ARTIGO **07** • **2023**

How does the increase of employee's qualifications relate to productivity? – An empirical assessment using firm-level data

Mónica Simões[1] and Sílvia Fonte Santa[2]

Abstract

The last decades have been characterized by a remarkable improvement of the Portuguese population' skills, following the implementation of a very comprehensive set of reforms and investments in education. However, empirical evidence of how these improvements in education relates to the productivity of firms has received a very small attention in the literature. As such, the main purpose of our paper is to contribute to the literature on the relationship between education and firms' productivity, focusing on workers with upper secondary education and workers with higher education.

The main results from our analysis suggest that an increase in workers' qualifications, regardless of type of education, is associated with higher firms' productivity (TFP). For all specifications, an increase of the share of workers with upper secondary education is associated with a larger increase to firms' productivity than a similar increase to the share of workers with only lower secondary education.

On a sector basis, we find that the effect is higher for the services sector when compared with manufacturing. We also find that increasing the workers with tertiary education, especially for the ones in Science and Technology occupations, is related to an even larger increase to TFP for firms in the services sector.

Keywords:* *Total Factor Productivity, Qualifications, Heterogeneity of labour force, Education reforms.

***JEL Classification*: J24.**

Gabinete de Planeamento, Estratégia, Avaliação e Relações Internacionais
Ministério das Finanças

Rua da Alfândega n.º 5A • 1100 – 016 Lisboa
www.gpeari.gov.pt

[1] GPEARI - Ministério das Finanças. The opinions expressed are those of the authors and not necessarily of the institutions. Any errors or omissions are the authors' responsibility.
[2] Banco de Portugal (the project was executed during the period the author was working at GPEARI).

1. Introduction

The last decades have been characterized by a remarkable improvement of the skill level of the Portuguese population, following the implementation of a very comprehensive set of reforms and investments in education (OECD, 2014 and OECD, 2020)[3]. One of the most relevant reforms was, in 2009[4], the extension of the age of compulsory education to 18 years old (or the conclusion of upper secondary degree). This decision occurred in tandem with several other reforms at the upper secondary level, especially regarding Vocational Education and Training (VET). For instance, in the past two decades there was a large increase in the offer of vocational education at public schools, the definition of a new quality assurance framework for VET courses, introduction of an academic credit System and several curriculum reforms aimed at improving the quality and the link with labour market' needs. During this period, the creation of the National Qualifications System (*Sistema Nacional de Qualificações*, SNQ, 2007) and the National Qualifications Framework (*Quadro Nacional de Qualificações*, QNQ, 2007) also took place.

At the same time, several programmes designed to promote educational success and reduce school failure and early school leaving[5] were also implemented, with solutions tailored to specific local needs. This includes the Education territory of Priority intervention, which was targeted at the most disadvantaged students. Additionally, in this period the introduction of universal free pre-primary education[6], as well as the implementation of programmes targeted at the adult population, such as the *Qualifica* programme also took place.

As a result of the implementation of this comprehensive set of reforms and investments, Portugal has been progressively reducing the level of low-skilled working age population from 81% in 2000 to 47.8% in 2019[7] (41% in 2021), in particular during the last decade. Following this, there was a large increase in the share of working age population with upper secondary and tertiary education, which reflects a more qualified labour force.

This improvement is most visible for the younger cohorts, as they are the target of most of the public policies. For instance, the share of young adults (aged 25 to 34 years old) with upper secondary education has increased by 19 p.p. in the past two decades (2001-2019), reducing the differential for the EU average to only 7 p.p. in 2019. Similarly, younger cohorts also display increasing shares of tertiary education. In 2001, only 14% of young adults had completed tertiary education, whereas in 2019 the share had increased to 37%, which is close to the European Union (EU) average (39%)[8]. Nevertheless, further progress regarding the skill level of the Portuguese population is still expected to occur in the coming years and decades, as the older workers, which are, on average, less qualified are replaced by younger and more qualified cohorts. This generational substitution effect, as well as additional measures to promote the upskilling of the adult population are still relevant in the current context, given that Portugal has the highest share of low skilled working age population amongst all countries in the EU due to the strong legacy effect.

From an economic perspective, these improvements to the skill level are expected to translate into higher productivity, and therefore higher potential growth. For instance, empirical evidence suggests a positive relationship between worker's education and the productivity of firms (Hægeland and Klette, 1997; Haltiwanger et al., 1999 and 2007; Hellerstein et al., 1999; Turcotte and Rennison, 2004; Galindo-Rueda and Haskle, 2005; Van Biesebroeck, 2011; Lebedinski and Vandenberghe, 2014; Rycx at al., 2015).

For Portugal, however, the evidence from the literature is scarce. Cima et al. (2022) find a positive and significant relationship between firm's

[3] A detailed description of the policies implemented in the last two decades, particularly in the most recent ones, can be found in OECD (2014), OECD (2020) and Capucha et al. (2021).

[4] The enrollment and attendance was only regulated in 2012.

[5] Examples of relevant programes: Programa de Combate ao Insucesso e Abandono escolar (2012), Programa Nacional de Promoção do Sucesso Escolar (2016), Programa dos Territórios Educativos de Intervenção Prioritária (1996, 2006, 2012). Measures to promote the autonomy of schools and the flexibility of the curriculum.

[6] For children with four years old or older.

[7] Data is also available for 2020 and 2021, however those figures might result from temporary effects associated with the Covid-19 pandemic.

[8] In 2021, the share of young adults with tertiary education has increased to 48%, corresponding to a value 6 p.p. higher than the EU average.

productivity and the average qualifications of workers, measured by a qualification index which is a composite indicator that incorporates information on worker's formal education (measured by years of schooling), age and unobserved ability, both in the manufacturing and services sectors. The authors also find that a worker's formal education is positive and significantly associated with firm's productivity, but the magnitude of the effect is very small when compared to the alternative specification. As in most countries, the literature for Portugal is focused mainly on wage returns, instead of productivity, and their potential spillovers (Portugal et al, 2018 and 2022; Martins and Jin, 2010).

The main purpose of our paper is, therefore, to contribute to the literature on the relationship between education and firms' productivity, focusing on the association with the increase of workers with upper secondary education and workers with higher education.

This paper is structured as follows. Section 2 reviews some of the relevant literature on the economic outcomes associated with skill improvements. Section 3 briefly describes the evolution of the qualification of Portuguese workers in the last decade. Section 4 describes the main data sources used in this paper, as well as the methodology used for our analysis. Section 5 presents some relevant statistics. Section 6 presents and discusses the empirical results. Section 7 describes the robustness analysis. Section 8 discusses some suggestions for future research. And finally, section 9 concludes.

2. Literature Review

The literature that extensively studies how different levels of education relates to wages find a positive and significant effect (Card, 1999, and Wilson and Briscoe, 2004, Martins and Jin 2010). Recent empirical evidence for Portugal also suggests a positive return on education of around 8.3 percent when following a simple Mincer type equation, and the presence of positive spillovers to their co-workers' wages for higher levels of education (Portugal et. al, 2022). However, some factors, such as heterogeneity in working conditions, labour adjustment costs, information asymmetry, social norms, discrimination, labour market regulations (minimum wages, collective bargaining, and unemployment benefits), among others, can cause a mismatch on the evolution of productivity and wages[9] (Rycx et al., 2015). As such, studying the impact on wages might be insufficient for a comprehensive assessment of the effects on productivity.

Another branch of the literature focuses on the macroeconomic impacts associated with rising education levels. Sianesi and Van Reenen (2003) through a literature review conclude that even though not all results are directly comparable, the evidence suggests that human capital increases the productivity of the economy instead of working solely as a signalling device of one's ability. However, the authors also identify some limitations regarding the need to consider differences in the quality of education across countries[10], reverse causality[11], level of countries' development, among others. Moreover, the length of available data can limit such analysis.

As such, some authors decide to study how differences to worker's education relates to the productivity of firms. This approach has the advantage of being able to take into account the heterogeneity in firm's productivity. Although the literature based on microdata is not extensive, it provides evidence suggesting a positive link between higher levels of education and productivity. For instance, some authors find that a higher proportion of workers that have attended university is positively associated with higher productivity of firms (Lebedinski and Vandenberghe, 2014, Rycx at al., 2015; Galindo-Rueda and Haskle, 2005; Hellerstein et al., 1999, Haltiwanger et al., 1999 and 2007, Turcotte and Rennison, 2004).

The evidence for lower levels of education is mixed. Some authors find that holding a secondary degree, as opposed to only having at most primary education, is positively and significantly associated to higher levels of productivity of firms (Haltiwanger et al., 1999 and 2007; Rycx at al., 2015), whereas other

[9] The human capital theory indicates that higher levels of education increase worker's productivity with positive effects on their wages (Becker, 1964). Further information on the Human Capital Theory is available in Fleischhauer (2007).

[10] Countries with the same average years of schooling, can have clear differences regarding the quality of their educational system.

[11] Faster growing countries invest more in education.

authors do not find the link to be significant (Lebedinski and Vandenberghe, 2014; Galindo-Rueda and Haskle, 2005). However, labour composition in terms of qualifications may change over time, as firms adjust them to better fit their need (Haltiwanger et al., 2007).

Using a slightly different approach, Hægeland and Klette (1997) look at how the productivity of Norwegian firms is associated with the level of qualifications of their employees while distinguishing these effects by experience and gender. Although they generally find a positive association between the two, for several of the specifications the results are not significant. In another perspective, Turcotte and Rennison (2004) also assess the interaction between human capital and technology use for Canadian firms and conclude that the largest productivity gains are associated with the combination of education, technology, and learning.

Some authors also use alternative measures of skills. Criscuolo et al. (2021), consider skill shares based on information from PIAAC, an OECD Survey of Adult Skills, which allows ranking the skills of the workers based on their occupation. The authors find evidence of a positive and significant effect on productivity associated with a higher share of skilled workers. Alternative measures of skills, such as the ones based on PIAAC test scores, might capture differences in both the quantity and quality of education, with evidence suggesting that later can be an important determinant of productivity dynamics (Égert et al., 2022).

For Portugal, Cima et al. (2022) find a positive and significant relationship between firms' productivity and the average qualifications of workers, measured by a qualification index which incorporates information on worker's formal education (measured by years of schooling), age and unobserved ability, both in the manufacturing and services sectors. The authors also find that a workers' formal education is positively and significantly associated with firms' productivity, but the magnitude of the effect is very small when compared to their preferred specification.

Some authors also analyse if the positive relationship between human capital and productivity can help explaining the large heterogeneity in firm's productivity that is observed even for narrowly defined sectors as shown in Syverson (2011). However, the evidence is not conclusive. For instance, Haltiwanger et al. (1999) find that labour composition (education, gender, age, foreign) helps to explain a large part of the differences in productivity across firms, but they also find that these characteristics are not good predictors of productivity growth. On the other hand, Galindo-Rueda and Haskle (2005) find that qualifications only explain a small portion of the productivity differences across firms, and Fox and Smeets (2011) suggest that using the level of education, as well, as additional measures for human capital quality, that can work as a proxy for ability, can better explain the differences in productivity. Similarly, Criscuolo et al. (2021) study how differences related to "Human Side of firms" can explain the productivity gap between the frontier and the median productivity firms. Amongst other factors, such as diversity of workers and managerial quantity and quality, they conclude that differences to the composition of workers account for 19% of the gap (with 6% associated with simultaneously increasing managers and workers' skills).

Another relevant point relates to the methodologies and theoretical approaches used to estimate these results. Most authors choose to use a production function approach with heterogeneous workforce for the analyses, with the latter being expressed as the quality of labour input. This quality of labour input can be further decomposed into different labour inputs, as to reflect different levels of education or experience of the labour force, amongst others. The different inputs are expressed as a share of total labour, and their effect on productivity can be directly estimated in the regressions using productivity as the dependent variable (Hægeland and Klette, 1997; Hellerstein et al., 1999; Galindo-Rueda and Haskle, 2005; Lebedinski and Vandenberghe, 2014; Rycx at al., 2015).

Adding to this, econometric models used vary greatly across the literature. On the one hand, some authors use simpler techniques such as maximum likelihood (Hægeland and Klette, 1997; Galindo-Rueda and Haskle, 2005), linear least squares methods, such as OLS (Haltiwanger et al., 1999; Turcotte and Rennison, 2004). On the other hand, other authors try alternative methods as to minimize potential heterogeneity bias and/or simultaneity bias. One of these approaches is the use of a fixed effects model in order to remove some of the potential heterogeneity bias (Lebedinski and Vandenberghe, 2014; Cima et al. 2022). Lebedinski and Vandenberghe (2014) also use an extension of the Ackerberg, Caves, and Frazer's intermediate approach, which incorporates fixed effects, as well as an alternative a system-GMM, which uses lagged labour inputs in levels and in first-differences to try to minimize both the heterogeneity and simultaneity bias. Rycx at al. (2015) also use a system-GMM in one of their specifications, and the Levinsohn and Petrin approach in the other.

