 0.09$). On the contrary, under-educated inventors filed a number of patents lower than inventors who are appropriately matched in the same occupation (-0.06), but more than well-matched inventors who hold the same amount of education ($0.21 - 0.06 > 0$). In the remaining part of the table, we include several control variables. Column 2 includes age and age^2, female dummy variable, foreign-born status, the number of inventors in the same area of work and a set of dummies for industry and technology fields of the patents. The inclusion of these controls do not affect the coefficients associated with the ORU variables, which remain highly significant. The coefficient of female is negative and statistically significant at the 1% level and indicates that women are less productive in terms of row number of patents compared to men. This result conforms to earlier evidence that women scientists may suffer from a productivity gap (Schettino *et al.*, 2013; Ding *et al.*, 2006).

As we expected, there exists a non-linear relationship between inventor productivity and age. The positive and significant sign of age and the negative coefficient for its square value indicate that the number of patents applied for increases less than proportionally with age (Frosch, 2011). The estimated coefficient associated with foreign-born inventors is negative, but it is statistically insignificant, implying that foreigner inventors are as productive as natives.

We also include the total number of inventors for each geographical area where the inventor works. The coefficient is positive and statistically significant, denoting that the areas with a higher concentration of inventors may be characterized by a greater knowledge transfer among inventors, which is likely to affect the creative performance of individual inventors (Giuri and Mariani, 2013).

Column 3 refines the previous check including the average productivity of inventors by geographical area. It is likely that inventors who work together in specific projects have frequent interactions, especially with more prolific inventors, and this may exert a positive effect on the productivity of all inventors (i.e. peer effects).

[7]In the Appendix 2 we have also reported the estimates considering inventors employed in a broad number of occupations (beyond sub-groups 21 and 31), with a number of observations sufficient to calculate the three ORU variables. Overall, the results are similar to those restricted to the two occupations related to the professionals and technicians (sub-groups 21 and 31).

Table 4: **Negative binomial regressions on total number of patents per inventor (2003-2010)**

	1	2	3	4	5	6	7	8	9	10	11	12
Dep: patent counts												
ORU vars.												
Req. edu	0.211***	0.185***	0.192***	0.184***	0.171***	0.155***	0.208***	0.182***	0.191***	0.181***	0.167***	0.152***
	(0.0230)	(0.0222)	(0.0226)	(0.0224)	(0.0225)	(0.0212)	(0.0308)	(0.0284)	(0.0292)	(0.0296)	(0.0284)	(0.0271)
Over-edu	0.0899***	0.0868***	0.0843***	0.0868***	0.0855***	0.0860***	0.0895***	0.0867***	0.0824***	0.0884***	0.0829***	0.0844***
	(0.0150)	(0.0145)	(0.0150)	(0.0146)	(0.0140)	(0.0142)	(0.0178)	(0.0167)	(0.0172)	(0.0169)	(0.0161)	(0.0164)
Under-edu	-0.0613***	-0.0550***	-0.0574***	-0.0501***	-0.0457***	-0.0412***	-0.0553**	-0.0488**	-0.0496**	-0.0429*	-0.0405*	-0.0358*
	(0.0144)	(0.0143)	(0.0143)	(0.0148)	(0.0141)	(0.0138)	(0.0244)	(0.0226)	(0.0234)	(0.0233)	(0.0222)	(0.0215)
Female		-0.323***	-0.392***	-0.228***	-0.334***	-0.332***		-0.335***	-0.450***	-0.227***	-0.343***	-0.342***
		(0.0602)	(0.0607)	(0.0632)	(0.0593)	(0.0574)		(0.0724)	(0.0729)	(0.0785)	(0.0716)	(0.0692)
Age		6.810***	7.620***	7.825***	6.346**	5.951**		6.121	7.620	7.199	6.215	5.405
		(2.584)	(2.604)	(2.595)	(2.561)	(2.475)		(5.397)	(5.445)	(5.527)	(5.257)	(5.210)
Age^2		-0.920***	-1.036***	-1.053***	-0.860**	-0.805**		-0.870	-1.081	-1.010	-0.882	-0.770
		(0.348)	(0.350)	(0.349)	(0.345)	(0.333)		(0.735)	(0.742)	(0.753)	(0.716)	(0.709)
Foreign-born		-0.0761	-0.0694	-0.0816	-0.0853	-0.112		0.0608	0.0586	0.0543	0.0560	0.0623
		(0.0718)	(0.0732)	(0.0726)	(0.0704)	(0.0717)		(0.153)	(0.147)	(0.167)	(0.155)	(0.148)
High-school grade							0.430**	0.456**	0.496***	0.405**	0.480***	0.468***
							(0.198)	(0.186)	(0.191)	(0.189)	(0.181)	(0.175)
Portfolio patent						0.149***						0.146***
						(0.0105)						(0.0127)
Firm size (omit: small size)												
Medium					0.232***						0.292***	
					(0.0761)						(0.0915)	
Large					0.550***						0.577***	
					(0.0692)						(0.0814)	
No. inventors (Area)		0.266*						0.271				
		(0.139)						(0.170)				
Avg. prod. (Area)			0.245*						0.249			
			(0.147)						(0.182)			
Avg. prod. (Industry)				-0.000226***						-0.000141**		
				(6.92e-05)						(6.43e-05)		
Other cont. vars		Ind. Tech.	Ind.	Tech.	Ind. Tech.	Ind. Tech.		Ind. Tech.	Ind.	Tech.	Ind. Tech.	Ind. Tech.
Constant	-2.377***	-14.79***	-17.16***	-16.24***	-14.13***	-13.01***	-4.840***	-15.52	-19.47**	-16.86*	-16.07*	-14.12
	(0.358)	(4.772)	(4.825)	(4.800)	(4.734)	(4.573)	(1.096)	(9.880)	(9.929)	(10.13)	(9.642)	(9.570)
Log pseudo lik.	-6524.16	-6431.15	-6460.85	-6470.76	-6391.59	-6328.95	-4592.34	-4507.11	-4539.1	-4548.93	-4480.92	-4440.69
Pseudo R2	0.02	0.034	0.0295	0.0281	0.0399	0.0494	0.023	0.0411	0.0343	0.0323	0.0467	0.0553
Observations	3,127	3,127	3,127	3,127	3,127	3,127	2,150	2,150	2,150	2,150	2,150	2,150

Source: Statistics Sweden and CIRCLE data on inventors. Coefficient results and robust standard errors are reported. Only inventors employed in 21 and 31 occupations (SSYK) are included. The dependent variable is the total number of patents applied for from 2003 to 2010. The control variables are: age, age^2, female dummy variable, foreign-born (omit. Swedish-born), high-school grade, firm's patent portfolio, the number of inventors distributed by geographical area (No. inventors), the average productivity of inventors (Avg. prod.) by geographical area and industry of work, 5 dummies for technology fields of the patents (Tech.), 3 dummies for firm size (omit: small size) and 23 dummies for industry of work (Ind.). *** *** p<0.01, ** p<0.05, * p<0.1.

