fondazione RODOLFO DEBENEDETTI

f R D B

XIV European Conference

The Wage Gap in the Transition from School to Work

Massimo Anelli, Sara de la Rica, Ainara de San Roman, Luca Flabbi, Giovanni Peri and Mauricio Tejada

The Wage Gap in the Transition from School to Work

This draft: May 20, 2012

By:

Giovanni Peri (UC Davis), Massimo Anelli (UC Davis), Sara de la Rica (University of the Basque Country), Ainara de San Roman (University of the Basque Country), Luca Flabbi (Georgetown University) and Mauricio Tejada (Georgetown University)

Summary of the Report:

In this report we analyze three fundamental aspects of the gender gap in labor market performance of highly educated individuals. First, is there discrimination between gender among employers and does this create a wage gap? Second, how relevant is the different role of men and women in raising children in affecting their labor market outcomes? Third, how relevant is the choice of college major in determining their occupation and wage and how different is it between genders?

In chapter 2 we use a structural model to separate the part of the wage difference between men and women in the US labor market that is due to productivity and the part due to prejudice/discrimination. We use very detailed data on US employers and employees to estimate the model. Interestingly we find that for all education groups prejudice explains a substantial part of the gender gap, although such share is decreasing from 1995 to 2005. As of 2005 largest part of gender differential in wages is explained by differences in productivity, except for the group of Masters and PhDs in which, surprisingly, prejudice plays a larger role in 2005 than earlier.

In Chapter 3 we analyze how the presence of children affects women labor market outcomes, relative to men, in particular their hours worked, hourly wage and their use of part-time. Using data on Spain we find that the very large use of part-time by women as of 2008 is explained in large part by the presence of children. Another interesting pattern we identify is that the gender gap in average wages in the nineties was mainly a result of family issues (having children penalized the woman), whereas by 2008, the average wage gap, has increased and it is not explained by family issues, but by other potential explanations (discrimination, segregation into different jobs, differences in the acquired skills).

This leads us to chapter 4 and 5 in which, using a novel database on high school career, college career and labor market outcomes for a panel of young Italians we analyze the role of choice of college major as determinant of gender gap in income. We find that the very large difference of choice of college major, in favor of humanities and away from math-engineering-business, explains one third of the gender gap in income and is very persistent. Moreover this choice does not seem driven by academic skill differentials that exist and are usually in favor or women. Non-academic skill differentials and psychological attitudes towards competition and towards the others may explain part of the choice. The impact of high school peers is found to be more relevant for women while the choice of a better partner (from the economic point of view) does not seem to drive the choice of college-major as people seem to marry within major rather than across.

Chapter 1: Gender Gap among college-educated: Three important factors

Giovanni Peri (UC Davis)
Massimo Anelli (UC Davis)

1.1 Introduction

Human resources and human talent is, by far, the most valuable economic factor available to each country. Even from a strictly economic point of view, giving women the same access to education, economic and political opportunities as men increases the possibility of a country to grow and to become wealthy. It doubles its human potential. Hence, gender inequality in economic and social opportunities and achievements is a topic of great relevance for many countries in the world and a reason of concern for some.

Economists and policy makers in poor countries have recently emphasized the "gender aspect" of development: empowering women has become a key aspect of several development policies. For instance, the World Development Report of 2012 is devoted to analyzing and dissecting gender disparity, in the world, with a special focus on developing countries, and proposing several policies to address it. The World Bank has developed a whole research agenda on "Gender and Development[1]". Developed countries are also monitoring their measures of gender gap very carefully and the goal of gender equality is explicitly stated in several Process and Strategy statements of companies and countries. Since 2006, for instance, the World Economic Forum has produced a yearly Global Gender Gap report, to monitor the progress that each country (rich and poor) is making in achieving gender equality in four key areas: health and survival, education, economic participation and opportunity, and political empowerment. Hence both scholars and policy makers are devoting increasing attention to this issue and to the causes of gender gaps in different domains and on the consequences of closing this gap.

[1] See the website at http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTGENDER/0,,menuPK:336874~pagePK:149018~piPK:149093~theSitePK:336868,00.html

We focus in this report on the gender gap in economic opportunity and we single out three countries, Italy, Spain and the United States. We choose these countries because we have a particularly good knowledge of the economies of these countries as we have studied them in the past, often in relation to issues related to the gender gap. They are all rich countries but they exhibit very different degrees of gender inequality in terms of economic access and income. In the Global Gender Gap report of 2011, the US was ranked 6th (from the top) in terms of gender inequality in earnings. Together with Northern European countries, the US is the one exhibiting the smallest gender gap in earnings. Still there is a very significant gender disparity in the US in access to highly paid and powerful positions both in the corporate world and in government (senior officers, CEO's, Board of Directors). While issues of gender inequality are also relevant among the disadvantaged (e.g. the poverty rate in the US is higher for women than for men) our study will mainly focus on gender inequality among highly educated. Hence in the US we will focus on discrimination and differences in productivity among highly educated women and men. Spain, on the other hand, has been a country with rapid economic growth in the last two decades, and similarly fast progress on gender inequality. The country was ranking 74th in the world relative to overall economic gender inequality (hence at the very middle of the distribution of 135 countries analyzed by the Global Gender Gap report). As we will document in Chapter 3 gender inequality in economic access and achievement is still significant in Spain, even when we focus only on the more educated. The different role of men and women in the family has affected heavily in the past the possibility of Spanish women to access the labor market. Only in the recent decades women in general and highly educated women in particular, have made progress in filling the economic gender gap.

Italy, finally, was ranked 90th in terms of gender differences in economic achievement by the Global Gender Gap report. This is the country with the largest gender income differences among OECD countries, except for Japan and Greece. While regional disparities in the gender Gap are still significant in Italy, and while the structure of the family is rapidly changing (the fertility rate in Italy is among the lowest in the world), there is concern that even highly educated Italian women do not have access to the same occupations and opportunities that men have. We will analyze in Chapter 4 and 5 these issues in greater detail.

The differences in two key measures of the gender gap across these three countries can be seen in Figure 1.

Italy exhibits a very substantial earning gap showing that women earn fully 46% less than man (a number that we will confirm in our analysis of Chapter 4 and 5) while in the US the income gap is only 12%. Similarly in the US Women have a participation rate equal to 88% of the rate of men while for Italy that figure is 70%. These countries, therefore, span a large part of the OECD range for many of the economic indicators of gender inequality. For instance the female to male ratio in participation rate was 96% in Sweden, the OECD country closest to equality and Italy was the OECD country with lowest female to male ratio of participation rate. Also, in terms of earned income the US is the second country in the female to male ratio after Norway (for which the ratio is equal to 1) and Italy is the second to the last (before Japan whose ratio is 0.51).

Considering the recent trends of gender gap in labor force participation, income and economic outcomes, points to a decrease in most countries. However, several rich countries, after some decades of progress in reducing those gaps find themselves with an amount of gender inequality that is stubbornly persistent and very hard to eliminate. On one hand most OECD countries have eliminated completely disparity in education between men and women. According to the Global Gender Gap report of 2011, the top 60 countries have a gender ratio in schooling achievement between women and men larger than or equal to 0.99. In the top 20 countries women outperform men in measures of schooling achievement. Figure 2 shows that for the three considered countries women have abundantly overtaken men in terms of tertiary school enrollment[2].

On the other hand access to top paid jobs, to high wages and to equal opportunities at the top of the income distribution are still much harder for women. Women are under-represented among senior officials, legislators and managers and also among science, technology and management professions. Hence most countries still exhibit a significant wage and income

[2] Other measures such as share of young individuals with a college education show a similar advantage of women.

gender gap and because men in those "high-power" position have very high wages, the gender gap is substantial even among highly educated.

In this report we analyze in some detail some of the possible explanations and determinants for the gender gap in economic performance and we focus specifically on highly educated individuals. While the country-focus of the report is due to our expertise and to the availability to us of interesting and novel data, the focus on highly educated is a choice. It is driven by the fact that we believe this is a crucial segment of the population in developed countries; it is growing in size and relevance for economic productivity and in large part it has been "on the winning side" of the large technological and economic transformation of the last decades. Highly and in particular college-educated individuals have become richer and healthier, their life expectancy is longer and their number has increased relative to the rest of the population. While gender gap in wage and career in less skilled manual workers can be considered as a remain of a mode of production and of institutions of the past, and hence as increasingly less relevant in the modern economies, gender gap among college-educated, among professionals, senior officials and managers reveals a tendency that may persist.

There are several studies that address the issue of gender gap at the top of the earning distribution and they focus on different aspects and determinants of it[3]. In this study we systematically address three possible determinants of gender gap in earnings and labor market performance for college-educated and we measure those using data. We use a combination of descriptive analysis, more structural and model-based simulations and empirical evidence to approach this issue. After a brief introduction, in each chapter, on the gender inequality relative to considered country, we focus on specific components and determinants of the gender gap in employment and wages.

In order to understand the gender gap among college educated and their labor market performance we need to go beyond the well-known fact (shown in Figure 2) that, in terms of

[3] For a review of the gender Gap literature see Bertrand (2010) and Blau and Kahn (2008) and reference thereof.

college enrollment and schooling achievements, women have caught up and are leading men in most of the developed countries. In fact, behind this convergence, there are some very persistent differences in field of study that, as we will see below, may be at the root of a significant part of the gap. Moreover the segment of highly educated women seems to differ significantly from that of men in terms of choices during their early career and the time devoted to the family and to children is still a relevant factor in determining the gap. Women with higher education are possibly using a large part of that human capital to raise and educate their children rather than to produce in the market. This is a very valuable use of their skills that we may not be still fully measuring in our models. Finally there may still be a degree of discrimination, protected by friction and hysteresis, in maintaining male privileges in some senior officer, director and manager positions. Among the small network of highly educated and powerful corporate and political leaders discrimination can still be an important part of the gender gap.

The three aspects we will be considering, therefore, are the role of prejudice and discrimination, the role of family and early career choice and the role of the choice of college major in determining the gender gap in labor market performances. Each of the three components will be studied specifically using one of our countries, and relative individual and aggregate data that are best suited to address that issue. Chapter 2 addresses the role of prejudice and discrimination looking at US data and developing a model with search frictions in the labor market, differences in productivity between men and women and "prejudiced" employers. The model is estimated on US data and it allows to separately identifying the effects of productivity and prejudice on wages and employment. Moreover we consider college educated women separate from non-college educated, and we can characterize the specific degree of prejudice towards them and if education increases or decreases such prejudice. In Chapter 3 we consider the role of family variables (marriage and children) and the early career of women and their effect on the gender gap in wage and employment. In this chapter we use Labor Force Survey data from Spain for the period 1994-2008 and we analyze how having children affects the employment and wage evolution of women and how much of the gender

gap can be explained by children. We also focus on college educated and we analyze the time evolution of such effect. Finally in Chapter 4 and 5 we present and use a novel database on Italian high school graduates, their choice of college major, their academic quality and their outcomes in terms of occupation, employment and income, between one and fifteen years after they have joined the labor force. The focus of those chapters will be the importance of the choice of Major, in determining the gender gap, and we analyze the determinants of choice of college majors considering academic ability, non-academic skills, family considerations and effects of peers and of teachers in high school.

1.2 The role of prejudice, differential productivity and labor market frictions on the gender gap

A first component of the gender gap in the labor market that we will try to quantify is the impact of prejudice and discrimination. We will be careful in distinguishing the two components and we will focus on only one possible source of prejudice: the distaste experienced by prejudiced employers when hiring women.

While explicit prejudice has been a part of economic theory for a long time[4], it is still very difficult to directly observe and measure it. One possible way to estimate its presence and impact is to infer explicit prejudice from the observed behavior of labor market agents, conditioning on a specific model of the labor market. A good candidate for such a model, which we will use in this chapter in an application to the US labor market, is a search-matching-bargaining model.

A search-matching-bargaining model is a good candidate for two main reasons. From a theoretical point of view, the presence of search frictions justifies the survival of prejudiced employers in equilibrium, as suggested by Heckman (1998) and Altonji and Blank (1999). Namely if prejudice is inefficient for the employers, because they hire an inefficiently low number of women, or they hire men instead, then in a perfectly competitive market with free entry it cannot survive in the long-run. Market frictions are needed for prejudice to survive in

[4] A theory of explicit prejudice ("taste discrimination") was first proposed by Becker in 1957 (see Becker (1971)) and has been very influential on the discrimination literature ever since (see Altonji and Blank (1999)).

equilibrium. Moreover, search models with matching and bargaining have been used in several empirical applications and have proved to be a good fit for data in empirical applications (Eckstein and van den Berg (2007)). Most importantly, in light of the applications performed in this chapter, Flabbi (2010a) shows that when Becker's taste discrimination is added to the framework, the model is able to separately identify the impact of explicit prejudice, differential productivity and gender-specific search frictions on labor market outcomes.

The separate identification of these three possible sources of gender differentials for different education groups is the first main contribution of the chapter. The literature attempting a similar exercise is very limited. Our work is most related to Flabbi (2010a) since we will use the same identification strategy. Flabbi (2010a) estimates a search-matching-bargaining model on a sample of white, college graduated extracted from the 1995 Current Population Survey. His results suggest that both prejudiced and unprejudiced employers wage discriminate and that half of the employers are prejudiced. As a consequence, two-third of the gender earning differential may be explained by prejudice while the rest is imputed to the lower female productivity. The first paper attempting a similar exercise was Bowlus and Eckstein (2002). Differently from Flabbi (2010a), they develop an equilibrium search model and use it to estimate race discrimination. Even if their identification strategy allows for it, their estimation strategy has some difficulty in separating the contribution of productivity, prejudice and search frictions on the observed race gap. Their results focus on showing the importance of explicit discrimination relative to productivity differences. In particular, about 56% of employers are estimated to be prejudiced and the disutility from hiring a black worker is estimated at about 31% of the productivity of a white worker. A recent contribution in the same line of research is Bartolucci (2011). The paper develops an equilibrium search model with rent splitting, on-the-job search and two-sided heterogeneity in productivity. Estimates obtained using German matched employer-employee data indicate that female workers are less productive and more mobile than male workers and have lower bargaining power. Estimates indicates that 65 percent of the gender gap is due to productivity differences, 17 percent to segregation, and 9 percent to search frictions.

The second main contribution of the chapter is using the estimated model to build an innovative welfare measure to evaluate returns to schooling and to perform two policy experiments. Since the overall welfare of labor market participants depends on their current labor market state but also on the labor market dynamics related to the transitions between labor market states and job offers, we need a summary measure able to take into account all these components if we want to assess market returns and the impact of policy experiments. We use the welfare measure proposed by Flinn (2002) which assigns to each labor market state occupied by the workers in steady state the corresponding utility value weighted accordingly to the equilibrium steady state distribution. Flinn (2002) contains an extended discussion of the welfare measure, provides a foundation for it and compares various alternative.

1.3 The role of family choice and early career in the gender gap

A second widely studied and very important component of the gender gap in the labor market performances is the different role of men and women in the family, with a specific importance of raising/educating the children, and the consequences of this role during their early working careers. In Chapter 3 we focus on this issue and we use data from Spain, to analyze how family characteristics, in particular the presence of children, have affected the labor supply decisions and the labor market performance of highly educated females in Spain during the last 20 years.

