Education Outcomes and the Labor Market

F. Obiols-Homs V. Sánchez-Marcos *

May, 2018

ABSTRACT

The quality of education appears to be negatively correlated with both the overeducation of workers at the tasks they perform and the unemployment rate across EU-15 countries, and positively correlated with the wage premium associated to tertiary education. We develop a model of the labor market with frictions to quantitatively investigate the impact of the education outcomes on the labor market. We show that both the ability of educated *and* non educated workers have sizable effects on the incentives of firms regarding the type of vacancies they open and also regarding the incentives of educated workers as of where to search for a job. Therefore education outcomes are relevant to understand the overeducation phenomena observed in the labor market. According to our quantitative analysis had the quality of education observed in Spain been similar to the European average then the overeducation rate would have been between 5 and 10 percentage points lower and the unemployment rate of the two types of workers would be reduced by 40%, but the tertiary education wage premium would be slightly smaller than in the benchmark economy.

KEYWORDS: OCCUPATIONAL-MISMATCH, TERTIARY EDUCATION WAGE PREMIUM, ABILITY
JEL CLASSIFICATION: I26, J21, J24

*F. Obiols-Homs Universitat Autònoma de Barcelona and Barcelona GSE; V. Sánchez-Marcos Universidad de Cantabria. An earlier version of this paper circulated under the title “Education, Occupation-Mismatch and Unemployment”. Part of this research was carried out while V. Sánchez-Marcos was visiting the Departament d’Economia i d’Història Econòmica at Universitat Autònoma de Barcelona and F. Obiols-Homs was visiting the IAE (CSIC). V. Sánchez-Marcos thanks the Spanish Ministry of Science and Technology for Grant ECO2015-64467-R (MINECO/FEDER). F. Obiols-Homs acknowledges the support of the Spanish Government through grants ECO2015-67602 and ECO2012-32392, and the financial support from the Spanish Ministry of Economy and Competitiveness through the Severo Ochoa Programme for Centers of Excellence in R&D (SEV-2015-0563).

1 Introduction

Across the EU-15 countries there is substantial variation in the skills of the adult population as measured by the average math score in the Programme for the International Assessment of Adult Competencies (PIAAC, 2013). As reported in the third column of Table 1 this figure ranges from 245 in Spain to 282 in Sweden and sizable differences remain even if only tertiary educated workers are considered (forth column). Interestingly, the average math score in the PIAAC is strongly and negatively correlated with two important labor market outcomes: the fraction of *mismatched* workers[1] (fifth column) and the unemployment rate of both tertiary educated and non educated workers (columns sixth and seventh). This negative correlation is also found if we restrict to the sub-sample of countries which share a similar fraction of tertiary educated workers. In particular, excluding Italy and Portugal, which report a fraction of tertiary educated workers clearly below the average of the EU-15, and Belgium and the UK, for which the different statistics are not measured with a comparable sample, the correlation of math scores with the fraction of *overeducated* workers is -0.84 and its correlation with the unemployment rate of both educated and non-educated workers is about -0.77. Furthermore, there is a strong and positive correlation of 0.7 between math scores and the wage premium to tertiary education.[2,3] Finally, there is empirical evidence supporting the importance of abilities as measured by the PIAAC to account for difference in income across countries, both at the aggregate level and at the individual level.[4]

The previous observations suggest that there may be a close connection between the outcomes of the education system, in terms of the quality of labor, and the degree of mismatch, the unemployment rate and the education wage premium. The purpose of this paper is to quantitatively investigate this connection by means of an equilibrium model. The literature based on equilibrium search to assess the effects of education policies on labor market outcomes is scarce. An important exception is Albrecht et al. (2009) which forecasts the long-run effects of a Swedish adult education program known as the *Knowledge Lift* implemented at the end of the nineties. This is an adult education program aimed at improving the skills of low-skilled workers towards the medium level. According to their analysis, as a result of the policy the fraction of vacancies tailored towards the medium-skill workers increases substantially. For each outcome measure, the treated gain most from the program. Those who have always been medium skilled also benefit, whereas those who remain low skilled suffer. We contribute to this literature and pose a search and matching model of the labor market à la Mortensen and Pissarides in which workers are heterogeneous in terms of their innate ability and in terms of their education. Our model,

[1] The notion of mismatch we adopt here is the definition of *vertical mismatch* proposed by Eurostat: individuals with at least tertiary education working in occupations for which the education requirement is lower. See the Appendix A for further details. In the literature this notion of mismatch is sometimes called *overeducation* or *over-qualification* (hence we will use it interchangeably) and there are several alternatives to measure it (see for instance Leuven and Oosterbeeck 2011 and the many references therein).

[2] In this calculation we exclude Denmark and Sweden because the labor markets in these countries are dramatically different to the markets in other European countries in relevant dimensions such as a large centralized bargaining, high female participation and social protection, amongst others.

[3] All the statistics provided in Table 1 are for 2007 in order to avoid the effect of the Great Recession after 2009 on labour market variables.

[4] Hidalgo-Cabrillana et al. (2017) find that differences in physical capital together with a broad measure of human capital that includes PIAAC ability account for 42% of the variance in output per worker, compared to only 27% when proxying human capital by average years of schooling only. At the individual level Hanushek et al. (2015) find that one-standard-deviation increase in numeracy skills as measured by the PIAAC is associated with a 18% percent wage increase among prime-age workers.

	Fraction Tert.	Tertiary Wage Prem.	PIAAC Scor.	Tert. PIAAC Scor.	Mismatch	Unemp. Below Tert.	Unemp. Tert.
Belgium[a]	31	1.43	276	310	20	0.07	0.03
Denmark	30	1.41	278	302	17	0.02	0.02
Germany	29	1.64	268	301	23	0.09	0,03
Ireland	31	1.60	254	285	30	0.07	0.03
Spain	29	1.51	245	278	34	0.10	0.05
France	26	1.66	254	295	18	0.06	0.04
Greece	24	-	-	-	-	0.04	0.04
Italy	12	2.09	249	280	13	0.05	0.03
Luxembourg	30	1.73	-	-	-	0.05	0.02
Netherlands	35	1.59	284	308	11	0.02	0.02
Austria	21	1.66	280	306	25	0.03	0.02
Portugal	11	2.55	-	-	-	0.06	0.04
Finland	31	1.66	280	305	13	0.06	0.03
Sweden	28	1.40	282	307	14	0.04	0.03
U.K.[b]	31	1.54	260	269	20	0.04	0.02
EU-15	27	1.68	268	295	19	0.05	0.03

Faction Tert.: fraction of tertiary educated workers from Education at a Glance OECD 2010, Table A1.1b. Tertiary Wage Prem: calculated from earnings as reported in Education at a Glance OECD 2010, Table A7.1 (page 127). PIAAC Scor.: PIAAC National Report. Chapter 3. Volume 1. Table 3.16 L y M. Mismatch: Eurostat Overeducation 2009. Table 8. Unemployment: Education at a Glance OECD 2010. Table A6.2b (male individuals aged 25 to 65).

Notes: (*a*) PIAAC data is for the Flemish Region only, (*b*) PIAAC is for England and North Ireland only.

Table 1: Several Statistics, EU-15

therefore, will be able to shed light on the relationship between the equilibrium allocation of ability and unemployment. The level of education of each worker is determined by an *education rule* which plays a twofold role: *selection*, it selects the abilities (i.e., workers) that receive higher education, and *quality*, it increases the effective ability of educated workers. The education rule is a shortcut to obtain an education outcome without the need to fully specify an education policy along the lines of those in place in actual economies. With this approach we focus our analysis on the impact of education outcomes in the labor market.

In the quantitative analysis we take as a benchmark the case of Spain. This choice is motivated by the fact that in Spain the average math score according to PIAAC is among the lowest in EU-15,[5] whereas the fraction of overeducated workers and the unemployment rate is among the highest in the EU-15. In addition, in Spain the wage premium to tertiary education is substantially smaller than in most of the EU-15 countries. We calibrate the model to mimic key observations of the Spanish economy in the mid 2000's and we conduct several counterfactual experiments to evaluate the effects of alternative education outcomes. For completeness we also

[5]According to Robles-Zurita (2017) the LOGSE (Spanish acronym for General Law of the Education System) reform passed in 1990 did not help to increase cognitive skills of the population, as measured by the PIAAC, despite an extension of compulsory years of education and postponement of the age of initial tracking into vocational and academic studies. In the Appendix A we provide a more detailed comparison of the education system and of the mismatch phenomena in Spain and in the EU-15 countries.

explore other possibilities related to differences in the sectors' productivity across countries.

Our findings support the view that enhancing the productivity of tertiary educated workers and implementing a more stringent selection of abilities into education would substantially reduce the fraction of overeducated workers. The intuition for this result comes from the basic mechanism at work in the AV family of models: an increase in the quality of educated workers produces an increase in its demand (and a reduction in the demand of the non educated). Thus it is worth to emphasize that the decrease in overeducation comes at the cost of an increase of the unemployment rate of non-educated workers. This result follows because the new education outcome mainly improves the productivity of educated agents relative to non educated workers. Once we target an education outcome such that the effective ability of *both* educated and non-educated workers is in line with what is observed in the average of EU-15, then we obtain a smaller, but still substantial reduction in overeducation, but a more notorious reduction in the unemployment rate of non-educated workers and a large increase in the Gross Domestic Product. The size of the reduction in overeducation depends however on the means by which the effective ability of educated workers is improved (selection vs quality). This result highlights the relevance of the distribution of abilities among the pool of workers searching for a job in a particular market to understand employment overeducation in equilibrium. We therefore view our results as suggesting that education outcomes have sizable effects on the labor market. We also explore the implications of a reduction (increase) in the productivity of the l-tech sector (h-sector) designed to meet the fraction of overeducated workers observed in the EU-15 countries. We find that such changes in productivity have a negative effect on the unemployment rate of ne-workers and a very modest effect on the wage premium to education. Thus the gap in terms of average skills for all workers between Spain and the EU-15 countries remains unexplained in this case.

A distinctive feature of our model is that there is a continuum of abilities. Both the distribution of abilities and the quality of labor are key variables that determine the profitability of vacancies posted by firms, which we assume that can be opened in a high and in a low-tech sector. Likewise, the education outcome of the rule determines the degree of competition among workers looking for jobs in each sector, which is relevant for them to choose where they would like to find a job. Thus with these assumptions the model is able to capture overeducation as educated workers accepting jobs in the low-tech sector. Since ability is continuously distributed, overeducation can happen to various degrees in an endogenous way. This is an important difference with respect to previous papers such us Albrecht and Vroman (2002) [AV] and Cuadras-Morató and Mateos-Planas (2013) [CMMP] in which mismatch is a binary event. Therefore our approach allows a more flexible mix of abilities and education than in previous papers which is convenient to undertake a meaningful quantitative analysis.[6]

In our model we abstract from job-to-job transitions and thus we focus on the persistent nature of overeducation. This is an important difference with respect to other papers in the literature in

[6] In this line of research see Blazquez and Jansen (2008) whom study the efficiency properties of equilibrium allocations in the AV model. More recently, CMMP introduce two education levels in the AV model and quantitatively study the effects of skill bias technological change (SBTC) in the US with respect to overeducation. See also Krusell et al. (2000) for an earlier application to the U.S. economy. The literature on the SBTC tries to account for mismatch and the skill premium by changes in the relative demand of educated workers. Our approach here is to asses the ability of changes in the relative supply and quality of skilled labor. Regarding education choices, Charlot and Decreuse (2010) study the efficiency in a similar model and show that overeducation (in the sense of too many individuals choosing to acquire education) arises since workers do not internalize the impact of their decision on the wage and employment perspectives of others.

