CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONÓMICAS, A.C.

EDUCATIONAL MISMATCH AND GENDER INEQUALITY IN THE LABOR MARKET

TESINA

QUE PARA OBTENER EL GRADO DE

MAESTRA EN ECONOMÍA

PRESENTA

KARLA PINEL VALERIO

DIRECTOR DE LA TESINA: DR. DAVID STRAUSS

CIUDAD DE MÉXICO JUNIO, 2017

To my family.

Agradecimientos

I would like to thank:

My advisor, David Strauss, for his invaluable help and support during the preparation of this work.
My lecturer, Fernanda Márquez-Padilla, for her comments and unconditional availability.
My family,you believe in me and taught me to work hard for what I want. Los amo hasta el cielo.

Contents

List of Tables

Chapter 1

Introduction

It is well known that, despite the decrease of its share in the Gross Domestic Income over time, wages still constitute a large source of household income in developing and developed countries[1]. Also, numerous studies such as Duncan and Hoffman (1981), Verdugo and Verdugo (1989) and Cohn et al. (2000) find that education has a positive effect on wages and is not reduced to an absolute level effect, but also that the returns on being overeducated or undereducated in the occupation of work are significant when determining them. Educational mismatch, defined as a situation when the educational attainment of the workers differs from the specific education required in an occupation is then an important variable to consider when studying labor market outcomes such as wages.

My main interest in this thesis is to estimate the returns of education and determine if there are differences in the returns to education between men and women considering educational mismatch variables. Then, I study the role they have determining the gender wage gap using a Oaxaca-Blinder Decomposition. In aggregate, this work studies gender differences in wages driven by differences in levels and returns to educational mismatch, but also occupation, experience, marital status, children, sector and other variables considering two samples: a whole

[1]For data series on this matter, please see the following link of the Federal Reserve Bank of St. Louis: https://fred.stlouisfed.org/series/W270RE1A156NBEA

representative sample of the U.S. labor force and a graduate[2] sample, following by studying their relative weights in the wage gap.

The principal results are as follows: there are significant differences in the returns of education and educational mismatch between men and women; however, they do not have a relative weight in determining the gender wage gap considering the whole sample, while the returns to required education significantly lower the wage gap for graduate workers. Differences in characteristics and payment to these characteristics such as experience, usual hours of work, marital status and children are the most important factors that drive the gender wage gap. This thesis complements literature on the topic such as the works of Boll (2014) and Salinas-Jimenez (2013) whose research focus on graduate German and Spanish workers respectively.

[2]Workers with a Master's or Doctoral Degree

Chapter 2

Literature Review

2.1 Returns to Education and Educational Mismatch

The role of education on improving people's conditions has been widely studied in economics and other social sciences. Considering labor outcomes, higher levels of education are strongly correlated with higher employment rates and earnings. According to Human capital theory (Becker, 1964), productivity is a function of variables such as education, skills, training and experience, and workers' wages are determined by the value of their marginal product. An empirical approach to test this theory was proposed by Mincer (1974) and the author studies the effects of a worker's educational attainment, experience among other variables on wages with the Mincer earnings equation. The results obtained with this approach suggest that the effects of education on wages are determined by the supply of human capital (i.e , the individual's educational level choice), giving no importance to job characteristics. Considering both sides of the story when studying the education effects on labor market outcomes such as salaries, the concept of educational mismatch arises.

Allen et al. (2001) give a general definition of the concept. Educational mismatch happens when the educational level or qualification of a worker does not match the ones required for the job given the job characteristics.[1] Educational mismatch can be vertical and/or horizontal. Kim et al. (2011) state that an educational mismatch is vertical when the employee's educational attainment (or level) is different from the one required for the job. A general example would be an individual holding a doctoral degree working in a job requiring a non-graduate level of education; this situation is usually defined as a positive mismatch (that leads to overeducation). A negative mismatch (leading to undereducation) example would be a worker with complete high-school working in a job that requires a bachelor's degree. A mismatch can (also) be horizontal when worker's field of study does not match the requirements of the job. Consider the case of an employee with an physics degree working in a job that requires medical knowledge. In this case, the quality of the match may be good (the field is exactly the one required by the job), weak (if it barely meets the occupation's requirement) and mismatch (like the previous example) (Robst, 2007).

The authors mention two key differences between vertical and horizontal mismatch. In the case of vertical mismatch, specific human capital of the worker can be preserved and developed in both positive or negative mismatch. This leads to a potential interest of the worker and/or the employer to form a mismatch. In the case of a positive mismatch, a worker may find the job as an opportunity to show his abilities to get a better one in the company or in another company. Also, the employer may take advantage of the extra knowledge of the worker. Now, consider a negative mismatch; in a situation of low supply of some specific skills, an employer may want to hire an undereducated worker and develop his knowledge through training and experience.In contrast, when a horizontal mismatch arises, it is more likely for the general human capital to play an important role. In some cases, similarities or common elements between the worker's profile and the job's requirements are much harder to define or do not exist. The second

[1]Even though the term required is ambiguous, it is widely used in literature on educational mismatch. A job's educational requirement should be interpreted as the usual education (level in the vertical case and qualification in the horizontal situation) in a specific occupation, not necessarily the ideal or desired education given the job characteristics. In this thesis, the required or usual education will be the mode years of schooling of the particular occupation.

difference between vertical and horizontal mismatch is the approaches to measure/identify the mismatch. Vertical mismatch is usually measured with a statistical approach (as deviations from the mode/mean years of education of a specific occupation) while horizontal mismatch is often identified with a subjective method (by asking directly to the worker if his/her field of study is relevant to his/her occupation).[2]Given the information available in the database used in this thesis, the vertical dimension of educational mismatch will be studied.