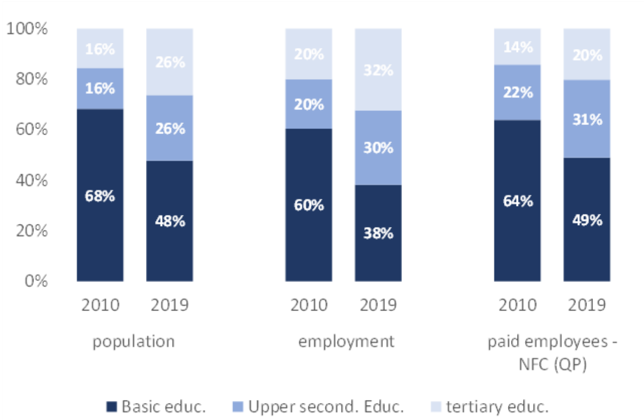
Most authors also control for several other important factors associated with productivity, with the most common being the share of women, share of part-time workers, share of workers with open ended contracts, and capital intensity[12]. Other controls include the share of workers by age groups (younger/older cohorts) or experience, share of foreign-born workers, share of workers with blue-collar jobs, firm size, industry and/or year fixed effects, amongst others.

3. Evolution of the qualifications of Portuguese workers in the period 2010-2019

For the past two decades, the improvement of Portuguese qualifications was under the agenda in order to address the overall low skill level of the Portuguese population as described in section 1. Therefore, we start to look to the main achievements of those policies across time.

Figure 1 pictures the evolution of the Portuguese qualifications for the working age population, total employment, and paid employees from non-financial corporations (NFC) for the period between 2010 and 2019. The improvement in the education levels is common across the groups considered and translates to a decrease in the share of individuals with the lowest levels of skills (lower secondary education or less), as well as an increase to both the share of individuals with upper secondary and tertiary education. However, there are also some important differences. For instance, the decrease in the shares of low skilled and the increase of those with tertiary education is lower for the employees from NFC when compared to total population and total employment. One possible explanation for this difference is that part of the workers with a higher level of qualifications are working in the public sector or are self-employed.

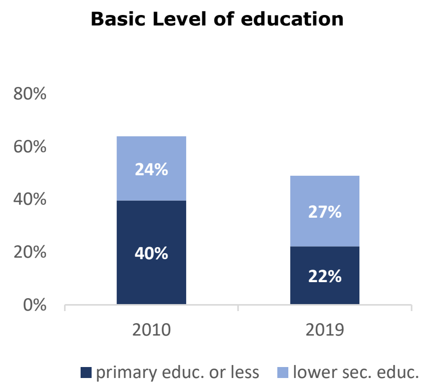
Figure 1 –Population and employment aged 25-64 years old by level of education – 2010 and 2019 (in percentage)



Source: own calculations using Eurostat (population, total employment) and Quadros de Pessoal (paid employees – NFC).

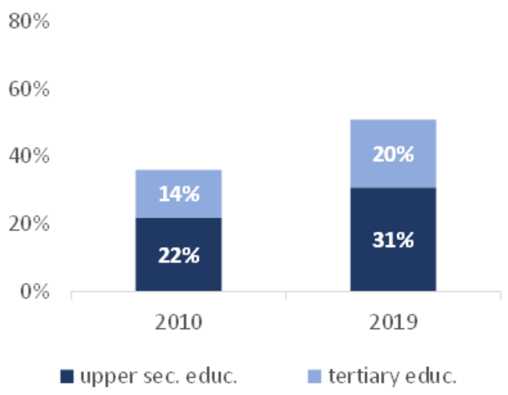
Additionally, Figure 2 shows that the decrease of the share of workers with lower secondary education or less is mostly driven by a drop in the share of the workers with only primary education or less – the group with the lowest level of skills.

Figure 2 –Qualifications of NFC's employees aged 25-64 years old, between 2010 and 2019 (in percentage)



[12] There is an extensive literature that studies the relationship between workers characteristics, such as age and gender and productivity (Hellerstein et al, 1999; Hellerstein and Neumark, 2007; Vanderberghe 2011a, 2011b).

Intermediate to Advanced level of education[13]

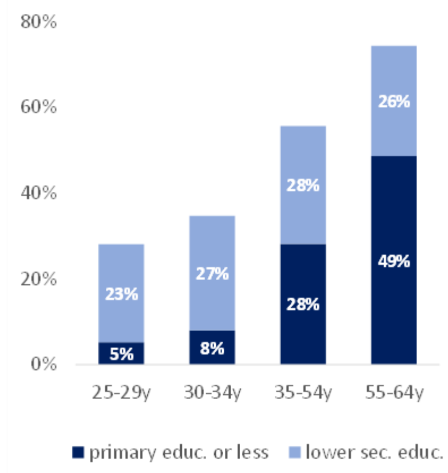


Source: own calculations using QP.

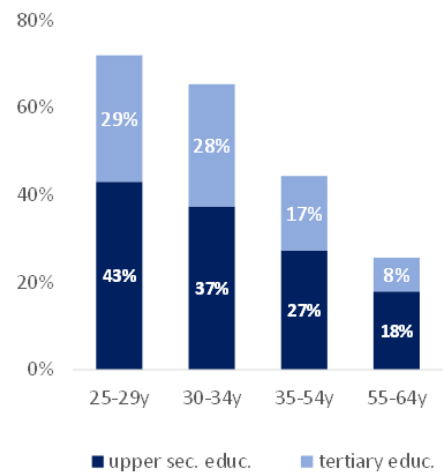
This large drop in the share of workers with primary education or less is mostly driven by the fact that less qualified cohorts progressively exiting the labour market are being replaced by younger and more qualified cohorts (Figure 3). As such, this intergenerational substitution effect is also responsible for a large increase in the share of workers with upper secondary education and tertiary education. The effects from upgrades to education level of those already in the labour market, changes to employment status, and shifts between sectors (public, private and self-employment) may also play a role explaining the dynamics between the two periods.

Figure 3 – Qualifications of NFC's employees by age group, 2019 (in percentage)

Basic Level of education



Intermediate to Advanced level of education



Source: own calculations using QP.

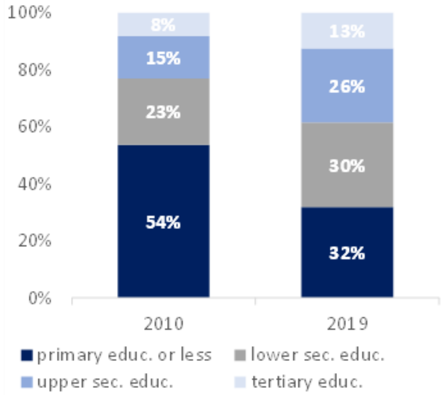
It is also relevant to understand the heterogeneity of these improvements across sectors of activity. In fact, Figure 4 shows important differences between manufacturing and services sector. Although both sectors show a similar decrease in the share of the workers with lower levels of education, the increase in the share of workers with upper secondary education is significantly more relevant in

[13] Basic level of education includes workers with lower secondary education or lower. Intermediate level of education includes workers with upper secondary education, and the Advanced level of education includes workers with tertiary education.

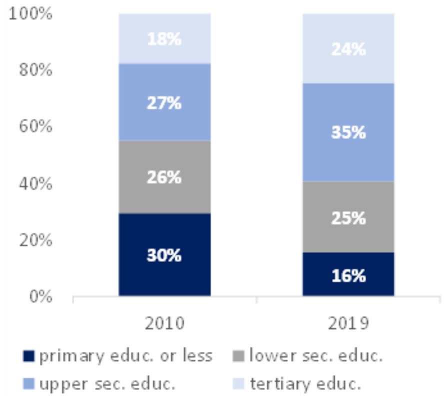
manufacturing (+11 p.p. between 2010 and 2019), when compared to the increase in the share of workers with tertiary education (+4 p.p.). In the services sector, the increase of the middle and advanced qualifications is more even (with an increase of 7 p.p. both in upper secondary and tertiary education shares). These differences also mirror, in part, distinct skill composition across sectors. For instance, the distribution of workers with different skills across these two sectors is not homogenous. Nevertheless, despite the year under analysis the manufacturing sector presents a higher share of workers without an upper secondary education degree (+21 p.p. when compared to the services sector in 2019), specifically a higher share of people with only primary education or less. Moreover, not only the shares of workers with upper secondary education and tertiary education are lower relative to services, they are also more concentrated in the first. This heterogeneity points to the importance to analyse these two sectors separately.

Figure 4 – Qualifications of NFC's employees between 2010 and 2019 (in percentage)

Manufacturing



Services



Source: own calculations using QP.

4. Data and methodology

In our analysis we use *Quadros de Pessoal* (QP), which is an employer-employee matched database. This census is conducted by the Portuguese Ministry of Employment, Solidarity and Social Security (MTSSS) each October to firms from the private sector with at least one employee (excludes public sector, self-employed and domestic servants). As it is a mandatory survey, this database includes information for most of the Portuguese employees working in the private sector.

The information at the firm-level from QP is complemented with information from *Sistema de Contas Integradas das Empresas* (SCIE), a database that includes economic and financial information on a firm-by-firm basis for all the Portuguese NFC reported in *Informação Empresarial Simplificada* (IES). The firm identifier is the same in both databases, which allows the merge of both databases. QP and SCIE are both made available to researchers by Statistics Portugal (INE).

This analysis uses information at the firm-level for the period 2010-2019 (the relevant period for the estimation is 2011-2019[14], as lags of the control variables are used). In terms of sector of activity, we consider firms from the Manufacturing (NACE 10-33), Utilities and Construction sectors (NACE 35-43), and Services sector (NACE 45-99), but excluding Financial and Insurance Activities (NACE 64-66), Public administration and defence (NACE 84), Activities of

[14] Information for the year 2020 for the QP and SCIE databases is already available; however, we have decided not to include it, as the data for the relevant variables may include effects related to the Covid-19 pandemic.

households as employers (NACE 97-98) and Activities of extraterritorial organizations and bodies (NACE 99). Firms with less than five employees and/or with negative values of gross value added (GVA) are excluded from the analysis due to restrictions of the methodology used[15]. Firms with negative or missing values for assets, turnover, number of employees and total remunerations, and negative values for Fixed Assets and Intangible Assets are also excluded. Moreover, to minimize error and noise we follow Criscuolo *et al.* (2021) and remove the entire firm if the annual growth rates of the firm's productivity is in the bottom or upper percentile of the distribution in one year[16]. After these adjustments and excluding observations for which no lag can be computed, the total sample size is 555 737 observations, corresponding to 111 599 firms over the period 2011-2019.

As the core of the analysis is performed at the firm-level, information on the employees (corresponding to 14 924 070 observations) is aggregated into ratios or averages. For this purpose, only data about paid workers is considered, and workers with missing or inconsistent information that could not be corrected are excluded.[17]

Methodology

In terms of methodology, we use a fixed effects model. The model follows the specification in equation (1), with the natural logarithm of Total Factor Productivity (TFP) for each firm *i* in period *t* as the dependent variable (y_{it}).

Total Factor Productivity (TFP) at the firm-level is calculated following the methodology proposed by Wooldridge (2009), which is implemented using the Stata package *prodest* (Rovigatti and Mollisi, 2016)[18]. This calculation takes place before the estimation of the main model.

The fixed effects model has the advantage of controlling for time invariant firm characteristics that can influence productivity and the explanatory variables. Indeed, the Hausman specification test confirms that the fixed effects model is more appropriate for our specification, when compared to random effects model[19].

$$y_{it} = \alpha + \beta' SQ_{kit-1} + \delta' X_{it-1} + \gamma C_t + +\tau_j + \theta_i + \varepsilon_{it} \quad (1)$$

The main explanatory variables (SQ_{kit-1}) are the share of workers by type of qualification *k*, for each firm *i* in period *t-1* (previous period), as specified in equation 2. The following shares of workers are considered: i) the share of workers with less than lower secondary education; ii) the share of workers with lower secondary education iii) the share of workers with upper secondary education; and iv) the share of workers with tertiary education[20].

$$SQ_{kit-1} = \frac{number\ of\ workers\ with\ qualtification\ k\ in\ firm\ i\ and\ period\ t-1}{total\ number\ of\ workers\ in\ firm\ i\ and\ period\ t-1} \quad (2)$$

We also include important control variables for each firm *i* in period *t-1* (X_{it-1}), namely tenure, and young workers, female workers and part-time workers as a percentage of total workers.

Tenure is defined as a quadratic function, using the average of the paid workers' tenure in firm *i*. The share of young workers corresponds to the ratio of workers with 35 or less years of age in firm *i* in period *t-1*, over the total number of workers in the same firm and period. Similarly, the shares of female and part-time workers correspond to the ratio of female and part-time workers respectively, over the total number of workers, in firm *i* in period *t-1*.

[15] The natural logarithm of Gross Value Added (GVA) is used to compute productivity measures, which excludes the possibility of using non-positive values. Furthermore, using firms with less than five employees would significantly increase the number of firms with null shares of employees with certain qualifications, and subsequently null variations of those shares.

[16] To reduce the risk of reporting errors and removes firms that had extraordinary revenues, or that underwent a very large restructuring.

[17] For more information, please refer to Appendix 1.