which overeducation is a way for workers to find their best match in the labor market and thus it is a transitory phenomena.[7] Our choice is motivated by the empirical evidence reported in Hidalgo-Pérez et al. (2015) suggesting that in Spain the fraction of mismatched college workers decreases very moderately with age (from 60% in the age group 30 to 34 to 50% in the age group 50 to 54).[8] This persistence of over-qualification is also consistent with the findings in Montalvo (2013) and Meroni and Vera-Toscano (2017), hence overeducation in Spain does not appear to be a transitory phenomena affecting just a reduced age-specific group of workers.[9] Our approach, therefore, can be seen as complementing previous work exploring the implications for transitory over-qualification phenomena when job-to-job transitions are allowed. In particular, Dolado et al. (2009) extend the model in AV and, among other things, they show that transitory skill mismatch by over-qualified workers is more harmful to the prospects of less-educated workers than permanent mismatch and that on-the-job search widens the wage differences among the highly educated workers. However, the quantitative work in Dolado et al. (2009) focuses on the U.S. and on a European average for which the possibility of transitory mismatch may be a more appropriate assumption compared to the Spanish case. Finally, the nature of mismatch at the center of our investigation is also different from the one stressed in other papers investigating mismatch as a result of frictions preventing sectoral and geographical adjustment of employment and its dynamics over the business cycle.[10]

Our work is also related to Dolado et al. (2000) exploring the importance of labor market institutions (job separation rate and the replacement rate) to understand the crowding-out of lower educated workers from their traditional entry jobs by higher educated workers. As emphasized above, our focus is instead on the importance of education outcomes to understand the aforementioned facts. An alternative explanation for the mismatch phenomena is provided by Marimon and Zilibotti (1999) who show that the more generous unemployment benefits in Continental Western Europe relative to the U.S. are able to explain the higher unemployment rates, the better quality of the matches between workers and jobs (i.e., the smaller occupation mismatch), and lower wage inequality observed in Europe than in the U.S. However, unemployment benefits in Spain are comparable to the ones in other European countries, and thus the argument runs counter to the higher overeducation observed in Spain.[11] Finally, our approach can also be seen as an alternative to the view that there are *demand factors* that may be able

[7]To obtain this sort of experimentation as an equilibrium outcome it is necessary to consider a model including heterogeneity -as we do- and asymmetric or incomplete information, a feature that our model abstracts from. Prominent examples in this line of research include Jovanovic (1979), Miller (1984) and more recently Papageorgiou (2014).

[8]These authors use a sample of the Social Security Records of the Spanish population (Muestra Continua de Vidas Laborales, MCVL) to explore the puzzling fall in the wage skill premium in Spain over the last decades. They also study the evolution of occupational mismatch among college graduated workers.

[9]Montalvo (2013) uses the Spanish School to Work Transition database to study these questions and finds that over-qualification is a very absorbing state since transition matrices show that the probability to continue overqualified after moving to a new job is 76%. Meroni and Vera-Toscano (2017) find evidence of a systematic trap into overeducation in Southern Europe using REFLEX survey.

[10]This notion of mismatch could be due for instance to workers looking for jobs in occupations that do not correspond to the field of education they have attended, known as *horizontal mismatch* in the statistics produced by Eurostat. Examples in this line of research include Sahin et al. (2014), Dvorkin (2013) and Guvenen et al. (2015).

[11]Unemployment benefits are multidimensional and thus it is not straightforward to choose the relevant dimensions of comparison. Stovicek and Turrini (2012) report evidence suggesting that Finland, the Netherlands and Spain are particularly generous in terms of replacement rates and duration of benefits in comparison with other European countries. For additional details on other OECD countries see also http://www.oecd.org/els/benefits-and-wages-statistics.htm.

to explain the relatively high unemployment and overeducation and the low wage premium to education observed in Spain. In fact, Díaz and Franjo (2016) use a version of the Neoclassical model of growth to report an inefficiently high investment rate in residential investment but too low in Investment Specific Technical Change. These authors conduct a growth accounting exercise and use a representative agent model in a frictionless economy which prevents them from addressing the main issues in our investigation.

The paper is organized as follows. In section 2 we develop the model economy that we use as framework for our analysis. In section 3 we discuss the calibration of the model to match relevant statistics of the labor market and education outcomes in Spain. In section 4 we undertake the quantitative analysis to assess the ability of different education policies to account for the differences between Spain and the average of the EU-15 countries in terms of labor market outcomes. We also explore some alternative explanations. Finally, section 5 concludes and the Appendix contains the details regarding the data and additional results that complement our analysis.

2 The Model

Time is continuous and in the economy there is a mass one of infinitely lived workers which are endowed with an ability level a. The key feature of our model is that ability is distributed according to a continuous density $\lambda(a)$ on a set of possible abilities A. We also assume that workers differ in their education level: some of them are *educated*, denoted e, and some of them are not, denoted ne. Thus, unlike ability, education is a discrete variable with only two mass points.

We think of the probability of each ability to receive education and of the effective ability after education as the education outcomes in the economy. Therefore, we assume there is a selection rule $\sigma(a) : A \to [0, 1]$, which indicates the fraction of agents with education amongst those with ability level a. We use $\mu(a) = \sigma(a)\lambda(a)$ to denote the fraction of (educated) e-agents with innate ability level a. Furthermore, we assume that innate ability is mapped into effective ability, $\tilde{a}_j$ for $j = e, ne$, as follows

$$\tilde{a}_j = \psi_j a, \tag{1}$$

with $\psi_e \geq \psi_{ne} = 1$. Thus it is natural to think of ψ_e as the quality of education.

In the production side of the economy there are firms/jobs that are either vacant or filled. These jobs differ in the minimum education requirement that a worker needs to satisfy to be able to successfully operate the corresponding technology. This means that there are firms with a technology such that ne-workers are unable to properly operate. We refer to these firms as *high-tech* firms, denoted h. Also, there are firms such that their technology can be operated by both educated and non educated workers, which we informally label as *low-tech* firms, and denote them by l. We denote by $y_{ij}(a)$ the output of a firm type $i = h, l$ employing a worker with education $j = e, ne$, and ability level $a \in A$. We assume that $y'_{ij}(a) > 0$, so that for all worker types and sectors output is larger the larger is the ability of the worker. Slightly abusing from notation, below we will denote by μ_{ij} the mass of j-educated agents that are employed or are looking for a job in the i-sector. Finally, creating a vacancy has a cost c_v and an employment relationship breaks up at exogenous rate δ_i. Once unemployed a worker receives unemployment benefits b.

We follow the Mortensen-Pissarides tradition and we assume that there are frictions in the labor market, such that both firms and workers need to spend some resources before a productive match can be formed. These frictions are captured by a matching function relating the number of new matches to the number of unemployed workers and to the number of outstanding vacancies. Hence, in this formulation of the labor market externalities due to congestion naturally arise and play an important role in shaping the equilibrium configuration. Notice that given the technological constraint about the education requirements, it is clear that *ne*-workers would never look for a job in the *h*-sector, hence in this sense the labor market is segmented by education. The assumptions on technology place no restriction on educated workers being able to operate the low-tech technology, and yet, we cannot rule out that the labor market be additionally segmented by ability: it is possible that some educated workers (presumably with low ability) choose to search jobs in the low-tech sector. This is the notion of overeducation that we study in this paper.[12] In order to better focus on this issue, we will assume that unemployed workers can only search for a job in one market, hence educated workers must choose beforehand whether to search for a job in the high or in the low sector. Likewise, a firm willing to create a vacancy needs to choose beforehand the sector in which it will be created.[13]

Given these assumptions the number of productive matches in sector $i = h, l$ is given by a constant returns to scale matching function $M(v_i, x_i)$ defined on the number of vacancies (v_i) and the mass of unemployed workers (x_i) participating in the corresponding market. The matching functions satisfy $M(v_i, x_i) = m(\theta_i)x_i$, where $\theta_i = v_i/x_i$ and $m(\theta_i) = M(\theta_i, 1)$. This means that the probability of an unemployed worker finding a vacancy, and the probability of a vacant position to be filled with an unemployed worker, are given respectively by $m(\theta_i)$ and $m(\theta_i)/\theta_i$.

2.1 The problem of a worker

Workers are assumed to be risk neutral and thus they maximize the present value of income: wages and unemployment benefits. We denote $w_{ij}(a)$ the wage of a worker type $j = e, ne$, with ability level a, who is matched to a firm in sector $i = h, l$, and we denote $W_{ij}(a)$ the value of this match. Similarly, $U_{ij}(a)$ stands for the value of searching for a job in sector $i = h, l$, for a type $j = e, ne$ worker with ability level a. The asset value of employment for a worker is given by:

$$rW_{ij}(a) = w_{ij}(a) + \delta_i(U_{ij}(a) - W_{ij}(a)), \tag{2}$$

for $i = h, l$, $j = e, ne$, all $a \in A$, and where r is the discount rate. The equation states the usual no arbitrage condition stressed in the literature: that the flow value of a type-j worker with ability level a who is employed in a type-i firm equals the sum of the flow return $w_{ij}(a)$ plus the expected instantaneous capital loss $\delta_i(U_{ij}(a) - W_{ij}(a))$ (from $W_{ij}(a)$ to $U_{ij}(a)$ which happens with probability δ). Likewise, the asset value of looking for a job in the i-sector for a worker with education level j and ability level a is given by

$$rU_{ij}(a) = b + m(\theta_i)\left\{W_{ij}(a) - U_{ij}(a)\right\}, \tag{3}$$

which has a similar interpretation to the previous one about value of employment ($m(\theta_i)$ is the arrival rate of a job offer to a worker in the i-sector). In the current environment overeducation

[12]See Herz and Van Rens (2011) for a related notion of potential inefficiency: the excessive unemployment above the level a planner would have chosen (*inefficient unemployment*).

[13]Saint-Paul (1996) and Cuadras-Morato and Mateos-Planas (2006) introduce similar assumptions.

may arise if for some ability level we have that an e-worker looks for (and accepts) jobs in the l-sector. That is, overeducation occurs when there is a subset $\tilde{A} \subseteq A$ such that $U_{he}(a) \leq U_{le}(a)$ for $a \in \tilde{A}$.

2.2 The problem of the firm

Firms create vacancies at a cost c_v irrespectively of the sector of operation, and we denote V_i for $i = h, l$ the value of a newly created vacancy that is not yet operative because it is vacant. The value of an operative match between a job in sector i and a worker type j and ability a is given by $J_{ij}(a)$, and it satisfies:

$$rJ_{ij}(a) = y_{ij}(a) - w_{ij}(a) + \delta_i[\max_{i' \in \{h,l\}} V_{i'} - J_{ij}(a)]. \tag{4}$$

This equation states that the flow value of an operative position equals the flow value of output $y_{ij}(a)$ net of labor cost, $w_{ij}(a)$, plus the expected change in its capital value: the match will be broken with probability δ_i and in that event the firm will be allowed to choose again the sector of operation ($\max_{i' \in \{h,l\}}$). After the optimal choice of sector of operation, i', the asset value will change in the amount $V_{i'} - J_{ij}(a)$.