Column 4 adds the average productivity of inventors distributed by industry of work. Accordingly, the industry dummies are left out from the regression in order to avoid correlation issues. The coefficient is negative and statistically significant which means that an increase in average patent productivity in the industry is associated with a lower individual inventor productivity. This is likely to reflect inventive rivalry between inventors working for firms in the same sector of activity. Alternatively, it may be that the larger is the number of patents achieved in a sector the more difficult it is to get new patents, due to diminishing technological opportunities of innovation activities (Jones, 1995b; Segerstrom, 1998; Bloom *et al.*, 2016). Column 5 and 6 include respectively the firm size where the inventor is employed and the firm's patent portfolio. More productive inventors work in larger firms in which are more intensively engaged in patenting activity (Schettino *et al.*, 2013). When we account for the extent of firm's patent portfolio, we observe a moderate decrease in the coefficient size for required education. Large firms have a higher strategic propensity to patent their inventions than small firms (Leiponen and Byma, 2009).

In the right-hand side of the table, we replicate the first set of regressions using high-school grade as additional control. The aim of this variable is to control for individual ability, and hence its effect could be captured by ORU variables. As discussed above, as long as individual inventor's ability is correlated with both patent productivity and educational mismatches (i.e. ORU), controlling for this effect may mitigate reverse causality problems, namely that more productive workers have less chances to be mismatched. The coefficient of our proxy for individual ability is positive and significant. Nonetheless, the inclusion of this variable does not affect our results on the relationship between educational mismatch and inventor productivity, which remain highly significant. The size of the coefficients reduces only marginally compared to the baseline values reported in column 1. By controlling for individual ability, we do observe a change in the effect of inventor characteristics, i.e. age and foreign-born background. This may reflect the correlation between these factors and high-school grade, or may be simply due to the sample composition which is now smaller than in our earlier estimates. Conversely, the impact of other controls remains unchanged after accounting for individual ability.

Table 5 presents our baseline results of the relationship between the ORU variables and the total number of forward citations received over the entire period of analysis. Column 1 shows the results of our ORU variables without the inclusion of controls. The mismatch indicators remain highly statistically significant and the size of coefficients for required education and under-education decrease only marginally, while the coefficient for over-education increases with respect to the estimates obtained with the number of patent counts. This means that educational mismatch is not particularly detrimental for the value of patent applications.

Individual ability seems more correlated with the quality-adjusted patents than the number of patent applications (see col. 8). This may explain why the coefficient for required education is smaller than above and the gap in the effects between over- and under-education is narrower. Indeed, in col. 12, the coefficient size of over-education is somewhat smaller than found in the last column of Table 4. Conversely, the impact of under-education is larger.

Finally, when patents are measured in terms of quality, the magnitude of the coefficient associated with firm size increases. This finding contrasts the empirical evidence which suggests that a more quality-oriented patenting activity should reduce the firms' strategic propensity to patent, in particular for large firms (Granstrand and Holgersson, 2012).

Table 5: **Negative binomial regressions on total number of forward citations (2003-2010)**

	1	2	3	4	5	6	7	8	9	10	11	12
Dep: nr. of citations												
ORU vars.												
Req. edu	0.196***	0.168***	0.179***	0.166***	0.151***	0.132***	0.182***	0.151***	0.165***	0.149***	0.133***	0.116***
	(0.0254)	(0.0241)	(0.0244)	(0.0250)	(0.0243)	(0.0227)	(0.0345)	(0.0311)	(0.0316)	(0.0333)	(0.0306)	(0.0291)
Over-edu	0.0988***	0.0867***	0.0850***	0.0894***	0.0847***	0.0849***	0.0929***	0.0799***	0.0763***	0.0850***	0.0751***	0.0766***
	(0.0170)	(0.0162)	(0.0165)	(0.0172)	(0.0156)	(0.0156)	(0.0198)	(0.0182)	(0.0186)	(0.0199)	(0.0175)	(0.0177)
Under-edu	-0.0763***	-0.0670***	-0.0712***	-0.0622***	-0.0568***	-0.0511***	-0.0824***	-0.0692***	-0.0726***	-0.0670***	-0.0596**	-0.0538**
	(0.0167)	(0.0157)	(0.0159)	(0.0168)	(0.0156)	(0.0150)	(0.0266)	(0.0240)	(0.0250)	(0.0251)	(0.0234)	(0.0223)
Gender		-0.290***	-0.347***	-0.180**	-0.305***	-0.298***		-0.298***	-0.409***	-0.181*	-0.310***	-0.305***
		(0.0748)	(0.0747)	(0.0794)	(0.0734)	(0.0706)		(0.0884)	(0.0877)	(0.0985)	(0.0868)	(0.0835)
Age		8.140***	9.148***	10.06***	7.657**	7.329**		8.501	10.25*	10.78*	8.884	7.960
		(3.010)	(3.019)	(3.086)	(3.049)	(2.876)		(5.692)	(5.727)	(5.919)	(5.499)	(5.400)
Age^2		-1.099***	-1.241***	-1.353***	-1.037**	-0.991**		-1.185	-1.428*	-1.483*	-1.236*	-1.109
		(3.010)	(3.019)	(3.086)	(3.049)	(2.876)		(5.692)	(5.727)	(5.919)	(5.499)	(5.400)
		-1.099***	-1.241***	-1.353***	-1.037**	-0.991**		-1.185	-1.428*	-1.483*	-1.236*	-1.109
		(0.405)	(0.406)	(0.415)	(0.410)	(0.387)		(0.776)	(0.781)	(0.807)	(0.750)	(0.736)
Foreign-born		-0.0441	-0.0411	-0.0550	-0.0535	-0.0838		0.117	0.114	0.0991	0.117	0.137
		(0.0807)	(0.0808)	(0.0828)	(0.0789)	(0.0787)		(0.149)	(0.142)	(0.159)	(0.150)	(0.149)
High-school grade							0.463**	0.537***	0.573***	0.467**	0.566***	0.552***
							(0.218)	(0.202)	(0.208)	(0.207)	(0.194)	(0.186)
Portfolio patent						0.168***						0.163***
						(0.0114)						(0.0136)
Firm size (omit: small size)												
Medium					0.236***						0.357***	
					(0.0849)						(0.0955)	
Large					0.597***						0.670***	
					(0.0783)						(0.0865)	
No. inventors (Area)		0.287*						0.309*				
		(0.163)						(0.184)				
Avg. prod. (Area)			0.374**						0.394**			
			(0.158)						(0.191)			
Avg. prod. (Industry)				-0.000202***						-0.000109*		
				(7.18e-05)						(6.14e-05)		
Other cont. vars		Ind. Tech.	Ind.	Tech.	Ind. Tech.	Ind. Tech.		Ind. Tech.	Ind.	Tech.	Ind. Tech.	Ind. Tech.
Constant	-1.951***	-16.92***	-20.06***	-19.96***	-16.21***	-15.17***	-4.419***	-19.88*	-24.76**	-23.31**	-21.02**	-18.76*
	(0.395)	(5.546)	(5.563)	(5.701)	(5.625)	(5.300)	(1.202)	(10.38)	(10.41)	(10.87)	(10.06)	(9.898)
Log pseudo lik.	-7122.1	-7017.87	-7041.23	-7083.25	-6975.21	-6903.08	-5005.26	-4905.14	-4933.01	-4975.12	-4874.14	-4830.26
Pseudo R2	0.0189	0.0333	0.0301	0.0243	0.0391	0.0491	0.0214	0.041	0.0356	0.0273	0.0471	0.0556
Observations	3,127	3,127	3,127	3,127	3,127	3,127	2,150	2,150	2,150	2,150	2,150	2,150