[18] In the method proposed by Wooldridge a production function is estimated in a single step GMM framework, with the intermediate inputs being used as a proxy for the unobserved productivity as in Levinsohn-Petrin (2003). For the estimation in Stata, we have used the package *prodest,* where TFP was computed using the Wooldridge methodology. We have used the natural logarithm of value added as the dependent variable, and the number of paid employees, intermediate consumption (sum of external services and supplies, and cost of goods sold, and materials consumed), and capital as inputs. Value added was deflated using a GVA deflator with an A82 sector classification from National Accounts and Capital was deflated using the Gross Capital Formation deflator from the National Accounts. The production function was estimated separately for 32 different sectors following OECD's STAN A38 industry classification (Table A.1., Appendix 2).

[19] Test results available in Appendix 2 (Table A.2)

[20] Usually the lower secondary level of education corresponds to nine years of education, the upper secondary level of education corresponds to twelve years, and tertiary education degrees to more than fifteen years of schooling. These values assume there is no repetitions of an academic year.

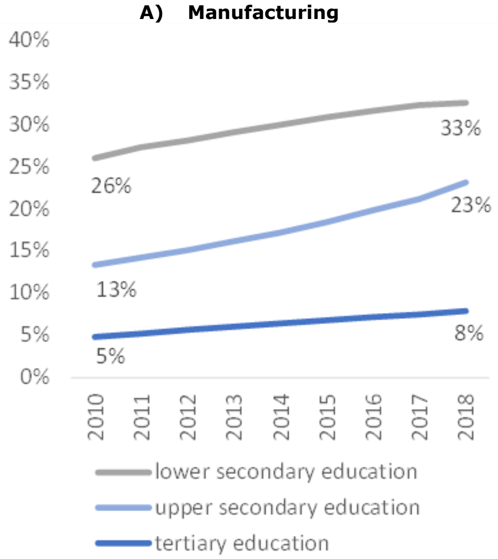
Additional controls for the economic cycle (C_t), using the estimates of the output gap (OG) for period t calculated by the European Commission, and for the firm size τ_j - micro, small, medium, and large firms - are also included. A micro firm is defined as a firm with less than 10 workers, a small firm as a firm with 10 to 49 employees, a medium firm as a firm with 50 to 249 workers, and lastly a large firm as a firm with 250 or more workers. We also use robust standard errors clustered at firm-level.

5. Descriptive Statistics

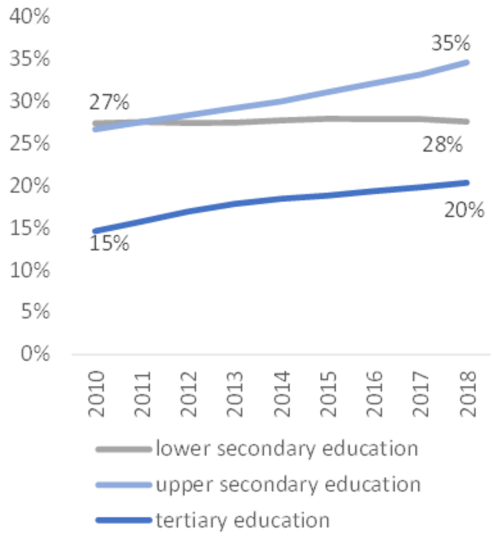
In line with section 3, we observe that in our sample, there is, on average, a consistent increase in the weight of the more qualified workers - workers with upper secondary or tertiary education over the years (Figure 5).

Figure 5 – Average firm's shares of workers by level of education in 2010-2018 in Manufacturing (A) and Services (B) (percentage, %)

A) Manufacturing



B) Services



Source: own calculations using QP and SCIE.

Note: Information for the total economy is available in Appendix 2 (Graph A.3.1.).

Moreover, the increase in the share of workers with upper secondary education is, on average, larger in manufacturing firms, especially for those firms with higher levels of productivity. While in the services sector a catching up process seems to occur as the increase is higher for firms with lower levels of productivity (Table A.3.1, Appendix 3).

Looking to the share of workers with tertiary education we have a different scenario, as the increase is higher in the services sector and for the most productive firms regardless of the sector (Table A.3.1, Appendix 3).

As for the workers with only lower secondary education, we see that their share on total workers has increased, on average, in the manufacturing sector, while it has remained relatively stable in the services sector.

These dynamics reflect the distribution of skills per sector, with firms in manufacturing having, on average, lower shares of workers with upper secondary education and tertiary education when compared with the ones from the services sector (Table 1). Moreover, the results also show that, on average, higher shares of workers with upper secondary education can be found in firms with a higher level of productivity from both sectors (Table A.3.2, Appendix 3). The same occurs for the shares of tertiary education. In the opposite direction, higher shares of workers with lower levels of education, such as lower secondary education, are, on average, associated with lower productivity firms.

Table 1 –Firm's shares of workers by level of education and by other relevant firm characteristics in the period 2010-2018 – mean and standard deviation

	Manufacturing (1)	Services (2)	Total (3)
Share workers w/ lower sec. education	**0.297**	**0.276**	**0.282**
	(0.255)	**(0.230)**	**(0.250)**
Share workers w/ upper secundary/post-sec. Education	**0.177**	**0.304**	**0.252**
	(0.191)	**(0.189)**	**(0.248)**
Share workers w/ tertiary education	**0.064**	**0.180**	**0.141**
	(0.130)	**(0.110)**	**(0.222)**
Average Tenure	8.409	6.554	6.881
	(4.545)	(5.926)	(5.402)
Share of Young workers, age ≤ 35 y.	0.312	0.370	0.346
	(0.225)	(0.213)	(0.244)
Share female workers	0.405	0.492	0.420
	(0.134)	(0.318)	(0.323)
Share part-time workers	0.012	0.037	0.029
	(0.074)	(0.054)	(0.098)
N. Observations	131729	350571	555737

Source: own calculations using QP and SCIE.

Note: Standard deviation reported in parentheses.

As for other indicators that will be considered under this analysis, the results indicate that tenure has, on average, increased for firms in both sectors, while the share of young workers has decreased. These dynamics are observed across all productivity levels, and in both sectors (Table A.3.1, Appendix 3). On the other hand, the share of female workers and part-time workers has, on average, remained more stable across time.

Despite this, there are important differences in terms of levels, with firms in manufacturing sector showing, on average, higher tenure, and lower shares of young, female and part-time' workers relative to the firms in the services sector (Table 1). Nonetheless, in both sectors, the most productive firms have, on average, workers with a higher tenure, as well as lower shares of female, and part-time workers (Table A.3.2, Appendix 3).

6. Empirical results and discussion

This section describes the results based on equation 1 by different levels of qualification. As expected, the results in Table 2 suggest that increasing the share of qualified workers in a firm (regardless of the type of education) is, on average, associated with a positive and significant effect in terms of TFP. At a sectoral level, the effect is, on average, positive and significant for both manufacturing and services and larger for the services sector (column 2).

As a large set of structural reform in the past was aimed to increase the number of young adults graduating from upper secondary education (e.g. the extension of the age of compulsory education to 18 years old or the conclusion of upper secondary degree), it is also important to understand the relationship between larger shares of upper secondary workers and productivity.

In fact, the results show that the effect on productivity of a larger share of upper secondary workers is, in fact, positive and significantly larger than the one associated with increasing the share of workers with only lower secondary education, with this difference being positive and significant for all specifications. Specifically, on average, a 1 p.p. increase of the share of workers with lower secondary education is associated with an average increase to firm's productivity between 0.104% and 0.133%, which compares with the effect of 0.187% to 0.223% for the share of workers with upper secondary education[21].

Another focus of analysis is the large set of reforms and policies aiming at supporting the graduation from tertiary educational level implemented in the last years. Indeed, the results show a positive relation between productivity and the share of these higher qualified workers, with a 1 p.p. increase in this share being associated with an average increase of firm's productivity between 0.190% and 0.267% (that compares with 0.187% and 0.223% for the share of workers with upper secondary education). However, the difference is only significant for services sector and total economy.

[21] In these specifications the share of workers with less that lower secondary education is the omitted share of qualifications

Table 2 – **Regression results for the main group of qualifications.**

		Manufacturing (1)	Services (2)	Total (3)
Share workers w/ lower sec. education (lag) - [S1]		**0.131*** (0.011)**	**0.133*** (0.009)**	**0.104*** (0.006)**
Share workers w/ upper secundary/post-sec. education (lag) - [S2]		**0.188*** (0.014)**	**0.223*** (0.009)**	**0.187*** (0.007)**
Share workers w/ tertiary education (lag) - [S3]		**0.190*** (0.029)**	**0.267*** (0.013)**	**0.221*** (0.010)**
Average Tenure (lag)		0.011*** (0.001)	0.015*** (0.001)	0.011*** (0.001)
Average Tenure^2 (lag)		-0.000*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)
Share of Young workers, age ≤ 35 y. (lag)		-0.076*** (0.011)	-0.045*** (0.006)	-0.063*** (0.005)
Share female workers (lag)		-0.081*** (0.017)	-0.050*** (0.009)	-0.069*** (0.008)
Share part-time workers (lag)		-0.110*** (0.037)	-0.083*** (0.011)	-0.094** (0.010)
Economic cycle		0.005*** (0.000)	0.015*** (0.000)	0.011*** (0.000)
Firm FE		Yes	Yes	Yes
Firm size FE		Yes	Yes	Yes
R^2		0.40	0.30	0.22
N. Observations		131729	350571	555737
Tests of hypothesis				
S1 = S2	F-statistic	17.15	119.42	150.64
	Prob>F	Rejected at 1%	Rejected at 1%	Rejected at 1%
S2 = S3	F-statistic	0.01	17.37	12.30
	Prob>F	Not Rejected	Rejected at 1%	Rejected at 1%

Source: own calculations using SCIE, QP, EC.

Notes: Dependent variable is the natural logarithm of total factor productivity. We use robust standard errors clustered at firm-level. Standard errors are reported in parentheses. Symbols *, **, and *** denote significance at 10%, 5%, and at 1% respectively.

Other variables included as controls also give some information that can be useful for the analysis. For instance, a higher level of average worker's tenure in a firm is associated with a positive and significant effect on productivity, although being non-linear (the quadratic term is negative, but close to zero). On the other hand, higher shares of young workers, female workers or part-time workers are, on average, associated with a negative effect to the firm's productivity. This outcome for control variables remains stable across the different specifications.

Econometric Tests

To ensure our model is correctly specified we performed several tests. For instance, we perform the variance inflation factor (vif) test, which measures the correlation and strength of correlation between the explanatory variables in a regression model. This test indicates that our specification does not have multicollinearity' issues. Our tests also indicate that is not possible to assume homoscedasticity, and for this reason we use in all the regressions robust standard errors which are clustered at the firm-level. Lastly, as it is common in this type of analysis we cannot reject the hypothesis of omitted variable bias, as there is a large set of factors that can influence productivity. We try to reduce this bias by using firm fixed effects to capture all those factors that are constant in time and that can influence the level of productivity of a firm. We also use lagged dependent variables to reduce the possibility of a simultaneity bias.

7. Robustness Analysis

To assess the robustness of our results we perform other types of assessment.

Firstly, we consider other measures of productivity like labour productivity and TFP computed using a different method.

Secondly, we explore other definitions for the main explanatory variables – the shares of workers by qualification – by considering other groups of qualification that encompass VET programmes; STEM courses[22], Scientific & engineers and the definition under QUEST III R&D model. Moreover, we include an alternative indicator to measure qualifications, which considers the mean years of schooling.

Lastly, we also include an additional variable that measures training costs supported by firms.

7.1. Alternative productivity measures

For this robustness check, we consider: i) labour productivity, which corresponds to the natural logarithm of GVA per paid employee[23], and ii) TFP calculated using the Levinsohn and Petrin (2003) methodology[24].

Table 3 – Regression results using an alternative productivity measure – labour productivity.

		Industry (1)	Services (2)	Total (3)
Share workers w/ lower sec. education (lag) - [S1]		**0.107*****	**0.106*****	**0.085*****
		(0.011)	**(0.009)**	**(0.006)**
Share workers w/ upper secundary/post-sec. education (lag) - [S2]		**0.144*****	**0.161*****	**0.139*****
		(0.014)	**(0.009)**	**(0.007)**
Share workers w/ tertiary education (lag) - [S3]		**0.145*****	**0.192*****	**0.164*****
		(0.029)	**(0.012)**	**(0.010)**
Average Tenure (lag)		0.015***	0.019***	0.015***
		(0.001)	(0.001)	(0.001)
Average Tenure^2 (lag)		-0.000***	-0.000***	-0.000***
		(0.000)	(0.000)	(0.000)
Share of Young workers, age ≤ 35 y. (lag)		-0.048***	-0.026***	-0.043***
		(0.012)	(0.006)	(0.005)
Share female workers (lag)		-0.082***	-0.041***	-0.058***
		(0.017)	(0.009)	(0.008)
Log Capital per worker, deflated (lag)		0.026***	0.021***	0.023***
		(0.002)	(0.001)	(0.001)
Share part-time workers (lag)		-0.101**	-0.088***	-0.096***
		(0.040)	(0.011)	(0.011)
Economic cycle		0.002***	0.011***	0.007***
		(0.000)	(0.000)	(0.000)
Firm FE		Yes	Yes	Yes
Firm size FE		Yes	Yes	Yes
R^2		0.31	0.20	0.24
N. Observations		131729	350571	555737
Tests of hypothesis				
S1 = S2	F-statistic	7.13	48.22	68.12
	Prob>F	Rejected at 1%	Rejected at 1%	Rejected at 1%
S2 = S3	F-statistic	0.00	9.49	7.08
	Prob>F	Not Rejected	Rejected at 1%	Rejected at 1%

Source: own calculations using SCIE, QP, EC.