The value of creating a vacancy in the h-sector satisfies:

$$rV_h = -c_v + \frac{m(\theta_h)}{\theta_h} \left\{ \max\{E_\mu[J_{he}(a)] - V_h, 0\} \right\}. \tag{5}$$

The flow value of creating a vacancy, rV_h, equals its cost of creation, $-c_v$, plus the expected change in the asset value due to filling the vacancy with a suitable worker. The flow probability of a vacancy being match with a worker in h-sector is $m(\theta_h)/\theta_h$. The max operator reflects the fact that it may not be profitable for a firm in the h-sector to offer a job to an educated worker (if her ability level is too low). Accordingly, E_μ in the expression above is the expectation conditional on meeting an educated worker as implied by the measure $\mu(a)$. We also have

$$\begin{aligned} rV_l = -c_v + \frac{m(\theta_l)}{\theta_l} \bigg\{ & \frac{x_{le}}{x_l} \left(\max\{E_\mu[J_{le}(a)] - V_l, 0\} \right) \\ & + \frac{x_{lne}}{x_l} \left(\max\{E_\mu[J_{lne}(a)] - V_l, 0\} \right) \bigg\}, \end{aligned} \tag{6}$$

for the case of a vacancy in the low-tech sector. In this case the flow value of opening a vacancy in the l-sector, rV_l, equals its cost of creation, $-c_v$, plus the expected change in its asset value, which depends of the type of the worker that meets the vacancy. In the previous expression $m(\theta_l)/\theta_l$ is the probability of a match between a vacancy and an unemployed worker. Also, x_{le} stands for the mass of educated unemployed workers searching for a job in the low-tech sector (x_{lne} is the corresponding number of non educated workers, and $x_l = x_{le} + x_{lne}$). Thus $(m(\theta_l)/\theta_l)(x_{le}/x_l)$ is the probability of meeting an e-worker who is searching in the l-sector, and $(m(\theta_l)/\theta_l)(x_{lne}/x_l)$ is the probability of meeting an unemployed ne-worker. As before, E_μ stands for the conditional expectations operator as implied by μ, the distribution of education and ability. Hence, if an e-worker meets a vacancy the asset value is expected to change in $E_\mu[J_{le}(a)] - V_l$ (provided that the vacancy is filled, and zero otherwise). If, however, the match involves the vacancy and a non educated worker then the asset value is expected to change in $E_\mu[J_{lne}(a)] - V_l$. Finally, in the equilibrium we consider we will assume free entry, so that $V_h = V_l = 0$ will hold.

2.3 Wage setting rule

We assume that once an unemployed worker is matched to a posted vacancy, the firm and the worker engage in a Nash bargaining process in order to split the surplus that the match may potentially create. Under these assumptions the wages satisfy

$$w_{ij}(a) = \text{argmax}\,(W_{ij}(a) - U_{ij}(a))^{\beta}\,(J_{ij}(a) - V_i)^{1-\beta}\,, \tag{7}$$

(where $\beta \in (0,1)$ represents the bargaining power of the workers), which is obtained by satisfying the FOC of the bargaining problem:

$$(1-\beta)(W_{ij}(a) - U_{ij}(a)) = \beta(J_{ij}(a) - V_i). \tag{8}$$

2.4 Stationary equilibrium

To simplify the exposition we introduce here an assumption that will also be useful in our quantitative analysis. In particular, we assume that the technology to produce goods is linear in ability:

$$y_{ij}(a) = y_i + \tilde{y}_{ij}\tilde{a}_j. \tag{9}$$

The term y_i captures the component of production that is sector-specific and unrelated to the ability of the worker operating the technology. The term $\tilde{y}_{ij}$ allows us to capture the fact that marginal productivity of ability may be both education and sector specific.[14] Subtituting Eq. (2), Eq. (3) and the expression for $J_{ij}(a)$ from Eq. (4) after imposing the free entry condition $V_i = 0$ in Eq. (7) we obtain:

$$w_{ij}(a) = \frac{\beta(r+\delta_i+m(\theta_i))y_{ij}(a) + (1-\beta)b(r+\delta_i)}{r+\delta_i+\beta m(\theta_i)}, \tag{10}$$

which after substituting Eq. (9) can be written as

$$w_{ij}(a) = w_i + w_{ij}a, \tag{11}$$

with

$$w_i = \frac{\beta(r+\delta_i+m(\theta_i))y_i + (1-\beta)b(r+\delta_i)}{r+\delta_i+\beta m(\theta_i)}, \tag{12}$$

and

$$w_{ij} = \frac{\beta(r+\delta_i+m(\theta_i))y_{ij}}{r+\delta_i+\beta m(\theta_i)}, \tag{13}$$

and where $y_{ij} = \tilde{y}_{ij}\psi_j$. Hence wages in each sector and for each type of worker are a linear function of the ability of the worker a (albeit they depend non linearly on the relevant θ_i). This characterization is useful because it allows us to write the asset value of unemployment in each sector also as a linear function of a: inserting Eq. (2) into Eq. (3) and rewriting produces:

$$rU_{ij}(a) - b = m(\theta_i)\frac{[w_{ij}(a) - b]}{r+\delta_i+m(\theta_i)}, \tag{14}$$

[14] We discuss in section 3.2 that by considering separately the effect of education on ability (by the term ψ_e), and the associated marginal productivity of effective ability in production (by the term $\tilde{y}_{ij}$) will help us to calibrate the model in a transparent way and consistently with the empirical observations on the distribution of ability and the average effective ability of tertiary educated workers.

which using the above expressions for wages can be written as:

$$rU_{ij}(a) - b = u_i + u_{ij}a, \tag{15}$$

with

$$u_i = m(\theta_i)\frac{\beta(y_i - b)}{r + \delta_i + \beta m(\theta_i)}, \tag{16}$$

and where

$$u_{ij} = m(\theta_i)\frac{\beta y_{ij}}{r + \delta_i + \beta m(\theta_i)}. \tag{17}$$

Thus with the linearity in a of $U_{ij}(a)$ there can be overeducation if the straight lines described by $U_{he}(a)$ and $U_{le}(a)$ cross for some $a \in A$ for the given θ's. For instance, in case of *positive assortative matching* there is a threshold $\bar{a}$ such that $U_{he}(\bar{a}) = U_{le}(\bar{a})$ and $U_{he}(a) > U_{le}(a) \iff a \geq \bar{a}$ (and thus $\tilde{A} = \{a \in A : a \leq \bar{a}\}$).[15] Of course, the model also allows the case of no overeducation (when $U_{he}(\bar{a})$ and $U_{le}(\bar{a})$ do not cross in the positive orthant). Figure 1 portrays an example with potential overeducation and a situation without it.

This characterization of the possibility of overeducation greatly simplifies the notion of stationary equilibrium.[16] Specifically, given a θ_i for $i = h, l$ we can use the value of unemployment for an e-worker in each sector using Equations (15)-(17) to pin down a value for $\bar{a}$. Given this value $\bar{a}$ then the joint distribution of education and ability determines the distribution of the labor force across sectors (i.e., μ_{ij}). The value $\bar{a}$, the distribution of the labor force and the stationary flow conditions of the labor market are then used used to determine the right hand side of Equations (5) and (6). Hence the values θ_i for $i = h, l$ constitute an equilibrium if the values of creating a vacancy are such that $V_i = 0$ (due to free entry). More precisely,

Definition: Given $\lambda(a)$ and $\sigma(a)$ implying $\mu(a)$, a Stationary Equilibrium consists of a list θ_h, θ_l such that:

i) $\bar{a}$ is determined as $U_{he}(\bar{a}) = U_{le}(\bar{a})$ when the value of unemployment is given by Eq. (15)-(17).
ii) The distribution of the labor force is consistent with $\mu(a)$ and the $\bar{a}$ implied by θ_h, θ_l:

$$\mu_{lne} = 1 - \int_A \mu(a)da, \ \mu_{le} = \int_{a\leq\bar{a}} \mu(a)da, \ \text{and} \ \mu_{he} = \int_{a\in A} \mu(a)da - \int_{a\leq\bar{a}} \mu(a)da. \tag{18}$$

where μ_{ij} stand for the mass of j-educated agents in the i-sector.
iii) Labor markets are stationary:

$$x_{he}m(\theta_h) = \bar{x}_{he}\delta_h, \ x_{le}m(\theta_l) = \bar{x}_{le}\delta_l, \ \text{ and } \ x_{lne}m(\theta_l) = \bar{x}_{lne}\delta_l, \tag{19}$$

where $\bar{x}_{ij}/x_{ij}$ stand for the mass of employed/unemployed j-educated agents in the i-sector and thus $\bar{x}_{ij} + x_{ij} = \mu_{ij}$.
iv) The following free entry conditions hold:

$$0 = -c_v + \frac{m(\theta_h)}{\theta_h}\left\{E_\mu[J_{he}(a)|a \geq \bar{a}]\right\}, \tag{20}$$

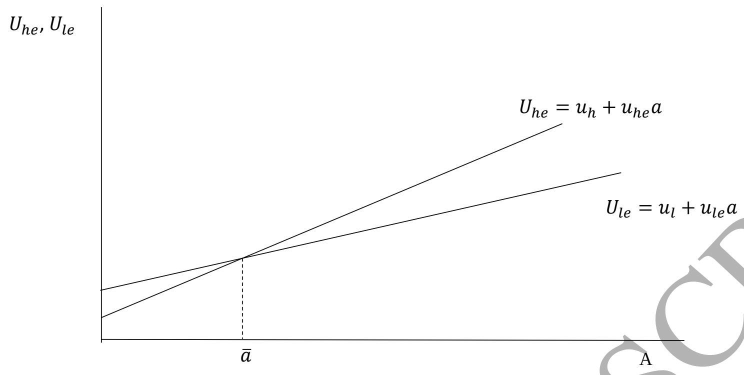
and

$$0 = -c_v + \frac{m(\theta_l)}{\theta_l}\left\{\frac{x_{le}}{x_l}\left(E_\mu[J_{le}(a)|a \leq \bar{a}]\right) + \frac{x_{lne}}{x_l}\left(E_\mu[J_{lne}(a)]\right)\right\}, \tag{21}$$

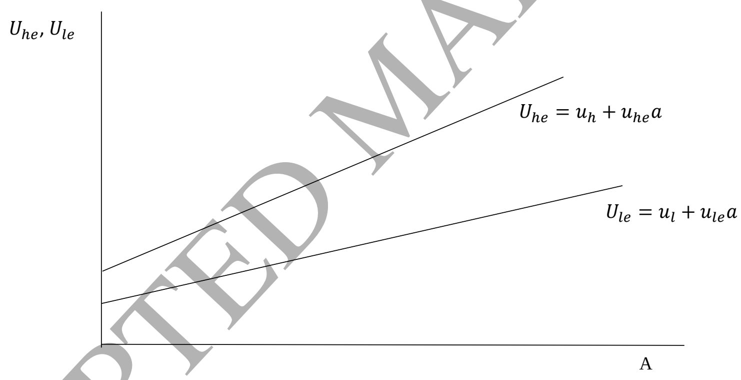
[15] The case of positive sorting is the one empirically relevant (see CMMP and the references therein), thus in our quantitative analysis we disregard negative sorting.

[16] We thank an anonymous referee for suggesting to us the approach to the definition of equilibrium.

Figure 1: Different types of equilibrium



A. Overeducation with positive sorting.



B. The case of no overeducation.

with

$$J_{ij}(a) = \frac{y_{ij}(a) - w_{ij}(a)}{r + \delta_i}, \tag{22}$$

and where $y_{ij}(a)$ satisfies Eq. (9) and $w_{ij}(a)$ satisfies Equations (11)-(13).

As previously noted in AV and in CMMP in similar models with a discrete number of ability levels there are three possible equilibrium configurations which are respectively characterized by (1) *ex-post segmentation*, when all educated workers work or look for jobs in the h-sector (remember that non educated workers can operate only the technology of the l-sector), (2) employment *overeducation* , which is observed when some educated workers look for and accept jobs in the l-sector, or (3) the case of multiple equilibria in which both types are simultaneously possible.[17]

With the linearity of the value of unemployment with respect to ability for an equilibrium with employment overeducation to exist it is necessary that in Equations (16) and (17), $u_h \leq u_l$ and $u_{he} \geq u_{le}$ (with at least one strict inequality). These two conditions can be summarized as

$$\frac{y_h - b}{y_{he}} < \frac{y_l - b}{y_{le}}. \tag{23}$$

In this case ability displays comparative advantage in the h-sector and thus educated but low-ability workers end up looking for jobs in the l-sector.[18] It is clear from the necessary condition in Eq. (23) for the existence of positive sorting that there are many parameter configurations that are compatible with such mismatch. Before we continue we briefly discuss the connections between the previous condition and the results in AV and CMMP.