Source: Statistics Sweden and CIRCLE data on inventors. Coefficient results and robust standard errors are reported. Only inventors employed in 21 and 31 occupations (SSYK) are included. The dependent variable is the total number of patent counts adjusted for the citations received up to 5 years after publication (counted over all the period of analysis). The control variables are: age, age^2, female dummy variable, foreign-born (omit. Swedish-born), high-school grade, firm's patent portfolio, the number of inventors distributed by geographical area (No. inventors), the average productivity of inventors (Avg. prod.) by geographical area and industry of work, 5 dummies for technology fields of the patents (Tech.), 3 dummies for firm size (omit: small size) and 23 dummies for industry of work (Ind.). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Robustness checks

In this section, we assess the robustness of our baseline results and investigate the pattern of our findings along different key inventors' characteristics. More specifically, we inspect whether the effect of educational mismatch changes between: 1) young and old inventors, 2) inventors working in metropolitan or urban and rural area and 3) industry of activity. Overall, our baseline results are found to be robust across different specifications. In Table 6 we distinguish the ORU variables among younger inventors (aged less than 45) and older inventors (aged equal or higher than 45) in relation to both the total number of patents counts and forward citations[8]. These results show that formal education matters more for younger and less experienced inventors. In fact, we observe that the coefficients associated with required education are higher for younger than older inventors. This may be explained by the fact that there are more graduates among youth than older inventors and younger inventors are more likely to be mismatched at the beginning of their career. Moreover, the opportunity cost of being over-educated is higher among younger inventors than older inventors. For the latter, the coefficients becomes insignificant when we include the high-school grade in the regression. For older inventors, instead, under-education covers an important role in terms of both quantity and quality of patents, whose coefficients are high and always significant.

Table 6: **Robustness checks: ORU variables distinguished by age (2003-2010)**

Dependent variable:	nr patents	nr patents	nr citations	nr citations	nr patents	nr patents	nr citations	nr citations
Req. edu (younger)	0.271***	0.254***	0.250***	0.227***				
	(0.0249)	(0.0299)	(0.0303)	(0.0355)				
Over-edu (younger)	0.0979***	0.0983***	0.109***	0.108***				
	(0.0206)	(0.0219)	(0.0217)	(0.0229)				
Under-edu (younger)	-0.0782***	-0.0444	-0.0895***	-0.0527				
	(0.0215)	(0.0323)	(0.0216)	(0.0340)				
High school grade		0.559**		0.615**		0.511		0.441
		(0.254)		(0.278)		(0.350)		(0.396)
Req. edu (older)					0.136***	0.119**	0.147***	0.121**
					(0.0249)	(0.0490)	(0.0303)	(0.0549)
Over-edu (older)					0.0466**	0.0240	0.0532*	0.0162
					(0.0204)	(0.0318)	(0.0287)	(0.0432)
Under-edu (older)					-0.0760***	-0.117***	-0.101***	-0.186***
					(0.0225)	(0.0368)	(0.0293)	(0.0474)
Constant	-3.383***	-6.408***	-2.883***	-6.134***	-1.131***	-3.799**	-1.078**	-3.108
	(0.386)	(1.386)	(0.475)	(1.509)	(0.383)	(1.659)	(0.449)	(2.063)
Log pseudo lik.	-4294.39	-3455.02	-4647.18	-3729.91	-2020.61	-991.99	-2243.89	-1113.53
Pseudo R2	0.018	0.022	0.018	0.022	0.0194	0.019	0.019	0.018
Observations	2,057	1,633	2,057	1,633	1,010	477	1,010	477

Source: Statistics Sweden and CIRCLE data on inventors. Coefficient results and robust standard errors are reported. Only inventors employed in 21 and 31 occupations (SSYK) are included. The dependent variable is the total number of patents applied and the total number of forward citations (2003-2010). *** p<0.01, ** p<0.05, * p<0.1.

Table 7 assesses the robustness of our baseline results by analyzing whether ORU variables vary across the geographical area where the inventor works. We hypothesize that mismatch should be lower in metropolitan areas than urban or rural areas given the presence of higher job opportunities in larger cities (Abel and Deitz, 2015). We find that the coefficients associated with the ORU variables are in general quite similar across areas, but those associated with required and over-education are slightly higher for inventors working in metro areas

[8]We have chosen to fix the threshold for younger and older inventors at 45 years. This allows us to have a number of observations sufficient to calculate the three ORU variables.