Several important contributions in the literature have analyzed the evolution of the role of college-educated women in the labor markets and their role within the family, especially focusing on the U.S. Goldin (2004a, 2004b), for instance, traces the demographic and labor force experiences of five cohorts of college graduates – from those born in 1900 to the last born in 1980, focusing on why career and family outcomes changed over time. She describes the path to the fast track that college graduate women have taken starting with the first cohort, born in 1900, and forced to choose between “family” and “career”, to the latest group, born in 1980 which has achieved success in combining career and family. Each generation built on the successes and frustrations of the previous ones. Each stepped into a society and a labor market with loosened constraints and shifting barriers. The road was not only long, but it has also been

winding. Some cohorts of college graduate women chose "family" whereas other chose "career". Only recently has a substantial group been able to grasp both at the same time.

In describing the trends that characterized some of the changes in women performance in the US during the 20th century, Goldin (2004b) emphasizes the speed at which the various transformations took place such as the one from jobs to careers and the one from early to later marriages. Moreover, most of these changes took place at the same time. The more likely factors to explain these rapid changes are, in Goldin's analysis: (1) government mandates, such as antidiscrimination policy in hiring and promotions, (2) change spurred by the resurgence of feminism that came on the heels of the Civil Rights movement in the 1960s and (3) the contraceptive innovation, known as the Pill, which gave young women freedom to plan their family and the ability to delay marriage and child-bearing and plan for a career.

These transformations have produced a significant degree of women advancement in the US, and equality seems within reach. Still family choices seem to penalize the labor market performance even of those among the most educated and highly skilled women. Goldin and Katz (2010) analyze the work and family choices among students graduating from one of the most elite institutions of higher U.S. education (Harvard). They explore the trade-offs between family and career, particularly for college graduate women. They find that male earnings are strongly and positively related to the number of children in the family, whereas female earnings are negatively related, especially for those having three or more. But the negative impact of children on women's earnings is entirely accounted for by hours worked. In fact, a *positive* relationship between children and earnings exists for women working full-time full-year. This suggests higher reservation wages for those with more children and positive selection into the labor force based on the number of children.

Similarly Bertrand et al (2010) study the careers of MBAs from a top US business school to understand how career dynamics differ by gender from 1990 to 2006. Three proximate factors account for the large and rising gender gap in earnings: differences in training prior to graduation, differences in career interruptions, and differences in weekly hours. The greater career discontinuity and shorter work hours for females are largely associated with motherhood. They find that the presence of children is the main contributor to the lesser job

experience, greater career discontinuity, and shorter work hours for female MBAs. Those mothers seem to actively choose jobs that are family friendly, and avoid jobs with long hours and greater career advancement possibilities.

In general most of the existing literature, relative to industrialized countries, has found that the presence of children is particularly important to explain the gap in hours worked. Women in families with children work significantly less than women in families without (see for instance Waldfogel, 1998; and Harkness and Waldfogel, 1999). Focusing specifically on Spain, the country considered in the analysis of Chapter 3 of our study, Molina and Montuenga (2009) use the Household Panel for 1994-2001 and find results confirming that there is clear evidence of a wage penalty for Spanish working-women with children. Specifically, the fact that there was a birth in the family during the current year implies a wage loss of 9% of women's wage. They use fixed effect estimation to control for the individual quality of women and effectively show the existence of the motherhood wage penalty in Spain. Specifically, the penalty is close to 9% if there is a birth in the family during the current year; 6% if there is one child living in the household, almost 14% if there are two, and more than 15% if there are three or more.

One interesting fact about the role of family on women career is that, possibly because of changing social norms, or because families have become much smaller, there seems to be a clear tendency to a reduction in this negative impact of children on the working hours of women. Considering Spain, and focusing on the college educated individuals (that will be the sample of our analysis) Figures 3 and 4, show the profile of participation to the labor market along the early years of a person's career. While we notice a negative female-male differential for all cohorts and ages, it is clear that the largest difference is for the age groups between 30 and 40, when the burden of raising children is highest. However, we also notice a drastic decrease in the gender gap moving from the older generations (born in 1960) to younger ones (born in 1975).

The role of family and children, therefore, while certainly changing in the recent decades, has usually penalized the labor participation of women. From the general evidence in industrialized

countries it seem to have had a stronger role on labor supply, rather than on wages. In chapter 3 we will analyze more systematically its role in affecting employment, hours worked and wages focusing on the interesting case of Spain, where the participation of educated women to the labor market seems increased substantially in the recent cohorts.

1.4 Gender gap in the choice of college majors and occupations

A third significant difference between men and women, certainly contributing to the gender gap in earnings and career, is the difference in their field of study while in college. This issue is what we tackle in Chapter 4 and 5. As we have already mentioned the educational gap between men and women in terms of years of schooling and level of tertiary education and enrollment has been fully eliminated, and in many OECD countries there is actually a positive gap in the achievement of women measured as years of schooling or graduation rates from college. However, another gap persists and it has been the subject of studies by economists in several countries: the gender-gap in the choice of college major and in the subsequent occupational choice. Women in many countries seem to shy away from the so called "STEM" majors (Science, Technology, Engineering and Management) and from the related occupations. In this section we draw on Flabbi (2010) in order to describe some of the general tendencies in the gender gap relative to the choice of major across OECD countries. We also provide some indication of the importance that the choice of major has in determining the choice of occupation.

We should emphasize that there is a substantial literature on the so called "Math-Science Gender gap"[5]. Economists, Sociologists and Scholars of high education have noticed, across countries, that women tend to avoid math and science intensive majors in college. Most of the studies are on the US and on the UK. There is also agreement that the "Science Gap" is an important component of the gender gap in earnings and it explains a substantial share of that, sometimes up to half of it (see for instance Paglin and Rufolo, 1990, and Brown and Corcoran, 1997). In chapter 4 and 5 we contribute new data, from Italy and new analysis to this debate. The data are rather unique as they contain very detailed information on the high school and

[5] See Xie and Shauman (2003) for an excellent overview of this topic.

college career of Italian students, who graduated from high school between 1985 and 2005, as well as measures of their labor market performance between one and fifteen years after they joined the labor force. We will explore how significant are the gender differences in the choice of college major and whether psychological, skill-based or preference based explanations are more consistent with the data. Here we review some of the cross-country evidence of the gender gap in the choice of major, to emphasize how Italy (whose data will be introduced in Chapter 4) is a rather average and representative case among OECD countries.

While there are several databases on the educational attainment of women and men across countries, there is much less cross-country comprehensive information on the choice of college majors. To produce some stylized facts we rely on the *Flexible Professional in the Knowledge Society* (REFLEX) data set. The REFLEX data set collects the result of a survey on graduates from tertiary education (equivalent to a College B.A. or a Master in most countries) who have about 5 years of experience since leaving higher education and obtained their degree in the academic year 1999/2000. The project covered 14 countries, all of them developed and most of them in Europe: Austria, Belgium (only Flanders), Finland, France, Germany, Italy, the Netherlands, Norway, Spain, UK, Czech Republic, Portugal, Japan and Estonia[6]. The focus on the REFLEX project was to study the "Flexible Professional in the Knowledge Society"; hence the data set is rich in variables describing the type of education and skills acquired by the individuals and their mapping into occupation choices and skills. The database is representative of the sampled cohort across countries. We will present, here, only statistics relative to the aggregate of all countries.

The first fact emerging from the data is that the proportion of graduates with a tertiary degree in 1999/2000 who are women was larger than 50% in all countries, except for Germany and Japan where the proportion was almost exactly 50%. Hence this database confirms that equality in tertiary graduation rates has been achieved in most developed countries and that education, as measured by years of schooling, cannot be a contributor to the gender gap. Figure 4 and 5 show, instead, prima facie evidence of the differences in choice of college

[6] Luca Flabbi has made the data and the summary statistics relative to this database available to this report.

majors, pooling all countries in the sample. The aggregate categories used to classify the college-majors are constructed in order to keep the classification consistent across countries. Figure 4 shows how men and women were distributed across college majors for people who graduated in year 2000.
Figure 5 shows the percentage of graduates of each of those college majors who are women.

Figure 4 clearly show that Engineering is the Major that attracts the largest share of men (more than 35%) while it attracts a very low share of women (barely above 10%). Conversely health and humanities are among the most popular majors among women while they attract a much lower percentage of men. Correspondingly Figure 5 shows that while 75-80% of the graduates in majors such as Education, Humanities and Health are women, only 25% of graduates in Engineering and 45% of graduates in Mathematics and Sciences are women. These tendencies are confirmed within each country. In particular there seem to be no exception, across the countries considered in the REFLEX study, to the fact that Science and Engineering major attract many more men, and Humanities many more women. Whatever the reason for this gender gap, it seems to be very pervasive and not linked to specific school policies but depending on generalized differences in abilities or in preferences across genders.

One possible reason for the gender gap in the choice of major is differential academic ability. If majors such as engineering and math-sciences are more demanding and men have a better academic performance than women, then a "vertical" differentiation of majors, according to the average quality of the admitted student may generate this pattern. The fact that Science majors are more demanding than the others is (mildly) supported by Figure 6. This figure shows the percentage of students graduating from secondary school with a grade within the top quartile of the national grade distribution, attending college in each of four categories of majors and separated between men and women.

The figure seems to support the idea that science majors are somewhat more selective and attract a higher percentage of top students. However, this selectivity is not very different for

men and women. Hence to explain the much lower percentage of women in Sciences it should be the case that the quality of women, as revealed by the grade in high school, is much lower than that of men. The two panels in Figure 7 below rule out that explanation. They show the percentage of men and women in each quartile of the (aggregate) distribution of high school exit-grades. In order to make the exit grades comparable across countries we consider only their relative distribution and we distinguish only between four quartiles within each country. The differences in distribution across grades are very small with women being somewhat more concentrated in the third quartile relative to men and less in the first and fourth. We notice, overall, that the high school grade distribution is quite similar between men and women and it does not show differences that would justify the large gap in enrollment in math-intensive majors observed in Figure 4.

Hence, in general among the considered countries, women of high academic quality choose in much larger proportions Humanities and Education, relative to Engineering and Science. As we will see in the in-depth analysis of Chapter 4, our micro-data from Italy fully confirm this tendency.

The important role of major choice in the labor market performance of individuals is due to the fact that different majors determine access to different occupations by providing students with different skills. If occupational differences are associated with different earning abilities and different career opportunities then this is a channel through which the choice of major will affect income significantly. All the considered individuals are highly educated and hence will have mainly access to occupations classified as “professionals and technicians” (according to the ISCO code). In fact more than 80% of college educated (men and women) are within this occupational group. Of the remaining college educated, around 5% are classified as “clerks” and another 5% as “senior officers”. Within the category of “professionals and technicians”, however, we see that, men and women have a very different distribution, certainly as a consequence of the choice of major. Figure 8 shows that more than 40% of men in the group work in professions classified as “physics mathematics and engineering” while only 12% of

women do. On the other hand almost 30% of women work in occupations classified as "teaching" while only 10% of men work in that group.

Summing up the stylized facts shown in this section, the link between college major and occupational choice seems very strong in all countries. This could be a crucial element to understand the gender gap. In particular three regularities seem pervasive. First highly educated men and women, in spite of their equal educational attainment (in terms of years of schooling) and comparable academic performance (in terms of grades) have a large and not decreasing gap in their choice of field of study (college major) and in the choice of occupation. Second, women exhibit systematically a negative gap in the probability of choosing math-intensive majors as Engineering and Mathematical sciences, while they show a systematic positive gap in the choice of Majors in the Humanities. Third, the college major seems to be strongly correlated with the choice of occupation. In particular "teaching" types of jobs are dominated by women, while "engineering" types of jobs are dominated by men. This difference contributes significantly to explain gender gap in career, wage and earnings. In chapter 4 and 5 of this report we will focus on micro-evidence on individual data from Italy to understand better the mechanisms and decompose this gender gap in the choice of college major.

1.5 The structure of this report

The remaining chapters of this report follow the structure anticipated in this introduction. We tackle, in turn, the analysis of different determinants of the gender gap beginning, in Chapter 2, with the role of prejudice and market frictions on the wage and employment gap, measured using US data. Then, in Chapter 3, we analyze the role of family and children in causing a gender gap in employment and wages, and we use data from Spain to estimate those effects. In Chapters 4 and 5 we focus on the role of choice of College Major in determining the gender gap and we analyze what are the causes of such different choices. We do this using a newly collected database on Italian high school graduates over the period 1985-2005 for which we have detailed information on high school performance, college career and some labor market

outcomes after up to 15 years from graduation. In order to describe the data and their characteristics in detail, and to perform also a more thorough statistical analysis we devote to this case Chapter 4 and 5. We devote Chapter 6 to summarize what we learned, to describe how likely it is that our results can be generalized and to distill some policy implications, with a special focus on the countries that we have considered.

Each chapter of the Report tries also to be self-contained, so that it can be read as an article. For this reason we have left to each chapter an initial description of the important features of the pre-labor market and of the labor-market gender gap, as measured in the country considered in that chapter. This way, the reader interested only in that particular case may find it easier to put the results into context. While this report focuses on specific countries we believe that the channels that we identify are very important, general and, in fact, crucial, in the understanding of gender gap in all developed countries. In particular countries such as Italy and Spain, where the disparities are still so large, require a full understanding of those channels and, possibly, new incentives and policies to reduce the gap. As stated at the beginning of the report, allowing women the same access and the same economic opportunity as men can be a very effective measure to reinvigorate the economies, bringing new human capital, new ideas and new skills. The effect of those policies can be the largest where those disparities are the biggest. This report aims at providing some new ideas and new evidence in support of those potential policies.

References for Chapter 1

Altonji, Joseph G., Blank, Rebecca M. (1999), “Race and gender in the labor market”, In: Ashenfelter, Orley C., Card, David (Eds.), *Handbook of Labor Economics*, vol. 3. Elsevier.

Bertrand, Marianne (2010), “New Perspectives on Gender”, In: *Handbook of Labor Economics*, Volume 4b, Elsevier.

Bertrand, Marianne, Goldin, Claudia, Katz, Lawrence F. (2010), “Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors”, *American Economic Journal: Applied Economics*, 228–255.

Blau, Francine D., Kahn, Lawrence M. (2008), « Women’s work and wages”, In: Durlauf, Steven N., Blume, Lawrence E. (Eds.), *The New Palgrave Dictionary of Economics*, second ed., Palgrave Macmillan.

Brown, Charles, Corcoran, Mary (1997), "Sex Based Differences in School Content and the Male-Female Wage Gap", *Quarterly Journal of Economics*, 99: 31-44.

Goldin, Claudia (2004a), “From the Valley to the Summit: The Quiet Revolution that Transformed Women’s Work”, *NBER Working Paper*, No. 10335.

Goldin, Claudia (2004b), “The Long Road to the Fast Track: Career and Family”, *Annals of the American Academy of Political and Social Science*, Vol. 596

Goldin, Claudia, Katz, Lawrence F. (2010), “Transitions: Career and Family Life Cycles of the Educational Elite”, *American Economic Review*: Papers & Proceedings 2008, 98:2, 363–369.

Harkness, Susan, Waldfogel, Jane (1999), “The Family Gap in Pay: Evidence from Seven Industrialized Countries”, *APPAM*, Washington, D.C.

Molina, Jose Alberto, Montuenga, Victor (2009), “The Motherhood Wage Penalty in Spain”, *J Family Economic Issues*, 30:237–251.