Remarks

Remark 1: In line with the results in CMMP, it is clear from Equation (23) that SBTC consisting in increasing y_{he} relative to y_{le} (the marginal productivity of ability in the h-sector relative to the l-sector) will favor the existence of employment overeducation. Hence our condition in Equation (23) offers a new insight for the existence of overeducation based on increased comparative advantage of higher ability e-workers in the h-sector.

Remark 2: The fact that the above sort of SBTC is able to give rise to employment overeducation is not possible in the AV model, in which search is undirected (there is a single labor market) and thus increasing y_{he} relative to y_{le} tends to reduce overeducation favoring an equilibrium with ex-post segmentation. Without disregarding the importance of undirected search, we notice that in our model with directed search a SBTC consisting in increasing y_h relative to y_l (that is, the sector specific parameter in the technology) will produce the same effects as in the AV model.

Remark 3: There may be employment overeducation as long as y_l is large relative to y_h. Thus, the costs of operating a vacancy stressed in CMMP as a necessary condition to generate overeducation appear to be irrelevant once production depends not only on the ability of the agent but also on the sector where she is (potentially) employed.

[17]We notice that in addition to these possibilities, in our model with a continuum of abilities we cannot rule out the possibility of multiple equilibria of the employment overeducation type: all that would be required are congruent expectations about θ_i. All in all, in our quantitative work in section 4 we numerically check that the equilibrium we find is in fact unique.

[18]See Sattinger (1975) for an early development of a sorting condition along this lines. If the opposite inequality holds then the equilibrium is characterized by *negative assortative matching*, and so high-ability workers would end up looking for jobs in the l-sector. For completeness, it is worth mentioning that the theoretical model admits additional forms of mismatch but they violate the assumption that $y_{ij} > 0$.

2.5 Discussion of the effect of the education rule

It is instructive to briefly discuss the potential effects of education as is represented in the model and to develop the intuition for the results in the following sections. In the current model the education outcome is the result of two different components: the selection rule of abilities to receive education and the labor productivity enhancement that education provides to the individual. These components are likely to shape the equilibrium outcomes and, in particular, have a sizable effect on the fraction of overeducated workers.

First, the fraction of overeducated workers depends on the distribution of education on the support A, i.e., on how the abilities to receive education are selected. For instance, for a given value of $\bar{a}$ as in Figure 1, there would be no overeducation if all educated workers happen to have a larger than $\bar{a}$. We capture the selection rule by $\sigma(a)$, the fraction of agents with ability level a that receive education. Consider the implementation of a *more stringent* selection rule such that more higher ability workers get education and fewer lower ability workers do so. The top panel of Figure 2 (in the figure $\Lambda(a)$ is the cdf of $\lambda(a)$) portrays an example along these lines: the resulting distribution of education with the new rule $\sigma(a)'$ First Order Stochastically dominates the distribution under the old rule $\sigma(a)$. It is clear that with a more selective rule overeducation -measured as the height of the function $\int \lambda(a)\sigma(a)da$ at a given $\bar{a}$- will unambiguously decrease irrespectively of the position of $\bar{a}$. The First Order Stochastic dominance property of a new selection rule is critical to reduce overeducation by improving selection. To see this, consider the lower panel in Figure 2. The new rule decreases the mass of educated agents at lower levels of ability, but it does not necessarily increase the mass at the higher ability levels. The figure shows that overeducation would decrease if the indifferent ability level is given by $\bar{a}$, but overeducation would increase if the indifferent ability level is $\bar{a}'$.

Second, the labor productivity enhancement parameter ψ_e that is embedded in y_{he} and in y_{le} is also likely to affect overeducation. In particular, the slopes of $U_{he}(a)$ and $U_{le}(a)$ in Equation (17) increase if ψ_e increases. With given θ_i's then the representation in Figure 1 for overeducation with positive sorting obtained when condition (23) holds implies that after an increase in ψ_e not only the value of unemployment in each sector will certainly increase for an e-worker, but also that $\bar{a}$ will shift to the left and thus overeducation will decrease.

Finally, changes in σ, ψ_e and ψ_{ne} will also have an effect on the degree of tightness in each sector, θ_i, which in turn will indirectly affect the determination of $\bar{a}$. In particular, it is clear from Equations (16) and (17) that increases in θ_i leading to larger $m(\theta_i)$ will increase both the intercept and slope of the $U_{ij}(a)$ curves in Equations (16) and (17), hence the final effect on $\bar{a}$ will depend on the relative shifts of the two curves. The intuition from the usual search and matching model (e.g. Diamond 1982, Mortensen 1982 and Pissarides 1990) suggests that θ_i is larger the larger is the profit to a firm of forming a match. Roughly speaking, θ_i is expected to be larger the higher is ψ_i as well as the easier is to form a match with a higher ability worker of the desired type. With this in mind Figure 3 offers a different insight of the effects of a more stringent selection rule. The new rule improves the average ability among educated workers and worsens it among non educated workers. The implication of this fact is that θ_h will tend to increase whereas θ_l will tend to decrease. These changes in θ_i's induce the corresponding upward and downward shifts represented in Figure 3. Notice that the level $\bar{a}$ decreases not only to $\bar{a}'$ but to $\bar{a}''$ precisely because improving selection for e-workers implies at the same time a worse selection for ne-workers.

The previous discussion suggests that the fraction of overeducated workers is jointly determined by the distribution of abilities among educated and non educated workers and by the position of the threshold level $\bar{a}$. The sort of policy changes that we explore in the following sections usually involve simultaneous shifts of both the distribution of abilities and of $\bar{a}$. Since sometimes these shifts operate in opposite directions a quantitative analysis is needed in order to assess their implications.

3 Quantitative Analysis

We fix functional forms and we discipline our quantitative exercise with a calibration of model parameters grounded on relevant statistics.

3.1 Functional forms

We assume that the matching functions are Cobb-Douglas of the form

$$M(v_i, x_i) = m_i v_i^{\eta} x_i^{1-\eta}, i = h, l, \tag{24}$$

where $\eta \in (0, 1)$ measures the vacancy-elasticity of the matching function. This assumption is in line with most of the quantitative literature about frictional labor markets (see for instance the closely related papers by AV and CMMP).

We assume that the distribution of ability $\lambda(a)$ is Pareto of parameters a_m and α, so that the density satisfies

$$\lambda(a) = \alpha \frac{a_m^{\alpha}}{a^{\alpha+1}} \tag{25}$$

if $a \geq a_m$ and zero otherwise. We require this density to have finite mean and variance hence we assume $\alpha > 2$. With respect to the education outcomes we explore the implications of a general selection rule such that for all $a \geq a_m$:

$$\sigma(a) = \sigma_0 + \sigma_1 \left(1 - \frac{a_m}{a}\right). \tag{26}$$

Notice that if $\sigma_1 = 0$, then the fraction of educated workers is the same for all ability levels, and that if $\sigma_1 > 0$, then the fraction of educated workers increases with the level of ability. Finally, notice that the function $\sigma(a)$ is bounded, strictly increasing and strictly concave. Under these assumptions we have that

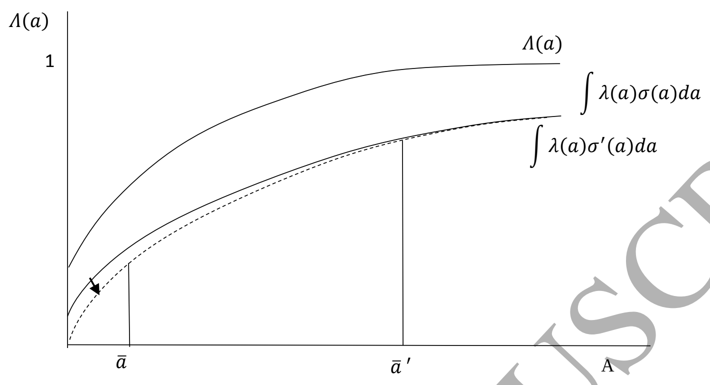
$$\mu(a) = \sigma(a)\lambda(a) = \frac{\mu_0}{a^{\alpha+1}} - \frac{\mu_1}{a^{\alpha+2}}, \tag{27}$$

where $\mu_0 = (\sigma_0 + \sigma_1)\alpha a_m^{\alpha}$, and that $\mu_1 = \sigma_1 \alpha a_m^{1+\alpha}$. The two-parameter family of selection rules is convenient because it allows us not only to control for the mass of educated agents, but also for their average ability.

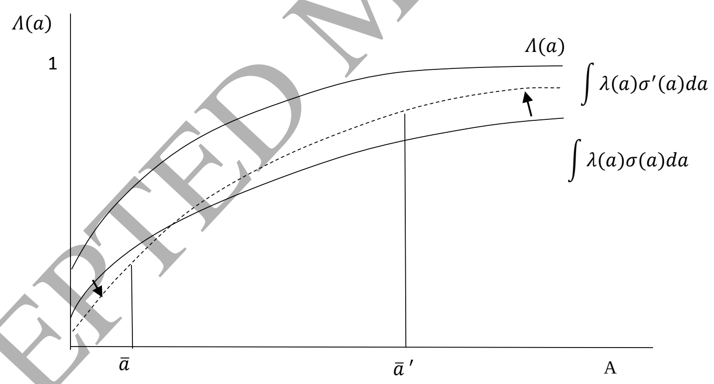
3.2 Calibration

In our calibration strategy there is a first block of parameter values that we borrow directly from existing studies in the related literature. This is the case of the worker's bargaining power

Figure 2: Selecting for education

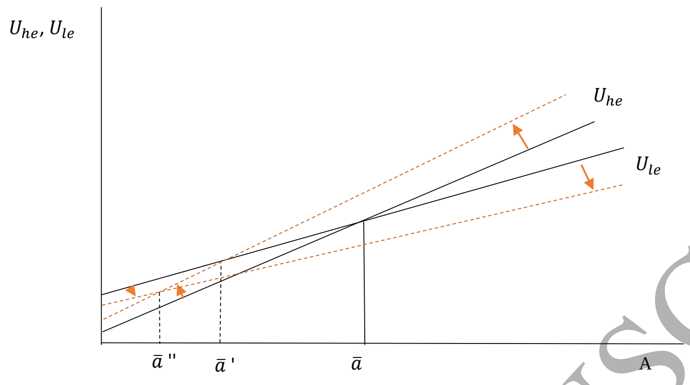


A. Unambiguous effect on overeducation.



B. Ambiguous effect on overeducation.

Figure 3: A more stringent selection



β, which we fix at 0.5, the parameters that govern the matching technology ($m_h = m_l = 1$, $\eta = 0.5$) and the quarterly interest rate r, which is set to 0.013. These are all the same as in AV. Hobijn and Sahin (2009) estimate a quarterly separation rate of 0.07 for the Spanish economy. Consistently with this estimate we fix $\delta_h = \delta_l = 0.07$. In addition we normalize ψ_{ne} to 1 and y_l is normalized so that the productivity of the workers with the smallest ability in the low-tech sector is equal to 1.