2.2 Measuring Vertical Educational Mismatch

In most research of this topic, three measures of vertical mismatch are used. Cedefop (2010) distinguishes them as follows:

- **Subjective method**: this method is based on self-reported information provided by the worker to determine if there is a match (or mismatch) between his educational attainment and the one required in the job. Verhaest and Omey (2006) separate this method into direct self-assessment and indirect self-assessment.In the first case, respondents are directly asked whether they feel over or undereducated for their job The authors give the following example: Do you have a level of education which is according to your own opinion too high, too low or appropriate for your job? In the case of indirect self-assessment, mismatch is determined by asking the individual what he feels is the required level of schooling for his job and comparing the actual educational attainment of the worker and his answer to the previous question.

- **Objective method**: this method is based on expert evaluation of required educational level for the job. Proper level of education for an occupation can be determined using ISCO classification.

- **Statistical method**: this method uses data on educational level distribution per occupation. There is a mismatch when the worker's education differs by more than one standard

[2]The following section of this chapter discusses other measures for vertical mismatch in detail.

deviation from the mean level of education in his occupation. Another statistical method is the one proposed by Cohn et al. (2000), that measures the educational mismatch relative to the mode instead of the mean. In this case, an overeducated worker has more years of education than the modal level of his occupation; on the contrary, an undereducated worker is less educated than the modal value of years of education in his occupation.

The three approaches have some limitations. In the case of the subjective method, the main weakness is that respondent's opinion on required education may differ from employer's. With the objective method, there is an issue with the level of disaggregation of occupations done by the researcher. The broader the definition of the occupation, the more heterogeneous workers it includes. The results obtained with the statistical approach lead to interpret that multiple occupations with the same mean or mode require the same educational attainment, which may not be the case. In this thesis, educational mismatch will be measured as as deviations from the mode years of education per occupation for two reasons. The first and main reason is the unavailability of information in the data to evaluate mismatch with a subjective or objective requirement such as direct questions like the one previously mentioned or an expert evaluation. Additionally, the statistical method provides a more standardized measure, making comparison to other works using a similar approach possible.

2.3 Educational mismatch and wages

The empirical approach in studying the effects of education on wages considering educational mismatch was introduced by Duncan and Hoffman (1981) with the ORU model. The model is an extension of the Mincer equation that separates the total years of education of an individual into three components: years of education required for an specific job, years of overeducation and years of undereducation.[3] Their main finding is that wages are not only determined by the level of education, but also by the returns on over and undereducation. They find a positive and significant return in the first case and, on the contrary, a significantly negative effect in the second.Unlike the methodology used in this work, the authors determine the mismatch using a (subjective) direct self-assessment measure.

Cohn et al. (2000) study the incidence and wage effects of over and underschooling in Hong Kong using data from the 1986 Hong Kong By-census and the 1991 Hong Kong Census. To measure the vertical mismatch, they use a variation of the statistical method proposed by Kiker et al. (1997) and used in this thesis. They define the variables SCHOOL as the years of education of an individual and adequate schooling (ADSCH) that is equal to the mode of SCHOOL for each occupation. So, a worker is matched (no vertical mismatch) if his SCHOOL equals the modal value of his occupation. A positive mismatch arises when a worker has a level SCHOOL greater than the modal level of schooling for their specific occupation.People with SCHOOL lower than the modal value of the SCHOOL for their occupation generate a negative mismatch.

Using a standard ORU wage equation, one for each gender, the authors find that overschooled (underschooled) workers are better (worse) remunerated than their matched co-workers. However, they perceive wages significantly lower (higher) than the wages they would receive in a job for which they are properly matched. Also, they found that the rates of return to a proper match and overeducation (undereducation) decline (rise) when labor market experience rises. Other similar research done by Bauer (2002), Hartog (2000) among others also focus on the effects on wages of a match, positive or negative mismatch. They find that wage losses due to vertical mismatches go from 13% to 19% of matched workers' wages.

[3]The authors use the subjective method to measure educational mismatch by asking the employees how much formal education is required to get a job like theirs. Having information on the actual educational attainment of the workers, Duncan and Hoffman define overeducation as the additional years of education

Studying the role of the returns to education considering educational mismatch in the gender wage gap is still not widespread. Salinas-Jimenez (2013) analyze the gender wage gap in Spain for 2006 and the role of educational mismatch determining it. The authors use a Oaxaca-Blinder Decomposition (OAB) proposed by Oaxaca (1973) and Blinder (1973) in separate for adequately matched, negatively and positively mismatched workers and by subgroups according to their education level (up to secondary or higher education) in order to determine if the wage gap is due to differences in characteristics between the sexes, differences in the payment to these characteristics or both. Their main finding is that most of the gender wage gap for the first group of workers is due to differences in returns but both effects play similar roles considering workers with higher education.

Boll (2014) proceeds to estimate the gender wage gap in Germany (for graduate workers) and weather education variables explain it analyzing an unbalanced panel for the years 1984-2010 using a OAB and find that the education plays a small role in determining the wage gap when considering educational mismatch. Both Boll and Salinas-Jimenez use separate models for men and women and use the Chow test to indicate that the estimates are different.

Chapter 3

Model

3.1 Returns to education

In order to estimate the returns to years of education, I use the standard framework based on Mincer (1974) with the earnings equation as follows:

$$ln(w_i) = \alpha + \beta S_i + \gamma_1 E_i + \gamma_2 E_i^2 + \delta X_i + \epsilon_i$$

where wages (w) are explained by years of schooling (S), labor market experience (E) and its squared expression (E^2), a vector of control variables (X) such as occupation, hours of work, sector (private or public), race, parenthood, marital status among others and the error term (ϵ). [1] The education component of this equation reflects the individual's choice of educational attainment, regardless the job in which he/she works.

[1] Age is excluded because the variables of schooling and experience uniquely identify it. Also, a gender dummy is not included because the equation will be used for each sex separately.