Notes: Dependent variable is the natural logarithm of GVA per paid worker. We use robust standard errors clustered at firm-level. Standard errors are reported in parentheses. Symbols *, **, and *** denote significance at 10%, 5%, and at 1% respectively. Tests of hypothesis refer to test on the equality of coefficients.

The results for labour productivity (Table 3) show that the coefficients related to qualifications have, on average, a lower magnitude when compared to those using TFP as a measure of productivity. Moreover, the results for the alternative measure of TFP (Table A.3.4., Appendix 3) display a similar magnitude confirming our main conclusions.

7.2. Workers' Qualifications

To stress the fact that different levels of qualifications should be considered we run our regression for four different groups of qualifications, which are summarized in the descriptive statistics under Table 4[25].

Group 1 breaks upper secondary education between general education and vocational education.

Group 2 focuses on the effects associated with human resources with tertiary education on core science and technology fields of study (STEM fields of study). As such, the share of workers with tertiary education is divided in two – share of workers with tertiary education in non-STEM fields of study and the share of workers with tertiary education in STEM fields of study.

Group 3 looks at the effects associated with human resources with tertiary qualifications that are actually employed in core Science and Technology occupations (Scientists and Engineers)[26]. Eurostat describes these human resources as "those people who conduct research, improve or develop concepts, theories and operational methods and/or apply scientific knowledge relating to fields which are covered by one of these occupations".

Group 4 considers the share of low skilled, medium skilled, and high skilled workers following the

[22] Includes natural sciences, mathematics, statistics, information and communication technologies, and engineering, manufacturing and construction field of study.

[23] GVA is deflated using the GVA deflator with an A82 sector classification from the National Accounts.

[24] Implemented in Stata using the package *prodest*. The Levinsohn and Petrin (2003) proposed methodology uses a Control Function approach with a two-step procedure with intermediate goods used to proxy productivity to address the simultaneity and selection bias (unbalanced panel). It uses a similar approach as Olley and Pakes (1996), but it uses materials instead of investment as the proxy variable. In this methodology, it is assumed that the materials' function is invertible, and that productivity is the only unobserved state variable.

[25] Additional statistics regarding growth rates, and information on the proportion of firms with null shares are available in Appendix 3 (Tables A.3.5. and A.3.6).

[26] Includes Science and engineering Professionals (ISCO-08 code 21), Health professionals (ISCO-08 code 22) and Information and Communication technology professionals (ISCO-08 code 25).

definition of the skill groups included in the QUEST III R&D model (Roger et al., 2008)[27].

Table 4 – Firm's shares of workers by subsets of the initial levels of qualifications – mean and standard deviation.

	Manufacturing (1)	Services (2)	Total (3)
Share workers w/ upper/post-secundary education (lag), by programme:			
general programme	0.126	0.226	0.252
	(0.165)	(0.242)	(0.248)
VET programme	0.050	0.078	0.141
	(0.107)	(0.157)	(0.222)
Share workers w/ tertiary education, by fields of study			
non-STEM fields study	0.037	0.136	0.098
	(0.077)	(0.214)	(0.183)
STEM fields study	0.028	0.044	0.044
	(0.068)	(0.135)	(0.119)
Share workers w/ tertiary education, by occupation			
other than Scientists & Engineers	0.048	0.130	0.098
	(0.089)	(0.205)	(0.177)
Scientists & Engineers	0.016	0.050	0.043
	(0.051)	(0.149)	(0.127)
Share workers under QUEST III R&D model			
medium-skilled workers	0.224	0.434	0.350
	(0.221)	(0.322)	(0.309)
high-skilled workers	0.016	0.050	0.043
	(0.051)	(0.149)	(0.127)
N. Observations	131729	350571	555737

Source: own calculations using QP and SCIE.

Note: Standard deviation reported in parentheses.

7.2.1. Vocational Education (Group 1)

As Vocational Education and Training (VET) were the focus of important policies in the past two decades, as to improve the offer and quality of VET courses (e.g. increasing the offer at public schools and implementing newer quality assurance framework for VET courses), it might be relevant to distinguish these types of programmes in our analysis.

For these purpose we divide the share of workers with upper secondary education into two groups: i) the share of workers with upper secondary education that have followed a general track programme, and ii) the share of workers with upper secondary education that have followed a VET programme. Table 5 reports the results of this new group of qualifications (Group 1).

Table 5 – Regression results for alternative groups of qualifications - group 1.

		Manufacturing (1)	Services (2)	Total (3)
Share workers w/ lower sec. education (lag) - [S1]		**0.131***** **(0.011)**	**0.133***** **(0.009)**	**0.104***** **(0.006)**
Share workers w/ upper secundary/post-sec. education (lag)				
by programme:				
general programme (lag) - [S2.1]		**0.193***** **(0.015)**	**0.225***** **(0.010)**	**0.191***** **(0.008)**
VET programme (lag) - [S2.2]		**0.175***** **(0.021)**	**0.216***** **(0.012)**	**0.176***** **(0.010)**
Share workers w/ tertiary education (lag) - [S3]		**0.190***** **(0.029)**	**0.267***** **(0.012)**	**0.221***** **(0.011)**
Average Tenure (lag)		0.011*** (0.001)	0.016*** (0.001)	0.011*** (0.001)
Average Tenure^2 (lag)		-0.000*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)
Share of Young workers, age ≤ 35 y. (lag)		-0.076*** (0.011)	-0.045*** (0.006)	-0.064*** (0.005)
Share female workers (lag)		-0.081*** (0.017)	-0.050*** (0.009)	-0.069*** (0.008)
Share part-time workers (lag)		-0.110*** (0.037)	-0.083*** (0.011)	-0.094*** (0.010)
Economic cycle		0.005*** (0.000)	0.015*** (0.000)	0.011*** (0.000)
Firm FE		Yes	Yes	Yes
Firm size FE		Yes	Yes	Yes
R^2		0.40	0.30	0.22
N. Observations		131729	350571	555737
Tests of hypothesis				
S1 = S2.1	F-statistic	16.71	115.57	149.43
	Prob>F	Rejected at 1%	Rejected at 1%	Rejected at 1%
S1 = S2.2	F-statistic	4.43	56.76	59.65
	Prob>F	Rejected at 5%	Rejected at 1%	Rejected at 1%
S2.1 = S2.2	F-statistic	0.76	0.92	3.10
	Prob>F	Not Rejected	Not Rejected	Not Rejected
S2.1 = S3	F-statistic	0.00	14.36	8.78
	Prob>F	Not Rejected	Rejected at 1%	Rejected at 1%
S2.2 = S3	F-statistic	0.19	16.08	15.22
	Prob>F	Not Rejected	Rejected at 1%	Rejected at 1%

Source: own calculations using SCIE, QP, EC.

Notes: Dependent variable is the natural logarithm of total factor productivity. We use robust standard errors clustered at firm-level. Standard errors are reported in parentheses. Symbols *, **, and *** denote significance at 10%, 5%, and at 1% respectively. Tests of hypothesis refer to test on the equality of coefficients.

Our results confirm that larger shares of both types of workers with upper secondary education (general programme vs. vocational programme) have, on average, a positive and significant association with productivity. The results also indicate that the effect on productivity associated with an increase of a 1 p.p. in the share of workers with upper secondary education remains significantly larger relative to the effect associate to an increase of the same magnitude in lower secondary education, for both types of programmes.

Furthermore, the difference in the effect associated to the general programme group relative to the vocational group of workers is not significant for any of the specifications. However, it is important to note that the results may reflect the large number of firms with a null share of workers with vocation education (65% - 66%) when compared to general education

[27] In QUEST III R&D model, low skilled workers include workers with lower secondary education or less, medium skilled workers include workers with upper secondary education, and tertiary education in non-STEM fields, and high skilled workers include workers with tertiary education in STEM occupations. The calibration of the model assumes that workers with STEM tertiary education are employed in core Science and Technology occupations (Scientists and Engineers).

(Table A.3.6., Appendix 3). Moreover, as previously mentioned, these results may not reflect or only partially reflect the effects of the recent policies implemented to these types of programmes. As the newer cohorts of the population, that have benefited from the most recent reforms to the VET programmes, enter the labour market and gradually increase their share on total workers, it will be easier to assess their effects on key indicators such as firm's productivity. Moreover, potential differences in the employability between the two groups are not considered in this analysis.

7.2.2. Tertiary education by field of studies: STEM education vs. non-STEM education (Group 2)

We also distinguish two subgroups of workers with tertiary education, those who have a degree in core Science and Technology fields of study (STEM field of study) and those who do not (group 2). Table 6 includes the results of such decomposition.

Table 6 – Regression results for alternative groups of qualifications - group 2.

		Manufacturing	Services	Total
		(1)	(2)	(3)
Share workers w/ lower sec. education (lag) - [S1]		**0.131*****	**0.133*****	**0.104*****
		(0.011)	**(0.009)**	**(0.006)**
Share workers w/ upper secundary/post-sec. education (lag) - [S2]		**0.188*****	**0.223*****	**0.188*****
		(0.014)	**(0.009)**	**(0.007)**
Share workers w/ tertiary education (lag) - [S3]				
by fiels of study:				
non-STEM fields study (lag) - [S3.1]		**0.216*****	**0.267*****	**0.238*****
		(0.035)	**(0.013)**	**(0.011)**
STEM fields study (lag) - [S3.1]		**0.152*****	**0.266*****	**0.179*****
		(0.045)	**(0.019)**	**(0.016)**
Average Tenure (lag)		0.011***	0.016***	0.011***
		(0.001)	(0.001)	(0.001)
Average Tenure^2 (lag)		-0.000***	-0.001***	-0.000***
		(0.000)	(0.000)	(0.000)
Share of Young workers, age ≤ 35 y. (lag)		-0.076***	-0.045***	-0.064***
		(0.011)	(0.006)	(0.005)
Share female workers (lag)		-0.082***	-0.050***	-0.070***
		(0.017)	(0.009)	(0.008)
Share part-time workers (lag)		-0.110***	-0.083***	-0.094***
		(0.036)	(0.011)	(0.010)
Economic cycle		0.005***	0.015***	0.011***
		(0.000)	(0.000)	(0.000)
Firm FE		Yes	Yes	Yes
Firm size FE		Yes	Yes	Yes
R^2		0.40	0.29	0.22
N. Observations		131729	350571	555737
Tests of hypothesis				
S1 = S2	F-statistic	16.57	119.41	152.04
	Prob>F	Rejected at 1%	Rejected at 1%	Rejected at 1%
S2 = S3.1	F-statistic	0.59	16.09	23.06
	Prob>F	Not Rejected	Rejected at 1%	Rejected at 1%
S2 = S3.2	F-statistic	0.86	5.67	0.38
	Prob>F	Not Rejected	Rejected at 5%	Not Rejected
S3.1 = S3.2	F-statistic	1.77	0.00	14.00
	Prob>F	Not Rejected	Not Rejected	Rejected at 1%

Source: own calculations using SCIE, QP, EC.

Notes: Dependent variable is the natural logarithm of Total Factor Productivity. We use robust standard errors clustered at firm-level. Standard errors are reported in parentheses. Symbols *, **, and *** denote significance at 10%, 5%, and at 1% respectively. Tests of hypothesis refer to test on the equality of coefficients.

As in our main regression, the results indicate that a higher share of workers with tertiary education is positively associated with the productivity of firms. However, the results also show that, for manufacturing and the total economy as a whole, the effect on productivity associated with an increase in the share of workers with tertiary education is not significantly different from the one associate to an increase in upper secondary education, for the two sub-groups considered. Only for services, this difference is positive and significant for both groups, meaning that an increase in the share of workers with tertiary education from either field of study (STEM and non-STEM) is associated with a significantly larger effect on productivity relative to a change of the same magnitude in the share of workers with upper secondary education.

Additionally, we find no evidence to support that the effect on productivity associated with a higher share of workers with a tertiary degree in STEM fields of study is larger than the one associated with a higher share of workers with a tertiary degree in non-STEM fields. It is important to take into consideration that a very large share of firms (75%) do not report to employ any worker with tertiary education in STEM fields of study (Table A.3.6, Appendix 3), which contributes to the low within variation associated to this indicator. Moreover, the average share of workers with this type of qualifications is also very small when compared to the other shares of workers under analysis. These two factors should be taken into consideration when interpreting these results.

7.2.3. Tertiary education by occupation: the role of Scientists and Engineers (Group 3)

We also distinguish workers with tertiary qualifications by their occupation, into those that are actually employed in core Science and Technology occupations (Scientists and Engineers), and those who are not (group 3).

The results show that a 1 p.p. increase in the share of workers with tertiary education employed in core S&T occupations (Scientists and engineers) is associated, on average, with an increase to the firm's productivity between 0.21% and 0.31% (Table 7).