Table 7: **Robustness checks: ORU variables distinguished by area of work (2003-2010)**

	1	2	3	4	5	6	7	8
	Metropolitan area				Not Metropolitan areas			
Dependent variable:	nr patents	nr patents	nr citations	nr citations	nr patents	nr patents	nr citations	nr citations
Req. edu	0.229***	0.189***	0.200***	0.186***	0.184***	0.194***	0.185***	0.173***
	(0.0295)	(0.0380)	(0.0342)	(0.0466)	(0.0357)	(0.0477)	(0.0371)	(0.0491)
Over edu	0.0946***	0.0850***	0.101***	0.0993***	0.0814***	0.0837***	0.0939***	0.0819***
	(0.0206)	(0.0235)	(0.0232)	(0.0272)	(0.0194)	(0.0231)	(0.0227)	(0.0241)
Under edu	-0.0609***	-0.0463	-0.0719***	-0.0918**	-0.0614***	-0.0467	-0.0821***	-0.0675**
	(0.0192)	(0.0325)	(0.0230)	(0.0370)	(0.0213)	(0.0335)	(0.0220)	(0.0336)
High school grade		0.563**		0.361		0.401*		0.582**
		(0.273)		(0.313)		(0.235)		(0.249)
Constant	-2.638***	-19.25	-1.984***	-3.823**	-2.001***	-4.510***	-1.845***	-5.071***
	(0.457)	(12.82)	(0.532)	(1.702)	(0.560)	(1.358)	(0.581)	(1.456)
Log pseudo lik.	-4006.24	-2865.07	-4388.82	-3137.12	-2506.79	-1719.02	-2718.1	-1854.25
Pseudo R2	0.02	0.02	0.018	0.02	0.019	0.02	0.02	0.02
Observations	1,881	1,312	1,881	1,312	1,246	838	1,246	838

Source: Statistics Sweden and CIRCLE data on inventors. Coefficient results and robust standard errors are reported. Only inventors employed in 21 and 31 occupations (SSYK) are included. The dependent variable is the total number of patents applied and the total number of forward citations (2003-2010). Areas of work are divided in metropolitan (i.e. Stockholm, Gothenburg and Malmö) and non metropolitan (i.e. urban and countrysides areas). *** p<0.01, ** p<0.05, * p<0.1.

Table 8: **Robustness checks: ORU variables distinguished by industry of work (2003-2010)**

	1	2	3	4	5	6	7	8
	Low-intensity R&D industries				High-intensity R&D industries/K-I Services			
Dependent variable:	nr patents	nr patents	nr citations	nr citations	nr patents	nr patents	nr citations	nr citations
Req. edu	0.234***	0.194***	0.194***	0.165***	0.196***	0.195***	0.195***	0.194***
	(0.0372)	(0.0477)	(0.0441)	(0.0591)	(0.0289)	(0.0365)	(0.0309)	(0.0405)
Over edu	0.116***	0.0837***	0.127***	0.123***	0.0696***	0.0617***	0.0778***	0.0704***
	(0.0262)	(0.0231)	(0.0273)	(0.0301)	(0.0165)	(0.0202)	(0.0213)	(0.0261)
Under edu	-0.0350	-0.0467	-0.0422	-0.0220	-0.0820***	-0.0963***	-0.102***	-0.136***
	(0.0233)	(0.0335)	(0.0265)	(0.0458)	(0.0184)	(0.0257)	(0.0213)	(0.0301)
High school grade		0.401*		0.378		0.504**		0.507**
		(0.235)		(0.391)		(0.201)		(0.235)
Constant	-2.767***	-4.510***	-1.971***	-3.711*	-2.123***	-5.051***	-1.913***	-4.826***
	(0.578)	(1.358)	(0.690)	(2.139)	(0.451)	(1.202)	(0.479)	(1.354)
Log pseudo lik.	-2818.25	-1719.02	-3055.4	-2158.36	-3697.05	-2584.67	-4060.88	-2840.13
Pseudo R2	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Observations	1,330	838	1,330	920	1,797	1,230	1,797	1,230

Source: Statistics Sweden and CIRCLE data on inventors. Coefficient results and robust standard errors are reported. Only inventors employed in 21 and 31 occupations (SSYK) are included. The dependent variable is the total number of patents applied and the total number of forward citations (2003-2010). We distinguish between high-intensity R&D industries/Knowledge-Intensive Services (chemicals; machinery and equipment n.e.c.; electrical and optical equipment; transport equipment; transport, storage and communication; financial intermediation; renting and business activities) and low-intensity R&D manufacturing industries (all the remaining industries). *** p<0.01, ** p<0.05, * p<0.1.

(without any substantial difference between the two types of indicators of patent performance).

In Table 8 we assess the robustness of our baseline results distinguishing by low-intensity and high-intensity R&D industries[9]. We observe that the size of coefficients associated with required education is quite similar among the two groups of industries. The coefficients of under-education are not significant in low-tech R&D industries which may suggest that in these industries - characterized by a relatively low level of technological sophistication - it is likely to apply for patents even with a level of education lower than that required. Over-education has a greater opportunity cost in low-tech industries than in high-tech industries. A higher level of education (compared to the modal education) acquires a greater importance in the low-tech industries where over-educated inventors are particularly productive[10]. In these industries R&D management is less efficient in allocating workers to the right R&D task, relying by far on learning-by-doing processes. In a specular way, under-education has a higher opportunity cost in high-tech industries and knowledge intensive services. These industries are characterized by a higher level of technological sophistication, hence under-educated inventors are mostly penalized in terms of inventive productivity. Moreover, in high-tech industries the level of abilities plays an important role. The coefficient of high-school grade is positive and statistically significant, but still it does not affect the coefficients associated with the ORU model.

6 Conclusions

This study is the first attempt to measure the effects of educational mismatch on inventor productivity. We have used original data on a sample of Swedish inventors from 2003 to 2010 and measured inventor productivity both in terms of patent counts and patent quality. The cross-section analysis has been adopted after having observed that the majority of inventors do not vary their level of mismatch over the period of analysis. Overall, the results suggest that over-educated inventors file a total number of patents higher than inventors who are appropriately matched in the same occupation ($\beta_2 > 0$), but perform poorly than well-matched inventors who hold a similar level of education ($\beta_2 < \beta_1$). Conversely, under-educated inventors file a total number of patents lower than inventors who are well-matched in the same occupation ($\beta_3 < 0$), but more than well-matched inventors who hold the same level of education ($\beta_1 + \beta_3 > 0$).

We control our baseline results for a set of variables related to both individual and employer characteristics and the pattern of ORU specification remains unchanged, even after a measure of individual ability is taken into account. We assess the robustness of our results by investigating the pattern and the significance of ORU model along some specific inventor characteristics. Our results show that formal education is more important for younger inventors as younger well-matched inventors are more productive than their older counterparts. Similarly, the opportunity cost of being over-educated is higher for younger than older inventors, while the opposite is true for under-educated inventors. We also find that ORU model is robust to controlling for geographical areas and industry of work. For the former, the differential effects between metropolitan and urban or rural areas are quite smaller, and for the latter, over-educated inventors appear to have a higher opportunity cost in low-tech industries, while under-educated are mostly penalized in high-tech industries, which are characterized by an

[9] The high-intensity R&D industries are: Chemicals and chemicals products (24), Machinery and equipment n.e.c. (29), Electrical and optical equipment (30-33), Transport equipment (34-35), Transport and storage (60-63), Post and telecommunications (64), Financial intermediation (65-67), Renting and business activities (71-74). The remaining industries are considered as low-intensity R&D industries (see Appendix 3).