To consider educational mismatch, I use the ORU framework proposed by Duncan and Hoffman (1981),[2] that decomposes the variable S into years of underschooling (US), years of required education (RS) years of overschooling (OS) and years of overschooling (OS), so the Mincer equation transforms to:

$$ln(w_i) = \alpha + \beta_U US_i + \beta_R RS_i + \beta_O OS_i + \gamma_1 E_i + \gamma_2 E_i^2 + \delta X_i + \epsilon_i$$

This equation allows us to determine if the effect of education on wages differ when considering educational mismatch. This means that, if a person is overeducated (undereducated), the returns of a year of overeducation (undereducation) may not be the same to the returns of a year of required education for a specific occupation. In fact, these years could have greater, smaller or even negative or effect on wages, implying that productivity levels and wages are flexible and dependent on the workers relative educational attainment, not just the nominal requirements of the job. In other words, wages are determined by the worker's education decisions (captured in the educational mismatch components which happen to be worker's educational level deviations from the required education in his/her occupation) and the occupation's side (captured in the required years' component) of the labor market.[3]In the alternative case that the returns of educational mismatch are zero, this would suggest that the worker's wage is fully determined by the job's education requirements.

[2]ORU notation states for overeducation, required education and undereducation

[3]As mentioned earlier, the years of schooling required in an occupation (RS) should be interpreted as the usual educational level of workers within the occupation (in this case, the mode years of schooling) and not necessarily the desired educational attainment for workers in the specific occupation.

3.2 Decomposing the gender wage gap

To determine the role of educational mismatch on the gender wage gap, I use the OAB standard decomposition method:

$$\overline{ln(w_M)} - \overline{ln(w_F)} = (\overline{X}_M - \overline{X}_F)\beta_F + (\beta_M - \beta_F)\overline{X}_F + (\overline{X}_M - \overline{X}_F)((\beta_M - \beta_F)$$

The left term of the equation denotes the difference in expected values of the male and female log wage income. The first term on the right hand side captures the endowment effect given by differences on *endowments* or characteristics between men and women.The second component on the right hand side reflects the coefficient effect driven by differences in the *returns* to these characteristics. And the third term is an interaction term that accounts the fact that both effects happen simultaneously. By construction, the endowment and the return effect sum up to the averaged gender wage gap. This effects (the total and the individual per variable) may be positive or negative.

Chapter 4

Data and Descriptive Statistics

4.1 Data

The data used in this thesis corresponds to a pooled cross-section of the American Community Survey (ACS) for the years 2010 and 2015. This mandatory survey is conducted every year by the U.S. Census Bureau and selects a random sample of households of about 3.5 million addresses per year. It provides detailed information on people's demographic, economic, and social characteristics, making the data very attractive for empirical research.

The analysis will consider a whole representative sample of the US labor force and a smaller sample with graduate workers only in order to analyze if the results change considering the highly educated workers. The observations used are restricted to individuals who are part of the labor force, with relevant information available such as educational attainment, wage income, work hours, and occupation. Self-employed workers are excluded because they lack wage income information. A total of 492 occupations are considered when measuring educational mismatch. For the whole sample, a total of 2,436,417 men and women of ages 18 to 66 are used in the analysis, where 50.5% of the observations are men and the remaining 49.4% are women. Considering only graduate workers, 265,658 observations are included; 52% of them are women and 48% men. The regressions for the graduate sample only consider occupations with a share of graduate workers greater than 10%.

Table 4.1 presents a brief description of the variables used in the analysis.

Table 4.1: Description of variables

Variable	Operational Definition
LNINCWAGE	Log of annual pre-tax wage and salary income for the previous year
S	Years of schooling
US	Years of underschooling w.r.t. the occupation's mode
RS	Mode years of schooling by occupation
OS	Years of overschooling w.r.t. to the occupation's mode
OCC	Categorical variable representing the individual's occupation
EXP	Potential years of work experience; EXP = (age – education – 5)
UHRSWORK	Number of hours worked per week for the previous year
PUBLIC	Dummy variable for sector of employment
MALE	Dummy variable for gender
WHITE	Dummy variable for race
MARRIED	Dummy variable for current marital status
CHILD	Dummy variable for present children residing with the individual
REGION	Region and division where the housing unit was located

Source: American Community Survey and own elaboration

4.2 Descriptive Statistics

Table 4.2 provide summary statistics on wages, years of education, educational mismatch variables, experience and usual hours of work per week.

Table 4.2: Summary statistics

Whole sample	Men				Women					
Variable	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max	Mean Diff.	t-test
Log wage income	10.503	1.064	1.386	13.397	10.121	1.066	1.386	13.397	0.38	-2.8e+02***
Years of education	13.7	2.92	0	23	14.039	2.771	0	23	–0.339	95.1590 ***
Years undereducated	0.67	1.488	0	21	0.62	1.418	0	11	0.05	-6.8967***
Years required	13.3	2.158	12	23	13.7	2.342	12	23	–0.4	112.9005***
Years overeducated	0.80	1.68	0	11	0.82	1.68	0	11	–0.02	4.0799***
Experience	23.011	13.057	–7	61	22.827	13.329	-7	61	0.184	-11.0921***
Usual hours worked per week	39.269	12.612	0	60	35.02	12.543	0	60	4.249	-2.7e+02***

Graduate sample	Men				Women					
Variable	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max	Mean Diff.	t-test
Log wage income	11.372	.866	2.303	13.397	10.915	.842	1.386	13.397	0.457	-1.4e+02***
Years of education	18.755	1.79	18	23	18.501	1.501	18	23	0.249	-40.0180***
Years undereducated	0.03	.437	0	5	0.03	.44	0	5	0	0.3365***
Years required	16.2	1.646	12	23	16.5	1.5	12	23	– 0.3	54.4392***
Years overeducated	2.45	2.044	0	11	1.95	1.896	0	11	0.5	-76.8704***
Experience	22.544	11.126	-7	43	20.576	11.23	-7	43	1.968	-45.6022***
Usual hours worked per week	43.904	9.929	0	60	40.322	10.714	0	60	3.582	-7.0568***