This effect is larger than the one estimated for an increase of the same magnitude to the share of workers with tertiary education in non-core S&T occupations (which on average is between 0.18% and 0.25%), although the difference is only significant for firms in the services sector.

As for the difference on the effect on productivity associated to a 1 p.p. increase in the share of workers with tertiary education relative one associate to an increase with same magnitude in share of workers with upper secondary education, the conclusions remain the same as the ones from the main regression (section 6).

Table 7 – Regression results for alternative groups of qualifications - group 3.

		Manufacturing (1)	Services (2)	Total (3)
Share workers w/ lower sec. education (lag) - [S1]		**0.131*** (0.011)**	**0.133*** (0.009)**	**0.104*** (0.006)**
Share workers w/ upper secundary/post-sec. education (lag) - [S2]		**0.188*** (0.014)**	**0.223*** (0.009)**	**0.187*** (0.007)**
Share workers w/ tertiary education (lag) - [S3]				
by occupation:				
other than Scientists & Engineers (lag) - [S3.3]		**0.183*** (0.031)**	**0.253*** (0.013)**	**0.217*** (0.011)**
Scientists & Engineers (lag) - [S3.4]		**0.212*** (0.054)**	**0.313*** (0.017)**	**0.232*** (0.015)**
Average Tenure (lag)		0.011*** (0.001)	0.016*** (0.001)	0.011*** (0.001)
Average Tenure^2 (lag)		-0.000*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)
Share of Young workers, age ≤ 35 y. (lag)		-0.076*** (0.011)	-0.046*** (0.006)	-0.064*** (0.005)
Share female workers (lag)		-0.080*** (0.017)	-0.049*** (0.009)	-0.069*** (0.008)
Share part-time workers (lag)		-0.110*** (0.037)	-0.083*** (0.011)	-0.094*** (0.010)
Economic cycle		0.005*** (0.000)	0.015*** (0.000)	0.011*** (0.000)
Firm FE		Yes	Yes	Yes
Firm size FE		Yes	Yes	Yes
R^2		0.40	0.29	0.22
N. Observations		131729	350571	555737
Tests of hypothesis				
S1 = S2	F-statistic	16.49	120.07	150.53
	Prob>F	Rejected at 1%	Rejected at 1%	Rejected at 1%
S2 = S3.3	F-statistic	0.02	7.07	8.06
	Prob>F	Not Rejected	Rejected at 1%	Rejected at 1%
S2 = S3.4	F-statistic	0.20	30.59	9.42
	Prob>F	Not Rejected	Rejected at 1%	Rejected at 1%
S3.3 = S3.4	F-statistic	0.26	13.19	1.03
	Prob>F	Not Rejected	Rejected at 1%	Not Rejected

Source: own calculations using SCIE, QP, EC.

Notes: Dependent variable is the natural logarithm of total factor productivity. We use robust standard errors clustered at firm-level. Standard errors are reported in parentheses. Symbols *, **, and *** denote significance at 10%, 5%, and at 1% respectively. Tests of hypothesis refer to test on the equality of coefficients.

Due to the very high number of firms with null or small shares of workers with STEM qualifications (either by field of study or occupation), a different analysis might be needed to properly assess the different impacts of this two types of qualifications.

7.2.4. QUEST III R&D classification (Group 4)

The results from group 3 suggest that the positive effect of a higher share of workers with tertiary education is mainly driven by the effect of the workers employed in core S&T occupations in the services sector. As such, we also estimate equation 1 using QUEST III R&D' classification of skills (Group 4), where high skilled workers correspond to workers with tertiary education in STEM fields (which are proxied by the share of workers with tertiary education in S&T occupations – Scientists and Engineers) and medium skilled workers correspond to workers with upper secondary education or tertiary education that are not considered as high-skilled.

The estimation results are presented in Table 8 and show that higher shares of medium and high skilled workers are, on average, associated with a positive and significant effect in terms of productivity, which is significantly higher for the high skilled workers in the services sector and for the total economy. These results are in line with the simulations from QUEST model (Appendix 4). However, as the results for the previous groups show, it might be relevant to distinguish between the effects of upper secondary education and tertiary education, as the differences are significant for the total economy and services sector.

It is also important to note that this outcome is not directly comparable to the ex-ante assessments of education policies included in past National Reform Programmes, which were estimated using the QUEST III R&D model[28]. In fact, the outcome of empirical studies like this can be important to assess the robustness of elasticities consider under a Dynamic Stochastic General Equilibrium model like QUEST.

[28] a New-Keynesian Dynamic Stochastic General Equilibrium model – DSGE - developed by the European Commission (Roger et al., 2008) and calibrated to the Portuguese economy

Table 8 – Regression results for alternative groups of qualifications - group 4.

		Manufacturing (1)	Services (2)	Total (3)
Share workers w/ lower sec. education (lag) - [S1]		**0.131*****	**0.133*****	**0.104*****
		(0.011)	**(0.009)**	**(0.006)**
Share medium-skilled workers (lag) - [S4]		**0.187*****	**0.228*****	**0.192*****
		(0.013)	**(0.009)**	**(0.007)**
Share high-skilled workers (lag) - [S5/S3.4]		**0.213*****	**0.304*****	**0.225*****
		(0.054)	**(0.017)**	**(0.015)**
Average Tenure (lag)		0.011***	0.016***	0.011***
		(0.001)	(0.001)	(0.001)
Average Tenure^2 (lag)		-0.000***	-0.001***	-0.000***
		(0.000)	(0.000)	(0.000)
Share of Young workers, age ≤ 35 y. (lag)		-0.076***	-0.046***	-0.063***
		(0.011)	(0.006)	(0.005)
Share female workers (lag)		-0.081***	-0.048***	-0.068***
		(0.017)	(0.009)	(0.008)
Share part-time workers (lag)		-0.110***	-0.083***	-0.094***
		(0.037)	(0.011)	(0.010)
Economic cycle		0.005***	0.015***	0.011***
		(0.000)	(0.000)	(0.000)
Firm FE		Yes	Yes	Yes
Firm size FE		Yes	Yes	Yes
R^2		0.40	0.29	0.22
N. Observations		131729	350571	555737
Tests of hypothesis				
S1 = S4	F-statistic	17.38	141.18	178.54
	Prob>F	Rejected at 1%	Rejected at 1%	Rejected at 1%
S4 = S5	F-statistic	0.23	24.55	5.50
	Prob>F	Not Rejected	Rejected at 1%	Rejected at 5%

Source: own calculations using SCIE, QP, EC.

Notes: Dependent variable is the natural logarithm of total factor productivity. We use robust standard errors clustered at firm-level. Standard errors are reported in parentheses. Symbols *, **, and *** denote significance at 10%, 5%, and at 1% respectively. Tests of hypothesis refer to test on the equality of coefficients.

The QUEST III R&D is a model particularly suitable to assess the impact of structural reforms and investments at the macroeconomic level, and it incorporates dynamic effects, as well as potential cumulative effects, with the cumulative effects inherent in an increase of the share of higher skills being potentially relevant, in particular in the case of STEM shares (Appendix 4). It should also be noted that our analysis includes marginal effects and not general equilibrium effects, which means it does not incorporate dynamic effects, as well as potential cumulative effects. Our analysis can also differ from the macroeconomic effect, as the results are estimated at the firm-level, with the universe of firms included in the analysis being restricted to non-financial corporations with positive values for Gross Value Added, to firms with five or more employees, and to firms not included in the public sector (a sector where important R&D agents might be present). However, our analysis has the important advantage that it considers the heterogeneity of the economic agents for the estimation. As we have seen, there are important differences across sectors, firms, and workers that influence productivity and should be accounted for when estimating the effects. Unfortunately, and as it is common in the literature, due to the inability to find a proper instrument in our database, our analysis does not allow us to infer a causal relation but an association, as it is performed in QUEST.

7.3. Years of education

Lastly, we include an alternative indicator to measure workers' qualifications – mean years of schooling. The results in Table 9 indicate that a higher average years of schooling of the workers in a firm is associated with a positive and significant effect on productivity, although non-linear (the quadratic term is negative, but close to zero, and insignificant for the total economy). This conclusion is in line with the findings for our main regression.

Table 9 – Regression results using mean years of schooling as a measure of worker's

	Manufacturing (1)	Manufacturing (2)	Services (3)	Services (4)	Total (5)	Total (6)
Average years schooling (lag)	**0.025*** (0.001)**	**0.049*** (0.009)**	**0.029*** (0.001)**	**0.050*** (0.005)**	**0.024*** (0.001)**	**0.029*** (0.004)**
Average years schooling, squared (lag)		**-0.001*** (0.001)**		**-0.001*** (0.000)**		**-0.000 (0.000)**
Average Tenure (lag)	0.011*** (0.001)	0.011*** (0.001)	0.015*** (0.001)	0.015*** (0.001)	0.011*** (0.001)	0.011*** (0.001)
Average Tenure^2 (lag)	-0.000*** (0.000)	-0.000*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
Share of Young workers, age ≤ 35 (lag)	-0.076*** (0.011)	-0.077*** (0.011)	-0.046*** (0.006)	-0.046*** (0.006)	-0.064*** (0.005)	-0.064*** (0.005)
Share female workers (lag)	-0.086*** (0.017)	-0.084*** (0.017)	-0.050*** (0.009)	-0.050*** (0.009)	-0.070*** (0.008)	-0.070*** (0.008)
Share part-time workers (lag)	-0.112*** (0.037)	-0.112*** (0.037)	-0.082*** (0.011)	-0.083*** (0.011)	-0.094** (0.010)	-0.094** (0.010)
Economic cycle	0.005*** (0.000)	0.005*** (0.000)	0.015*** (0.000)	0.015*** (0.000)	0.011*** (0.000)	0.011*** (0.000)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm size FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.42	0.41	0.30	0.31	0.23	0.23
N. Observations	131729	131729	350571	350571	555737	555737

Source: own calculations using SCIE, QP, EC.

Notes: Dependent variable is the natural logarithm of total factor productivity. We use robust standard errors clustered at firm-level. Standard errors are reported in parentheses. Symbols *, **, and *** denote significance at 10%, 5%, and at 1% respectively. Tests of hypothesis refer to test on the equality of coefficients.

7.4 Training costs

In this section, we add an additional variable to our specification - training costs supported by firms measured by the expenditure on training costs over total payroll – to analyse if it can be important to explain differences in firm's productivity. The results in Table 10 show that changes to the new variable do not seem to have a significant effect on firm's productivity, with the effects associated with our main explanatory variables and control variables remaining stable.

Table 10 – Regression results controlling for training costs supported by firms

		Manufacturing (1)	Services (2)	Total (3)
Share workers w/ lower sec. education (lag) - [S1]		**0.131*** (0.011)**	**0.133*** (0.009)**	**0.104*** (0.006)**
Share workers w/ upper secundary/post-sec. education (lag) - [S2]		**0.188*** (0.014)**	**0.223*** (0.009)**	**0.187*** (0.007)**
Share workers w/ tertiary education (lag) - [S3]		**0.190*** (0.029)**	**0.267*** (0.013)**	**0.221*** (0.011)**
Average Tenure (lag)		0.011*** (0.001)	0.015*** (0.001)	0.011*** (0.001)
Average Tenure^2 (lag)		-0.000*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)
Share of Young workers, age ≤ 35 (lag)		-0.076*** (0.011)	-0.045*** (0.006)	-0.064*** (0.005)
Share female workers (lag)		-0.081*** (0.017)	-0.050*** (0.009)	-0.069*** (0.008)
Share part-time workers (lag)		-0.110*** (0.037)	-0.083*** (0.011)	-0.094** (0.010)
Training costs (lag)		**-0.045 (0.143)**	**0.086 (0.088)**	**0.087 (0.073)**
Economic cycle		0.005*** (0.000)	0.015*** (0.000)	0.011*** (0.000)
Firm FE		Yes	Yes	Yes
Firm size FE		Yes	Yes	Yes
R^2		0.41	0.26	0.22
N. Observations		131729	350571	555737
Tests of hypothesis				
S1 = S4	F-statistic	16.49	119.35	150.53
	Prob>F	Rejected at 1%	Rejected at 1%	Rejected at 1%
S4 = S5	F-statistic	0.00	17.39	12.32
	Prob>F	Not Rejected	Rejected at 1%	Rejected at 1%

Source: own calculations using SCIE, QP, EC.

Notes: Dependent variable is the natural logarithm of total factor productivity. We use robust standard errors clustered at firm-level. Standard errors are reported in parentheses. Symbols *, **, and *** denote significance at 10%, 5%, and at 1% respectively. Tests of hypothesis refer to test on the equality of coefficients.

8. Suggestions for future research

Overall, the results point to the important conclusion that, on average, the increase in workers' qualifications (regardless of type of education) is, in fact, associated with an increase in productivity in both manufacturing and services. Unfortunately, and as it is common in the literature, due to the inability to find a proper instrument in our database, our analysis does not allow us to infer a causal relation but only an association. As new information becomes available in the next years, it is possible that a proper instrument can be found. Moreover, the new data may also allow to extend the period of analysis. This extension may help to capture additional effects that could not be assessed in shorter datasets such as ours (period 2010-2019), as well as to capture the effects coming from polices implemented in more recent years (such as new VET programmes or the extension of the age of compulsory education).