[10] Contrary to this finding, Mahy *et al.* (2015) find that the productivity effect of over-education in Belgium over the period 1999-2010 is stronger among firms with a higher share of highly-educated workers and active in high-tech and knowledge-based industries.

intense technological competition and a rapid obsolescence of technology knowledge. In conclusion, it appears that investment in education raises the inventor productivity, but the skill content and specific job requirements may limit the inventor to fully utilize her abilities.

References

Abel, J. R. and Deitz, R. (2015). "Agglomeration and job matching among college graduates". *Regional Science and Urban Economics*, **51**:14–24.

Aghion, P., Howitt, P., Brant-Collett, M., and García-Peñalosa, C. (1998). *Endogenous Growth Theory*. MIT press.

Ahlin, L. and Ejermo, O. (2017). "Mobility as a Learning Mechanism: Evidence from Inventor Data". *mimeo*.

Allen, J. and Van der Velden, R. (2001). "Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search". *Oxford Economic Papers*, **53**(3):434–452.

Azoulay, P., Ding, W., and Stuart, T. (2007). "The determinants of faculty patenting behavior: Demographics or opportunities?" *Journal of Economic Behavior & Organization*, **63**(4):599–623.

Bauer, T. (2002). "Educational mismatch and wages: a panel analysis". *Economics of Education Review*, **21**(3):221–229.

Becker, G. S. (1962). "Investment in Human Capital: A Theoretical Analysis". *Journal of Political Economy*, **70**(5):9–49.

Björklund, A., Lindahl, M., and Sund, K. (2003). "Family background and school performance during a turbulent era of school reforms". *Swedish Economic Policy Review*, **10**:111–136.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2016). "Are Ideas Getting Harder to Find?" *Manuscript, Stanford University, Palo Alto*.

Boualam, B. *et al.* (2014). "Getting a first job: quality of the labor matching in French cities". *Institut d'Economie et Econométrie, Université de Genève*.

Crespi, G. A., Geuna, A., and Nesta, L. (2007). "The mobility of university inventors in europe". *The Journal of Technology Transfer*, **32**(3):195–215.

Dahlstedt, I. (2011). "Occupational Match". *Journal of International Migration and Integration*, **12**(3):349–367.

Desjardins, R. and Rubenson, K. (2011). "An analysis of skill mismatch using direct measures of skills". *OECD Education Working Papers, OECD Publishing, Paris*, (63):1–88.

Ding, W. W., Murray, F., and Stuart, T. E. (2006). "Gender differences in patenting in the academic life sciences". *Science*, **313**(5787):665–667.

Duncan, G. J. and Hoffman, S. D. (1981). "The incidence and wage effects of overeducation". *Economics of Education Review*, **1**(1):75–86.

Ejermo, O. and Kander, A. (2011). "Swedish business research productivity". *Industrial and Corporate Change*, **20**(4):1081–1118.

Flisi, S., Goglio, V., Meroni, E. C., Rodrigues, M., and Vera-Toscano, E. (2017). "Measuring occupational mismatch: overeducation and overskill in europeevidence from piaac". *Social Indicators Research*, **131**(3):1211–1249.

Frosch, K. H. (2011). "Workforce age and innovation: a literature survey". *International Journal of Management Reviews*, **13**(4):414–430.

Giuri, P. and Mariani, M. (2013). "When distance disappears: inventors, education, and the locus of knowledge spillovers". *Review of Economics and Statistics*, **95**(2):449–463.

Giuri, P., Mariani, M., Brusoni, S., Crespi, G., Francoz, D., Gambardella, A., Garcia-Fontes, W., Geuna, A., Gonzales, R., Harhoff, D., *et al.* (2007). "Inventors and invention processes in europe: Results from the PatVal-EU survey". *Research Policy*, **36**(8):1107–1127.

Granstrand, O. and Holgersson, M. (2012). "The anatomy of rise and fall of patenting and propensity to patent: the case of Sweden". *International Journal of Intellectual Property Management*, **5**(2):169–198.

Griffith, R. and Macartney, G. (2014). "Employment protection legislation, multinational firms, and innovation". *Review of Economics and Statistics*, **96**(1):135–150.

Griliches, Z. (1998). "Patent statistics as economic indicators: a survey". In "R&D and productivity: the econometric evidence", pages 287–343. University of Chicago Press.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools". Working Paper 8498, National Bureau of Economic Research.

Harhoff, D., Scherer, F. M., and Vopel, K. (2003). "Citations, family size, opposition and the value of patent rights". *Research Policy*, **32**(8):1343–1363.

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

Hersch, J. (1991). "Education match and job match". *The Review of Economics and Statistics*, pages 140–144.

Hoisl, K. (2007). "Tracing mobile inventors - The causality between inventor mobility and inventor productivity". *Research Policy*, **36**(5):619–636.

Hoisl, K. (2009). "Does mobility increase the productivity of inventors?" *The Journal of Technology Transfer*, **34**(2):212–225.

Hoisl, K. and de Rassenfosse, G. (2017). "Knowledge fit and productivity gains from mobility". *Working paper in progress.*

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2018). "The Allocation of Talent and US Economic Growth". Working paper, National Bureau of Economic Research.

Hunt, J. and Gauthier-Loiselle, M. (2010). "How much does immigration boost innovation?" *American Economic Journal: Macroeconomics*, **2**(2):31–56.

Jones, C. I. (1995b). “R & D-based Models of Economic Growth”. *Journal of Political Economy*, pages 759–784.

Jones, C. I. (2011). “Misallocation, Economic Growth, and Input-Output Economics”. NBER Working Papers 16742, National Bureau of Economic Research, Inc.

Joona, P. A., Gupta, N. D., and Wadensjö, E. (2014). “Overeducation among immigrants in Sweden: incidence, wage effects and state dependence”. *IZA Journal of Migration*, **3**(1):1.

Jovanovic, B. (2014). “Misallocation and Growth”. *American Economic Review*, **104**(4):1149–71.

Jung, T. and Ejermo, O. (2014). “Demographic patterns and trends in patenting: Gender, age, and education of inventors”. *Technological Forecasting and Social Change*, **86**:110–124.

Kaiser, U., Kongsted, H. C., and Rønde, T. (2015). “Does the mobility of R&D labor increase innovation?” *Journal of Economic Behavior & Organization*, **110**:91–105.

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

Karakaya, G., Plasman, R., and Rycx, F. (2007). “Overeducation on the belgian labour market: evaluation and analysis of the explanatory factors through two types of approaches”. *Compare*, **37**(4):513–532.