*** $p < 0.001$

Source: American Community Survey and own elaboration

As can be observed, the log wage income for men is significantly higher than for women in both samples. Also, we can see that the averages of experience and usual hours of work are higher for men (for both samples). A priori, this may tell us that differences in these variables could explain part of the gender wage gap (also for variables such as education and years required for the graduate sample). In contrast, variables like years required in the graduate sample plus years of education and years overeducated may help to reduce the gender wage differentials.[1]

Table 4.3: Educational attainment distribution by gender

Educational Attainment (%)	Men	Women
Incomplete high school	8.3	5.5
Complete high school	52.4	48.9
Professional Degree	7.9	10.6
Bachelor's Degree	20.1	22.1
Master's Degree	9.8	11.8
Doctoral Degree	1.5	1.2
Total (%)	100	100

Source: American Community Survey and own elaboration

Table 4.3 presents the distribution of educational attainment by gender. As we can see, around 50% of men and women in the sample have complete high school which corresponds to twelve years of education. The incidence of incomplete high school is higher for men (around 50% higher relative to women). Additionally, women more often hold a bachelor's degree (10% more than men) and graduate degrees (15% more than men).

[1]It is important to highlight that women having more required years than men (on average) tells us that women are more educated than men, not that educational requirements on women are higher.

Table 4.4: Educational mismatch distribution by gender

	Whole sample		Graduate Sample	
	Men	Women	Men	Women
Negative Mismatch	15.8	17.7	0.6	0.7
Match	60.5	59.0	30.3	46
Positive Mismatch	23.7	23.3	69.1	53.3
Total	100	100	100	100
Observations	1232686	1203731	127791	137867

Source: American Community Survey and own elaboration

Table 4.4 denotes the distribution of educational mismatch and the average years of under, required and overshcooling by gender. The fist thing to notice is that incidence of educational mismatch (positive or negative) is significant for both men and women. For the whole sample, negative mismatch is more incident on women (12% higher). Also, the proportion of overeducated workers is higher for both men and women relative to the undereducated (28% and 50% more for women and men respectively).

Comparing the graduate sample distribution to the whole sample distribution, we can see two things: differences in educational mismatch between genders are greater when accounting graduate workers only and the incidence of educational mismatch is even higher for this subsample given the increase in the proportion of workers with positive mismatch.[2] 46% of women are properly matched in their jobs, while properly matched men represent 30% of total and the proportions of overeducated workers more than double for both men and women relative to the whole sample.

[2]This is not surprising given that the sample includes the highly educated workers only.

Chapter 5

Results

5.1 Returns to education

Table 5.1: Estimated returns to education

	Whole sample		Graduate Sample	
Log wage income	Men	Women	Men	Women
i) Mincer equation				
Years of education	0.0571***	0.0572***	0.0331***	0.0337***
	(196.98)	(189.36)	(27.50)	(25.93)
R^2	0.56	0.58	0.39	0.41
Chow Test	0.248***	0.723***	1.96***	1.769***
ii) ORU equation				
Years undereducated	-0.0606***	-0.0695***	-0.0317***	-0.0380***
	(-129.99)	(-139.38)	(-6.00)	(-5.06)
Required years of education	0.221***	0.189***	0.119***	0.149***
	(67.10)	(25.83)	(9.82)	(5.61)
Years overeducated	0.0544***	0.0489***	0.0332***	0.0341***
	(134.00)	(121.10)	(26.86)	(25.47)
R^2	0.56	0.58	0.39	0.41
Chow test	0.3375***	1.0245***	0.520***	0.963***
Observations	1232686	1203731	127791	137867

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5.1 documents the results of the Mincer and ORU equations that estimate the returns to education on wages (the latter considers educational mismatch variables) for men and women considering the whole sample and a subsample with graduate workers only.[1]

Looking at the results derived with the Mincer model, we observe a positive and highly significant return to an additional year of schooling for men and women in both samples. However, the female sex is rewarded with a higher pay relative to males in the graduate sample with each year of education giving an increase of 3.37% to women and 3.31% to men. As previously mentioned, this model only reflects the individual's education decisions, telling one part of the story in wage determination. The ORU model presented gives us more disaggregated results.

On average, an additional year of undereducation yields a negative return of 6.06% and 6.95% for men and women in the whole sample while 3.17% and 3.8% respectively in the graduate sample. A year of overeducation pays in the two samples, being this reward of 5.44% and 4.89% for men an women in the whole sample while 3.32% and 3.41% considering gradutate workers only. Finally,if the worker's occupation required an additional year of education, the pay will is high for this year. For women in the whole sample, the wage increases 18.9% while graduate women earn 14.9% more. In the case of men, the increases for the whole and graduate sample correspond to 22.1% and 11.9% respectively.[2]

Three important results must be highlighted from the same table: matched graduate and positively mismatched females win by receiving a higher return for each year of required and over schooling than their male counterparts; however, they lose in a negative mismatch situation relative to men. Finally, the returns to overeducation are smaller than the returns to required education, which implies that overeducated workers earn more than their co-workers, but less

[1]Appendix 1 and 2 document the results in detail

[2]Chow tests are estimated to affirm that the estimations between men and women are different. The statistics obtained are highly significant

than properly matched workers with the same educational level.

The results on education variables are consistent with the ones obtained by Duncan and Hoffman (1981), Cohn et al. (2000), Boll (2014) and Salinas-Jimenez (2013) as we can see that productivity levels and wages are flexible and depend on the worker's relative educational attainment. In contrast, Verdugo and Verdugo (1989) a negative pay to overeducation. The reason for this swap is that their specification differs from the one in Duncan & Hoffman's paper in three aspects. First, they measured required schooling with the mean statistical method. Second, they included dummies for being over or undereducated instead of a year level variable. Finally, they control for completed years of schooling in the regression. Given this notation, their results are to be interpreted differently: a negative return to overeducation (undereducation) implies that the worker earns less (more) than a properly matched individual with the same educational level.