Second, we fix the parameters that govern the distribution of innate ability. In particular we identify a_m and α by targeting the mean and dispersion in the PISA scores (Science) for Spain in 2006, which are respectively 4.88 and 0.19 (targeting the mean in the PISA score is simply a normalization criteria). We therefore fix $a_m = 4.10$ and $\alpha = 6.3$ and these two parameter values pin down the distribution of innate ability.[19]

Third, we calibrate the parameters governing the education rule σ_0, σ_1 and ψ_e. One important target for the identification of ψ_e is the relative roles played by the *selection of abilities that receive education* and by the *quality of education* in shaping the score of tertiary e-workers relative to ne-workers. Although the evidence on this is scarce (see for instance Fang 2006 and Hendricks and Leukhina 2014), the results in Fang (2006) suggest that about two thirds of the wage premium in the US is accounted for by productivity enhancement of college attendance. In our analysis we pursue this target. Since ψ_{ne} is normalized to 1, we need to fix $\psi_e = 1.15$.[20] Note that if ψ_e was assumed to be 1 the ability gap between tertiary educated and non educated workers would all be due to *selection* into education alone. With this parameter fixed we select σ_0 and σ_1 to target the fraction of individuals with tertiary education according to European Union Survey of Income and Living Conditions (EU-SILC) in 2007, which is 0.31, and the mean score in math test of individuals with tertiary education *relative* to non educated individuals according to the results from PIAAC (2013), which is 1.2.[21] Within the model this is equivalent to an average *effective ability* of 5.77 for tertiary educated workers and of 4.81 for non educated individuals. Table 2 contains this second set of parameter values and relevant data targets.

Finally we calibrate c_v, y_h, $\tilde{y}_{he}$, $\tilde{y}_{le}$, $\tilde{y}_{lne}$ and b to match specific targets of the Spanish labor market, which are reported in Table 3. In particular, we restrict parameter values to be consistent with: (i) the incidence of unemployment across education groups, (ii) the tertiary education wage premium and, finally, (iii) the degree of inequality in the labor market within each education category (thus we restrict the equilibrium to be consistent with the coefficient of variation of wages, denoted CV, according to the level of education of the workers). These statistics provide us with the information needed to calibrate the parameters that govern the importance of ability to determine productivity and wages in each sector and the relative wages across different education groups. To this end we use microdata from EU-SILC (2007) and find that $CV_e = 0.38$ and $CV_{ne} = 0.27$.[22] For the sake of consistency we use this same database to calculate

[19]Cubas, Ravikumar and Ventura (2013) also proxy the distribution of talent in several countries using the distribution of PISA scores.

[20]In the benchmark economy the tertiary educated wage premium is 1.45. In the absence of the productivity enhancement by the education system that we assume (i.e. if $\psi_e = 1$) the compute a wage premium equal to 1.15. Therefore productivity enhancement accounts for the remaining wage premium up to 1.45, which is 66% (30/45) of the whole wage premium in the benchmark economy.

[21]Since effective ability is an outcome of the education rule the corresponding parameters are calibrated by targeting statistics of the distribution of education and math scores in the adult population who have already completed their education.

[22]Our sample consists of male individuals aged 25 to 54. Wages correspond to full-time workers after trimming the bottom and top 5% of the distribution in each education group.

the tertiary education wage premium (1.44) and the unemployment rates by education.[23] The 9% unemployment rate of non educated workers and the 4% of the tertiary educated workers are used in the identification process.[24] Finally, an important target in the calibration is the wage of educated relative to non educated workers conditioning for those who are mismatched. According to Hidalgo et al. (2015) using the Muestra Continua de Vidas Laborales (which is a sample of Social Security Administration records) the ratio of the average wage of mismatched college to non-college workers is about 1.15. Since our focus here is on tertiary educated instead of college educated individuals it is appropriate to target a smaller value and thus we pursue a 10% premium.[25]

To determine the equilibrium in the numerical simulations we proceed iteratively: given initial guesses for θ_i we find the implied wages and a potential threshold level $\bar{a}$. With this information we integrate the values of active matches and check if the free entry conditions are close to zero, and we iterate on the θ_i until these conditions are approximately satisfied. Once a candidate equilibrium with overeducation is found we check that no other equilibrium can be found nearby: we restart the algorithm from many different initial conditions and check that we always converge to the same candidate. Furthermore, we also check that there is no equilibrium with ex-post segmentation so that the equilibrium with employment overeducation is unique. To rule out this possibility we solve for equilibrium assuming that $\tilde{A} = \emptyset$ and check that $U_{he}(a_m) < U_{le}(a_m)$. Thus, the employment overeducation equilibrium reported in the following tables appears to be unique.

Parameter	
a_m	4.10
α	6.30
σ_0	0.25
σ_1	0.42
Target	
PISA mean score science (OECD 2006)	4.88
PISA standard deviation to mean (OECD 2006)	0.19
Fraction of workers with tertiary educ. (EU-SLIC 2007)	0.31
Average skills tertiary educ. relative to non educ. (PIAAC 2013)	1.2

Table 2: Calibrated Parameters and Targets I

[23] Note that the wage premium, the unemployment rates and the fraction of tertiary educated workers for Spain that we target according to EU-SILC data are slightly different from the figures with OECD data reported in Table 1.

[24] Although the full set of parameters affects each of the equilibrium outcomes, it is reasonable to think the unemployment rates are especially relevant for the identification of c_v and b.

[25] In spite of the negative sign calibrated for y_h the productivity of the lowest ability worker in the high-tech sector is positive.

Parameters			
$y_l = -0.76$	$\psi_{ne} = 1$	$\beta = 0.5$	$\eta = 0.5$
$m_h = m_l = 1$	$\delta_h = \delta_l = 0.07$	$r = 0.013$	
$c_v = 0.13$	$\psi_e = 1.15$	$y_h = -1.53$	$\tilde{y}_{he} = 0.59$
$\tilde{y}_{le} = 0.45$	$\tilde{y}_{lne} = 0.43$	$b = 1.25$	

Targets	Data	Model
Unemp. rate dropouts (EU-SLIC 2007)	0.09	0.09
Unemp. rate tertiary educ. (EU-SLIC 2007)	0.04	0.05
Tertiary educ. wage premium (EU-SLIC 2007)	1.44	1.45
Tertiary educ. wage premium, mismatched (Hidalgo et al. 2014)	1.10	1.09
CV of wages, tertiary ed. (EU-SLIC 2007)	0.38	0.37
CV of wages, non educated (EU-SLIC 2007)	0.27	0.25
Frac. of wage premium accounted for by selection (Fang, 2006)	0.33	0.33

Table 3: Calibrated Parameters and Targets II

3.3 Benchmark

In this section we assess the suitability of our benchmark economy to perform counterfactual analysis. The equilibrium outcome of our main interest is the size of employment overducation that the model economy endogenously generates. Interestingly, this fraction is equal to 33%, a figure that is very close to that reported by Eurostat in 2009.[26] Of course, the comparison between the model and the data in this respect is not straightforward because the model and the data do not necessarily capture the same notion of overeducation. In particular, as we argued in the Introduction, our model produces persistent overeducation whereas in the data employment overeducation could also include temporary phenomena. However, according to the empirical evidence we refer to in the Introduction in Spain overducation is very persistent over the lifetime and therefore its transitory component is expected to be modest. Thus our calibration is consistent with the case of Spain, but our counterfactual analysis below will remain silent about overeducation of a different nature that may well be relevant in other EU-15 countries.

As a validation exercise of our benchmark economy we explore the implications of reducing the fraction of tertiary educated workers to the 21% observed in the beginning of the 90's.[27] In Table 4 we compare the labor market outcomes of our benchmark economy with those in this counterfactual economy (in order to do that we set $\sigma_0 = 0.15$, instead of the $\sigma_0 = 0.25$ in our benchmark). We obtain that the fraction of overeducated workers decreases from 0.33 to 0.30 and the tertiary education wage premium increases from 1.45 to 1.48 (an increase in the

[26] Unfortunately, the EU-SILC data does not provide information on the quality of the job match for each worker, so we cannot compute the fraction of overeducated workers in this data set. The figure provided by Eurostat is calculated as those with tertiary education who hold a job beneath their educational level.

[27] Starting in 1993 Spain went through a very deep recession with dramatic consequences on the labor market. For this reason our comparison will be done with statistics prior to that year.

	Benchmark	*begining* − 90*s*
Fraction of educated workers	0.31	0.21
Average skills tertiary educ. relative to non educ. (PIAAC 2013)	1.20	1.21
Unemp. rate, educated	0.04	0.05
Unemp. rate, non educated	0.09	0.10
Frac. of educated, overeducated	0.33	0.30
Education wage premium	1.45	1.48
Education wage premium, overeducated	1.09	1.09
CV of wages, educated	0.37	0.39
CV of wages, non educated	0.26	0.26
GDP	1.39	1.33

Table 4: Changing tertiary education attainment

unemployment rate is observed, but it is negligible). This evolution is consistent with empirical evidence for Spain. In particular, according Pijoan-Mas and Sánchez-Marcos (2010) the tertiary education wage premium decreased from around 1.65 in 1993 to about 1.50 in 2000. Furthermore Hidalgo et al. (2015) report a substantial increase of about 10 percentage points in the fraction of overeducated workers using the MCVL. Finally, the unemployment rate of educated workers was around 7% and around 10% for non-educated workers according to the Labor Force Survey. In other words, the recent expansion of the educational attainment of the population in Spain could account to some extent for the decrease in the tertiary education wage premium and the increase in the fraction of employment overeducation.

4 Counterfactuals

In this section we conduct several counterfactual exercises to assess the impact of alternative education rules on labor market outcomes and its ability to account for the differences observed between Spain and the EU-15 countries. In general there are multiple alternatives to implement a given education goal: for instance to increase the average ability of educated workers one could increase it in all levels of abilities by means of a better quality of education (larger ψ_e) or one could implement a more stringent selection of abilities to receive education (higher σ_1 relative to σ_0). In the quantitative investigation we report below we discipline the analysis by restricting the fraction of educated workers to be the same and the ability of the different types of workers to be that observed in the EU-15 countries. We also analyze as alternative explanations differences in the productivity of the two sectors. In our analysis we mainly focus on the consequences for the figures stressed in the Introduction: the fraction of overeducated workers, the unemployment rate of each education group and the tertiary education wage premium.

4.1 Improving tertiary education outcomes

In our first exercise we evaluate the effects of improving the education outcomes of tertiary education. To this end, the first column of Table 5 reproduces the benchmark situation that was introduced above. For completeness and to easy the comparison, in the second column we report a summary of observations for EU-15 from Table 1.[28]

In the third column of Table 5 (S) we provide the statistics of an economy in which the *selection* of abilities is more stringent (in the First Order Stochastic dominance sense): fewer lower ability workers receive tertiary education but a larger fraction of higher ability workers do so. In particular, the parameters of the education rule are selected such that the fraction of e-workers is the same as in the benchmark, but their average skills increased half the way towards the EU average (from 5.77 to 5.92). This is implemented with a σ_0 smaller than in the benchmark and equal to 0.2 and with a larger σ_1 and equal to 0.8. A first implication of this policy is that the average effective labor productivity is larger for educated workers and lower for non educated workers (see that average skills are larger for educated and lower for non educated workers), which translates into a larger market value of education. Given this, the ratio of vacancies in the high sector relative to that in the low sector increases (from 1.03 to 1.67, not reported in the table because there is not empirical counterpart) and in the new equilibrium θ_h is higher and θ_l is lower than in the benchmark. It follows that the fraction of overeducated workers goes down to 0.24 and the wage premium of education increases up to 1.54 (as a matter of fact, the value of $\bar{a}$ decreases from 4.45 in the benchmark to 4.39). Notice that this is the result of both a compositional effect (since now the mass of e-workers among higher ability types is larger), and the endogenous response of the e-workers that are now more prone to search for a job in the h-sector. In other words, by improving selection, overeducation decreases because of the combined effect of the improvement in the h-sector and the worsening in the l-sector. The better selection shifts e-workers from the l to the h-sector but otherwise has no significant effects on their unemployment rate, whereas the unemployment rate of ne-workers increases from 9% to 12%. Finally, there is a slight decrease in GDP. The reason is that in spite of the increase in output in the h-sector due to the higher employment rate of e-workers, it cannot counter balance the reduction in the output of the l-sector due to the larger unemployment rate of ne-workers.