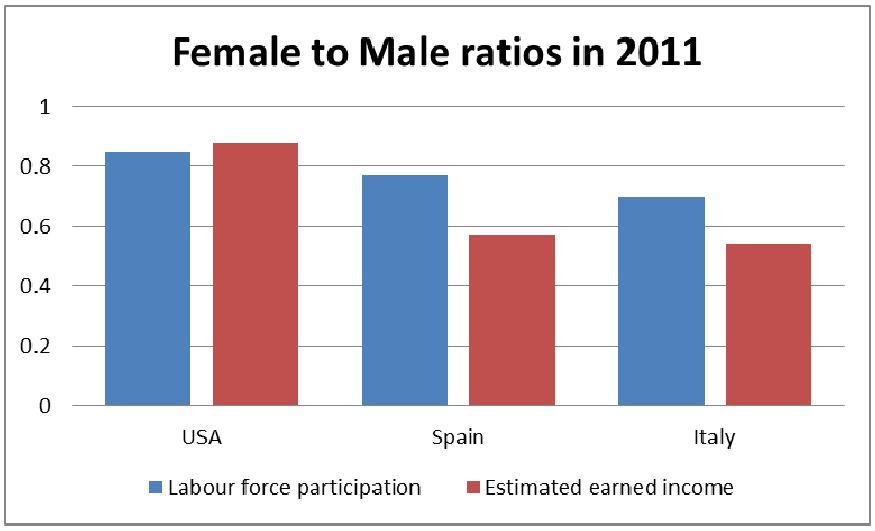
Paglin, Morton, Rufolo, Anthony M. (1990), "Heterogeneous Human Capital, Occupational Choice, and Male-Female Earnings Differences", *Journal of Labor Economics*, 8: 123-144.

Xie, Yu, Shauman, Kimberlee A. (2003), "Women in Science: Career Processes and Outcomes", Cambridge, MA: Harvard University Press.

Waldfogel, Jane (1998), “Understanding the ‘Family Gap’ in Pay for Women with Children”, *Journal of Economic Perspectives*, v. 12: 137-156.

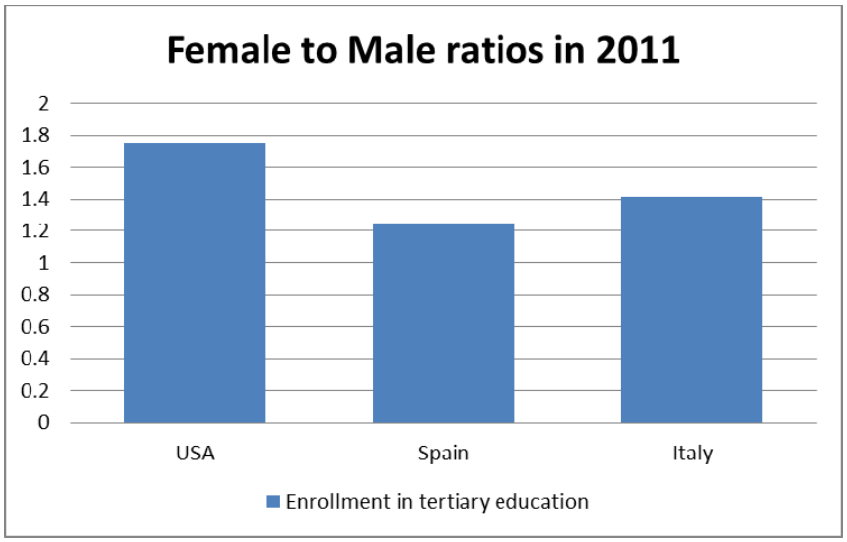
Harkness, Susan and Jane Waldfogel. (1999) “The Family Gap in Pay: Evidence from Seven Industrialized Countries” *APPAM*, Washington, D.C.

Figure 1



Note: Source is the Global Gender Gap Report 2011

Figure 2



Note: Source is the Global Gender Gap Report 2011

Figure 3:

Life Cycle of Labor force participation rates, Males and Females.

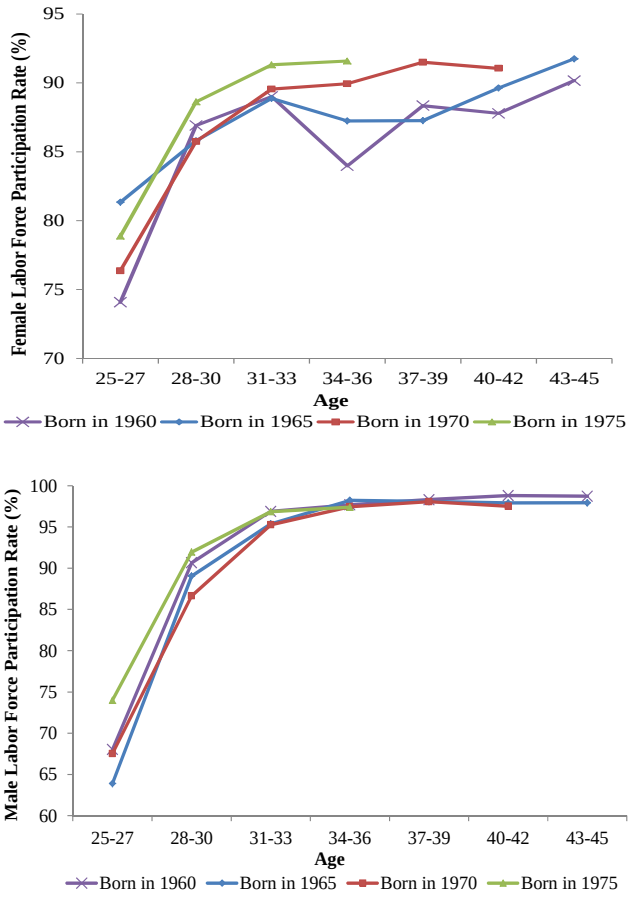


Figure 4

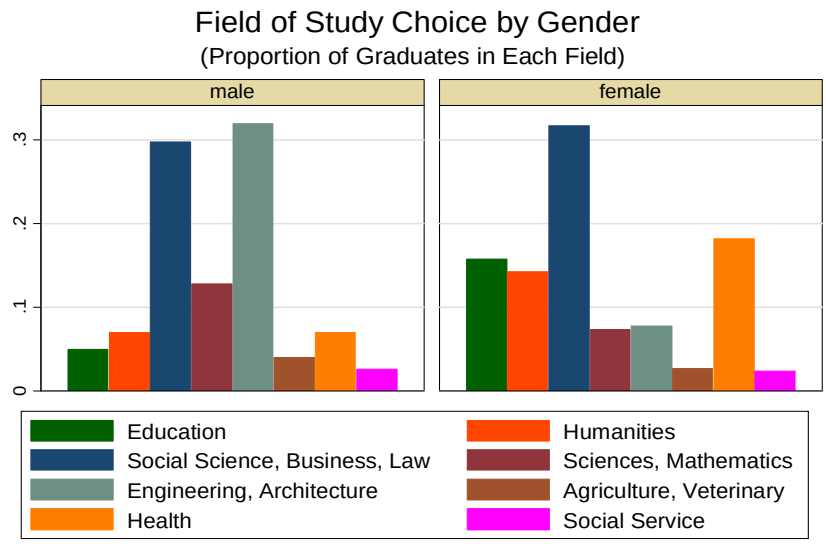


Figure 5

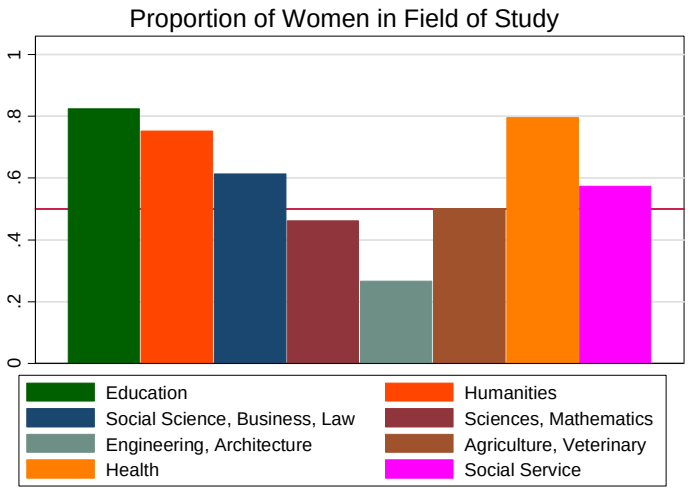
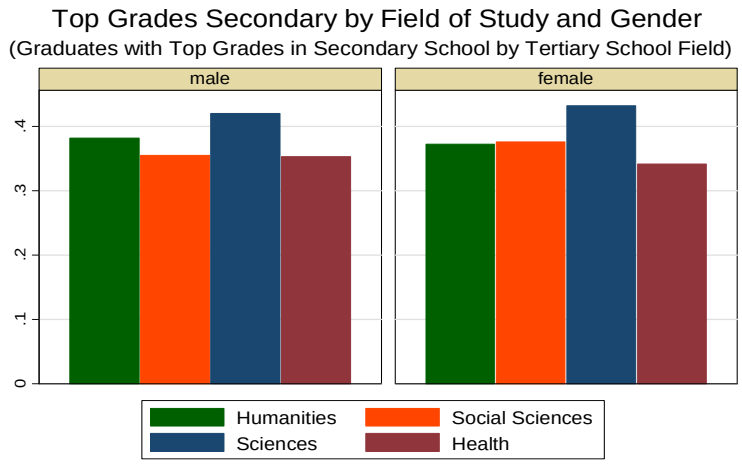
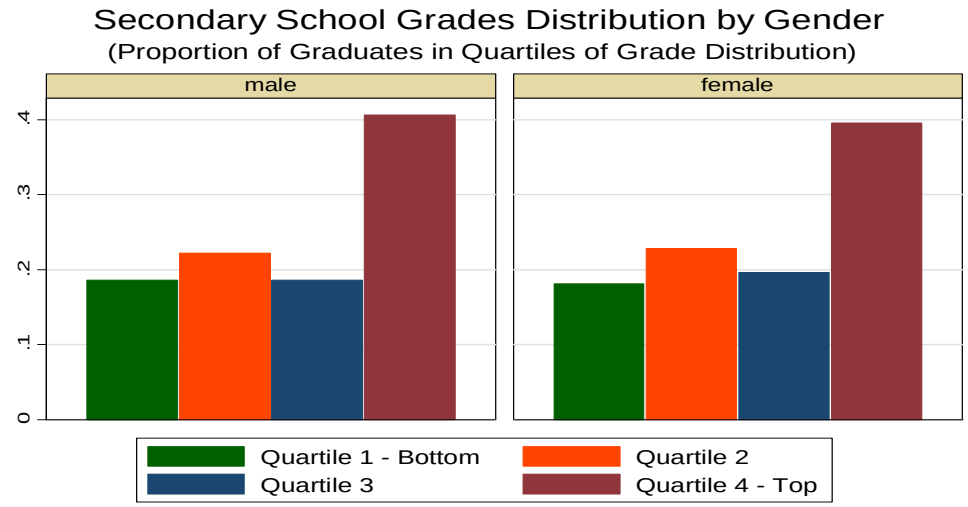


Figure 6



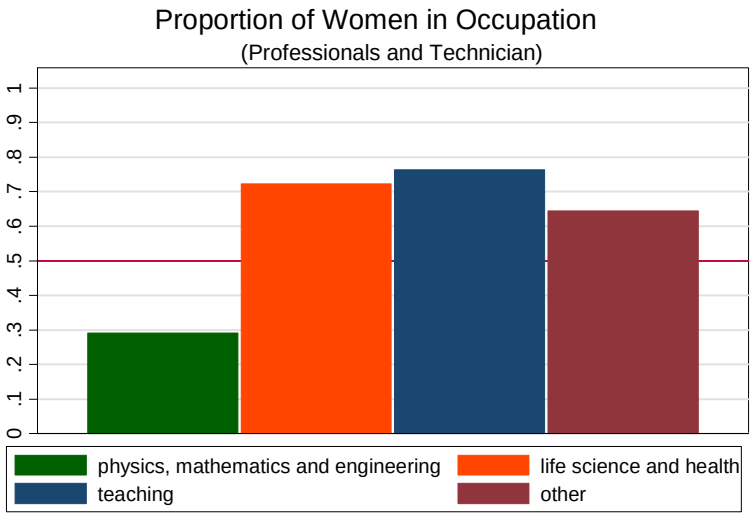
Notes: Top Grades means top grade available or in the top 25% of country-specific distribution.

Figure 7



Notes: Grade distribution computed by country on the joint sample of males and females.

Figure 8



Notes: First job after graduation.

Chapter 2: Gender Gap in Labor Market Outcomes in the US: The Impact of Employers' Prejudice

Luca Flabbi (IDB, Georgetown University, and IZA)
Mauricio Tejada (Georgetown University)

2.1 Introduction

The chapter proposes three contributions. First, it provides descriptive evidence on gender differentials by education in the US labor market over the last twenty years. Second, it uses the structural estimation of a search model of the labor market to identify and quantify the impact of employers' prejudice on labor market gender differentials. Third, it connects both the descriptive and analytical findings to recent policy interventions in the US labor market and performs some policy experiments.

For all the analysis in the chapter, we use the *Annual Social and Economic Supplement* (ASES or March supplement) of the Current Population Survey (CPS)[7].

2.1.1 Descriptive evidence

We organize the descriptive evidence following the decision process of an individual deciding to supply labor in the market. First, we look at the outcomes of education decisions. Education decisions constitute the most important component of pre-labor market human capital and they influence not only future performance in the labor market but also the decision to participate in the labor market itself. It is also a choice and process where gender asymmetries are present but evolving quickly. We provide evidence on the overall quantity of education acquired and on one aspect of the "quality" of education acquired: the field of study. All the evidence on labor market outcomes will be correlated to these previous education decisions. The main result on pre-labor market characteristics is that women acquire more college education than men, reinforcing a trend started with the generation born in 1959. A lot of asymmetry by gender persists in the choice of field of study.

[7] A detailed description of the data and the estimation sample is contained in the Appendix.

Second, we look at the decision to supply labor in the market both with respect to the *extensive* margin (the participation decision) and to the *intensive* margin (the hours supply decision). We correlate this evidence with education and age and we look at the evolution over-time based both on survey year and on cohort of birth. Together with the labor supply evidence, we also present some evidence on unemployment, focusing on unemployment dynamics. The main result from this section is that women supply less labor than men both on the extensive and on the intensive margins. The gender gap is not alleviated by education: the gender gap in the hours worked per week is actually exacerbated by education since it is larger on the College graduates sample than on the overall sample.

Third, we look at gender earnings differentials in the labor market. We compute both the raw differential and the differential conditional on standard human capital characteristics. We also provide evidence on the gender gap at different percentile of the earnings distribution and on very high-skilled occupations (CEOs and General Managers) to assess the magnitude of the so-called *glass-ceiling* effect. Results show that women earn about 20% less on average than men. The gap has been fairly stable in the last 10/15 years, after a period of significant reduction in the 1970s and 1980s. One reason for the persistent gender gap in recent years, in particular among skilled workers, is the large differential at the top of the earnings distribution, a possible indication of glass-ceiling effects.

2.1.2 The Impact of Employers' Prejudice

In the second part of the chapter, we want to investigate the source of the observed gender differentials. We give priority to gender differentials in the labor market – both in terms of wage differentials and of labor market dynamic - but we will attempt to draw some inference also on gender differentials in pre-labor market characteristics. We will be able to make distributional considerations and evaluate some determinants of the glass-ceiling but we will abstain from an analysis of labor supply determinants.

We will focus our analysis on three main determinants of gender differentials: productivity differences; employers' prejudice; and search frictions. These three determinants constitute a

quite exhaustive list of the possible explanations proposed in the literature to account for the observed gender gap[8]. The difficulty, from a quantitative point of view, is to separately identify the contribution of these three components on the observed differential. The methodology we have chosen to use exploits the structure of a specific model – a search-matching-bargaining model with employers taste discrimination – to separately identify and then estimate these three components.