As mentioned in the literature review, Cohn et al. (2000), Salinas-Jimenez (2013) and Boll (2014) work with independent models for men and women instead of including gender interaction variables. Using a Chow Test, they determine that the coefficients of over, under and required years of education differ for both models giving men advantage for the mismatch variables and for women in the required years. Therefore, my results are consistent with the negative mismatch estimates but differ for the required years' reward for the whole sample (being greater for men) and the returns to overeducation for the graduate sample (being greater for women).

5.2 Decomposition of the gender wage gap

Table 5.2: Blinder-Oaxaca Decomposition

	Whole Sample	Graduate Workers
Gender wage gap	0.363***	0.459***
	(259.40)	(135.50)
Endowments	0.192***	0.208***
	(175.07)	(96.40)
Coefficients	0.194***	0.252***
	(183.30)	(79.65)
Interaction	-0.0238***	-0.000417
	(-37.55)	(-0.23)
i) Endowments		
Years undereducated	-0.00359***	0.0000204
	(-22.81)	(0.33)
Required years of education	-0.0433***	-0.0120***
	(-109.63)	(-19.30)
Years overeducated	-0.000812***	0.0203***
	(-6.03)	(25.35)
ii) Coefficients		
Years undereducated	0.00796***	0.000412
	(22.14)	(1.50)
Required years of education	-0.0217**	-0.381***
	(-2.58)	(-9.03)
Years overeducated	0.00130**	-0.0221***
iii) Interaction		
Years undereducated	0.000660***	-0.00000591
	(15.98)	(-0.32)
Required years of education	0.000518**	0.00708***
	(2.58)	(8.90)
Years overeducated	-0.0000199*	-0.00926***
	(-2.40)	(-8.92)
Observations	2566722	269627

t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: American Community Survey and own elaboration

The results of the Oaxaca (1973) and Blinder (1973) wage gap decomposition are presented in Table 5.3.[3]. The mean gender wage gap amounts to 36.3% and 45.9% for the whole sample

[3]Appendix 3 documents this results in detail. Appendix 4 is a Blinder-Oaxaca decomposition with a Mincer model as base included as a robustness check due to the large coefficient effect of required years of education in the graduate sample.

and graduate workers only respectively. Also, 19.2 percent point (pp) of the total are due to differences in endowments or characteristics between men and women, while 19.4pp refers to different coefficients or remuneration of this characteristics and -2.38pp to the interaction of this factors. In general terms, the effects are split 50-50.

5.2.1 Endowment effect

For the whole sample, the three education components have a significant and negative effect on the gender wage gap, meaning that differences in these characteristics between men and women reduce it. In other words, women tend to be more educated than men and, in the opposite case, the gender wage gap would be even greater.

In detail, women earn on average 4.33% more than men because they work more often in jobs with higher required education. They also win 0.36% (0.08%) more than their counterparts because they tend to work more frequently in jobs where their negative (positive) deviations to the occupation's education requirements are smaller (greater) than men (i.e, they tend to be undereducated less often. This is consistent with the facts that women, on average, have more years of required education and overeducation than men and less years of undereducation. [4]

Considering graduate workers only, the years of undereducation component is not significant and the years of overeducation have a positive significant effect on the gender wage gap in contrast with the whole sample. In this case, women earn 1.2% more than men because they work more often in jobs that require more years of education; they also receive 2% less than their male counterparts because they work more frequently in jobs where their positive deviations to the occupation's requirements are relatively lower.

[4]This can be corroborated with the statistics collected in Table 4.2

My results contrast with those of Boll (2014) who finds that only required years of education explain the gap. In this case, Salinas-Jimenez (2013) proceed with individual models for every component of education and also find very small effects of differences in endowments in determining the wage gap.

Apart from endowment differences in education, other variables play a significant role. For example, men earn 24.7% more because they work more hours per week than women. [5]

5.2.2 Coefficient effect

The coefficient component reflects differences in the returns to characteristics between men and women, so a positive effect of a variable means that women receive a smaller pay for each unit, while a negative sign means that the pay to this variable reduces the gender wage gap. Considering the whole sample, the educational mismatch components have positive signs while the required years of education reduces the gender wage gap. Men earn 0.79% and 0.13% more tan women because they receive a higher returns to under and overeducation. In contrast, women earn 2.17% more because their pay on required years of education is higher

Graduate workers have negative and significant effects of on the years of required and over education. Graduate women earn 2.21% and 38.1% more than men because the return for overeducation and required education is higher for them. This means that women have a great incentive to work in jobs that require more education because the pay is significantly greater for them.Boll (2014) finds insignificant price effects for the three education components, so no comparisons can be made.

Differences in returns to other variables reinforce or reduce the gap, as in the endowment effect. Significant factors that reinforce the gap because of the price effect are: experience,

[5] Appendix 3 enlists the complete Oaxaca-Blinder Decomposition

married status, children present and white race. In contrast, experience squared and usual hours worked per week reduce the gap. This factors explain a greater part of the wage gap compared to t mismatch variables.

Women earn 6.3% and 3% less than men because they are payed less when they get married or have children respectively (the percentages are similar for the graduate sample). This may reflect some family sharing decisions that benefit men in terms of income. Also, men earn 35% more than women because the pay for experience is significantly higher for them in the first years; however, as this years increase, women receive better compensations. Additionally, women earn 25.8% more than men because they receive a higher remuneration for each hour of work.[6]

5.2.3 Total educational mismatch effect

In order to determine the role of educational mismatch on the gender wage gap, I add the endowment, coefficients and interaction effects of this variables. In aggregate, the years of undereducation explain 1.38pp of the gender wage gap for the whole sample while the effects for graduate workers are insignificant.

The total effect of required years is negative, which means that the gap would be bigger if this variable aggregate effect did not exist. Relative to the wage gap, this component represents -17.7pp of the total, meaning that it would be 17.7% bigger in its absence. The gender wage gap would be 84% greater in the absence of the total effect of years of required education. Finally, the ratio of the total overeducation effect and the wage gap (multiplied by 100) is 0.13pp for the whole sample and -2.4pp for the graduate subsample.