A possible extension of our analysis that could be relevant to explore is to distinguish the effect of newer and older cohorts of workers with upper secondary education in each track - vocational and general education - as the effect may present relevant differences across the years. These differences may reflect the effects of the very large set of policies implemented in the last years, such as changes to the curriculum, and the large investments and reforms to vocational education, that can lead to changes in the quality of education across the years. For example, as the newer VET courses provide, in theory, an education that is more tailored to the needs of the labour market, as oppose to general education, they can be seen has having a superior quality. This distinction could be an important addition to our work, because although we have distinguished the effect between vocational and general education in upper secondary education, we did not take into account potential differences in the quality of education across the years for any of the qualification groups.

Moreover, it would also be interesting to assess the robustness of our results to the inclusion of other types of education, such as on-job training or informal education. These other types of education can potentially improve workers' skill level with positive impact on their productivity and therefore on the firm's value added. Ideally, these differences in skills and ability should be incorporated, but in these databases this information is not available. As such, to minimize a potential omitted variable bias, we use firm fixed effects to capture differences in skills that remain fixed over time, and we also include information on training expenditure as a robustness check.

Additionally, our analysis has the important advantage of including potential effects from frictions, namely wage rigidities, and inefficiencies in resource allocation of the most qualified workers (mismatch). In a simple exercise (Annex 5) we compare the qualifications of workers with the qualifications of their specific occupation, with the results pointing to the existence of potential inefficient combinations of competences and occupations. As such, to better inform the design and implementation scheme of public policies aimed at promoting a more efficient combination of skills and occupations, it would be important to analyse in more detail those potential inefficiencies in the allocation of the more qualified workers in the future.

Moreover, considering the high number of null shares of workers with some qualifications, it could also be relevant to understand if firms that employ a worker with tertiary education for the first time are the ones that benefit more that those firms that already employ workers with this type of qualifications.

Finally, management quality can also influence the capacity of firms to use their human resources more efficiently, including the more qualified. As such, it could also be important to test how productivity relates to the qualifications of the managers.

9. Conclusions

The main purpose of our paper is to study the relationship between the increase in the shares of workers with higher qualifications and the productivity of Portuguese firms. For this purpose, we use firm and employee level data from *Quadros de Pessoal* complemented with information from *Sistema de Contas Integradas das Empresas*. The firm identifier is the same in both databases, which allows for combinations of the information.

The results from our analysis suggest that an increase in workers' qualifications (regardless of type of education) is associated with an increase in firms' total factor productivity, in both manufacturing and services sectors. We find that the effect is of a higher magnitude for the services sector. We also find that these conclusions are robust to the use of alternative measures of productivity, measures of qualifications, and the inclusion of the expenditure on training made by firms each year.

Our findings also reveal that when comparing the effect of increasing the firm's share of workers with upper secondary with the increase of the share of workers with only lower secondary education, we find the first is associated with a significantly larger effect on productivity. This difference exists not only for the sample as a whole, but also for the manufacturing and services sectors. We also find that this difference is robust when considering different types of study tracks - general and vocational education.

In what regards tertiary education the conclusions differ between sectors. On one hand, for the services sector the effect on productivity is significantly larger than the one associate with the share of workers with only upper secondary education. One the other hand, we do not find the same evidence for the manufacturing sector and total sample. A similar conclusion can be obtained when distinguishing the firm's share of workers with tertiary education by field of studies and by occupation more focused on Science and Technology activities from those that are not.

Our findings also reveal that increasing the share of workers with tertiary education is associated with an increase to productivity of greater magnitude when compared with an increase of the share of workers with upper secondary, particularly regarding workers in core Science and Technology occupations, such as Scientists and engineers. This outcome does not apply for the manufacturing sector, and as such, future analysis on possible explanations would be very relevant.

Considering this evidence, and the fact that Portugal, in 2021, had the largest share of working age population with low qualifications (lower secondary education or less) of all European Union countries, it is utmost important to continue implementing public policies that promote higher qualifications for the Portuguese population. These policies should focus on the level of upper secondary education and tertiary education, and also policies targeted to the adult population, the group with the lowest level of qualifications.

Moreover, future research should analyse in more detail the inefficiencies in the allocation of more qualified workers. This would better inform the design and implementation scheme of public policies aimed at promoting a more efficient combination of skills and occupations. In addition, it could also be relevant to understand if firms that employ a worker with tertiary education for the first-time benefit more that those firms that already employ workers with this type of qualifications. Lastly, it might also be relevant to distinguish the effect across different generation of workers between vocational and general education considering the large investments and reforms allocated to vocational education in recent years.

Bibliography

Capucha, L., Pereira, C., Godinho, R., Capucha, A. R., Simões, A., Ferreira, G., Henriques, J. P., Sebastião, J., Martins, S. & Seabra, T. (2021). Avaliação do Contributo do PT2020 para a Promoção do Sucesso Educativo, Redução do Abandono Escolar Precoce e Empregabilidade dos Jovens. Available at https://www.adcoesao.pt/wp-content/uploads/relatorio_final_contributo_pt2020_educacao.pdf.

Cima, J., Pimenta, A. C., Portela, M. & Silva, M. (2022). Qualificações dos trabalhadores e produtividade das empresas. Banco de Portugal Working Paper number 4.

Criscuolo, C., Gal, P., Leidecker, T., & Nicoletti, G. (2021). The human side of productivity: Uncovering the role of skills and diversity for firm productivity. OECD Productivity Working Papers, No. 29, OECD Publishing, Paris.

Égert, B., de la Maisonneuve, C. & Turner, D. (2022). A new macroeconomic measure of human capital exploiting PISA and PIAAC: Linking education policies to productivity. OECD Economics Department Working Papers, No. 1709, OECD Publishing, Paris.

Fleischhauer, K. J. (2007). A review of human capital theory: Microeconomics. University of St. Gallen, Department of Economics Discussion Paper, (2007-01).

Fox, J. T., & Smeets, V. (2011). Does input quality drive measured differences in firm productivity?. *International Economic Review*, *52*(4), 961-989.

Fundação José Neves (2022). Estado da Nação 2022. Digital version available at https://joseneves.org/pt/estado-da-nacao-2022.

Rovigatti, G. & Mollisi, V. (2016). PRODEST: Stata module for production function estimation based on the control function approach. Statistical Software Components S458239, Boston College Department of Economics, revised 28 Jul 2020.

Galindo-Rueda, F., & Haskel, J. (2005). Skills, workforce characteristics and firm-level productivity: evidence from the matched ABI/employer skills survey. IZA Discussion Paper No. 1542.

Hægeland, T., & Klette, T. J. (1997). Do higher wages reflect higher productivity?: education, gender and experience premiums in a matched plant-worker data set. Discussion Papers No. 208, Statistics Norway, Research Department.

Haltiwanger, J. C., Lane, J. I., & Spletzer, J. (1999). Productivity differences across employers: The roles of employer size, age, and human capital. *American Economic Review*, *89*(2), 94-98.

Haltiwanger, J. C., Lane, J. I., & Spletzer, J. R. (2007). Wages, productivity, and the dynamic interaction of businesses and workers. *Labour Economics*, *14*(3), 575-602.

Haskel, J., Hawkes, D. D., & Pereira, S. C. (2005). Skills, human capital and the plant productivity gap: UK evidence from matched plant, worker and workforce data. Worker and Workforce Data (November 2005).

Hellerstein, J. K., Neumark, D., & Troske, K. R. (1999). Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equations. *Journal of labor economics*, *17*(3), 409-446.

Hellerstein, J. K., & Neumark, D. (2007). Production function and wage equation estimation with heterogeneous labor: Evidence from a new matched employer-employee data set. In *Hard-to-measure goods and services: Essays in honor of Zvi Griliches* (pp. 31-71). University of Chicago Press.

Lebedinski, L., & Vandenberghe, V. (2014). Assessing education's contribution to productivity using firm-level evidence. *International Journal of Manpower, 35(8)*, 116-1139.

Martins, P. S & Jin, J. (2010). Firm-level social returns to education. *Journal of Population Economics, 23*, 539–558.

OECD (2014). "Education Policy Outlook: Portugal". OECD Publishing, Paris.

OECD (2020). "Education Policy Outlook: Portugal". OECD Publishing, Paris.

Portugal, P., Reis, H., Guimarães, P. & Cardoso, R. (2022). Human Capital Spillovers and Returns to Education. Banco de Portugal Working Paper number 19.

Portugal, P., Reis, H., Guimarães, P. & Cardoso, R. (2018). The returns to schooling unveiled. Banco de Portugal Working Paper number 5.

Roeger, W., Varga, J. & in 't Veld, J. (2008). Structural reforms in the EU: a simulation-based analysis using the QUEST model with endogenous growth. European Economy Economic Paper 351.

Rycx, F., Saks, Y., & Tojerow, I. (2015). Does education raise productivity and wages equally? The moderating roles of age, gender and industry. *The Moderating Roles of Age, Gender and Industry*.

Sianesi, B., & Reenen, J. V. (2003). The returns to education: Macroeconomics. *Journal of economic surveys*, *17*(2), 157-200.

Syverson, C. (2011). What determines productivity?. *Journal of Economic literature*, *49*(2), 326-65.

Turcotte, J., & Rennison, L. W. (2004). The link between technology use, human capital, productivity and wages: firm-level evidence. *International productivity monitor*, 25-36.

Van Biesebroeck, J. (2011). Wages equal productivity. Fact or fiction? Evidence from sub Saharan Africa. *World Development*, *39*(8), 1333-1346.

Vandenberghe, V. (2011a). Firm-level Evidence on Gender Wage Discrimination in the Belgian Private Economy. *Labour*, *25*(3), 330-349.

Vandenberghe, V. (2011b). Boosting the employment rate of older men and women. *De Economist*, *159*(2), 159-191.

Wilson, R. A., & Briscoe, G. (2004). The impact of human capital on economic growth: a review. *Impact of education and training. Third report on vocational training research in Europe: background report. Luxembourg: EUR-OP*, *23*.

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. Economics letters, 104(3), 112-114.

Appendix 1 – Data Cleaning

Information regarding workers' characteristics is obtained from *Quadros de Pessoal*. Given that our analysis is performed at the firm-level, we have to aggregate workers' information into ratios or averages. As such, we tried to preserve as much information as possible, and adjusted some of the information whenever possible.

To clean our dataset, we first identified which information was missing and classified which observations may had inconsistent information.

Regarding the missing information on the key variables used in our analysis, we followed two approaches. The first consisted on trying to find if, for the worker that had an observation with missing information, the database had sufficient information in past or subsequent years that could help us to infer it. If it had such information, we replaced the missing data by that information. For example, if for a given year we had no information for the age of the worker, but we had that information for the previous year, we replaced the missing information with the age in the previous year plus one. The second approach consisted on dropping the observations from the database that had not sufficient information to infer its correct value.

Regarding the observations with inconsistent information, the following steps were followed. First, whenever we had observations for the same worker and firm that were duplicated (workers had the same characteristics, but the information appeared in the database more than once), we only included the observation with the highest remuneration. Second, if firms had only one of these issues: levels of qualifications that decreased over time, age and tenure that did not progress in line with time, we substituted, whenever possible, the inconsistent information using data from previous or subsequent years. If this correction was not possible to compute, the observation was dropped from the dataset.

Appendix 2 – Additional information on database and methodology

Table A.2.1. – Sectors of activity

Sector	Description	NACE 2 digit codes
1*	Agriculture, forestry and fishing	1-3
2*	Mining and quarrying	5-9
3	Manufacturing of food products, beverages and tobacco	10-12
4	Manufacturing of textiles, wearing apparel, leather and related products	13-15
5	Manufacturing of wood and paper products, and printing	16-18
6	Manufacturing of coke and refined petroleum products	19
7	Manufacturing of chemicals and chemical products	20
8	Manufacturing of basic pharmaceutical products and pharmaceutical preparations	21
9	Manufacturing of rubber, plastics and other non-metallic mineral products	22-23
10	Manufacturing of basic metals and fabricated metal products, except machinery and equipment	24-25
11	Manufacturing of computer, electronic and optical products	26
12	Manufacturing of electrical equipment	27
13	Manufacturing of machinery and equipment n.e.c.	28
14	Manufacturing of transport equipment	29-30
15	Manufacturing of furniture; other manufacturing; repair and installation	31-33
16	Electricity, gas, steam and air conditioning supply	35
17	Water supply; sewerage, waste management and remediation activities	36-39
18	Construction	41-43
19	Wholesale and retail trade, repair of motor vehicles and motorcycles	45-47
20	Transportation and storage	49-53
21	Accommodation and food service activities	55-56
22	Publishing, audiovisual and broadcasting activities	58-60
23	Telecommunications	61
24	IT and other information services	62-63
25*	Financial and insurance activities	64-66
26	Real estate activities	68
27	Legal and accounting activities; etc	69-71
28	Scientific research and development	72
29	Advertising, market research; other professional, scientific and technical activities	73-75
30	Administrative and support service activities	77-82
31*	Public administration and defence; compulsory social security	84
32	Education	85
33	Human health activities	86
34	Residential care and social work activities	87-88
35	Arts, entertainment and recreation	90-93
36	Other service activities	94-96
37*	Activities of households as employers; undifferentiated activities of households for own use	97-98
38*	Activities of extraterritorial organizations and bodies	99

Table A.2.2. – Hausman Test Results (H0: difference in coefficients is not systematic)

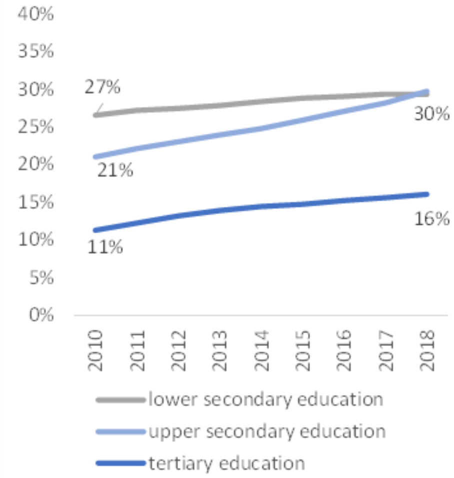
	Manufacturing	Services	Total
Chi2	10331	18842	30787
Prob>Chi2	0.0000	0.0000	0.0000

Source: own calculations using QP and SCIE.