We estimate the model over two time periods (1995 and 2005) and three education levels (Master and PhD; College; and High School). Results suggest that the gender gap in wage offers is smaller than the gender gap in accepted wages on the high education samples but larger on the low education samples. The gender gap in productivity is estimated to be relatively small on the College and High School sample but it is increasing over time. The productivity gap for Master and PhD holders is quite large and increasing over time. The gender gap between workers employed at prejudiced employers and workers employed at unprejudiced employers is larger the higher the education level. The gender gaps in unemployment rates are relatively small in all years and education groups.

We use the separate identification and estimation of the three main determinants of gender differentials to decompose their impact on the gender wage gap. We find that prejudice has a significant impact in explaining the wage gap but the impact is decreasing over time and it becomes smaller than the impact of productivity on all education groups in 2005. Master and PhD graduates are the exception to the trend: they experience a stronger impact of prejudice in 2005 than in 1995. Thanks to a decomposition of the gender wage gap at different points of the distribution, we also estimate that the Master and PhD sample shows some evidence of glass-ceiling effects.

2.1.3 Policy Experiments and Policy Implications

[8] Two main determinants are left out: gender asymmetries in household production and differential preferences with respect to job amenities. The first issue is the subject of Chapter 3 of this book, with an application to Spain. For a review, see Waldfogel (1998). The second issue is partially addressed by Chapters 4 and 5 of this book, with an application to Italy and a focus on the relationship between education choices and occupation choices. For an example within the search and matching literature, see Flabbi and Moro (2012).

The descriptive section of the paper shows that gender differentials are by no means limited to average wage differentials, the variable often used in the literature to summarize the "gender gap" in the labor market. Gender differences concern the shape of the entire wage distribution, the labor market dynamics across labor market states and the schooling choice preceding the entrance in the labor market. To really judge the overall welfare of labor market participants and to compare welfare across schooling groups and time periods, it is necessary to build an indicator able to take into account all these different elements. Thanks to the estimates of the structural model, we are able to propose and compute such indicator.

First, we use the indicator to compute a "welfare return" to schooling, i.e. the welfare differentials enjoyed by labor market participants at different level of schooling. We find that in 1995 female returns were higher than male returns, providing a possible explanation for the higher level of education acquired by women. However, the returns are estimated to be lower in 2005, a result mainly due to the presence of prejudiced employers in the labor market.

Second, we perform policy experiments that mimic the major policy interventions implemented in the US labor market: Equal Pay policies and Affirmative Action policies. In the equal pay policy, we impose that wage schedules cannot be set conditioning on gender. In the affirmative action policy, we implement an employer subsidy to hire women. The Equal Pay policy is effective in redistributing welfare from men to women but it is never enough to completely close the gender gap. It is more effective for lower education levels than for higher ones and it has larger impacts in 1995 than in 2005. The affirmative action policy has a modest but positive impact on closing the gender gap in welfare. Despite the modest impact, the policy is promising because it is frequently able to close the gap without reducing overall welfare and it targets better the group that is not showing a positive evolution in closing the gap over time: the Master and PhD graduates sample.

2.2 Descriptive Evidence

2.2.1 Gender Differentials in Pre-Labor Market Characteristics

Figure 1 shows the gender gap (women - men) in the percentage of college graduates by gender. To look at the evolution over-time, we report the cross-sectional evidence based on survey year in Figure 1a and the cohort evidence obtained by pooling the survey years data together in Figure 1b.

The most important result is that the gender gap has been shrinking over the last twenty years and it has actually become positive starting with the generation born in 1959. A positive gap means that women acquire more college education than men. This is a well-established empirical fact, which is becoming increasingly common among OECD countries (OECD (2008); Flabbi (2011)). It is also evidence that has prompted an increasing literature trying to explain why the gap in education is positive while the gender gap in earnings remains negative[9].

In Figures 2 and 3 we focus on the top education levels: Masters and PhDs. We report only the cohort-based evidence on the pooled sample but the cross-sectional evidence generates a similar picture. At the master level there is no evidence of a systematic gap between men and women: there is some men's advantage at the beginning of the period but in the last 10 years the two levels have been very similar. The overall proportion of Master degree graduates is also fairly stable over the last 10 years, while the decrease by cohort is probably due to demographics: in the US educational system it is pretty common to go back to school to acquire a Master degree after a few years of work experience therefore generations born after the mid-seventies may have not yet completed their graduate school decisions. The PhD evidence, instead, shows an interesting pattern: the gender gap is decreasing on recent generations but it is not attaining the stable positive gap shown on college graduates.

In Figure 4 we look at one dimension of the quality of education: field of study. One possible explanation for the puzzle of a positive education gap together with a negative gender wage gap is that women may choose fields of study correlated with lower wages. While evidence on

[9] The most complete explanations proposed so far focus on the return to education on the marriage market (Chiappori, Iyigun, and Weiss (2009) and GE (2011)). For a different explanation based on job amenities, see Flabbi and Moro (2012). For international comparisons, see Becker, Hubbard and Murphy (2010).

this correlation may have different sources, Figure 4 shows that the asymmetries on fields of study choices are substantial. If the favorite field for both men and women is "Business and Law", almost 20% of men choose "Engineering" while less than 5% of women do. At the same time, almost 15% of women choose "Humanities" compared with less than 8% of men. As a result, many fields see an imbalance in terms of gender distribution: looking at Figure 4, we see that the proportion of women in the "Education" and "Medical" fields is much higher than the proportion of women in the population (the red horizontal line). The opposite is true for "Science" and "Engineering". The red horizontal line reports the overall proportion of women in the population giving an immediate pictures of which field of study see an over- or an under-representation of women.

We borrow the terminology of Flabbi and Moro (2012) to propose just one possible correlation between field of study choice, occupational choice and job characteristics[10]. Figure 5 reports the correlation between the occupational choice of women and the degree of "flexibility" in the job by field of study choice. (In this Table flexibility simply means the possibility of working less than 30 hours a week.) What we find is that the field of study least favored by women ("Engineering") is characterized by the lowest degree of flexibility while the two relative most favored fields ("Education" and "Medical" professions) are characterized by the highest degree of flexibility.

To summarize, the evidence on pre-labor market characteristics shows that women acquire more college education than men, reinforcing a trend started with the generation born in 1959. Women acquire less graduate education than men but the differential is shrinking. Where a lot of asymmetry by gender persists is in the choice of field of study, choice that may be correlated to desirable future job characteristics.

2.2.2 Gender Differentials in Labor Supply

[10] For some recent country-specific evidence on this correlation in European Countries, see Beffy, Fougère and Maurel (2009) and Chevalier (2011).

We look at the labor supply evidence by presenting separate results for the overall sample and for the sample of College graduates. In this way, we attempt to isolate specific trends on highly educated individuals.

Figures 6 report evidence on the first labor supply decision: supply labor in the market or not. This extensive margin reports a negative gender gap: men systematically participate in the labor market more than women both on the overall sample and on the college sample (Figure 6a) but the differential has been shrinking over time. Education makes a difference: the differential is smaller on the college sample and it has uniformly been smaller over the entire twenty years period under consideration.

Age is also a relevant determinant of participation rates since it is strongly correlated with fertility. We have divided the sample in three age groups roughly describing a period before children (younger than 30); with young children (30-45 years old) and with older children (older than 45). What we find is that participation decisions are very sensitive to age but the education differential remains significant. Evidence by cohort (Figure 6a) confirms this view because it seems to reflect more life-cycle patterns than major breaks across different generations.

Employment rates are reported in Figure 7 following the same structure used in Figure 6. Figures 7a reports no gender gap in employment rates among college graduates while a small positive gap exists on the overall sample. Notice the big drop in employment rates during the "Great Recession" of the last three years. Employment rates are systematically lower for younger workers and for individuals without a college degree. Interestingly, the lack of a gender gap in employment for the college graduates is a quite old phenomenon, involving cohorts going as far back as 1940 (Figure 7b).

To look at some measure of labor force dynamics, we concentrate on the probability of leaving the unemployment state. Figure 8 reports Kaplan-Meier estimates of the hazard rate out of unemployment. A first striking result is how much more difficult it is to find a job during the recent recession: 2010 is clearly an outlier when compared with the other years. A second

interesting result is that the impact of the recession is pretty homogenous on all four education/gender groups, in particular a college degree does not seem to cushion the severity of the crisis. Differences among the other years are smaller and the fact that college graduates systematically take more time to find a job suggests that a lower hazard rates also reflect "pickier" workers, i.e. workers waiting in unemployment for better job opportunities to come along. Differences between men and women are larger on the college sample than on the overall sample but they do not exhibit a strong trend over time.

Figure 9 focuses on the intensive margin of the labor supply by reporting weekly hours worked. To give an idea of the over-time change we report the first year available (survey year 1985, Figure 9a) and the last one (survey year 2010, Figure 9b). As expected, about 50% of the sample declares to work 40 hours per weeks. What is interesting is that most of the remaining population of male workers is concentrated to the right of the 40 hours spike (i.e. it is working more than 40 hours per week) while most of the female workers is concentrated on the left. This gender difference is not alleviated by education, it is actually exacerbated by education because male college graduates are more likely to work more than 40 hours per week than no college graduates. It is also a gender differential fairly stable over time.

We aggregate some of the distributional variation in hours worked to generate Figure 10. Figure 10 reports the incidence of part-time for men and women and shows a very large gender gap: women are more than twice more likely to work part-time than men. The gap is not significantly reduced when focusing on college graduates because the incidence of part-time is smaller both on men and women with college.

To summarize, the evidence on labor supply shows that women participate less than men in the labor market but when they do they obtain similar employment rates. The intensive margin of the labor supply shows a large gender gap, mainly due to the much larger incidence of part-time among female workers than among male workers. This gender gap is not alleviated but

actually exacerbated by education since it is larger on the college sample than on the overall sample.

2.2.3 Gender Differentials in the Labor Market

Figure 11 reports estimates of the gender earnings gap from 1981 to 2011. We estimate the gap as the coefficient of a dummy equal to 1 if the individual is a woman in an OLS regression of log hourly earnings. The top panel (Figure 11a) reports results from a specification including simply a constant and the dummy woman: it is therefore an estimate of the raw differential at the mean, unconditional on any observables. The bottom panel (Figure 11b) reports results from a specification including a constant, the dummy woman, three educational dummies, age linear and squared, two race dummies, a dummy for marital status and a dummy for presence of children younger than 18: it is therefore an estimate of the differential conditional on standard human capital and demographic characteristics. Dotted lines describe the 95% confidence interval.

As it is well documented, the gender gap in wage and earnings is persistent in the US labor market[11]. Figure 11a shows a significant convergence for all the 1980s, following a trend started in the previous decade (not shown). In the following decade, instead, the convergence slows down and then stops completely in the mid-1990s. The trend in the last ten years is less clear, with periods of minor convergence followed by period of small divergence. The most recent year available (2011) reports the smallest gender earnings gap ever, breaking the 20% mark on the unconditional differential for the first time ever. Figure 11b shows a very similar evolution over time but usually at a slighter lower level, implying that human capital differences do not explain the gender gap in earnings (at least under the very simple mincerian specification we use).

In Figure 11, we also report periods of recession to show how the gender gap is reacting to the business cycle, following a literature correlating the cycle, the change in inequality and the gender gap (Fortin and Lemieux (2000); Biddle and Hamermesh (2012)). We find that recent

[11] For over-time evidence in the US, see Eckstein and Nagypal (2004); Blau and Kahn (2006), and Flabbi (2010).

recessions had a very different impact than earlier recessions: in the 1981 and 1990 recessions the gap decreased while in both the last two recessions the gap increased.

There is a large literature pointing out that the crucial sources of the gender gap in high-income economies may be concentrated at the top of the earnings distribution[12] and the top of the hierarchical ladder at the firm[13] (the so called *glass-ceiling* hypothesis). Figure 11c looks at the gender gap at different points of the unconditional earnings distribution. Results do not support the presence of a glass ceiling in the 1990s but they are consistent with the presence of a glass ceiling in the last 10/15 years, with a particularly large spike in 2000.

Figure 12 reports some evidence on the gender asymmetries in the top hierarchical ladder at the firm. If we look at the CPS data we have used so far, the differentials are large: among CEOs and General Managers, only 20% are women, compared with a presence of women in the labor force which is very close to 50% (Figure 12). Overall, the proportion of men in Management Occupations is about 13% compared with about 8% for women. These differentials are essentially constant over-time, even if the time span we are forced to consider is much smaller than in previous figures because a change in the CPS definition of occupations does not make the values comparable over the entire three decades we are considering in this study. The data collected by Gayle, Golan and Miller (2009) look at the top executive in the CHECK firm. Thanks to their data, they are able to look deeper in the still relative large category of "CEOs and General Managers" reported by the CPS. Based on job titles and transitions across occupations, they build seven rankings within the top executives (7^{th} is the lowest and 1^{st} the highest). What we see it is very low presence of women overall (no more than 6/7%), with the proportion decreasing as the ranking is increasing.

To summarize, the evidence on gender earnings differential shows that women earn about 20% less on average than men, even when controlling for standard human capital and demographic characteristics. The gap shows a significant reduction from the 35% levels of the early 1980s but

[12] See Albrecht, Björklund, and Vroman (2003); Blau and Kahn (2006) and Bertrand, Goldin and Katz (2009).

[13] See Bertrand and Hallock (2001); Albanesi and Olivetti (2008); Bertrand, Goldin and Katz (2010); and Gayle, Golan and Miller (2009).

has remained fairly stable in the last 10/15 years. One reason for the persistent gender gap in recent years, in particular among skilled workers, is the large differential at the top of the earnings distribution.

2.3 The Impact of Employers' Prejudice

2.3.1 The Search-Matching-Bargaining Model

If explicit prejudice has been a part of economic theory for a long time[14], it is still very difficult to directly observe and measure. One possible way to gauge its presence and impact is to infer explicit prejudice from differential behaviors of labor market agents, conditioning on a parsimonious model of the labor market. A good candidate for such a model is a search-matching-bargaining model.

The model is a good candidate both for theoretical and empirical reasons. From a theoretical point of view, the presence of search frictions justifies the survival of prejudiced employers in equilibrium, as suggested by Heckman (1998) and Altonji and Blank (1999). From an empirical point of view, search models with matching and bargaining have been used in many empirical applications and have proved to have a good data fit (Eckstein and van den Berg (2007)). Most importantly, Flabbi (2010a) shows that when Becker's taste discrimination is added to the framework, the model is able to separately identify the impact of explicit prejudice, differential productivity and gender-specific search frictions on labor market outcomes.

2.3.1.1 Environment

The model's environment is as follows[15]. The economy is populated by four types of agents infinitely lived: two types of workers (men and women) and two types of employers (prejudiced and unprejudiced). Unemployed workers are looking for jobs and employers with unfilled vacancies are looking for workers to fill them. Search frictions are present in the market so that meetings may take time before they actually happen. Once an employer and a worker meet,

[14] A theory of explicit prejudice ("taste discrimination") was first proposed by Becker in 1957 (see Becker (1971)) and has been very influential on the discrimination literature ever since (see Altonji and Blank (1999)).
[15] A formal description of the model is relegated in Appendix 2.

they observe the productivity that their specific match will produce. Upon observing the match-specific productivity value and their types, employers and workers engage in Nash-bargaining over the wage. Workers' utility functions are linear in wages and there is no disutility from working. Employers' utility is linear in profit and in the intensity of discrimination. The intensity of discrimination is defined as the disutility from hiring women that affect prejudiced employers (Becker (1971)).