[6]This results are corroborated by adding a gender interaction to the married, children and usual hours of work variables and can be found in Appendix 7

Chapter 6

Conclusions

When studying labor market results such as wages and the impact of education returns , it is important to consider educational mismatch. In our sample, 40% of the workers are mismatched according to the statistical method used in this thesis to measure it. Some of the results obtained in this thesis complements previous work regarding differences in returns to education, whereas it is a required, over or under year of education. The coefficients for educational mismatch are highly significant, which reinforces the fact that wages are flexible and not determined by both sides of the labor market.

In the ORU model, I find that women carry greater losses relative to men when they are negatively mismatched in both samples, but the graduate women win when they are overeducated relative to men. On the other hand, matched men in the whole sample are the winners relative to women while the opposite result happens for graduate workers only.

Parallel, the gender wage gap is a persistent phenomenon in both developed and developing countries. My main finding is that, in general, educational mismatch level differences between men and women exist. However, their magnitude is relatively small to other factors and does not fully explain the gender wage gap. The only relevant effect of mismatch is the coefficient effect for required years of education. The gender wage gap for graduate workers would be significantly higher in the absence of this effect. Other variables such as usual hours of work per week, experience, children and marital status play a significant role in the wage gap.

References

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

Battu, H., Belfield, C. and Sloane, P. (2000). How well can we measure graduate overeducation and its effects? National Institute Economic Review 171: 82–93.

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

Becker, G. (1964). Human Capital Theory, Chicago, University of Chicago Press.

Blinder, A. S. (1973): Wage Discrimination: Reduced Form and Structural Estimates, The Journal of Human Resources 8 (4): 436-455.

Boll, C., Leppin, J.S.(2014). Equal matches are only half the story.Why German female graduates earn 27 % less than males. HWWI Policy Paper, No. 75. HWWI, Hamburg.

BUCHEL, F., DE GRIP, A., MERTENS,A. (eds), Overeducation in Europe, Edward Elgar,Cheltenham.

Cedefop (2010), The skill matching challenge–Analysing skill mismatch and policy implications, Publications Office of the European Union, Luxembourg.

Cohn, E. and Ng, C. Y. (2000). Incidence and wage effects of underschooling and Overschooling in Hong Kong. Economics of Education Review 19: 159–168.

Croce, G. and Ghignoni, E. (2012) ‘Demand and Supply of Skilled Labour and Overeducation in Europe: a CountryLevel Analysis”, Comparative Economic Studies, 54(2): 413-439

Cutillo, A. and G. DiPietro (2006), “The Effects of Overeducation on Wages in Italy: a Bivariate Selectivity Approach”, International Journal of Manpower, Vol. 27, No. 2, pp. 143-168.

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

Franzini M, Raitano F (2011) Few and underutilized? Over-education of Italian graduates. In: Mandrone E (ed) L’intermediazione pubblica, privata ed informale. ISFOL, Roma

Freeman, R. B. (1976). The Overeducated American. NewYork: Academic Press.

Goldin, Claudi (1990). Understanding the Gender Gap: an economic history of American women, New York, Oxford University Press.

Groot, W. and van den Brink, H. (2000). Overeducation in the labour market: a meta analysis. Economics of Education Review 19(2): 149–158.

Hartog, J. and Oosterbeek, H. (1988). Education, allocation and earnings in the Netherlands: overschooling? Economics of Education Review 7(2): 185–194.

Hersch, J. (1991) Education match and job match. Review of Economics and Statistics 73, 140-144.

Jackman, R., Layard, R. and Savouri, S. (1991). 'Mismatch: a framework for thought.' In Padoa-Schioppa (1991).

Kelly, E., O'Connell, P.J. and Smyth, E. (2010). 'The economics returns to field of study and competencies among higher education graduates in Ireland', Economics of Education Review, vol. 29(4), pp. 650-657

Kiker, B., Santos, M. and Mendes de Oliveiria, M. (1997). Overeducation and undereducation: evidence for Portugal. Economics of Education Review 16(2): 111–125.

Kim, H-K., Seung A., Kim J.(2011), Vertical and Horizontal Education-Job Mismatches in the Korean Youth Labor Market: A Quantile Regression Approach, "Sogang University Working Paper"

Leuven, E., and Oosterbeek, H. (2011), 'Overeducation and Mismatch in the Labor Market'. In Hanushek, E., Machin, S., and Woessman, L. (ed.), Handbook of the Economics of Education, 4, 283-326.

McGuinness, S. and Bennet, J. (2006). Overeducation in the graduate labour market: a quantile regression approach. Forthcoming, Economics of Education Review.

Oaxaca, R. (1973). Male-female wage differentials in urban labour markets, International Economic Review, 14, 693–709.

Rasovec, T., Vavrinová, T. (2014). Skills and Educational Mismatch in the Czech Republic: Comparison of Different Approaches Applied on PIAAC Data. National Training Fund, Prague, Czech Republic.

Robst, J. (2007). Education and job match: The relatedness of college major and work, Economics of Education Review, 26(4), 397-407.

Rubb, S. (2004). Overeducation in the labor market: a comment and re-analysis of a meta analysis. Economics of Education Review 22: 621–629.

Salinas-Jimenez, M., Rahona-López, M., Murillo-Huertas, I.(2013). Gender Wage Differentials and Educational Mismatch: An Application to the Spanish case. Applied Economics 45(30), 4226-4235 .

Sicherman, N. (1991). “Overeducation” in the labour market. Journal of Labor Economics 9(2): 101–122.

Tsang, M. (1987). The impact of underutilisation of education of productivity: a case study of the U.S. Bell companies. Economics of Education Review 11: 239–234

U.S. Bureau of Economic Analysis, Shares of gross domestic income: Compensation of employees, paid: Wage and salary accruals retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/W270RE1A156NBEA, April 15, 2017.