Kerr, W. R. and Lincoln, W. F. (2010). “The supply side of innovation: H-1b visa reforms and US ethnic invention”. *Journal of Labor Economics*, **28**(3):473–508.

Kiker, B. F., Santos, M. C., and De Oliveira, M. M. (1997). “Overeducation and Undereducation: Evidence for Portugal”. *Economics of Education Review*, **16**(2):111–125.

Korpi, T. and Tåhlin, M. (2009). “Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000”. *Labour Economics*, **16**(2):183–193.

Lanjouw, J. O. and Schankerman, M. (2004). “Patent quality and research productivity: Measuring innovation with multiple indicators”. *The Economic Journal*, **114**(495):441–465.

Leiponen, A. and Byma, J. (2009). “If you cannot block, you better run: Small firms, cooperative innovation, and appropriation strategies”. *Research Policy*, **38**(9):1478–1488.

Leuven, E. and Oosterbeek, H. (2011). “Overeducation and mismatch in the labor market”. *Handbook of the Economics of Education*, **4**:283–326.

Levels, M., van der Velden, R., and Allen, J. (2014). “Educational mismatches and skills: new empirical tests of old hypotheses”. *Oxford Economic Papers*, **66**(4):959–982.

Liu, T., Mao, Y., and Tian, X. (2016). “The Role of Human Capital: Evidence from Patent Generation”. *Kelley School of Business Research Paper No. 16-17*.

Lucas, R. E. (1988). “On the mechanics of economic development”. *Journal of Monetary Economics*, **22**(1):3–42.

Mahy, B., Rycx, F., and Vermeylen, G. (2015). “Educational mismatch and firm productivity: Do skills, technology and uncertainty matter?” *De Economist*, **163**(2):233–262.

Mariani, M. and Romanelli, M. (2007). “Stacking” and “picking” inventions: The patenting behavior of european inventors”. *Research Policy*, **36**(8):1128–1142.

McGoldrick, K. and Robst, J. (1996). “Gender differences in overeducation: A test of the theory of differential overqualification”. *The American Economic Review*, **86**(2):280–284.

McGowan, M. A. and Andrews, D. (2015a). “Labour Market Mismatch and Labour Productivity: Evidence from PIAAC Data”. OECD Economics Department Working Papers 1209, OECD Publishing, Paris.

McGowan, M. A. and Andrews, D. (2015b). “Skill Mismatch and Public Policy in OECD Countries”. OECD Economics Department Working Papers 1210, OECD Publishing, Paris.

McGuinness, S. (2006). “Overeducation in the labour market”. *Journal of Economic Surveys*, **20**(3):387–418.

Miguélez, E. and Moreno, R. (2013). “Skilled labour mobility, networks and knowledge creation in regions: a panel data approach”. *The Annals of Regional Science*, **51**(1):191–212.

Montobbio, F., Fassio, C., and Venturini, A. (2015). “Innovation and human capital: Age, skills and ethnicity. a study on france, germany and uk”. *DRUID, Rome.*

Nakajima, R., Tamura, R., and Hanaki, N. (2010). “The effect of collaboration network on inventors’ job match, productivity and tenure”. *Labour Economics*, **17**(4):723–734.

Nelson, R. R. and Phelps, E. S. (1966). “Investment in Humans, Technological Diffusion, and Economic Growth”. *The American Economic Review*, **56**(1/2):69–75.

OECD (2013). “OECD Skills Outlook 2013: First Results from the Survey of Adult Skills”. Technical report, OECD Publishing, Paris.

OECD (2015). “Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development”. *The Measurement of Scientific, Technological and Innovation Activities. OECD Publishing, Paris.*

OECD (2016). “Getting Skills Right: Assessing and Anticipating Changing Skill Needs”. Technical report, OECD Publishing, Paris.

Onishi, K. and Nagaoka, S. (2012). “Life-cycle productivity of industrial inventors: Education and other determinants”. *RIETI Discussion Paper Series 12-E-059.*

Orlando, M. J. and Verba, M. (2005). “Do only big cities innovate? technological maturity and the location of innovation”. *Economic Review-Federal Reserve Bank of Kansas City*, **90**(2):31.

Quintini, G. (2011). “Over-Qualified or Under-Skilled: A Review of Existing Literature”. *OECD Social, Employment and Migration Papers. OECD Publishing, Paris*, (121):47.

Rahko, J. (2017). “Knowledge spillovers through inventor mobility: the effect on firm-level patenting”. *The Journal of Technology Transfer*, **42**(3):585–614.

Romer, P. M. (1990). “Endogenous Technological Change”. *Journal of Political Economy*, **98**(5):S71–102.

Schettino, F., Sterlacchini, A., and Venturini, F. (2013). “Inventive productivity and patent quality: Evidence from Italian inventors”. *Journal of Policy Modeling*, **35**(6):1043–1056.

Segerstrom, P. S. (1998). “Endogenous Growth without Scale Effects”. *The American Economic Review*, **88**(5):1290–1310.

Squicciarini, M., Dernis, H., Criscuolo, C., *et al.* (2013). “Measuring Patent Quality: Indicators of Technological and Economic Value”. OECD Science, Technology and Industry Working Papers 2013/03, OECD Publishing, Paris.

Toivanen, O. and Väänänen, L. (2016). “Education and invention”. *Review of Economics and Statistics*, **98**(2):382–396.

Tong, X. and Frame, J. D. (1994). “Measuring national technological performance with patent claims data”. *Research Policy*, **23**(2):133–141.

Trajtenberg, M. (1990). “A penny for your quotes: patent citations and the value of innovations”. *The Rand Journal of Economics*, pages 172–187.

Trajtenberg, M. and Shalem, R. (2009). “Software patents, inventors and mobility”. *Working paper*.

Verdugo, R. R. and Verdugo, N. T. (1989). “The impact of surplus schooling on earnings: Some additional findings”. *Journal of Human Resources*, pages 629–643.

Verhaest, D. and Omey, E. (2006). “Measuring the incidence of over- and undereducation”. *Quality and Quantity*, **40**(5):783–803.

Whittington, K. B. and Smith-Doerr, L. (2005). “Gender and commercial science: Womens patenting in the life sciences”. *The Journal of Technology Transfer*, **30**(4):355–370.

Zheng, Y. (2017). “How Immigrants Invent: Evidence from Sweden”. *Lund: Printed in Sweden by Media-Tryck, Lund University*.

Zheng, Y. and Ejermo, O. (2015). “How do the foreign-born perform in inventive activity? Evidence from Sweden”. *Journal of Population Economics*, **28**(3):659–695.