While unemployed, workers receive an instantaneous utility (or disutility) flow that takes into account search costs, unemployment benefits and other utility benefits and costs correlated with the state of unemployment. While a vacancy is unfilled, employers sustain no cost and receive no benefit. The model is dynamic and an exogenous and common discount rate is assumed to hold in the economy.

The technology used to produce the homogenous good produced in the economy is constant returns to scale with labor as it's the only factor of production. Therefore, the total output at a given employer is the sum of the productivity levels of all his/her matched employees.

2.3.1.2 Equilibrium

Given this environment, workers have a very simple decision to make: accept or reject the match with a given employer. They make this decision by balancing the flow benefit of receiving a wage higher than the current utility of unemployment with the expected benefit of receiving a potentially better offer in the future. Since the present discounted value of being unemployed does not depend on a given wage but only on the entire expected wage offers distribution while the present discounted value of being employed at a given employer is increasing in the wage received[16], there will exist a wage at which the worker is indifferent between accepting the job or remaining unemployed. We call this value the *reservation wage*. A similar argument holds for the employer.

However, wages are not simply posted by the employer but are the result of a bargaining process between the worker and the employer upon observing types and match-specific productivity. We assume a Nash-Bargaining solution to this problem. Such a solution has the

[16] See Appendix 2 for analytical expressions of these present discounted values.

property that the worker and the employer will always agree to a match when the match is producing a surplus and they agree to share the surplus according to their respective bargaining weight and their outside options[17]. The analytical expressions of the resulting wage schedules make this concept clear.

First, look at the wage of a man working for a prejudiced or unprejudiced employer (the employer's type has no impact on male workers) with a match-specific productivity equal to *x*:

$$w_M(x) = \rho U_M + \alpha(x - \rho U_M) \qquad (1)$$

The expression states that the wage guarantees the worker his outside option (ρU_M) plus a portion of the surplus ($x - \rho U_M$) equal to the his bargaining weight (α). The bargaining weight captures factors related to the bargaining strength of workers with respect to employers: the higher the weight, the higher the wage at given productivity. The outside option of the worker is the present discounted value of being unemployed: we denote this value with ρU_M and we formally characterize this expression in Appendix B. The higher the outside option's utility, the higher the wage at given productivity because the worker will have a better state to go back to if the match is not realized. Finally, $(x - \rho U_M)$ is the surplus generated by the match because it is the difference between what is produced in the match (x) and what is lost if the match is realized (ρU_M). Notice that the employer does not loose anything if the match is realized because the cost of keeping a vacancy open is zero.

Second, look at the wage of a woman working for an unprejudiced employer with a match-specific productivity equal to *x*:

$$w_{WN}(x) = \rho U_W + \alpha(x - \rho U_W) \qquad (2)$$

The expression is exactly equal to equation (1) with the difference that the outside option is allowed to be different. We use subscript M to denote men and subscript W to denote women. Notice also that the wage equation has two subscript: W to denote women and N to denote

[17] See Appendix 2 for analytical expressions of the surplus and the Nash-Bargaining product.

unprejudiced employers. This is necessary because the female wage schedules are employer's type-specific in equilibrium.

Third, look at the wage of a woman working for a prejudiced employer with a match-specific productivity equal to *x*:

$$w_{WP}(x) = \rho U_W + \alpha(x - d - \rho U_W) \qquad (3)$$

The expression is different from (2) because the surplus is reduced by the disutility that the prejudiced employers receive when hiring women (*d*). The expression makes clear that, as a result of the bargaining process, the cost of prejudice is shared by both the employer and the worker: the higher the discrimination intensity *d*, the lower the wage at same productivity.

We are now ready to state the *equilibrium decision rules* resulting from the model:

1. The optimal decision rules are *reservation values rules* and both workers and employers agree on what these reservation values are. The reservation value rule in this case means that the match will be realized (i.e. both workers and employers agree to enter a job relationship governed by wage equations (1)-(3)) if the match-specific productivity is higher than the *reservation productivity value*. The wages corresponding to these reservation productivity values are the *reservation wages*.
2. The reservation productivities are different between men and women and they are different between women accepting to work for a prejudiced employer and women accepting to work for an unprejudiced employer. We denote them with *x** and they are defined as follows:

$$x_M^* = \rho U_M \qquad (4)$$

$$x_{WN}^* = \rho U_W \qquad (5)$$

$$x_{WP}^* = \rho U_W + d \qquad (6)$$

3. The reservation wages are worker's type-specific but not employer's type-specific. We denote them with *w** and they are defined as follows:

$$w_M^* = \rho U_M \qquad (7)$$

$$w_W^* = \rho U_W \qquad (8)$$

The structure of the *equilibrium* has some interesting *implications* about the impact of prejudice on labor market outcomes:

1. Everything else equal, the presence of prejudiced employers makes the present discounted value of participating in the market (*U*) lower for women than men (See Proposition 1 in Flabbi (2010a)).
2. Wage discrimination is present at prejudiced employers: women working at prejudiced employers receive lower wages than men working at prejudiced with the *same* level of productivity. This is easy to see by comparing wage equation (1) with wage equation (3): for given *x,* women earn lower wages because of the negative impact of d (the direct effect of prejudice) and because the women's outside option is lower than the men's outside option (the equilibrium or "spillover"[18] effect of prejudice).
3. Wage discrimination is also present at unprejudiced employers: women working at unprejudiced employers are also receiving lower wages for same productivity. The effect results from comparing equation (1) and (2): if women outside option are lower (as stated in the first equilibrium), then unprejudiced employers wage discriminate due to the spillover effects even if they do not have any prejudice against women. This is an interesting result that allows to make a clear distinction between explicit prejudice and wage discrimination.
4. Partial segregation arises in equilibrium, that is women are overrepresented at unprejudiced employers and underrepresented at prejudiced employers. This is an important result to explain, or at least be consistent with, the segregation observed in labor market data. We emphasize that we obtain "partial" segregation as opposed to complete

[18] For a formal definition of this spillover effect, see Definition 4 in Flabbi 2010a.

segregation. Complete segregation is a starker result which is at odds with the recent empirical evidence since it implies that prejudiced employers never hire women. This is the setting of the previous, and still most influential, model merging search frictions with taste discrimination: Black (1995).

2.3.2 Estimation and Identification

Search and matching models have been extensively studied and implemented. The identification theory is laid out by Flinn and Heckman (1982): they show that under an appropriate parametric assumption the crucial structural parameters of the model are identified from data on unemployment durations and accepted wages.

To the Flinn and Heckman's result, we have to add the identification of the prejudiced parameters. Flabbi (2010a) shows that, under the same parametric assumptions imposed by Flinn and Heckman (1982), the proportion of prejudiced employers and the disutility they receive from hiring women are identified. This is a useful result because it allows for the separate identification of the prejudiced parameters, the gender-specific productivity parameters, and the gender-specific search frictions parameters. One parameter that is difficult to identify is the Nash-Bargaining weight (Flinn (2006)). We do not attempt to identify it and we simply impose a standard assumption in the literature: symmetric Nash-Bargaining, i.e. workers and employers have the same Nash-Bargaining weight which is therefore set to be equal to 1/2.

Estimation is performed by maximum likelihood: after a first stage in which an order statistic (the minimum observed wage) is used to obtain a strongly consistent estimator of the reservations wages, the maximization of the resulting concentrated likelihood delivers estimates of all the remaining structural parameters. The analytical expression for the maximum likelihood estimator is provided in Appendix B.

2.3.3 Results

2.3.3.1 Estimation Results

Table 1 and 2 shows some relevant predicted values obtained from our estimation results. The Maximum Likelihood Estimates of the structural parameters are reported in Table A2 of

Appendix B. Table 1 focuses on the cross-sectional distribution of productivity and wages. The Accepted Wages distribution corresponds to the observed wage data and it is the measure conventionally used to compute the gender wage gap. The wage offers distribution is not directly observed and we are able to predict it thanks to the model structure: it indicates the wage offers actually received by men and women before they decide if accepting them or not. In many respects, the wage offers distribution represents a better measure to gauge the actual disadvantage or wage gap experienced by women because it avoids the selection bias due to gender differences in reservation wages[19]. The productivity distribution is also unobserved and it is the true primitive distribution in the model. Finally, another unobserved component that we are able to recover thanks to the structural estimates is the assignment of women to prejudiced and unprejudiced employers.

The gender gap in accepted wages is in line with the descriptive evidence, ranging from 26% to 20% overall. The gap is decreasing on the High School sample, stable on the College sample and actually increasing on the Master and PhD sample. The evidence of the gender wage gap over time - for example Eckstein and Nagypal (2004), Blau and Kahn (2006) and Flabbi (2010b) – report a stable or decreasing gap but they do not focus on schooling level as high as Master and PhD. This relative disadvantage of very high skilled women is a robust finding throughout the chapter.

The gender gap in wage offers is smaller than the gender gap in accepted wages at high education levels and larger at low education levels. This is evidence consistent with high skilled women being relative more choosy than similarly educated men. Two possible sources of this behavior are: 1) gender and education-specific preferences for job amenities and 2) gender asymmetries in household-level decisions. An example of the first is the result in Flabbi and Moro (2012): women with a college degree value the job amenity “work flexibility” more than women with an High School degree. An example of the second is the result in Flabbi and Mabli

[19] This is one of the advantage of obtaining structural paramaters estimates. The first paper estimating a search-matching-bargaining model (Eckstein and Wolpin (1995)) makes a similar argument in the context of returns to schooling estimation.

(2012): once we take into account that labor market decisions are taken at the household level, gender differentials in wage offers are estimated to be smaller than in an individual search model. Both elements are ignored in the version of the model we estimate.

The gender gap in productivity is relatively small in the College and High School sample but it is increasing over time. The productivity gap for Master and PhD holders is quite large and increasing over time. As a result, as we will see in the next section, differential in productivity are responsible for most of the gap we observe on the highest education group. However, since we ignore important factors such as preferences for job amenities and gender asymmetries in household-level decisions, we should regard this result with caution. Productivity is introduced in the model in a very reduced form way and it may well be the residual component that absorbs dynamic not explicitly modeled in our framework.

The gender gap between workers employed at prejudiced employers and workers employed at unprejudiced employers is larger the higher the education level. For example, the gender gap in wage offers at prejudiced employers on the Master and PhD sample in 2005 is equal to 55% while at unprejudiced employers is equal to 80%. This 25 percentage points difference decreases to about 7 percentage points on lower education levels. This result shows that even if the overall impact of prejudice on the high skilled sample is smaller than the impact of gender-specific productivity, it still has major impact in generating wage discrimination.

Table 2 shows some evidence on labor market dynamics: the hazard rates out of a given labor market state and the proportion of workers in each labor market states in equilibrium. The overall hazard rate out of unemployment is higher for women, a result explained both by higher arrival rates of offers and by lower reservation wages (see Table A2 in the Appendix). The gender gaps in unemployment rates are relatively small. Larger but not big differences are observed in the distribution of men and women between prejudiced and unprejudiced employers. Notice that we report the overall employment by employer type so when the proportion of prejudiced employers is extremely high (as in High School sample in 1995) most

of employed workers must work for them. This only *partial* segregation result implies that policies imposing quotas by employer (as common in some Affirmative Action policies, see Section 2.4.3) would not be very effective in reducing wage discrimination and the impact of prejudice on labor market outcomes.

2.3.3.2 Gender Wage Gap Decomposition

Table 3 reports a decomposition of the gender wage gap at different points of the accepted wages distribution. The gap is decomposed in the three sources of gender differentials assumed in the model: productivity, search frictions and prejudice. We perform the decomposition by taking into account equilibrium effects, i.e. by taking into account that changing the labor market environment induces individual agents to adjust their behavior.

The procedure we implement is the following. To isolate the impact of productivity, we impose that all the other differences between men and women do not exist. In particular, we assume that there are not prejudiced employers in the economy and that men and women face the same search frictions. Given this new environment, we compute the new optimal decision rules and we obtain new accepted wages distributions. On this counterfactual accepted wages distributions, we compute average accepted wages for men and women and we take the ratio of women values over men values. These ratios are reported in the Table. For example, the first row of college graduates in Table 3 states that if the only difference between men and women was the differential productivity we estimate, then the observed wage differential at the mean of the entire distribution would be much smaller than the one observed in the data: 7.8% as opposed to 21.2%. To isolate the impact of prejudice, we follow the same procedure: we fix productivity and search frictions of women equal to those of men but we let the proportion of prejudiced employers and their disutility to be equal to the one we estimate. We recompute the equilibrium and we obtain the statistics on the counterfactual wage distribution. Same exercise is done to isolate the impact of search frictions: the only parameters allowed to be different are the arrival rate of offers and the job termination rate. The “all parameters” exercise is a sort of goodness of fit: we generate wage offers distributions from an environment with all the parameters set equal to the point estimates.

Differences in productivity are the most important factor in explaining the wage gap for the top education level in 1995: productivity differentials alone will generate the entire differential we observe at the mean. This strong impact becomes smaller in 2005 and it is significantly smaller on the other education groups. The impact of search frictions always plays in favor of women: we would actually observe a reverse gender wage gap if the only differences between men and women in the labor market were due to search frictions[20]. Finally, the impact of prejudiced employers is very strong for High School and College graduated in 1995: prejudice is the most important factor in explaining the wage gap for these years and education groups. The impact of prejudice becomes smaller (and less important than the impact of productivity) over time for both education groups. Instead, the top education group (Master and PhD) shows the opposite trend: smaller impact in 1995 and stronger impact in 2005.

The top education group also shows a different behavior in terms of wage gap at top percentiles. This evidence is important to link back to the glass-ceiling issue we mentioned in the descriptive section. Master and PhD graduates in 2005 are the only group exhibiting evidence of glass-ceiling, i.e. a wage gap increasing as we move toward the top of the accepted wages distribution. This increasing wage gap is captured by the model but with a much smaller magnitude than in the data.

In conclusion, most of the results confirm what found in Flabbi (2010b): prejudice has still a significant impact in explaining the wage gap but the impact is decreasing over time, reaching a level smaller than the impact of differential productivity on all education groups in 2005. The disturbing exception to the trend, missed by Flabbi (2010b) since it considers only college graduates, is the top education group: Master and PhD graduates experience a stronger impact of prejudice in 2005 than 1995. This is also the education group showing evidence of glass-ceiling in 2005, confirming the view that sees the glass-ceiling as the remaining obstacle to reach gender equality in the labor market. However, the model does not have strong prediction in terms of the sources of the glass ceiling because it is able to only qualitatively match the gap

[20] The positive impact of search frictions is mainly driven by the higher arrival rate of job offers to women (see Table A2 in the Appendix).

at different percentiles of the distributions but it is unable to provide a precise quantitative fit[21].

2.4 Policy Implications and Policy Experiments

It has long been recognized that the presence of discrimination and prejudiced behavior generates inefficiencies and negative externalities[22] therefore presenting an opportunity for policy interventions. The United States has a relative long tradition of anti-discriminatory laws targeting the labor market. They can be broadly separated between *Equal Employment Opportunity* policies and *Affirmative Action* policies, even if the difference between the two types of policies tends to be starker in theory than in practice (Holzer and Neumark (2006)). The results of the Descriptive Section of the paper confirm systematic differences in labor market outcomes for men and women. The results of the more quantitative section of the paper suggest the presence of explicit prejudice against women and of a significant amount of wage discrimination and segregation. The main policy implications is therefore that policy interventions could be justified. More specific policy implications can be drawn by simulating policy interventions exploiting the estimates of the labor market environment we generated in Section 2.3. This is the objective of this last section of the paper. Before articulating specific policies, though, we have to define the welfare measure we want to use to evaluate them. In doing so, we also want to say something about pre-labor market decisions, which - as shown in the descriptive section of the model - are also significantly influenced by gender.