Verdugo, R. and Verdugo, N. (1989). The impact of surplus schooling on earnings. Journal of Human Resources 24(4): 629–643.

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

Appendix 1: Estimated returns to education with the Mincer equation

	Whole sample		Graduate Sample	
	Men	Women	Men	Women
Number of years of education	0.0571***	0.0572***	0.0331***	0.0337***
	(196.98)	(189.36)	(27.50)	(25.93)
Experience	0.0521***	0.0362***	0.0560***	0.0421***
	(241.06)	(173.30)	(71.64)	(60.99)
Experience squared	-0.000794***	-0.000512***	-0.000968***	-0.000763***
	(-179.36)	(-117.21)	(-56.04)	(-47.75)
Usual hours worked per week	0.0408***	0.0498***	0.0303***	0.0397***
	(557.89)	(756.01)	(134.59)	(216.20)
Public	0.0550***	0.0651***	-0.0527***	0.0926***
	(24.71)	(34.51)	(-10.13)	(21.65)
Married	0.149***	0.0511***	0.136***	0.0458***
	(88.20)	(36.34)	(25.61)	(11.44)
Children	0.0629***	0.0146***	0.0775***	0.0436***
	(39.11)	(10.20)	(16.49)	(10.58)
White	0.0545***	0.00287	0.0376***	0.00745
	(33.45)	(1.84)	(7.46)	(1.62)
Northeast region	0.0154***	0.0311***	0.0388***	0.0977***
	(7.77)	(15.79)	(6.89)	(18.75)
Midwest region	-0.0942***	-0.106***	-0.102***	-0.0754***
	(-49.41)	(-56.26)	(-17.27)	(-13.70)
South region	-0.0970***	-0.121***	-0.0530***	-0.0817***
	(-57.48)	(-71.73)	(-10.45)	(-17.10)
Constant	7.911***	7.625***	8.938***	8.547***
	(947.15)	(662.82)	(321.02)	(268.01)
Observations	1232686	1203731	127791	137867

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix 2: Estimated returns to education with the ORU equation

	Whole sample		Graduate Sample	
	Men	Women	Men	Women
Years undereducated	-0.0606***	-0.0695***	-0.0317***	-0.0280***
	(-129.99)	(-139.38)	(-6.00)	(-5.06)
Required years of education	0.221***	0.189***	0.119***	0.149***
	(67.10)	(25.83)	(9.82)	(5.61)
Years overeducated	0.0544***	0.0489***	0.0332***	0.0341***
	(134.00)	(121.10)	(26.86)	(25.47)
Experience	0.0521***	0.0360***	0.0560***	0.0421***
	(240.87)	(172.47)	(71.65)	(61.02)
Experience squared	-0.000793***	-0.000507***	-0.000969***	-0.000763***
	(-178.95)	(-116.03)	(-56.05)	(-47.78)
Usual hours worked per week	0.0408***	0.0497***	0.0303***	0.0397***
	(557.98)	(755.92)	(134.62)	(216.22)
Public	0.0547***	0.0640***	-0.0522***	0.0926***
	(24.56)	(33.96)	(-10.05)	(21.65)
Married	0.149***	0.0512***	0.136***	0.0457***
	(88.22)	(36.44)	(25.63)	(11.43)
Children	0.0634***	0.0157***	0.0775***	0.0436***
	(39.40)	(11.00)	(16.50)	(10.60)
White	0.0539***	0.00157	0.0378***	0.00755
	(33.10)	(1.01)	(7.51)	(1.65)
Northeast region	0.0149***	0.0297***	0.0386***	0.0976***
	(7.52)	(15.13)	(6.86)	(18.75)
Midwest region	-0.0950***	-0.109***	-0.102***	-0.0755***
	(-49.77)	(-57.38)	(-17.29)	(-13.72)
South region	-0.0974***	-0.122***	-0.0531***	-0.0817***
	(-57.68)	(-72.37)	(-10.48)	(-17.10)
Constant	5.167***	7.911	7.472***	6.518***
	(97.18)	(0.01)	(37.86)	(15.05)
Observations	1306890	1259832	129694	139933

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix 3. Oaxaca-Blinder Decomposition: ORU model

	Whole Sample	Graduate Workers
Group 1 (Male Workers)	10.43***	11.35***
	(10559.04)	(4602.41)
Group 2 (Female workers)	10.06***	10.90***
	(10151.14)	(4687.38)
Difference (Gender wage gap)	0.363***	0.459***
	(259.40)	(135.50)
Endowments	0.194***	0.222***
	(175.84)	(99.07)
Coefficients	0.196***	0.223***
	(184.37)	(69.44)
Interaction	-0.0267***	0.0151***
	(-41.65)	(7.68)
Endowments		
Years undereducated	-0.00359***	0.0000204
	(-22.81)	(0.33)
Required years of education	-0.0433***	-0.0120***
	(-109.63)	(-19.30)
Years overeducated	-0.000812***	0.0203***
	(-6.03)	(25.35)
Experience	0.00763***	0.0827***
	(11.60)	(35.23)
Experience squared	-0.000961*	-0.0655***
	(-2.24)	(-31.08)
Usual hours worked per week	0.247***	0.161***
	(318.20)	(89.80)
Public	0.00171***	0.0168***
	(19.79)	(24.55)
Married	0.00484***	0.00797***
	(49.68)	(15.46)
Children	-0.000635***	0.00206***
	(-10.30)	(10.19)
White	0.000362***	0.0000364
	(10.09)	(0.54)
Northeast region	-0.000199***	-0.00102***
	(-9.72)	(-7.41)
Midwest region	0.000359***	0.000792***
	(6.28)	(4.66)
South region	0.000544***	0.000246
	(7.31)	(1.54)
group(occ)	-0.0187***	0.00868***
	(-53.91)	(17.37)