Note: Using the chosen specification, the initial hypothesis that the firm-level effects are adequately modelled by a random-effects model is rejected for manufacturing, services and total economy

Appendix 3 – Additional information on descriptive statistics and results

Figure A.3.1. – Average firm's shares of workers by levels of education in 2010-2018 – Total Economy (percentage, %)



Source: own calculations using QP and SCIE.

Table A.3.1. – Difference* in firm's shares of workers by levels of education and other relevant firm characteristics between two consecutive years: breakdown by sector and productivity groups (mean and standard deviation)**

	S. workers w/ lower secundary educ.	S. workers w/ upper/ post secund. educ.	S. workers w/ tertiary educ.	Average Tenure ***	S. of Young workers, age ≤ 35 y.	S. female workers	S. part-time workers
Manufacturing	**0.005**	**0.011**	**0.003**	**0.391**	**-0.015**	**0.001**	**0.000**
	(0.100)	**(0.082)**	**(0.045)**	**(1.340)**	**(0.104)**	**(0.072)**	**(0.039)**
by productivity percentile:							
Firms with low levels of productivity	0.008	0.009	0.002	0.570	-0.019	0.000	0.000
(below 25th percentile)	(0.119)	(0.088)	(0.042)	(1.394)	(0.123)	(0.086)	(0.045)
Firms with medium-low levels of productivity	0.006	0.011	0.002	0.444	-0.016	0.001	0.000
(between 25th and 50th percentiles)	(0.108)	(0.087)	(0.043)	(1.396)	(0.113)	(0.076)	(0.042)
Firms with medium-high level of productivity	0.004	0.011	0.003	0.351	-0.014	0.002	0.000
(between 50th and 75th percentiles)	(0.096)	(0.085)	(0.048)	(1.357)	(0.100)	(0.070)	(0.038)
Firms with high levels of productivity	0.001	0.012	0.004	0.225	-0.012	0.001	0.000
(above 75th percentile)	(0.073)	(0.070)	(0.048)	(1.193)	(0.077)	(0.054)	(0.031)
Services	**-0.001**	**0.008**	**0.005**	**0.396**	**-0.018**	**0.001**	**0.001**
	(0.115)	**(0.116)**	**(0.077)**	**(1.361)**	**(0.130)**	**(0.096)**	**(0.072)**
by productivity percentile:							
Firms with low levels of productivity	0.002	0.010	0.002	0.440	-0.017	0.001	0.001
(below 25th percentile)	(0.154)	(0.142)	(0.077)	(1.494)	(0.159)	(0.126)	(0.102)
Firms with medium-low levels of productivity	-0.001	0.008	0.005	0.442	-0.020	0.001	0.001
(between 25th and 50th percentiles)	(0.123)	(0.125)	(0.083)	(1.409)	(0.139)	(0.100)	(0.073)
Firms with medium-high level of productivity	-0.002	0.008	0.005	0.392	-0.019	0.001	0.000
(between 50th and 75th percentiles)	(0.104)	(0.111)	(0.078)	(1.347)	(0.124)	(0.088)	(0.061)
Firms with high levels of productivity	-0.003	0.005	0.007	0.326	-0.017	0.002	0.000
(above 75th percentile)	(0.075)	(0.088)	(0.071)	(1.212)	(0.099)	(0.071)	(0.049)
Total Economy	**0.001**	**0.009**	**0.004**	**0.407**	**-0.018**	**0.001**	**0.000**
	(0.113)	**(0.106)**	**(0.070)**	**(1.350)**	**(0.124)**	**(0.089)**	**(0.064)**
by productivity percentile:							
Firms with low levels of productivity	0.004	0.010	0.002	0.492	-0.019	0.001	0.001
(below 25th percentile)	(0.145)	(0.125)	(0.071)	(1.464)	(0.150)	(0.113)	(0.089)
Firms with medium-low levels of productivity	0.001	0.009	0.004	0.453	-0.020	0.001	0.000
(between 25th and 50th percentiles)	(0.120)	(0.113)	(0.073)	(1.392)	(0.132)	(0.092)	(0.065)
Firms with medium-high level of productivity	0.000	0.009	0.004	0.391	-0.019	0.001	0.000
(between 50th and 75th percentiles)	(0.106)	(0.103)	(0.071)	(1.344)	(0.119)	(0.082)	(0.055)
Firms with high levels of productivity	-0.001	0.007	0.006	0.313	-0.016	0.002	0.000
(above 75th percentile)	(0.079)	(0.084)	(0.066)	(1.210)	(0.095)	(0.067)	(0.045)

Source: own calculations using QP and SCIE.

Note: * difference (D) of share of workers or other relevant firm characteristics between two consecutive years ($DX=X_t-X_{t-1}$). ** Firms are divided into four groups according to their position in the productivity distribution. *** Difference in the average tenure (in years) between two consecutive years. Standard deviation reported in parentheses.

Table A.3.2. – Firm's shares of workers by levels of education and by other relevant firm characteristics in the period 2010-2018: breakdown by sector and productivity groups* (mean and standard deviation)

	S. workers w/ lower secundary educ.	S. workers w/ upper/ post secund. educ.	S. workers w/ tertiary educ.	Average Tenure (years)	S. of Young workers, age ≤ 35 y.	S. female workers	S. part-time workers
Manufacturing	**0.297**	**0.177**	**0.064**	**8.409**	**0.312**	**0.405**	**0.012**
	(0.255)	**(0.191)**	**(0.130)**	**(4.545)**	**(0.225)**	**(0.134)**	**(0.074)**
by productivity percentile:							
Firms with low levels of productivity	0.304	0.112	0.022	7.259	0.291	0.592	0.014
(below 25th percentile)	(0.270)	(0.178)	(0.071)	(5.930)	(0.230)	(0.367)	(0.066)
Firms with medium-low levels of productivity	0.304	0.148	0.037	8.410	0.311	0.408	0.014
(between 25th and 50th percentiles)	(0.247)	(0.186)	(0.084)	(6.091)	(0.224)	(0.327)	(0.059)
Firms with medium-high level of productivity	0.303	0.194	0.063	8.709	0.317	0.330	0.012
(between 50th and 75th percentiles)	(0.223)	(0.192)	(0.103)	(5.982)	(0.212)	(0.265)	(0.049)
Firms with high levels of productivity	0.280	0.242	0.125	9.103	0.327	0.314	0.008
(above 75th percentile)	(0.174)	(0.175)	(0.137)	(5.555)	(0.187)	(0.229)	(0.040)
Services	**0.276**	**0.304**	**0.180**	**6.554**	**0.370**	**0.492**	**0.037**
	(0.230)	**(0.189)**	**(0.110)**	**(5.926)**	**(0.213)**	**(0.318)**	**(0.054)**
by productivity percentile:							
Firms with low levels of productivity	0.336	0.256	0.086	5.391	0.387	0.563	0.061
(below 25th percentile)	(0.296)	(0.281)	(0.189)	(5.115)	(0.276)	(0.308)	(0.154)
Firms with medium-low levels of productivity	0.288	0.305	0.151	6.460	0.376	0.508	0.040
(between 25th and 50th percentiles)	(0.270)	(0.277)	(0.240)	(5.378)	(0.269)	(0.329)	(0.118)
Firms with medium-high level of productivity	0.269	0.312	0.193	6.693	0.368	0.465	0.029
(between 50th and 75th percentiles)	(0.247)	(0.260)	(0.265)	(5.241)	(0.252)	(0.321)	(0.092)
Firms with high levels of productivity	0.227	0.332	0.268	7.403	0.351	0.449	0.025
(above 75th percentile)	(0.204)	(0.229)	(0.276)	(5.035)	(0.224)	(0.283)	(0.080)
Total Economy	**0.282**	**0.252**	**0.141**	**6.881**	**0.346**	**0.420**	**0.029**
	(0.250)	**(0.248)**	**(0.222)**	**(5.402)**	**(0.244)**	**(0.323)**	**(0.098)**
by productivity percentile:							
Firms with low levels of productivity	0.319	0.200	0.069	5.867	0.348	0.507	0.046
(below 25th percentile)	(0.288)	(0.258)	(0.161)	(5.317)	(0.265)	(0.349)	(0.132)
Firms with medium-low levels of productivity	0.292	0.242	0.113	6.790	0.350	0.428	0.032
(between 25th and 50th percentiles)	(0.265)	(0.260)	(0.204)	(5.508)	(0.257)	(0.340)	(0.102)
Firms with medium-high level of productivity	0.281	0.263	0.148	7.034	0.349	0.385	0.024
(between 50th and 75th percentiles)	(0.243)	(0.247)	(0.229)	(5.429)	(0.241)	(0.315)	(0.080)
Firms with high levels of productivity	0.246	0.292	0.219	7.626	0.339	0.378	0.020
(above 75th percentile)	(0.201)	(0.220)	(0.247)	(5.213)	(0.214)	(0.278)	(0.070)

Source: own calculations using QP and SCIE.

Note: * Firms are divided into four groups according to their position in the productivity distribution. Standard deviation reported in parentheses.

Table A.3.3. – Proportion of firms with null values for the firm's shares of workers by levels of education and other relevant firm characteristics in the period 2010-2018

	Manufacturing (1)	Services (2)	Total (3)
Workers w/ lower sec. education	**0.13**	**0.25**	**0.22**
Workers w/ upper secundary/post-sec. Education	**0.29**	**0.21**	**0.26**
Workers w/ tertiary education	**0.56**	**0.47**	**0.49**
Young workers, age ≤ 35 y.	0.11	0.12	0.12
Female workers	0.11	0.08	0.13
Part-time workers	0.90	0.82	0.84
N. Observations	131729	350571	555737

Source: own calculations using QP and SCIE.

Table A.3.4. – Regression results using an alternative productivity measure – TFP calculated using the Levinsohn and Petrin methodology

		Manufacturing (1)	Services (2)	Total (3)
Share workers w/ lower sec. education (lag) - [S1]		**0.132*****	**0.132*****	**0.104*****
		(0.011)	**(0.009)**	**(0.006)**
Share workers w/ upper secundary/post-sec. education (lag) - [S2]		**0.190*****	**0.221*****	**0.186*****
		(0.014)	**(0.009)**	**(0.007)**
Share workers w/ tertiary education (lag) - [S3]		**0.193*****	**0.266*****	**0.221*****
		(0.029)	**(0.012)**	**(0.011)**
Average Tenure (lag)		0.011***	0.016***	0.011***
		(0.001)	(0.001)	(0.001)
Average Tenure^2 (lag)		-0.000***	-0.001***	-0.000***
		(0.000)	(0.000)	(0.000)
Share of Young workers, age ≤ 35 y. (lag)		-0.076***	-0.046***	-0.063***
		(0.011)	(0.006)	(0.005)
Share female workers (lag)		-0.081***	-0.049***	-0.068***
		(0.017)	(0.009)	(0.008)
Share part-time workers (lag)		-0.114***	-0.083***	-0.094**
		(0.037)	(0.011)	(0.010)
Economic cycle		0.005***	0.015***	0.011***
		(0.000)	(0.000)	(0.000)
Firm FE		Yes	Yes	Yes
Firm size FE		Yes	Yes	Yes
R^2		0.41	0.30	0.23
N. Observations		131729	350571	555737
Tests of hypothesis				
S1 = S2	F-statistic	17.15	116.92	147.44
	Prob>F	Rejected at 1%	Rejected at 1%	Rejected at 1%
S2 = S3	F-statistic	0.01	18.04	13.31
	Prob>F	Not Rejected	Rejected at 1%	Rejected at 1%

Source: own calculations using SCIE, QP, EC.

Notes: Dependent variable is the natural logarithm of total factor productivity. We use robust standard errors clustered at firm-level. Standard errors are reported in parentheses. Symbols *, **, and *** denote significance at 10%, 5%, and at 1% respectively.