2.4.1 Welfare Measure and Returns to Schooling

The overall welfare of labor market participants depends on their current labor market state and, if employed, on their current wage. However, it also depends on the labor market dynamics related to the transitions between labor market states, the movements over the wage distributions and the durations in each state. A summary measure of overall workers'

[21] One possible reason for this poor fit are the imprecise estimates of the prejudiced parameters: probably due to the small sample size, the estimate do the disutility suffered by the prejudiced employers hiring Masters and PhDs is very imprecise. See Table A2 in the Appendix, parameter *k*.

[22] An example of negative externality generated by the formal model presented in this paper is the spillover effect inducing unprejudiced employers to wage discriminate.

welfare should then go beyond the comparison of wage gaps that we presented in Section 2.3.3.2.

One relative straightforward way to proceed is assigning to each labor market states occupied by workers in steady state the corresponding utility value (i.e. the wage if the worker is employed and the flow utility of unemployment if the worker is unemployed) and then averaging out these utilities values according to the equilibrium steady state distribution. This summary measures takes into account both the cross-sectional and the dynamic components of the labor market: the first is captured by the utility values associated to the labor market states and the second by the distribution over them since the distribution is directly related to durations and transitions probabilities[23].

In the first column for each year in Table 5, we present the gender differential over this welfare measures. Not surprisingly, since we have already seen that they face a worse labor market, women enjoy less welfare than men in all years and all schooling levels. In 1995, the differentials is about 25% with no much variation across education groups. In 2005, the differential is higher the higher the education level but the differences are still very small.

We also exploit this welfare measure to address some determinants of education decision. We have seen that women acquire more education than men even if the usual returns to schooling do not seem to suggest a higher return for women in the labor market. We think we can contribute to the debate by computing the returns based on our welfare measure. We can also perform the counterfactual experiments of computing what the returns would be if there were no prejudiced employers in the labor market. The results are reported in Table 4 where we compute for each year the gender-specific return of each schooling level. The first column states that completing a Master of PhD increase welfare on average by about 19.1% with respect to simply completing college. The returns increase to 59.9 with respect to completing only High School and they are about 34% when comparing college with High School.

The comparison of men and women leads to one of the most interesting result of our analysis. In 1995, female returns are higher than male returns. This result may provide an explanation

[23] An in depth discussion of this and similar welfare measures is in Flinn (2002). For the analytical expression of the welfare measure we use here, see Flabbi (2010a), Appendix A.5, Definition 5.

for the empirical puzzle found in the descriptive section where we observe women with higher education levels but lower hourly wages than men. By simply looking at cross-sectional wages, we are ignoring that women may have more to gain in terms of the overall labor market dynamic by acquiring additional education. Women completing Masters and PhDs receive a 22.5% return with respect to college and a 63.8% return with respect to High School; men receive, respectively, a 15.1% return and a 54.6% return. However, the ranking is reversed in 2005, implying that these incentives may be coming to an end. The counterfactual exercise in which we compute the same measures eliminating prejudiced employers (and taking into account equilibrium effects) suggests a potentially more constructive alternative: if prejudice may actually be eliminated in a market similar to the one we estimate in 2005, then the welfare returns to schooling would be equalized between men and women and the current distorted incentives could be eliminated.

2.4.2 Equal Employment Opportunity Policies

Equal Employment Opportunity policy interventions date back at least to the *Civil Rights Act* of 1964 which made it unlawful for an employer to "fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual with respect to his compensation, terms, conditions or privileges or employment, because of such individual's race, color, religion, sex, or national origin" (Section 703 (a)). The Act also established a specific institutional body to implement the law: the *Equal Employment Opportunity Commission* (EEOC). The role of the EEOC have been progressively expanded by subsequent legislation and the Commission is now responsible to enforce all the federal statutes prohibiting discrimination.

Equal Pay policies - i.e. policy specifically aiming at eliminating pay discrimination - are an active part of the equal employment opportunity policy agenda. The first Act signed by President Obama into Law after his inauguration is an example of equal pay policy. The *Lilly Ledbetter Fair Pay Act* of 2009 is a federal statue amending the *Civil Rights Act* of 1964 and stating that the 180-day statute of limitation for filing an equal-pay lawsuit regarding pay discrimination resets with each new discriminatory paycheck. Another related policy initiative is the *Paycheck Fairness Act,* which updates and strengthens the *Equal Pay Act* of 1963 to ensure better

protection against sex-based pay discrimination. The Act has the objective of preventing retaliation against workers who voluntarily discuss or disclose their wages and it allows women to receive the same protections for sex-based pay discrimination that are currently available to those subject to race or ethnicity-based discrimination[24].

In the context of an economic model, an equal pay policy can be defined as any policy that imposes restrictions on the wage determination with the objective of equalizing differentials among clearly identified groups. A simple implementation of such a policy within the Search-Matching-Bargaining model of the previous section would be to require each employer to pay the same wage to workers with identical productivity. Enforcement of such a policy can be guaranteed by assuming that the public authority responsible of enforcing (say, the EEOC) has the possibility of observing the match-specific value of productivity. Clearly, this is a very strong assumptions since the measures used to proxy productivity are often quite limited. An alternative way to think this equal pay policy is requiring that gender cannot be observed when wages and hiring are decided. Very limited examples of such a policy have been implemented in practice. Blind auditions to hire musicians implemented by some of the major US orchestras are probably the most well known (Goldin and Rouse (2000)).

We impose the requirement that each employer has to pay the same wage to workers with identical productivity by interpreting the Nash-Bargaining wage schedules defined in equations (1)-(3) as a reduced form sharing rules. As a result, offered wages are the average between the wages that would have been offered without the policy, where the average is over the respective proportions of men (m) and women in the population. The new wage equations are reported in Appendix B.

Results of the policy are summarized in Table 5. For each year, the first column of Table 5 reports the Benchmark model, the second the Equal Pay policy experiment and the third the Affirmative Action policy experiment that we will discuss in the next section. The Benchmark model is the model simulated using the point estimated obtained by our estimation procedure.

[24] The *Paycheck Fairness Act* was recently reintroduced in the 112th Congress after having twice passed the US House of Representatives but falling two votes short of a Senate vote on its merits in the 111th Congress.

The table reports the average welfare values by gender, year and education normalized with respect to the men's average welfare value of the appropriate year-education cell. For example, looking at the first column in 2005 we observe that the average welfare of MA and PhD women is 76% of the average welfare of MA and PhD men. The value increases as education decreases, reaching about 80% on the High School graduates sample.

The Equal Pay experiment is effective in redistributing welfare from men to women but it is never enough to completely close the gender gap. The closest to closing the gap we get is with the High School sample in 1995: on this sample, the gap decreases from about 26% to about 4%. In general, the equal pay policy is more effective in reducing the gap at low education levels than at high education levels. Due to the equilibrium effects and the presence of spillover, the policy has the potential to generate average net gains, i.e. a situation where the average benefit received by women is higher than the average loss experienced by men. By looking at the overall average welfare value, we observe that this is the case only for the 1995 College and High School sample.

In conclusion, the policy impose a very strong requirement in terms of wage determination but it does not seem to generate very large effects. It is a policy more effective for lower education levels than for higher ones and it has larger impacts in 1995 than in 2005. In two cases, the equal pay policy is able to generate net gains, i.e. the average male loss is more than compensated by the average female gain.

2.4.3 Affirmative Action Policies

Affirmative Action policies in the labor market officially starts in the US with the 1961 *Kennedy Executive Order #10925* that mandates "affirmative action" to avoid discrimination by race in the labor market. The 1967 *Johnson Executive Order #11375* extends its application to cover women. In the legislative and policy debate an Affirmative Action policy is any anti-discrimination policy that requires proactive steps (Holzer and Neumark 2000). In the economic literature, an Affirmative Action policy is frequently described as a "quota" policy, i.e. a system of exogenously imposed numerical yardsticks for minority in hiring, federal contracting or

school enrollments. A quota system definition was not mentioned in the original Presidential Executive orders but, given its convenience in providing objective measures and targets, was introduced in the subsequent regulations governing the executive orders. For example, the 1968 Department of Labor Regulations governing the 1967 Johnson executive order require explicitly to identify “underutilization” of women and minority. The quota system definition is also the definition most frequently enforced by the Equal Employment Opportunity Commission (EEOC).

The difference between "proactive steps" and "exogenously imposed quota" seems to inform a lot of the debate on affirmative action in the labor market and in education. Holzer and Neumark (2000) provide an extensive review of the economic and public policy literature and conclude that the difference is crucial both in terms of effectiveness and in terms of political viability of the policies. Their overall conclusion is that affirmative action in the US has offered "significant redistribution toward women and minorities, with relatively small efficiency consequences". Donohue and Heckman (1991) focus on the impact of the Civil Rights legislations on labor market outcomes of African-Americans. They also broaden the definition of affirmative action beyond a simple quota system and conclude that the policies had a significant role in improving labor market outcomes. The two most recent Supreme Court opinions about affirmative action - Grutter v. Bollinger and Gratz v. Bollinger - were delivered on June 24, 2003 and stress the unconstitutionality of explicit quota policies bit the admissibility of proactive policies. Discussing, respectively, the admission policy to the College and Law School of the University of Michigan, Justice O'Connor states in the majority opinion that “a race-conscious admission program cannot use a quota system” but a "narrowly tailored plan system" in which "race or ethnicity" may be considered "a 'plus' in a particular applicant's file" constitutes a legitimate affirmative action policy. In conclusion, the tendency of the legislation and the public policy debate has been to push affirmative action policies away from rigid and exogenous quota target toward other proactive steps that could endogenously generate similar outcomes.

In line with this debate, we propose an affirmative action policy which is not a quota policy but a proactive step in the form of a subsidy. The policy is defined as a flow subsidy received by an

employer for each woman hired. The subsidy is paid by a lump-sum tax on workers. The new wage equations are reported in Appendix B.

A subsidy does not impose any predetermined quota but by being offered only for hiring women, definitely make gender a "a 'plus' in" some "applicant's file", as stated in the Supreme Court opinion. The impact of the policy is magnified by the spillover effects: not only the presence of a flow subsidy has a direct positive impact on women wages because firms receive additional revenue from hiring them but also has an indirect positive impact because it increases women's outside option. We fix the subsidy to be equal to 5% of men's average wage in the corresponding year and education group.

The results of the subsidy policy are reported in the third column of each year in Table 5. The policy is less effective in closing the gap than the equal pay policy is but at the same time it is generating more net gains. Net gains are realized in both 1995 and 2005 on the College sample and in 1995 on the Master and PHD sample. Another advantage of the policy is that it can be calibrated more precisely by increasing or decreasing the subsidy and its costs can be distributed in a variety of ways by changing the structure of the tax necessary to support it. In this respect, the lump-sum tax implemented here is the least distortionary for the economy but it is also the most costly for men. If even under this scheme, three cases out of six generate positive gains then an affirmative action policy structured as a subsidy should be seen as the policy with the larger potential for success among the two policies considered here. It is also the policy that generates the largest impact on the group that the previous evidence has shown to be the most problematic: women with top education skills, supplying labor in the most recent years.

In conclusion, an affirmative action policy structured as a relative modest subsidy provided to employers that hire women has a modest but positive impact in closing the gender gap in welfare. Despite the modest impact, the policy is promising because it is frequently able to close the gap without reducing overall welfare and it is effective in targeting the most problematic education and demographic group (the 2005 Master and PhD sample).

2.5 Summary and Conclusion

2.5.1 Summary of Results

Three main results emerge from the descriptive evidence on gender differentials by education in the US labor market over the last twenty years.

1. Women acquire more college education than men, reinforcing a trend started with the generation born in 1959. The proportion of women with a Master or PhD degree is still smaller than the proportion of men but the differential is shrinking. As shown for other countries in other chapters of the book, a lot of gender asymmetry by gender persists in the choice of field of study at same education level.
2. Women participate less than men in the labor market but when they do, they obtain similar employment rates. The intensive margin of the labor supply shows a large gender gap, mainly as a result of the larger incidence of part-time work among female workers. The gender gap in labor supply is not reduced but actually magnified by additional education.
3. The gender earnings gap is about 20%, even after controlling for standard human capital and demographic characteristics. The gap shows a significant reduction over time (it was about 35% in the early 1980s) but has remained fairly stable in the last 10/15 years. One reason for the persistent gender gap in recent years among highly-educated workers is the large differential at the top of the earnings distribution, an evidence often correlated with "glass-ceiling" effects.

We use a search-matching-bargaining model to investigate some of the sources of the observed gender differentials. We focus on gender differentials with the objective of isolating the impact of three determinants of gender gaps: productivity differences; employers' prejudice; and search frictions. Four main results emerge from the analysis.

1. Prejudice has a significant impact in explaining the wage gap but the impact is decreasing over time and it becomes smaller than the impact of productivity on all education groups in 2005.

2. Search frictions actually favor women, thanks to the higher frequency with which they receive job offers.
3. Master and PhD graduates are an exception to the decreasing impact of prejudice over time: they experience a stronger impact of prejudice in explaining the wage gap in 2005 than in 1995.
4. We also decompose the impact of the three sources of the gender wage gap at different point of the distribution: the data show evidence of glass-ceiling effects on the Master and PhD sample in 2005, a result that only the equilibrium effects of all three factors together is able to partially explain.

We use the structural estimates of the model to build a welfare measure to evaluate returns to education and two policy interventions. The first policy is an equal pay policy imposing one wage at same productivity, the second is an affirmative action policy providing incentives to hire women. Three main results emerge from the exercise:

1. In 1995, female returns to schooling estimated using our welfare measure are higher than male returns. This new result provides a rationale for the apparent empirical puzzle of women acquiring more education than men even if they are paid less for these skills in the labor market. In 2005, female returns are estimated to be lower than male returns, implying that women attitudes toward education could change, reversing the positive education gap. A counterfactual exercise suggests that if prejudice were to be completely eliminated in 2005, then welfare returns to schooling would be fully equalized between men and women.
2. The equal pay policy redistributes welfare from men to women but it is not able to fully close the gender gap. Given the strong requirement in terms of wage determination imposed by the policy, we judge it is not very effective. Still, the policy is able to generate net gains on the College and High School samples in 1995.
3. The Affirmative Action policy has a smaller impact on closing the gender gap than the equal pay policy but it is more likely to generate net welfare gains. The impact of the policy is increasing in the education level of the worker.

2.5.2 Conclusions

We think the evidence provided in this chapter indicates that gender gaps in labor market outcomes are far from being settled issues. On top of the traditional issues of lower participation rates and gender gaps in wages, new issues that are likely to become more relevant in the future include:

1. Convergence in education levels is not enough to close gender gaps in the labor market.
2. Evidence of "glass-ceiling" effects. The evidence includes a marked underrepresentation of women in top positions at the firm and a larger gender wage gap at the top of the wage distribution.

The conclusion of our analytical contribution is that prejudice may still have a role in explaining the evidence. Even if the magnitudes of our effects is conditional on a highly stylized model, we characterize in some details at least one scenario were the possibility of the presence of prejudiced employers in the labor market has substantial effects. In particular, it is responsible for the reversal of the return to schooling ranking in recent years and it may explain up to 44% of the gender wage gap of the top education group (Master and PhD) in 2005.