Zwick, T., Frosch, K., Hoisl, K., and Harhoff, D. (2017). “The power of individual-level drivers of inventive performance”. *Research Policy*, **46**(1):121 – 137.

Appendix 1

Table 9: **Levels of education imputed into years of education**

Categories	Levels of Education	Years of education
10	Lower secondary education (up to 9 years)	**7**
20	Lower secondary education (9-10 years)	**9**
31	Upper secondary education (less than 2 years)	**10**
32	Upper secondary education (2 years)	**11**
33	Upper secondary education (3 years)	**12**
41	Post-secondary education (less than 2 years)	**13**
52	Post-secondary education (2 years)	**14**
53	Post-secondary education (3 years)	**15**
54	Post-secondary education (4 years)	**16**
55	Post-secondary education (5 years)	**17**
60	Other postgraduate program	**18**
62	Licentiate program (half of PhD)	**19**
64	Doctoral program (4 years completed)	**21**

Source: The Swedish Educational Terminology (SUN) equivalent to ISCED97

Appendix 2

Tables 10 and 11 show the results of the effects of educational mismatch on inventor's productivity including all inventors employed in a set of occupations, but limited to those occupations for which was available a number of observations that allow us to calculate the ORU variables (over-, required and under-education): corporate managers (12), managers of small firms (13), scientists and engineers - professionals (21), business professionals and others (24), technicians in science and engineering (31), life science and health technicians (32), other associate professionals (34), office clerks (41) and plant and machine operators (82). Overall, the coefficients estimated are similar to those restricted to 21 and 31 occupations.

Appendix 3

Sector list according to according to the Swedish Standard Industrial Classification (SNI1992 and SNI2007, built on ISIC Rev. 3): Agriculture, Hunting and Forestry (01-05), Mining and Quarrying (10-14), Food, beverages, tobacco (15-16), Textiles, leather and related products (17-19), Wood and products of wood and cork (20), Pulp, paper and paper products, printing and publishing (21-22), Coke, refined petroleum products and nuclear products (23), Chemicals and chemicals products (24), Rubber and plastic products (25), Other non-metallic mineral products (26), Basic metals and fabricated metal products (27-28), Machinery and equipment n.e.c. (29), Electrical and optical equipment (30-33), Transport equipment (34-35) and Manufacturing n.e.c. (36-37), Electricity, gas and water supply (40-41), Construction (45), Wholesale and retail trade, Repair of motor vehicles, motorcycles and Personal and household goods (50-52), Hotel and restaurants (55), Transport, storage and communication (60-64), Financial Intermediation (65-67), Real estate, renting and business activities (70-74), Public administration and defence, compulsory social security (75), Education (80), Health and social work (85) and remaining sectors (90-93, 95, 99).

Table 10: **Negative binomial regressions on total number of patents per inventor (2003-2010)**

Variables	1	2	3	4	5	6	7	8	9	10	11	12
Dep: patent counts												
ORU vars.												
Req. edu	0.248***	0.219***	0.231***	0.213***	0.210***	0.191***	0.246***	0.215***	0.233***	0.209***	0.204***	0.185***
	(0.0171)	(0.0162)	(0.0163)	(0.0169)	(0.0165)	(0.0161)	(0.0221)	(0.0195)	(0.0201)	(0.0210)	(0.0200)	(0.0196)
Over-edu	0.0977***	0.0942***	0.0935***	0.0922***	0.0942***	0.0957***	0.0980***	0.0922***	0.0918***	0.0927***	0.0905***	0.0937***
	(0.0165)	(0.0159)	(0.0164)	(0.0159)	(0.0152)	(0.0159)	(0.0195)	(0.0180)	(0.0185)	(0.0184)	(0.0173)	(0.0181)
Under-edu	-0.0570***	-0.0431***	-0.0475***	-0.0417***	-0.0355***	-0.0317**	-0.0415*	-0.0242	-0.0260	-0.0251	-0.0179	-0.0123
	(0.0137)	(0.0138)	(0.0139)	(0.0142)	(0.0137)	(0.0133)	(0.0231)	(0.0218)	(0.0226)	(0.0223)	(0.0214)	(0.0207)
Gender		-0.305***	-0.364***	-0.232***	-0.315***	-0.316***		-0.352***	-0.440***	-0.268***	-0.357***	-0.357***
		(0.0563)	(0.0571)	(0.0582)	(0.0553)	(0.0539)		(0.0679)	(0.0691)	(0.0717)	(0.0666)	(0.0647)
Age		6.548***	6.867***	7.846***	6.176**	5.410**		7.376	8.073	9.280*	7.013	5.797
		(2.429)	(2.466)	(2.463)	(2.410)	(2.345)		(5.082)	(5.186)	(5.202)	(4.976)	(4.922)
Age^2		-0.885***	-0.934***	-1.058***	-0.835**	-0.728**		-1.048	-1.151	-1.304*	-0.993	-0.825
		(0.327)	(0.332)	(0.332)	(0.324)	(0.316)		(0.692)	(0.706)	(0.708)	(0.677)	(0.669)
Foreign-born		-0.0941	-0.0805	-0.0936	-0.106	-0.128*		0.0172	0.00368	0.0147	0.0118	0.0189
		(0.0681)	(0.0692)	(0.0686)	(0.0669)	(0.0680)		(0.141)	(0.139)	(0.147)	(0.144)	(0.135)
High school grade							0.458**	0.499***	0.514***	0.437**	0.516***	0.503***
Patent Portfolio						0.145***						0.147***
						(0.00947)						(0.0114)
Firm size (omit: small size)												
Medium					0.242***						0.257***	
					(0.0678)						(0.0825)	
Large					0.539***						0.555***	
					(0.0644)						(0.0749)	
No. inventors (Area)		0.333**						0.138				
		(0.138)						(0.173)				
Avg. prod. (Area)			0.238**						0.172			
			(0.119)						(0.148)			
Avg. prod. (Industry)				-0.000222***						-0.000149**		
				(7.52e-05)						(6.12e-05)		
Other cont. vars		Ind. Tech.	Ind.	Tech.	Ind. Tech.	Ind. Tech.		Ind. Tech.	Ind.	Tech.	Ind. Tech.	Ind. Tech.
Constant	-2.957***	-14.96***	-16.45***	-16.66***	-14.52***	-12.69***	-5.609***	-18.58**	-20.86**	-21.14**	-18.40**	-15.68*
	(0.262)	(4.496)	(4.580)	(4.569)	(4.467)	(4.344)	(1.072)	(9.314)	(9.496)	(9.556)	(9.129)	(9.056)
Log pseudo lik.	-7518.56	-7409.16	-7445.57	-7452.04	-7364.53	-7299.23	-5270.62	-5177.03	-5207.68	-5216.32	-5144.54	-5100.29
Pseudo R2	0.02	0.04	0.03	0.03	0.04	0.05	0.02	0.04	0.04	0.03	0.05	0.06
Observations	3,559	3,559	3,559	3,559	3,559	3,559	2,444	2,444	2,444	2,444	2,444	2,444