In the fourth column of Table 5 (Q_e) we report the statistics of an economy in which the same education outcome for e-workers achieved in S is implemented but by only increasing ψ_e up to 1.18 (hence selection is like in the benchmark case). Note in particular that in this case the average ability of a ne-worker remains the same as in the benchmark economy. The increase in ψ_e has a direct effect on the productivity of e-workers which tends to increase wages of e-workers. In addition the increase in labor productivity makes more attractive for firms to create vacancies in the h-sector, hence there is an additional indirect effect that increases the probability of finding a job in the h-sector. These two effects explain the shift of e-workers from the l-sector to the h-sector ($\bar{a}$ decreases from 4.45 in the benchmark to 4.33) and the increase in the wage premium. The shift of e-workers from the l to the h-sector worsens the average ability of workers in the l-sector which in turn reduces the incentives to open vacancies in that sector. That is, improving the productivity of e-workers has an indirect negative effect on the prospects of ne-workers. This effect explains the increase in the unemployment rate of ne-workers.

[28] GDP for EU-15 is obtained by applying the proportion $\text{GDP}_{\text{EU-15}}/\text{GDP}_{\text{Spain}}$ observed in the per capita data in Purchasing Power Standards from Eurostat, which is 1.22, to the the Spanish GDP predicted by the model.

Interestingly, although columns S and Q_e follow different strategies to implement the same average ability of e-workers, they have a similar impact on the fraction of overeducated workers. It is worth noticing here that if we consider an economy with the same selection rule as in economy S and in which the average ability of ne-workers is kept as in the benchmark economy the fraction of overeducated workers would be higher than in S.[29] Therefore, we conclude that with two different policy rules producing the same average ability for e and ne workers, in one combining selection and quality and in the other improving only quality, overeducation will be smaller in the second. We come back to this issue below.

In the fifth column of Table 5 (S+Q_e) we report the results under an education outcome that matches the average ability of e-workers observed in EU-15. This is achieved by combining the more stringent selection rule (reported in the third column) with a convenient increase in the labor productivity of e-workers (the improvement due to *selection* is therefore implemented as before, and the improvement due to *quality* is introduced by increasing ψ_e up to 1.19). We notice that the increase in the labor productivity of e-workers that we implement here resembles the Skill Biased Technological Change (SBTC) stressed in the literature: it is equivalent to an increase in the productivity component of education in the h-sector (although a SBTC would not alter the observed ability of e-workers). The general picture that emerges is that the direct effect due to the increase in the productivity of e-workers reinforces the previous indirect effects of an improved selection: it increases even more the average skills of e-workers and so the market value of education also increases. We then observe a more notorious increase in the wage premium of education up to 1.62, a higher ratio of vacancies in the h-sector relative to the l-sector of 2.28, a reduction in employment overeducation to 0.11 (in particular $\bar{a}$ decreases from 4.45 to 4.23) and a slight reduction in the unemployment rate of e-workers which then delivers an increase in GDP. As an illustration of the new equilibrium that is reached, Figure 4 in Appendix C represents the change in the value of $\bar{a}$ originated by the more stringent selection rule and by shift in the slope of the U_{ij} curves do to the changes in θ_e.

In view of the previous results we conclude that policies that improve the average quality of e-workers (either indirectly by implementing a more restrictive selection or by directly increasing the effective productivity of the workers that obtain education) would move the Spanish economy closer to the EU-15 average, except for the higher unemployment rate among ne-workers that it implies. All in all, the previous combination of policies is able to reproduce the average skill of tertiary educated workers in EU-15 but it misses the same statistic for the non-educated. In particular, the more stringent selection increases the fraction of low ability workers that are not educated, hence their average ability is reduced from 4.81 to 4.75. This finding is relevant because in Spain the average skills of both e-workers and ne-workers are lower than in EU-15, but the difference is larger for ne-workers. We address this issue in the following subsection.

[29] In concrete we computed the equilibrium of an economy with the same selection as in S but we increased ψ_{ne} to keep the same average ability of ne-workers as in the benchmark economy (or as in Q_e). The higher ψ_{ne} increases labor productivity of ne-workers *and* makes more attractive to create vacancies in the l-sector, which in turn makes more attractive to look for jobs in that sector. Hence there is an indirect effect on θ_l such that the probability of finding a job in the l-sector is now larger. Thus the previously indifferent e-worker is now better off by looking for a job in the l-sector. In that economy overeducation is larger than in S (0.29 instead of 0.24), hence also larger than in Q_e.

	Bench.	EU-15	S	Q_e	$S{+}Q_e$
Fraction of educated workers	0.31	0.27	0.31	0.31	0.31
Average skills, educated	5.77	6.13	5.92	5.92	6.13
Average skills, non educated	4.81	5.33	4.75	4.81	4.75
Unemp. rate, educated	0.05	0.03	0.05	0.04	0.04
Unemp. rate, non educated	0.09	0.05	0.12	0.10	0.13
Frac. of educated, overeducated	0.33	0.19	0.24	0.24	0.11
Education wage premium	1.45	1.68	1.54	1.51	1.62
GDP	1.39	1.69	1.38	1.42	1.41

S: the *selection* of abilities is more stringent, $S{+}Q_e$: the *selection* of abilities is more stringent and the productivity of *e-workers* is larger.

Table 5: Counterfactuals: changing tertiary education outcomes

4.2 Improving the quality of educated and non educated workers

In the first two columns of Table 6 we repeat for convenience the results of the benchmark economy and the European averages. In the third column of Table 6 we keep the selection and labor productivity of e-workers as in column $S + Q_e$ in Table 5 and additionally we adjust the labor productivity of ne-workers to match the average observed in EU-15 (this is column $S{+}Q_{e+ne}$, and in the model this amounts to fix $\psi_{ne} = 1.12$). Improving the quality of ne-workers increases their market value. Hence as a consequence of this policy we observe sizable reduction in their unemployment rate and in the wage premium to education. The redistribution towards the non educated workers that this policy implies in terms of relative wages is in line with the results in Albrecht et al (2009). This is in contrast with the results in the previous section in which the quality of education improved *only* for tertiary educated workers. In addition, some of the lower ability but e-workers that under the previous policies preferred to look for jobs in the h-sector (albeit longer unemployment spells), now find more profitable to look for jobs in the l-sector in which wages are higher than before (and in which) an unemployment spell is shorter due to a larger vacancy creation: relative vacancy creation now reduces to 0.59, substantially smaller than the 2.28 observed under the $S + Q_e$ policy). As a consequence, not only non educated workers but also mismatched educated workers benefit from a lower unemployment rate. All in all, overeducation is 5 percentage points smaller than in the benchmark case but larger than in the $S{+}Q_e$ policy ($\bar{a}$ slightly shifts to the left, from 4.45 in the benchmark to 4.43). That is, in this economy there is an indirect effect through a higher θ_l that counterbalances the effects previously identified in the $S + Q_e$ economy. The reduction in the unemployment rates of both educated and non educated workers explains the 19% increase in GDP. Under this policy, therefore, we conclude that the model approaches the EU-15 average in several dimensions but it still produces a wage premium of tertiary education that is smaller than in the benchmark economy. Figure 5 in Appendix C provides a graphical exposition of the new equilibrium.

Finally, in the fourth column of Table 6 we keep selection as in the benchmark economy and explore the effects of an improved labor productivity of education for both e-workers and ne-workers (the column labeled Q_{e+ne}). Thus one could alternatively think of this experiment

	Bench.	EU-15	S+Q_{e+ne}	Q_{e+ne}
Fraction of educated workers	0.31	0.27	0.31	0.31
Average skills, educated	5.77	6.13	6.13	6.13
Average skills, non educated	4.81	5.33	5.33	5.33
Unemp. rate, educated	0.05	0.03	0.03	0.03
Unemp. rate, non educated	0.09	0.05	0.05	0.05
Frac. of educated, overeducated	0.33	0.19	0.28	0.23
Education wage premium	1.45	1.68	1.37	1.37
GDP	1.39	1.69	1.65	1.65

S+Q_{e+ne}: the *selection* of abilities is more stringent and all workers are more productive, Q_{e+ne}: all workers are more productive.

Table 6: Counterfactuals: changing education system outcomes

as an improvement in the overall productivity of workers, which could be originated in the education system or not. In this exercise we keep $\sigma_0 = 0.25$ and $\sigma_1 = 0.42$ as in the benchmark case and we fix $\psi_e = 1.22$ and $\psi_{ne} = 1.11$ to match the average labor productivity in EU-15. Notice that in order to match the EU-15 statistics ψ_e is larger than in the policy reported in column $S + Q_e$ of Table 5. All the statistics move in the same direction than in the case of the policy explored in the third column. However, the decrease in overeducation with respect to the benchmark is larger in this scenario (the critical ability level $\bar{a}$ shifts to the left, from 4.45 in the benchmark to 4.32). This is exactly the result that we expected after our discussion in the previous section of alternative ways of improving the education outcomes (in terms of average abilities) for all workers. As we explained, overeducation is smaller when quality, rather than quality and selection, is used to achieve the goal. Figure 6 in Appendix C graphically illustrates the new equilibrium that is reached.

It is clear from the previous tables that in Spain the fraction of e-workers is slightly larger than in EU-15, thus it is natural to ask if this fact is quantitatively relevant to account for the differences observed in labor market outcomes. We explored this possibility by reducing σ_0 to match the fraction of tertiary educated workers in EU-15. In this case it was observed a small reduction in overeducation, to 0.32. If we instead implemented the reduction of the fraction of tertiary educated workers by reducing only the value of σ_1, then the fraction of overeducated workers would be even larger (0.38). The reason is that in that scenario the role of individual's ability to be selected in tertiary education is played down. Hence the differences between Spain and the EU-15 do not seem to be accounted for by the differences in the fraction of educated workers in Spain.

4.3 Alternative explanations

It is often argued that the expansion of the housing sector that fueled the most recent boom of the Spanish economy may be responsible for some of the misbehavior of the labor market with respect to other developed countries. In this section we try to *remove* the effect of the housing

	Bench.	EU-15	low-tech	high-tech
Fraction of educated workers	0.31	0.27	0.31	0.31
Average skills, educated	5.77	6.13	5.77	5.77
Average skills, non educated	4.81	5.33	4.81	4.81
Unemp. rate, educated	0.05	0.03	0.05	0.04
Unemp. rate, non educated	0.09	0.05	0.14	0.10
Frac. of educated, overeducated	0.33	0.19	0.19	0.19
Education wage premium	1.45	1.68	1.47	1.47
GDP	1.39	1.69	1.33	1.40

low-tech: the productivity of this sector is lower, high-tech: the productivity of this sector is larger.

Table 7: Counterfactuals: changing sector productivity

boom in the 2000's and explore the implications for the equilibrium under a relatively less productive l-sector. In order to discipline our exercise here we select the overall productivity of the l-sector to target the fraction of overeducated workers in the EU-15 countries (we need a 3.7% lower value of y_l to achieve the fraction of overeducated workers in the EU-15 countries). In the third column of Table 7 we report the result of this exercise. The reduction in the productivity of the l-sector produces a small increase in the tertiary education wage premium from 1.45 to 1.47. Furthermore, the unemployment rate is higher under these circumstances for the ne-workers, going up from 9% to 14% . Therefore, although a relatively higher productivity of the l-sector could be responsible for the higher incidence of overeducation in Spain with respect to EU-15, it barely accounts for differences in the tertiary education wage premium. More importantly, this widens the gap in terms of the unemployment rate of ne-workers between Spain and the EU-15 countries.