Table A.3.5. – Difference* in firm's shares of workers by levels of education between two consecutive years – mean and standard deviation

	Manufacturing (1)	Services (2)	Total (3)
Share workers w/ upper/post-secundary education (lag), by programme:			
general programme	0.008	0.006	0.006
	(0.079)	(0.116)	(0.105)
VET programme	0.003	0.002	0.002
	(0.056)	(0.084)	(0.075)
Share workers w/ tertiary education, by fields of study			
non-STEM fields study	0.002	0.004	0.003
	(0.039)	(0.077)	(0.066)
STEM fields study	0.001	0.000	0.001
	(0.031)	(0.047)	(0.047)
Share workers w/ tertiary education, by occupation			
other than Scientists & Engineers	0.002	0.004	0.003
	(0.042)	(0.076)	(0.066)
Scientists & Engineers	0.000	0.001	0.001
	(0.024)	(0.049)	(0.045)
N. Observations	131729	350571	555737

Source: own calculations using QP and SCIE.

Note: * difference (D) of share of workers or other relevant firm characteristics between two consecutive years ($DX=X_t-X_{t-1}$).

Table A.3.6. – Proportion of firms with null values for the firm's shares of workers by other levels of education in the period 2010-2018

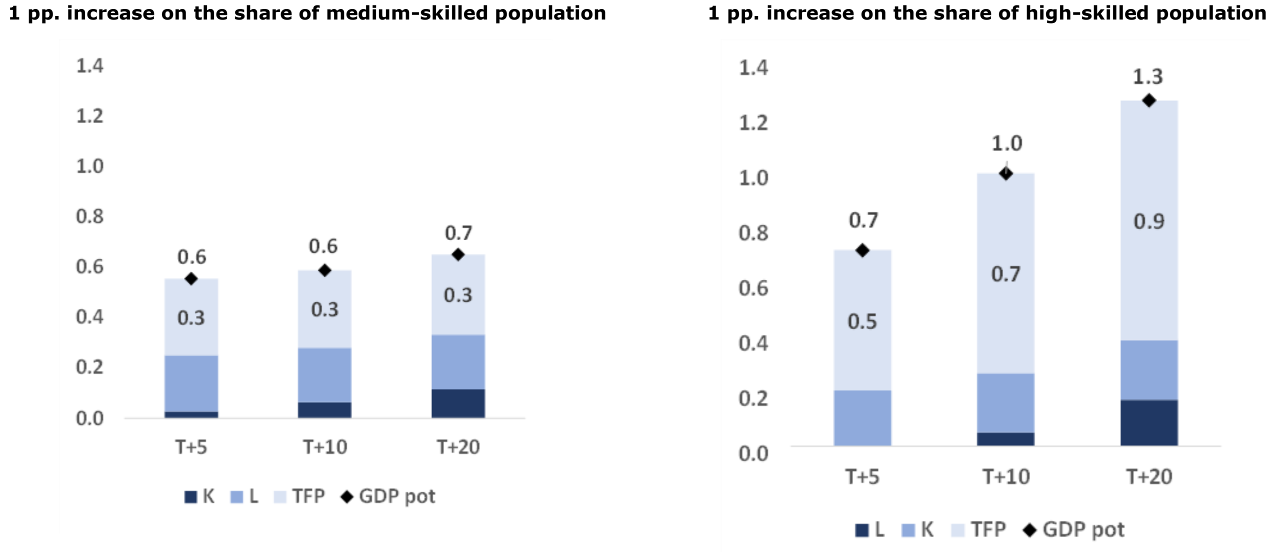
	Manufacturing (1)	Services (2)	Total (3)
Share workers w/ upper/post-secundary education (lag), by programme:			
general programme	0.39	0.32	0.37
VET programme	0.65	0.65	0.66
Share workers w/ tertiary education, by fields of study			
non-STEM fields study	0.64	0.52	0.58
STEM fields study	0.73	0.80	0.75
Share workers w/ tertiary education, by occupation			
other than Scientists & Engineers	0.60	0.52	0.57
Scientists & Engineers	0.80	0.82	0.78
N. Observations	131729	350571	555737

Source: own calculations using QP and SCIE.

Appendix 4 - Increasing skills levels in population and productivity: an analysis using QUEST III R&D

From a macroeconomic perspective, improvements to the level of education are expected to lead to increasing levels of productivity, and potential growth. An estimation of the potential macroeconomic impact (ex-ante assessment) of these improvements to the skill structure can be obtained by using the QUEST III R&D model (Roger et al., 2008), a New-Keynesian Dynamic Stochastic General (DSGE), developed by the European Commission and calibrated to the Portuguese economy. The model includes three levels of qualifications for the population – low, medium, high skilled[29], for which are assumed increasing levels of efficiency, employability, and wage.

Figure A.4.1. – Potential macroeconomic impact of simulations of two shocks using QUEST III R&D model (in percentage)



Source: own calculations using QUEST III R&D model.

Note: the impact corresponds to the changes to the level of Gross Domestic Product (GDP), labour (L), capital (K) and Total Factor Productivity (TFP) when compared to their level in a baseline scenario (no policy change scenario

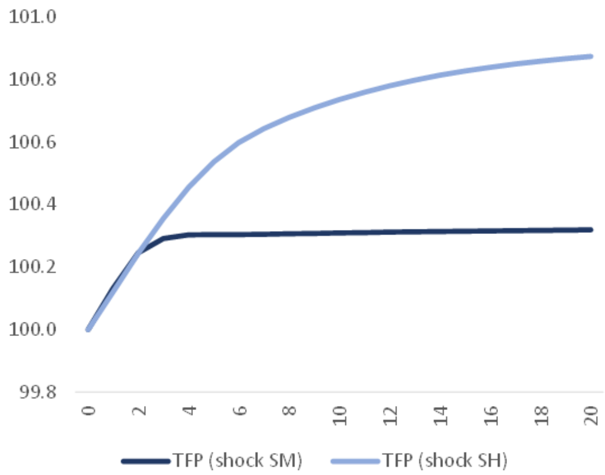
A simulation of an increase of the share of medium-skilled population by 1 p.p. (Figure A.4.1) indicates that the potential macroeconomic impact is positive, leading to a potential TFP increase of 0.3% in the long-run, relative to the baseline scenario, which is mostly explained by increases to the efficiency of the workforce. The estimated potential impact is higher if instead we consider an increase (by 1 p.p.) of the share of high-skilled population, with the results of such improvement in skills pointing to an increase in TFP of 0.5% in the medium-run (T+5), that converges to 0.9% in the long-run. This higher and increasing impact reflects higher levels of efficiency in the workforce, as well as a higher level of activity in the Research and Development (R&D) sector.

These results should be interpreted as potential impacts, as the model assumes an efficient economy without frictions in the long-run, which constitutes an important limitation regarding the use of this type of models. As part of these assumptions may not be verified in practice, it is important to perform additional empirical analysis (ex-post assessment) not only to understand the forecast capacity of the model but also to estimate the shock that better captures the impact of the reform in the model. One way to accomplish this is to analyse in more detail the relationship between improvements to the skill's composition of the workforce and productivity at the

[29] In QUEST III R&D model, low skilled workers include workers with lower secondary education or less, medium skilled workers include workers with upper secondary education, and tertiary education in non-STEM fields, and high skilled workers include workers with tertiary education in STEM occupations. The calibration of the model assumes that workers with STEM tertiary education are employed in core Science and Technology occupations (Scientists and Engineers).

firm-level, while controlling for other factors that may influence productivity dynamics, such as the economic cycle, capital intensity, and other structural factors. It is important to take into account that this empirical analysis would give the impact on a firm-level and not for the economy as a whole. Moreover, QUEST model does not consider other factors that may determine the economic cycle and that are not control for in the empirical regression.

Figure A.4.2. – Potential impact on productivity of an increase by 1 p.p. on the share of medium-skilled pop. (SM) and of an increase by 1 p.p. on the share of high-skilled pop. (SH) - simulation using QUEST III R&D model (Index, T0=100)



Source: own calculations using QUEST III R&D model.

Note: the impact corresponds to the changes to the level of TFP when compared to their level in a baseline scenario (no policy change scenario), expressed as an index.

Appendix 5 - Qualification of a worker vs. Qualification associated to the worker's specific occupation

In this simple exercise, we compare the qualification of the worker with the qualification associated with his/her specific occupation, we can see that this inefficiency does indeed exist. For this purpose, we follow the methodology by Fundação José Neves (2022) and define overqualification whenever the number of years of schooling of a worker is above the years of schooling associate with a certain occupation, with the latter being calculated as the mode of years of schooling for each occupation in a three-year period (year of analysis and the two previous years).

To simplify, for Figures A.5.1 and A.5.2 we have considered only 3 groups of qualifications – workers with basic education or less, with upper secondary education or with tertiary education. This exercise indicates that almost 70% of young workers with upper secondary can be classified as overqualified in manufacturing, compared to over 30% in the services sector. Moreover, close to 50% of young workers with tertiary can be classified as overqualified in both sectors. This inefficiency seems to be stronger for the younger cohorts (from 25 to 34 years old) relative to the rest (over 35 years old). Given these results, it seems there is evidence that we may not have an efficient combination between competences and occupations, and that it is important that those effects are being captured in our analysis.

Figure A.5.1. – Workers with upper secondary educ. in occupations in which the average worker has Basic educ. or less (in percentage)

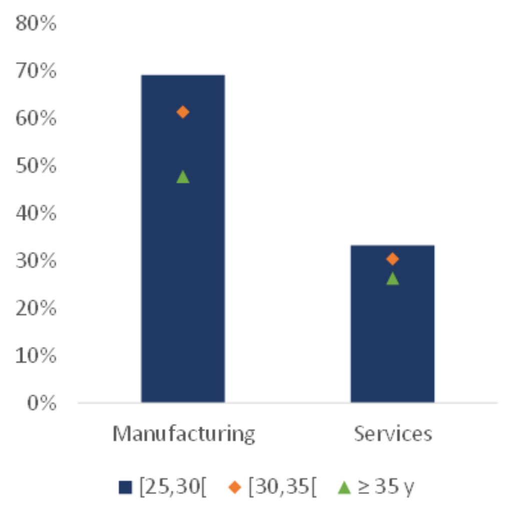
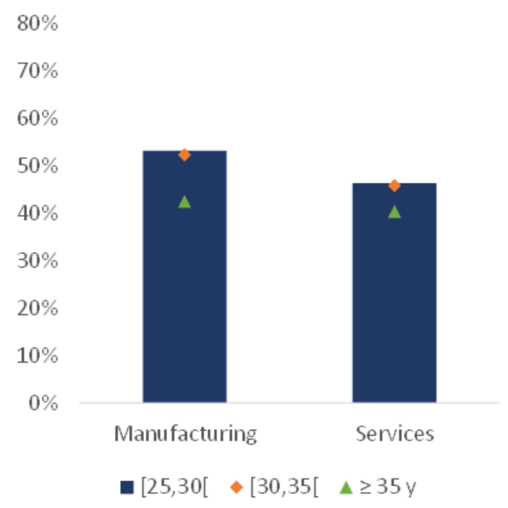


Figure A.5.2. – Workers with tertiary educ. in occupations in which the average worker has upper secondary educ. or less (in percentage)



Source. Own calculations using QP, SCIE.

Figure A.5.3. – Over qualification of workers with upper secondary education (in percentage)

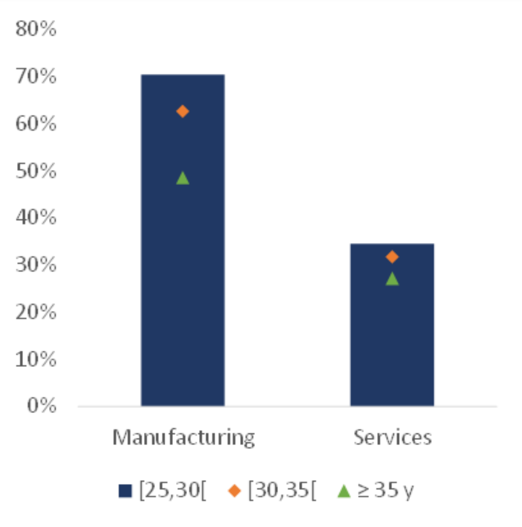
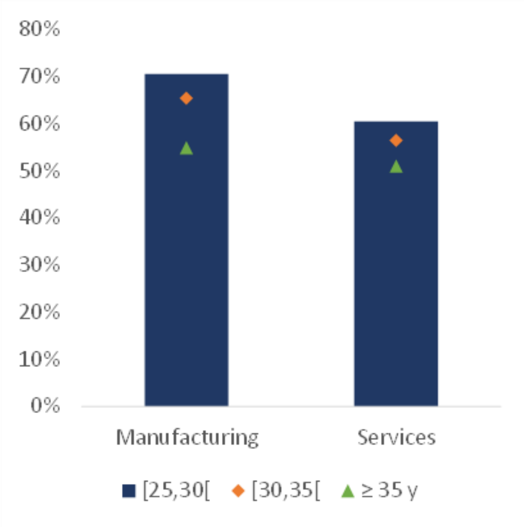


Figure A.5.4. – Over qualification of workers with tertiary education (in percentage)



Source. Own calculations using QP, SCIE.