If prejudice is still important, then policy interventions may be effective in attaining both efficiency and welfare gains. We use our model to evaluate an equal pay policy and an affirmative action policy. Among the policies we consider, we favor an affirmative action policy structured as a relative modest subsidy provided to employers for hiring women. We favor this policy because it is frequently able to close the gender gap without reducing overall welfare and because it is effective in targeting the group that should take center stage in the future debate about gender differentials: high-skilled, high-earners workers.

References for Chapter 2

Albanesi, Stefania, Olivetti, Claudia (2006), "Gender and Dynamic Agency: Theory and Evidence on the Compensation of Female Top Executives", *WP2006-061, Boston University - Department of Economics*.

Albrecht, James, Björklund, Anders, Vroman, Susan (2003), "Is There a Glass Ceiling in Sweden?", *Journal of Labor Economics*, Vol. 21, No. 1, pp. 145–77.

Becker, Gary S., Hubbard, William H., Murphy, Kevin (2010), "Explaining The Worldwide Boom in Higher Education of Women", *Journal of Human Capital*, 4(3): 203-241.

Beffy, Magali, Fougère, Denis,. Maurel, Arnaud (2009), "Choosing the Field of Study in Post-Secondary Education: Do Expected Earnings Matter?", *IZA Discussion Paper* No. 4127.

Bertrand, Marianne, Goldin, Claudia, Katz, Lawrence (2010), "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors", *AEJ: Applied Economics*, 2: 228-255.

Bertrand, Marianne, Hallock, Kevin (2001), "The Gender Gap in the Top Corporate Jobs", *Industrial and Labor Relations Review*, 55(1): 3-21.

Biddle, Jeff, Hamermesh, Daniel (2012), "Wage Discrimination over the Business Cycle", *IZA Discussion Paper*, No. 6445.

Black, Dan (1995), "Discrimination in an Equilibrium Search Model," *Journal of Labor Economics*, 13(2) (1995), 309–34.

Blau, Francine, Kahn, Lawrence (2006), "The US Gender Pay Gap in the 1990s: Slowing Convergence", *Industrial and Labor Relations Review,* 60(1).

Chevalier, Arnaud (2011), "Subject Choice and Earnings of UK Graduates", *IZA Discussion Paper*, No. 5652.

Chiappori, Pierre-André, Iyigun, Murat, Weiss, Yoram, (2009), "Investment in schooling and the marriage market", *American Economic Review*, 99:5, 1689–1713.

Eckstein, Zvi, Nagypal, Eva (2004), "The evolution of US earnings inequality: 1961-2002", *Federal Reserve Bank of Minneapolis Quarterly Review*, 28 (2).

Eckstein, Zvi, van den Berg, Gerard (2007), "Empirical Labor Search: a Survey", *Journal of Econometrics*, 136: 531-564.

Eckstein, Zvi, Wolpin, Kenneth (1995), "Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model", *The Review of Economic Studies*, 62(2): 263-286.

Flabbi, Luca (2011), "Gender Differences in Education, Career Choices and Labor Market Outcomes on a Sample of OECD Countries", OECD Background Paper for: World Bank, *World Development Report 2012*, World Bank: Washington, DC.

Flabbi, Luca (2010a), "Gender discrimination estimation in a search model with matching and bargaining", *International Economic Review* 51 (3), 745–783.

Flabbi, Luca (2010b), "Prejudice and Gender Differentials in the U.S. Labor Market in the Last Twenty Years", *Journal of Econometrics*, 156: 190-200.

Flabbi, Luca, Mabli, James (2012), "Household Search or Individual Search: Does it Matter? Evidence from Lifetime Inequality Estimates", *Georgetown University, Department of Economics WP* 12-02.

Flabbi, Luca, Moro, Andrea (2012), "The Effect of Job Flexibility on Women Labor Market Outcomes: Estimates from a Search and Bargaining Model", *Journal of Econometrics*, 168: 81–95.

Flinn, Christopher (2006) "Minimum wage effects on labor market outcomes under search, bargaining and endogenous contact rates", *Econometrica* 73, 1013–1062.

Flinn, Christopher (2002), "Interpreting Minimum Wage Effects on Wage Distributions: A Cautionary Tale", *Annales d'Economie et de Statistique*, 67/68, 309-355.

Flinn, Christopher, Heckman, James (1982), "New methods in analyzing structural models of labor market dynamics", *Journal of Econometrics* 18, 115–168.

Fortin, Nicole M., Lemieux, Thomas (2000), "Are women's wage gains men's losses? A distributional test", *American Economic Review Papers & Proceedings* 90 (2), 457:460.

Gayle, George-Levi, Golan, Limor, Miller, Robert A. (2009), "Are There Glass Ceilings for Female Executives?", *mimeo*, Tepper School of Business, Carnegie Mellon University.

Ge, Suqin (2011), "Women's college decisions: how much does marriage matter?", *Journal of Labor Economics*, Vol. 29, No. 4, pp. 773-818.

Goldin, Claudia, Rouse, Cecilia (2000), "Orchestrating Impartiality: the Impact of "Blind" Auditions on Female Musicians", *American Economic Review*, 90(4), 715-741.

Holzer, Harry, Neumark, David (2006), "Affirmative Action: What Do We Know?", *Journal of Policy Analysis and Management*, 25(2): 463-490.

Holzer, Harry, Neumark, David (2000), "Assessing Affirmative Action", *Journal of Economic Literature*, 38(3): 483-568.

OECD (2008), *Education at a Glance 2008*, OECD Publications: Paris.

Waldfogel, Jane (1998), "Understanding the 'Family Gap' in Pay for Women with Children", *Journal of Economic Perspectives*, v. 12: 137-156.

Figure 1a: Gender Gap in Percentage of College Graduates by Year

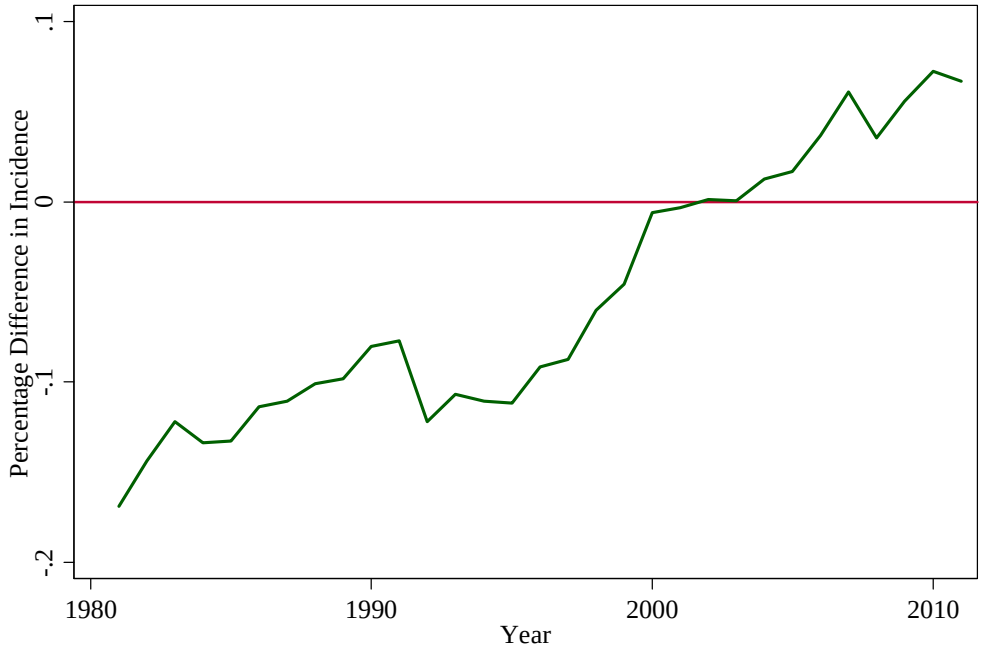


Figure 1b: Gender Gap in Percentage of College Graduates by Cohort

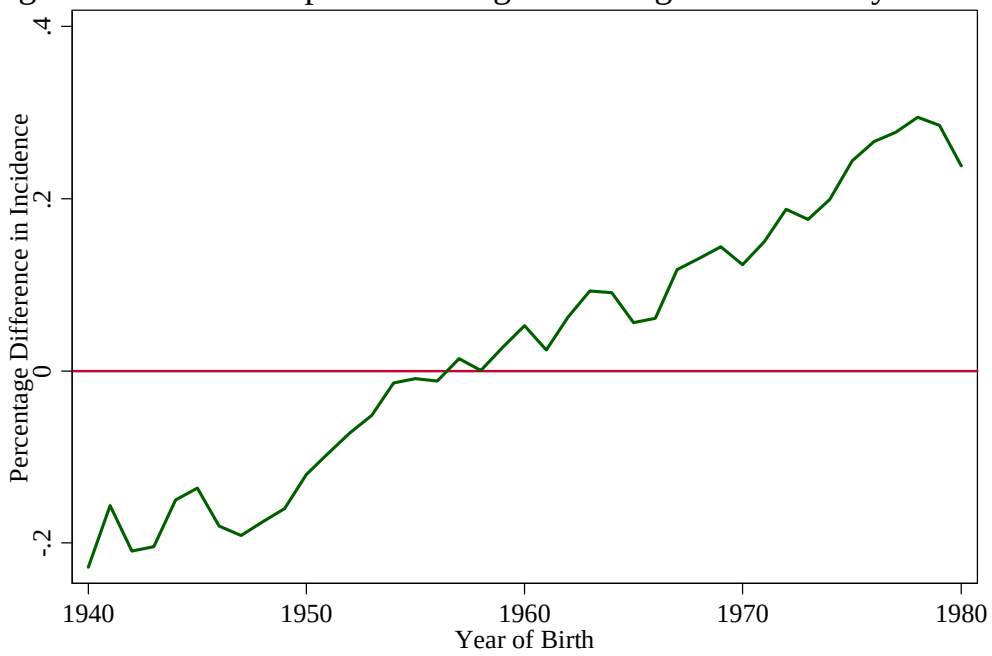


Figure 2: Gender Gap in Percentage of Master Degree Graduates by Cohort

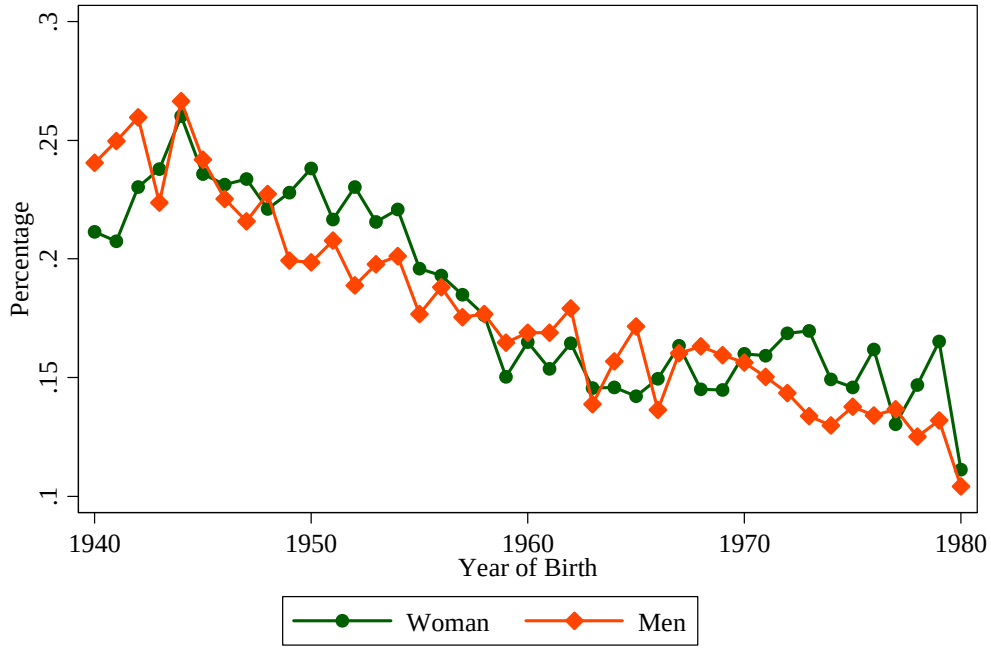


Figure 3: Gender Gap in Percentage of Doctoral Degree Graduates by Cohort

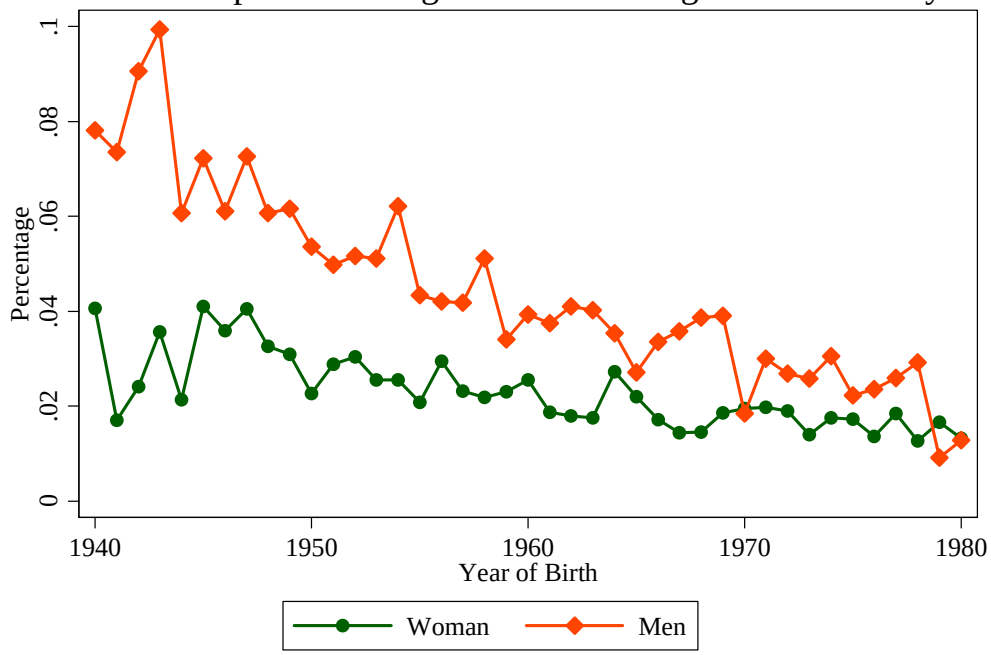
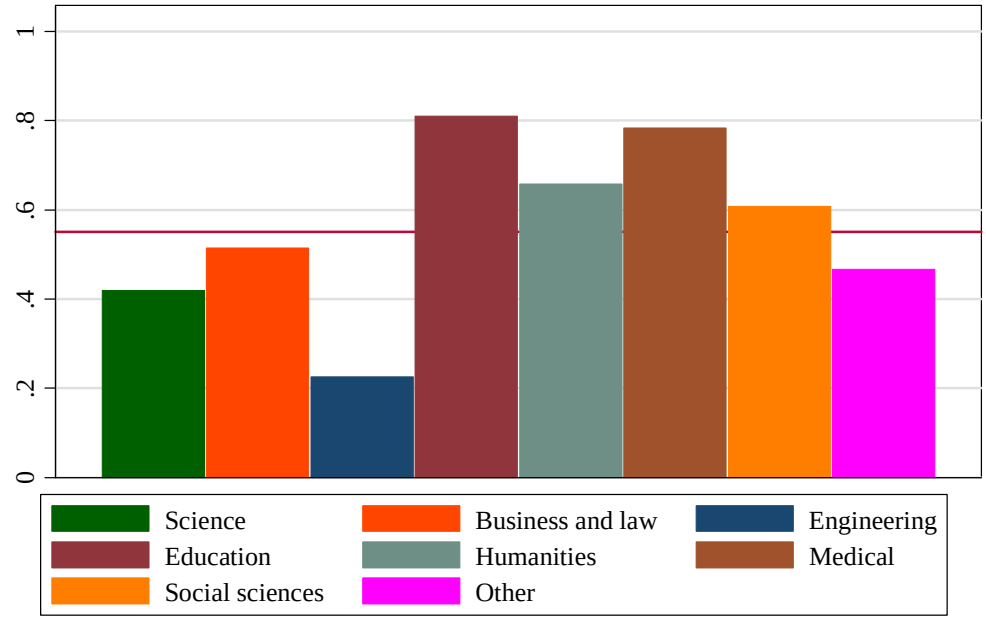
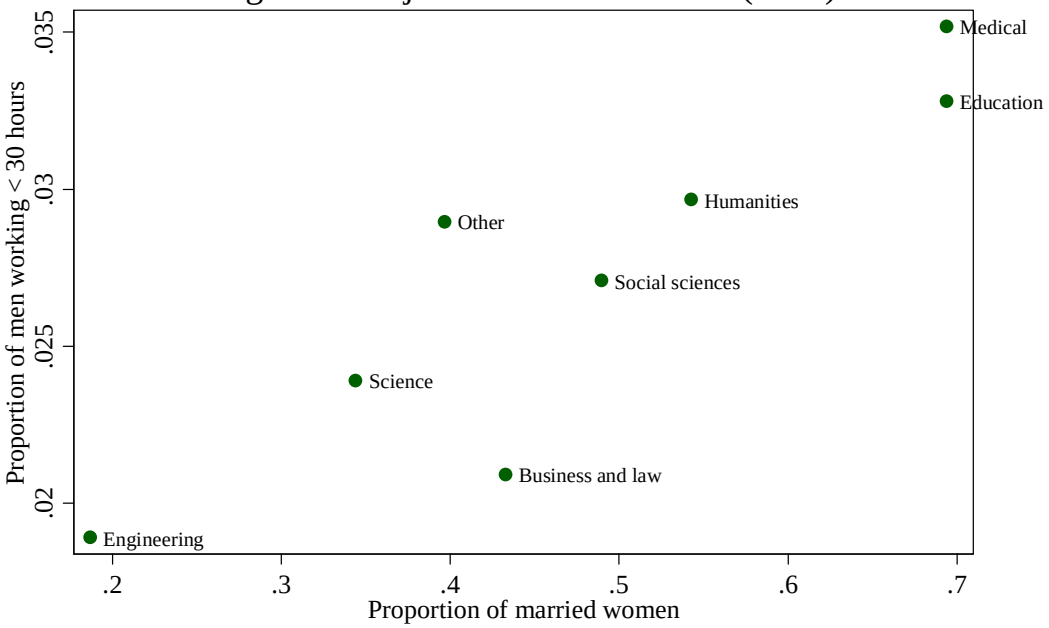