Finally, we extend the previous analysis with a brief exploration of the effects of a lower labor productivity in the h-sector in Spanish firms relative to the EU-15 average as is sometimes stated in informal debates (for a formal account of facts along these lines see Palazuelos and Fernandez 2009 and the references therein). Specifically, in the last column of Table 7 we show the implications of having a more productive h-sector, in which again we design the exercise to target the fraction of overeducated workers observed in the EU-15 countries (we need a 2.4% higher value of y_h to achieve the fraction of overeducated workers in the EU-15 countries). As it can be seen, this would move the tertiary wage premium and the unemployment rate of e-workers in the right direction, but the effect would be rather modest. Furthermore, the unemployment rate among ne-workers remains at a relatively high level.

4.4 Summary of the results

According to our analysis, changing the education rules to improve the ability of all workers to meet the EU-15 standards would reduce the unemployment rate of both e and ne-workers to the EU-15 levels and it would reduce the fraction of overeducated workers between 5 to 10 percentage points. However, the education wage premium would move only slightly and in the opposite direction to what is observed in EU-15 countries. In contrast, if only the education outcomes of tertiary educated workers are improved, the fraction of overedcuated workers may be more than halved and the wage premium can be increased up to 1.62. In this case however, the unemployment rate of non educated workers would increase. We take these results as supporting the view that differences in education outcomes are able to account for a sizable fraction of the differences in overeducation and in the unemployment rate of Spain with respect to the EU-15 countries. However, our calibration exercises suggest that education outcomes by themselves cannot account simultaneously for the differences in overeducation, unemployment and wage premium to education, hence there may be other aspects that are missing in our model that are relevant to provide a comprehensive understanding of the labor markets.

We extended the previous analysis by considering an scenario in which the productivity of the l-tech sector (h-sector) is reduced (increased) to meet the fraction of overeducated workers observed in the EU-15 countries. We find that such changes in productivity have a negative effect on the unemployment rate of ne-workers and a very modest effect on the wage premium to education. Furthermore, the gap in terms of average skills for all workers between Spain and the EU-15 countries would remain unexplained in this case. In Appendix B we report the results from additional exercises in which we combine the effects of the education policy along the lines in subsection 4.2 with the reduction (increase) in the low-tech (high-tech) sector productivity in the subsection 4.3. We find that the unemployment rates and the fraction of overeducated workers are closer to the EU-15 statistics, but the wage premium to education is smaller than in the benchmark case, hence in this dimension the economy worsens with respect to the EU-15 average.

5 Conclusions

We develop an equilibrium model of the labor market with frictions in which workers are heterogeneous in terms of ability and education. We depart from existing models in that we assume that education does not only represent a barrier for non educated workers to obtain jobs in technologically advanced firms, but it also increases labor productivity of educated workers in the less advanced sector. Furthermore we consider a continuum of ability levels which allows us to address the question of how differences in the composition of educated workers affects firms' incentives to open different types of vacancies.

We perform a quantitative analysis in order to illustrate the implications of alternative education outcomes on employment overeducation, unemployment and on tertiary education wage premium. We discipline our model by calibrating the parameter values to match significant facts of the Spanish economy. The results of these counterfactual experiments suggest that the differences observed in the equilibrium labor market between Spain and the average of the EU-15 countries would be smaller had Spain implemented a more selective education rule (improve the ability mix of the educated workers), and/or if the education system was able to increase labor productivity *of both* educated and non educated workers. In particular, according to our quantitative analysis had the quality of education observed in Spain been similar to the European average then overeducation would have been between 5 and 10 percentage points lower and the unemployment rate of the two types of workers would be reduced by 40%. However, the tertiary education wage premium would be slightly smaller than in the benchmark economy.

From a policy perspective it is important to emphasize that improving education outcomes of only higher educated will effectively help to reduce overeducation, but at the cost of higher unemployment among non educated workers and of more inequality in the wages across educated and non educated workers.

Our analysis shows that there are significant effects of different education outcomes on the unemployment rates, overeducation rates and on the the wage premium to education, but also, that education outcomes alone are not able to account simultaneously for the discrepancies between Spain and the EU-15. Thus there must be other issues that are relevant for the labor market and that are missing in the current model. An interesting extension along these lines would be to include *transitory mismatch* (sectoral, geographical or/and due to experimentation along the lines indicated in the Introduction) and reevaluate the role of education in that richer model. Related to this, we also find that different education policies have sizable effects on the relative size of the sectors and on GDP. Our model remains silent with respect to how education is financed and thus it is not possible to investigate the optimality or efficiency of education policies. Extending the model to explicitly account for the cost and financing of education is another promising line for future research.

We conclude with additional extensions of our work that are worth investigating. First, the model studied in this paper belongs to a broad class in which multiple equilibria are possible. Thus from the theoretical perspective it would be valuable to have a characterization of the conditions under which such a multiplicity arises and under which the equilibrium is unique. Second, in regards to the quantitative analysis, our model could be extended to consider education choices at the individual level. Currently the fraction of educated workers is purely determined as the result of a particular education rule. Since in our framework there are incentives to complete tertiary education even for those individuals who will end up working in the l-sector, then allowing for the choice of the education level will not necessarily eliminate overeducation. It would be interesting to quantify the effects of changes in the quality of education (say in terms of additional units of efficient labor) and compare them with the implications of more stringent requirements (in terms of minimal ability) to be allowed to complete tertiary education. These extensions are left for future work.

References

Albrecht, J., and S. Vroman, 2002: *"A Matching Model with Endogenous Skill Requirements,"* International Economic Review (43), pp. 283-305.

Albrecht, J., van den Berg, G.J. and S. Vroman, 2009: *"The aggregate labor market effects of the Swedish Knowledge Lift program,"* Review of Economic Dynamics 12 (1), pp. 129-146.

Blazquez, M. and M. Jansen, 2008: *"Search, mismatch and unemployment,"* European Economic Review 52, pp. 498-526.

Charlot, O., and B. Decreuse, 2010: *"Over-education for the rich, under-education for the poor: A search-theoretic microfoundation,"* Labour Economics, vol. 17(6), pp. 886-896.

Ciccone, A. and W. Garcia-Fontes, 2009: *"The quality of the Catalan and Spanish education systems: A Perspective from PISA,"* IESE, WP 810.

Cuadras-Morató, X., and X. Mateos-Planas, 2006: *"Skill bias and unemployment frictions in* *the US labor market 1970-1990,"* International Economic Review, Vol. 47, pp. 129-160.

Cuadras-Morató, X., and X. Mateos-Planas, 2013: *"Overeducation and skill-biased technical change,"* The B.E. Journal of Macroeconomics (Advances), 13 (1), pp. 105-122.

Cubas, G., Ravikumar, B., and G. Ventura, 2016: *"Talent, labor quality and economic development,"* Review of Economic Dynamics, Vol. 21, pp.160-181.

Diamond, P.A., 1982: *"Wage determination and efficiency in search equilibrium,"* Review of Economic Studies, 49, pp. 217-227.

Díaz, A., and L. Franjo, 2016: *"Capital goods, measured TFP and growth: The case of Spain,"* European Economic Review, 83, pp. 19-39.

Dolado, J. J., Felgueroso, F., and J. F. Jimeno, 2000: *"Youth labour markets in Spain: Education, training and crowding-out,"* European Economic Review, pp. 943-958.

Dolado, J.J.,Jansen, M. and J.F. Jimeno (2009): *"On the job search in a matching model with heterogenous jobs and workers,"* The Economic Journal, 119, pp. 200-228.

Dvorkin, M., 2013: *"Sectoral shocks, reallocation and unemployment in a model of competitive labor markets,"* Technical report.

Eurostat, The Bologna Process in higher education in Europe, 2009.

Eurostat, Statistics explained, students in tertiary education, 2012, YB15.png.

Fang, H., 2006: *"Disentangling the college wage premium: estimating a model with endogenous education choices,"* International Economic Review, 47(4), pp. 1151-1185.

Guvenen, F., Kuruscu, B., Tanaka, S. and D. Wiczer, 2015: *"Multidimensional skill mismatch,"* NBER Working Papers 21376.

Hanushek, E.A., Schwerdt, G., Wiederhold, S. and L. Woessmann (2015): *"Returns to skills around the world: Evidence from PIAAC,"* European Economic Review 73, pp. 103130.

Hendricks, L., and O. Leukhina, 2014: *"The return to college: selection bias and dropout risk,"* CESifo Working Paper Series, 4733.

Herz, B., and T. van Rens, 2016: *"Accounting for mismatch unemployment,"* WP, previous version IZA WP 8884.

Hidalgo-Cabrillana, A., Kuehn, Z. and C. Lopez-Mayan, 2017: *"Development accounting using PIAAC data"*, SERIEs, 8 (4), pp 373-399.

Hidalgo-Pérez, M., Jimenez-Martín, S., and F. Felgueroso, 2015: *"The puzzling fall of the wage skill premium in Spain,"* The Manchester School. Forthcoming.

Hobijn, B. and A. Sahin, 2009: *"Job-finding and separation rates in the OECD,"* Economic Letters.

Jovanovic, B., 1979: *"Job Matching and the Theory of Turnover,"*, Journal of Political Economy, 87, pp. 972-990.

Krusell, K., Ohanian, L. E., Ríos-Rull, J. V., and G. L. Violante, 2000: *" Capital-skill complementarity and inequality: A macroeconomic analysis,"* Econometrica, vol. 68 (5), pp. 1029-53.

Leuven, E., and O. Oosterbeeck, 2011: *"Overeducation and mismatch in the labor market,"* IZA Discussion Paper 5523.

Marimon R., and F. Zilibotti, 1999: *"Unemployment vs. mismatch of talents: Reconsidering unemployment benefits,"* The Economic Journal, 109 (April), pp. 266-291.

Meroni, E. C. and E. Vera-Toscano, 2017: *"The persistence of overeducation among recent graduates,"* Labour Economics 48, pp. 120143.

Miller, R.A., 1984: *"Job matching and occupational choice,"* Journal of Political Economy, 92, pp.1086-1120.

Montalvo, J., 2013: *" Temporary or permanent overqualification: evidence from young workers in the Spanish labor market,"* mimeo.

Mortensen, D.T., 1982: *"The matching process as a noncooperative bargaining game,"* in *The economics of information and uncertainty*, University Chicago Press and NBER, pp. 233-258.

Naylor, R. A., and J. Smith, 2009: *"Ability Bias, Skewness and the College Wage Premium,"* mimeo.

Nieto, S., 2014: *"Overeducation, skills and wage penalty: Evidence for Spain using PIAAC data,"* Institut de Recerca en Economia Aplicada Regional i Pblica. Document de Treball 2014/06.

OECD, 2007: *"Science competencies for tomorrow's world."*

OECD, 2010: *"Educatation at a glance."*

Palazuelos, E. and R. Fernández, 2009: *"Labor productivity: A comparative analysis of the European Union and United States, 1994-2007,"* Political Economy Research Institute, WP 208.

Papageorgiou, T., 2014: *"Learning your comparative advantages,"* Review of Economic Studies, 81, pp. 1263-1295.

PIAAC, 2013: *"Programa internacional para la evaluación de las competencias de la población adulta."*

PISA, 2006: Report of the OECD.

Pissarides, C.A., : *"Equilibrium unemployment theory,"* Cambridge M.A., Basil Blackwell.

Pijoan-Mas, J. and V. Sánchez-Marcos, 2010: *"Spain is different: falling trends in inequality,"* Review of Economic Dynamics, 13 (1) pp. 157-178.

Robles-Zurita, J.A. (2017): *"Cognitive skills and the LOGSE reform in Spain: evidence from PIAAC,"* SERIEs, 8 (4), pp 401415.

Sahin, A., Song, J., Topa, G. and G.L. Violante, 2014: *"Mismatch unemployment,"* American Economic Review, 11, pp. 3529-3564.

Sattinger, M., 1975: *"Comparative Advantage and the Distributions of Earnings and Abilities,"* Econometrica (43), pp. 455-468.