Source: Statistics Sweden and CIRCLE data on inventors. Coefficient results and robust standard errors are reported. The dependent variable is the total number of patents applied for from 2003 to 2010. Inventors employed in 12, 13, 21, 24, 31, 32, 34, 41 and 82 occupations (SSYK) are included. These occupations are related to corporate managers (12), managers of small firms (13), scientists and engineers - professionals (21), business professionals and others (24), technicians in science and engineering (31), life science and health technicians (32), other associate professionals (34), office clerks (41) and plant and machine operators (82). The control variables are: age, age^2, female dummy variable, foreign-born (omit. Swedish-born), high-school grade, firm's patent portfolio, the number of inventors distributed by geographical area (No. inventors), the average productivity of inventors (Avg. prod.) by geographical area and industry of work, 5 dummies for technology fields of the patents (Tech.), 3 dummies for firm size (omit: small size) and 23 dummies for industry of work (Ind.). Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 11: **Negative binomial regressions on total number of forward citations (2003-2010)**

Variables	1	2	3	4	5	6	7	8	9	10	11	12
Dep: nr. of citations												
ORU vars.												
Req. edu	0.226***	0.195***	0.207***	0.190***	0.184***	0.162***	0.245***	0.189***	0.208***	0.187***	0.177***	0.154***
	(0.0249)	(0.0225)	(0.0227)	(0.0241)	(0.0226)	(0.0218)	(0.0374)	(0.0246)	(0.0249)	(0.0257)	(0.0243)	(0.0238)
Over-edu	0.102***	0.0887***	0.0897***	0.0881***	0.0882***	0.0898***	0.105***	0.0810***	0.0829***	0.0829***	0.0783***	0.0820***
	(0.0175)	(0.0169)	(0.0173)	(0.0175)	(0.0162)	(0.0166)	(0.0275)	(0.0190)	(0.0192)	(0.0201)	(0.0182)	(0.0187)
Under-edu	-0.0786***	-0.0613***	-0.0660***	-0.0609***	-0.0529***	-0.0472***	-0.0961***	-0.0432*	-0.0454*	-0.0476*	-0.0362	-0.0289
	(0.0154)	(0.0149)	(0.0152)	(0.0157)	(0.0147)	(0.0142)	(0.0352)	(0.0230)	(0.0239)	(0.0244)	(0.0225)	(0.0215)
Gender		-0.264***	-0.309***	-0.177**	-0.281***	-0.280***		-0.336***	-0.406***	-0.246***	-0.346***	-0.343***
		(0.0751)	(0.0744)	(0.0804)	(0.0727)	(0.0699)		(0.0823)	(0.0822)	(0.0900)	(0.0800)	(0.0775)
Age		7.278**	7.614***	9.108***	6.708**	5.946**		10.15*	10.59*	13.32**	9.733*	8.151
		(2.931)	(2.941)	(3.054)	(2.963)	(2.812)		(5.336)	(5.438)	(5.598)	(5.182)	(5.092)
Age^2		-0.983**	-1.032***	-1.225***	-0.908**	-0.801**		-1.424*	-1.489**	-1.850**	-1.363*	-1.145*
		(0.396)	(0.397)	(0.413)	(0.400)	(0.380)		(0.727)	(0.742)	(0.764)	(0.706)	(0.694)
Foreign-born		-0.0643	-0.0528	-0.0573	-0.0749	-0.0997		0.0450	0.0325	0.0265	0.0393	0.0611
		(0.0771)	(0.0780)	(0.0795)	(0.0756)	(0.0753)		(0.138)	(0.134)	(0.142)	(0.138)	(0.137)
High school grade							0.400	0.568***	0.572***	0.483**	0.593***	0.576***
							(0.299)	(0.193)	(0.200)	(0.198)	(0.186)	(0.179)
Patent Portfolio						0.163***						0.163***
						(0.0104)						(0.0123)
Firm size (omit: small size)												
Medium					0.198**						0.247***	
					(0.0780)						(0.0931)	
Large					0.557***						0.594***	
					(0.0748)						(0.0871)	
No. inventors (Area)		0.361**						0.111				
		(0.155)						(0.185)				
Avg. prod. (Area)			0.306**						0.243			
			(0.126)						(0.151)			
Avg. prod. (Industry)				-0.000177***						-0.000127**		
				(6.83e-05)						(5.26e-05)		
Other cont. vars		Ind. Tech.	Ind.	Tech.	Ind. Tech.	Ind. Tech.		Ind. Tech.	Ind.	Tech.	Ind. Tech.	Ind. Tech.
Constant	-2.415***	-15.82***	-17.52***	-18.52***	-14.97***	-13.12**	-4.989***	-23.49**	-25.46**	-28.35***	-23.28**	-19.81**
	(0.389)	(5.395)	(5.425)	(5.645)	(5.470)	(5.184)	(1.613)	(9.728)	(9.907)	(10.24)	(9.462)	(9.330)
Log pseudo lik.	-8213.69	-8091.93	-8120.29	-8161.44	-8046.27	-7966.32	-3567.43	-5641.39	-5665.4	-5702.34	-5606.4	-5100.29
Pseudo R2	0.02	0.03	0.03	0.03	0.04	0.05	0.02	0.04	0.04	0.03	0.05	0.05
Observations	3,559	3,559	3,559	3,559	3,559	3,559	1,479	2,444	2,444	2,444	2,444	2,444

Source: Statistics Sweden and CIRCLE data on inventors. Coefficient results and robust standard errors are reported. The dependent variable is the total number of patent counts adjusted for the citations received up to 5 years after publication (counted over all the period of analysis). Inventors employed in 12, 13, 21, 24, 31, 32, 34, 41 and 82 occupations (SSYK) are included. These occupations are related to corporate managers (12), managers of small firms (13), scientists and engineers - professionals (21), business professionals and others (24), technicians in science and engineering (31), life science and health technicians (32), other associate professionals (34), office clerks (41) and plant and machine operators (82). The control variables are: age, age^2, female dummy variable, foreign-born (omit. Swedish-born), high-school grade, firm's patent portfolio, the number of inventors distributed by geographical area (No. inventors), the average productivity of inventors (Avg. prod.) by geographical area and industry of work, 5 dummies for technology fields of the patents (Tech.), 3 dummies for firm size (omit: small size) and 23 dummies for industry of work (Ind.). Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.