	Whole Sample	Graduate Workers
coefficients		
Years undereducated	0.00796***	0.000412
	(22.14)	(1.50)
Required years of education	-0.0217**	-0.381***
	(-2.58)	(-9.03)
Years overeducated	0.00130**	-0.0221***
	(2.61)	(-8.98)
Experience	0.354***	0.296***
	(49.06)	(12.53)
Experience squared	-0.195***	-0.0966***
	(-42.45)	(-6.79)
Usual hours worked per week	-0.258***	-0.320***
	(-70.93)	(-25.49)
Public	0.0157***	-0.0594***
	(30.26)	(-21.57)
Married	0.0626***	0.0657***
	(50.39)	(13.63)
Children	0.0301***	0.0216***
	(28.78)	(6.53)
White	0.0531***	0.0305***
	(28.53)	(5.11)
Northeast region	-0.00301***	-0.0122***
	(-5.25)	(-6.00)
Midwest region	0.00217***	-0.00707***
	(3.30)	(-4.08)
South region	0.0125***	0.0117***
	(13.34)	(4.36)
group(occ)	0.000805	0.0365***
	(0.39)	(6.94)
Constant	0.134***	0.659***
	(12.84)	(14.27)

	Whole Sample	Graduate Workers
Interaction		
Years undereducated	0.000660***	-0.00000591
	(15.98)	(-0.32)
Required years of education	0.000518**	0.00708***
	(2.58)	(8.90)
Years overeducated	-0.0000199*	-0.00926***
	(-2.40)	(-8.92)
Experience	0.00298***	0.0282***
	(11.31)	(12.08)
Experience squared	-0.000496*	-0.0145***
	(-2.24)	(-6.70)
Usual hours worked per week	-0.0315***	-0.0284***
	(-69.48)	(-24.67)
Public	-0.00384***	0.0223***
	(-29.04)	(21.01)
Married	0.00611***	0.0114***
	(42.96)	(13.35)
Children	-0.00265***	0.00172***
	(-26.23)	(6.20)
White	0.00143***	-0.0000315
	(23.31)	(-0.54)
Northeast region	0.0000987***	0.000838***
	(4.86)	(5.17)
Midwest region	-0.0000316**	0.000266**
	(-2.93)	(3.12)
South region	-0.000152***	-0.0000952
	(-6.44)	(-1.46)
group(occ)	0.000167	-0.00431***
	(0.39)	(-6.88)
Observations	2566722	269627

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Source: American Community Survey and own elaboration

Appendix 4. Blinder-Oaxaca Decomposition: Mincer model

	Whole sample	Graduate sample
Group 1 (Male workers)	10.43***	11.35***
	(10559.05)	(4602.43)
Group 2 (Female workers)	10.06***	10.90***
	(10151.14)	(4687.41)
Difference	0.363***	0.459***
	(259.40)	(135.50)
Endowments	0.184***	0.222***
	(166.05)	(100.25)
Coefficients	0.198***	0.223***
	(184.38)	(70.14)
Interaction	-0.0183***	0.0139***
	(-28.50)	(7.25)
Endowments		
Years of education	-0.0328***	0.00905***
	(-96.90)	(22.62)
Experience	0.00764***	0.0827***
	(11.60)	(35.22)
Experience squared	-0.000953*	-0.0655***
	(-2.24)	(-31.08)
Usual hours worked per week	0.251***	0.161***
	(318.92)	(89.84)
Public	-0.000176*	0.0166***
	(-2.07)	(24.97)
Married	0.00545***	0.00799***
	(53.13)	(15.51)
Children	-0.000794***	0.00206***
	(-12.70)	(10.21)
White	0.000519***	0.0000362
	(13.93)	(0.54)
Northeast region	-0.000200***	-0.00103***
	(-9.71)	(-7.42)
Midwest region	0.000363***	0.000791***
	(6.28)	(4.66)
South region	0.000539***	0.000246
	(7.31)	(1.54)
group(occ)	-0.0471***	0.00852***
	(-138.95)	(17.44)

	Whole Sample	Graduate Workers
coefficients		
Years of education	-0.114***	-0.317***
	(-22.04)	(-9.73)
Experience	0.354***	0.298***
	(48.68)	(12.66)
Experience squared	-0.192***	-0.0981***
	(-41.47)	(-6.89)
Usual hours worked per week	-0.268***	-0.321***
	(-73.40)	(-25.61)
Public	0.0105***	-0.0611***
	(20.25)	(-22.70)
Married	0.0614***	0.0656***
	(48.99)	(13.62)
Children	0.0304***	0.0215***
	(28.78)	(6.52)
White	0.0473***	0.0302***
	(25.25)	(5.07)
Northeast region	-0.00362***	-0.0123***
	(-6.26)	(-6.07)
Midwest region	0.00139*	-0.00719***
	(2.10)	(-4.15)
South region	0.0116***	0.0118***
	(12.23)	(4.37)
group(occ)	0.0327***	0.0339***
	(19.02)	(6.57)
Constant	0.227***	0.579***
	(30.59)	(15.29)

	Whole Sample	Graduate Workers
Interaction		
Years of education	0.00319***	-0.00435***
	(21.54)	(-9.45)
Experience	0.00298***	0.0285***
	(11.31)	(12.20)
Experience squared	-0.000487*	-0.0147***
	(-2.24)	(-6.81)
Usual hours worked per week	-0.0328***	-0.0285***
	(-71.80)	(-24.77)
Public	-0.00258***	0.0229***
	(-19.88)	(22.05)
Married	0.00599***	0.0114***
	(42.08)	(13.34)
Children	-0.00267***	0.00172***
	(-26.23)	(6.19)
White	0.00127***	-0.0000312
	(21.42)	(-0.54)
Northeast region	0.000119***	0.000848***
	(5.62)	(5.21)
Midwest region	-0.0000203*	0.000271**
	(-2.00)	(3.15)
South region	-0.000141***	-0.0000955
	(-6.30)	(-1.46)
group(occ)	0.00679***	-0.00401***
	(18.97)	(-6.52)
Observations	2566722	269627

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Source: American Community Survey and own elaboration