Stovicek, K. and A. Turrini (2012): *"Benchmarking Unemployment Benefits in the EU,"* IZA Policy Paper 43.

6 Appendix A: A closer look to EU-15

As indicated in the Introduction the notion of mismatch adopted in the current paper is the definition of *vertical mismatch* proposed by Eurostat: a worker is considered to be occupational mismatched if her educational attainment is at least ISCED 5, but her occupation is not considered to be ISCO 1, 2 or 3. ISCED stands for International Standard Classification of Education. Levels 0 to 4 include education between pre-primary school and upper-secondary education. Levels 5 and 6 are tertiary education levels (respectively, not leading/leading to an advanced research qualification). ISCO stands for International Standard Classification of Occupations. Categories 1, 2 and 3 include legislators, senior officials, managers, professionals, technicians and associate professionals. Categories 4 to 9 include clerks, service workers, etc., to elementary occupations.

In what follows we provide a more detailed account of the facts that motivate our research as mentioned in the Introduction.

The Statistical Book of Eurostat corresponding to the Bologna Process in Higher Education in Europe (2009) reports the distribution of tertiary students in the ISCED levels 5A, 5B, and 6 as a percentage of all tertiary students in private and public institutions for the period 2001 to 2006 (see the Table 8).[30] It is clear from the table that the differences between Spain and the average EU-15 in the mid 2000's are remarkably small. Hence, the explanation for the higher overeducation observed in Spain is not due to a disproportionately large/small fraction of students involved in scientific/academic activity.

	2001			2002			2003			2004			2005			2006		
ISCED	5A	5B	6	5A	5B	6	5A	5B	6	5A	5B	6	5A	5B	6	5A	5B	6
EU-15	78	18	4	78	19	4	79	18	4	79	19	4	81	17	4	81	17	4
Spain	86	11	3	84	12	4	83	13	4	82	14	4	82	14	4	82	13	4

Source: UIS, UOE (The Bologna process in higher education in Europe 2009, Table 0 p. 189).

Table 8: Distribution of students in higher ISCED levels as a percentage of all tertiary students, 2001-06

Second, the distribution of the population across fields of specialization in Spain is similar to the average of the EU countries, hence the higher fraction of overeducated workers in Spain was not due to a higher concentration of workers in certain fields of specialization. Roughly speaking the fraction of workers in "Humanities", "Education", "Agriculture", "Health" and "Social sciences" is similar in Spain to the average of the EU countries (see Table 9). There are only moderate differences in the fraction of workers in "Science" (about 14% in the EU in contrast to 19% in Spain) and in "Social sciences" (about 32% in the EU in contrast to 29% in Spain). Therefore we conclude that the phenomena of occupational mismatch is not due to compositional differences in terms of the fraction of educated workers in each field of specialization.

[30] ISCED level 5A are tertiary programs that are largely theoretically based and are intended to provide sufficient qualifications for gaining entry into advanced research programs and profession with high skills requirements. Programs in ISCED 5B are typically shorter than those in 5A and focus on occupationally specific skills geared for entry into the labor market, although some theoretical foundations may be covered. Level ISCED 6 is reserved for tertiary programs which lead to the award of an advanced research qualification (they typically require the submission of a thesis or dissertation of publishable quality (see the Statistical book of Eurostat pp. 239-240 for further details).

	EU*	Spain
Education	4	5
Humanities and Arts	8	7
Social sciences, Business and Law	32	29
Science, Mathematics and Computing	14	19
Engineering, Manufacturing and Construction	28	27
Agriculture and Veterinary	3	4
Heath and Welfare	7	7
Services	3	1

Source: REFLEX 1999-2000. EU* includes Portugal, Spain, Italy, France, Switzerland, Austria, Germany, Netherlands, Belgium, United Kingdom, Norway, Sweden and Finland.

Table 9: Distribution of Graduated Individuals Across Fields of Education

Next, we report in Table 10 the fraction of workers aged 25 to 34 who are considered to be overeducated by field of education. The incidence of mismatch by field of specialization in Spain is higher than the European average (with the sole exception of Agriculture and Veterinary). It is clear that the average fraction of overeducated workers across fields of specialization is substantially higher in Spain than in EU-15, and also that overeducation is not a phenomenon concentrated in a very specific subset of fields. In EU-15 the highest fraction of overeducated workers is found in Services (48) and it is followed by Agriculture (35) and Social Sciences (26). In Spain the highest fraction is found in Services (64) and it is followed by Engineering field (50) and Social Science (44). Both in Spain and in the average of the EU-15 the lowest fraction of overeducated workers is found in Health fields (11 in UE-15 in contrast to 27 in Spain) and in Education (11 in EU-15 and 28 in Spain). The largest gap between Spain and the UE-15 (more than double) is found in Education field and it is followed by Health and Welfare.

	EU-15	Spain
Education	11	28
Humanities and Arts	25	37
Social sciences, Business and Law	26	44
Science, Mathematics and Computing	14	28
Engineering, Manufacturing and Construction	21	50
Agriculture and Veterinary	35	35
Health and Welfare	11	27
Services	48	64

Source: Eurostat, EU-LFS, 2003-2007 (The Bologna process in higher education in Europe 2009, Table D5.C p. 229).

Table 10: Percentage of Workers Vertical Mismatched, aged 25-34 by Field of Education

Finally, one may wonder about the comparability of tertiary educated workers in terms of the official number of years of education across countries. In Table 11 we can see that for the selected sample of countries there are noticeable differences in the distribution of years in primary, secondary and high school. However, looking specifically at tertiary education the differences seem rather small: in Spain higher education starts a year before than in other countries, but it takes one more year (together with Germany) to complete college education. Given this, we would find difficult to justify the lower performance of tertiary educated workers in Spain in terms of PIAAC scores simply by the smaller number of years of education.

	Formal school	Prim.+sec. +high s.	Voc. educ. starts	Univ. starts	Univ. educ. (min. years)
Austria	6	4+4+4	14/15	18	3+
Belgium	6	6+2+4(+1)	14	18	3+
Denmark	6	11+2	16	19	3+
Finland	7	9+3	16	19	3+
France	6	5+4+3	15	18	2+
Germany	6	4+6+3	16	19	4+
Grece	6	6+3+3	15	18	4+
Italy	6	5+3+5	14	19	3+
Ireland	4	8+6	15	18	3+
Luxembourg	6	6+3+4	15	19	5+
Netherlands	4	8+3+3	16	18	3+
Portugal	6	6+ 3+3	15	18	3+
Spain	6	6+4+2	15	18	4+
Sweden	7	9+3	16	19	3+
UK	5	6+3+2(+2)	16	18	3+

Source: Eurydice, The structure of the European Education systems 2009-10.

Table 11: European education systems

7 Appendix B: Simultaneous technology and policy changes

In this Appendix we report the results of combining each of the education reforms in Table 6 with the reduction (increase) in the low-tech (high-tech) sector productivity in Table 4. The results are reported in Tables 12 and 13.

8 Appendix C: Graphical exposition

Figure 4 portrays the effect of improving the selection rule (orange curves) and of enhancing the productivity of educated workers (green curves). The first change induces the shifts indicated with the orange arrows due to the fact that θ_h increase and θ_l decreases. In this case $\bar{a}$ would decrease to $\bar{a}'$. The additional improvement due to the increase of ψ_e increases the slope of the U_{he} and U_{le} curves, which reinforces the initial effects on U_{he} and counterbalances them on U_{le}. These effects are represented by the green arrows and $\bar{a}$ would further decrease from $\bar{a}'$ to $\bar{a}''$.

Figure 5 represents the effects of improving labor productivity of non educated workers. This improvement induces an increase in θ_l which then results in the upward shift of the U_{le} curve (the orange arrow). In the new equilibrium it is observed an increase in overeducation up to $\bar{a}'''$ (relative to the previous situation), but still it is smaller than in the initial equilibrium ($\bar{a}$).

Finally Figure 6 represents the improvement in labor productivity of both educated and non

Figure 4: Equilibrium after improving selection and ψ_e

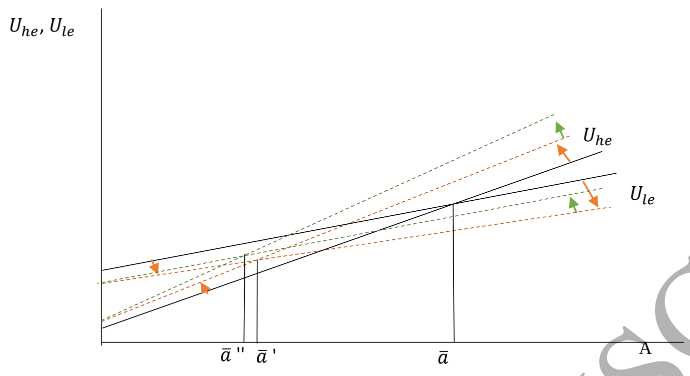
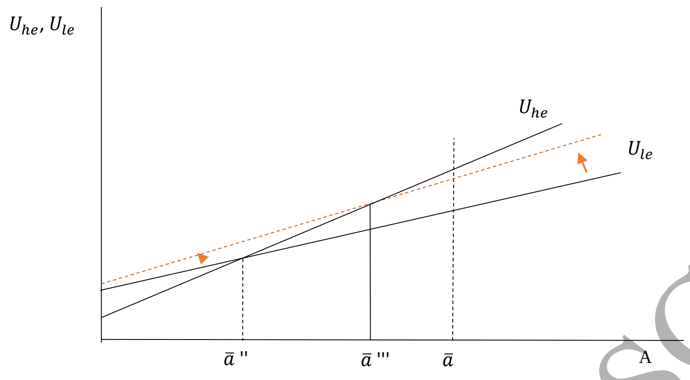


Figure 5: Equilibrium with higher ψ_{ne}, after improving selection and ψ_e



	Bench.	EU-15	low-tech $+S+Q_{e+ne}$	low-tech $+Q_{e+ne}$
Fraction of educated workers	0.31	0.27	0.31	0.31
Average skills, educated	5.77	6.13	6.13	6.13
Average skills, non educated	4.81	5.33	5.33	5.33
Unemp. rate, educated	0.05	0.03	0.03	0.03
Unemp. rate, non educated	0.09	0.05	0.05	0.05
Frac. of educated, overeducated	0.33	0.19	0.17	0.10
Education wage premium	1.45	1.68	1.39	1.39
GDP	1.39	1.69	1.63	1.63

Table 12: Counterfactuals: combinations I

	Bench.	EU-15	high-tech $+S+Q_{e+ne}$	high-tech $+Q_{e+ne}$
Fraction of educated workers	0.31	0.27	0.31	0.31
Average skills, educated	5.77	6.13	6.13	6.13
Average skills, non educated	4.81	5.33	5.33	5.33
Unemp. rate, educated	0.05	0.03	0.03	0.03
Unemp. rate, non educated	0.09	0.05	0.05	0.05
Frac. of educated, overeducated	0.33	0.19	0.14	0.05
Education wage premium	1.45	1.68	1.39	1.39
GDP	1.39	1.69	1.66	1.66

Table 13: Counterfactuals: combinations II

educated workers, starting with the same selection rule it was in place in the benchmark equilibrium. These changes induce an increase in the slope of U_{he} as well as an increase in θ_h and θ_l (as now both types are more productive than in the previous cases). The shift in U_{he} is larger because ψ_e is larger than before, which in turn induces a larger increase in θ_e. We summarize the shifts in the black arrows, which delivers a decrease in overeducation. It is clear from this graphical analysis that the final effect on equilibrium overeducation depends on the relative strength of each of the effects discussed above. Hence the quantitative analysis is needed to properly asses the net effects.

Figure 6: Equilibrium after improving ψ_e and ψ_{ne}