Figure 4: Proportion of Women in Field of Study (2002)



Note: 30-45 years old.

Figure 5: Majors with Flexible Jobs (2002)



Note: 30-45 years old.

Figure 6a: Participation Rates by Gender and Education Level

(Percentage of Relevant Group in Adult Civilean Population)

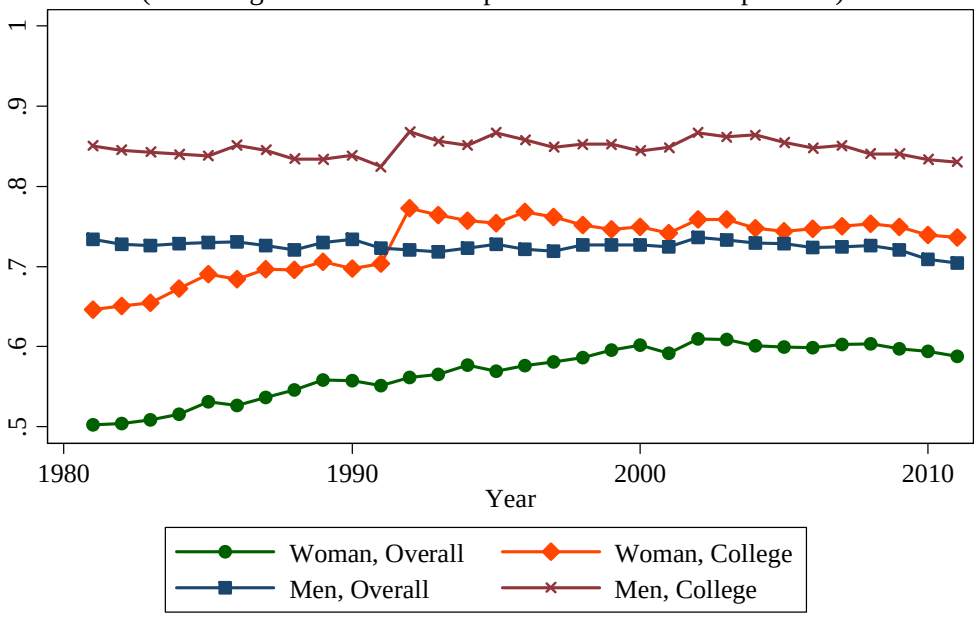


Figure 6b: Participation Rates by Cohort, Gender and Education Level

(Percentage of Relevant Group in Adult Civilean Population)

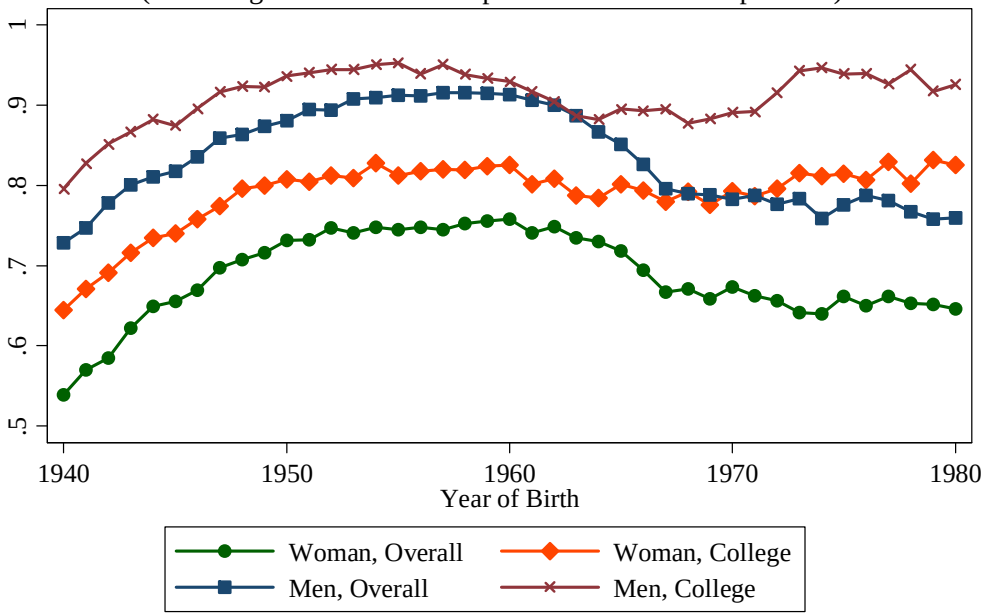


Figure 7a: Employment Rates by Gender and Education Level

(Percentage of Relevant Group in Labour Force)

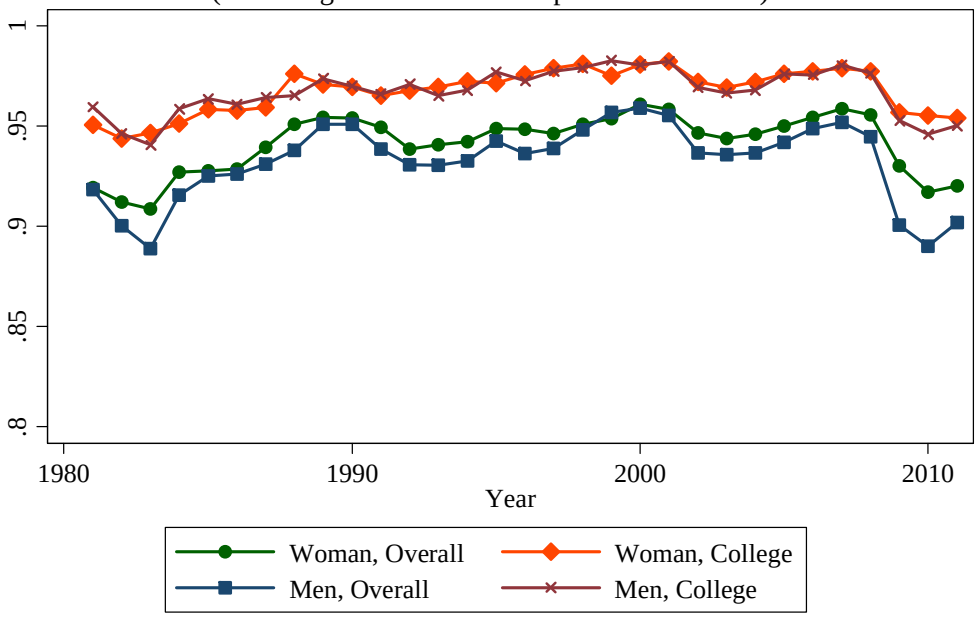


Figure 7b: Employment Rates by Cohort, Gender and Education Level

(Percentage of Relevant Group in Labour Force)

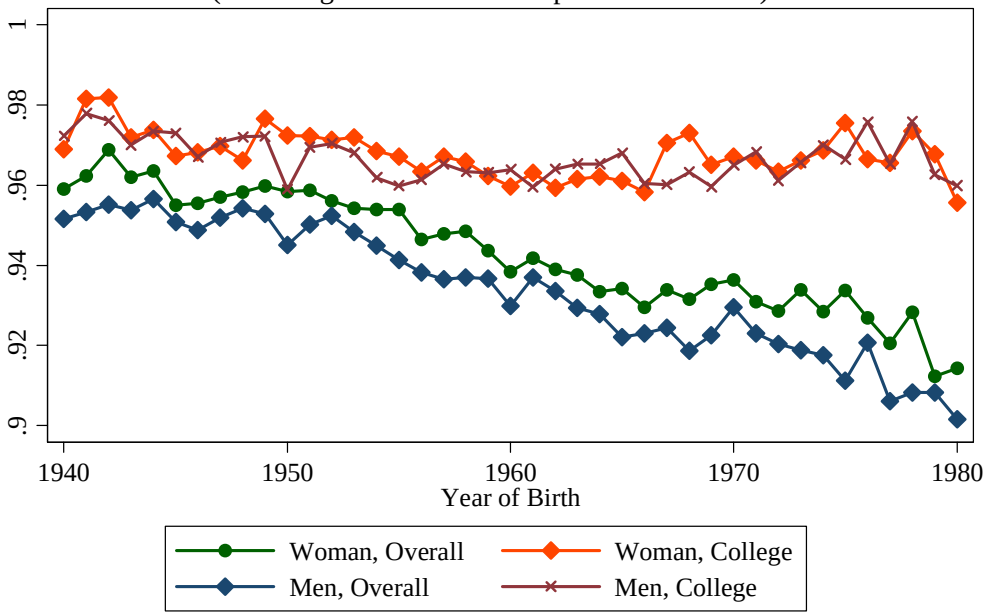
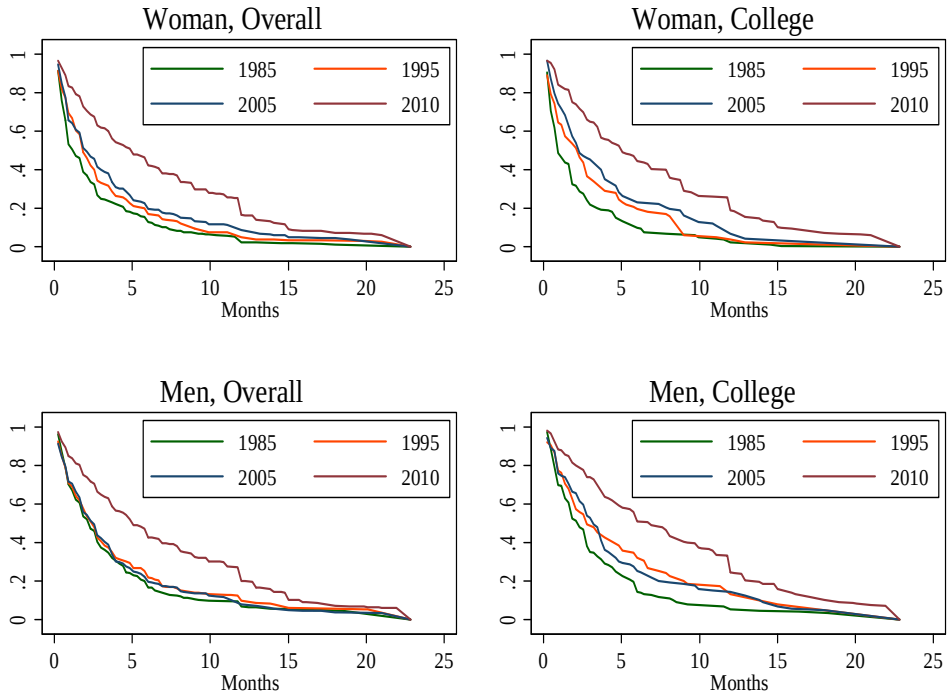


Figure 8: Survival Function in Unemployment by Gender and Education Level



Note: Kaplan-Meier Estimator

Figure 9a: Distributional aspects of Hours Worked by Gender and Education Level (1985)

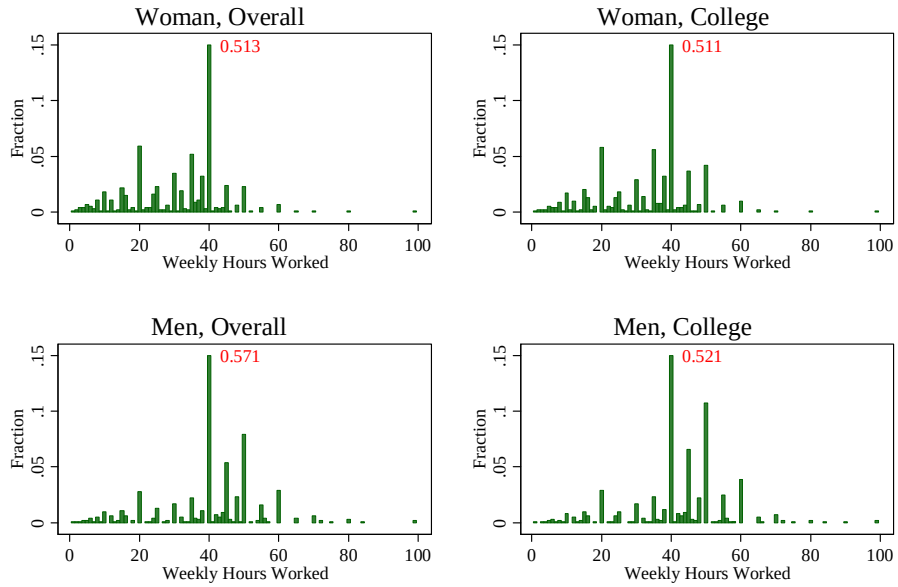


Figure 9b: Distributional aspects of Hours Worked by Gender and Education Level (2010)

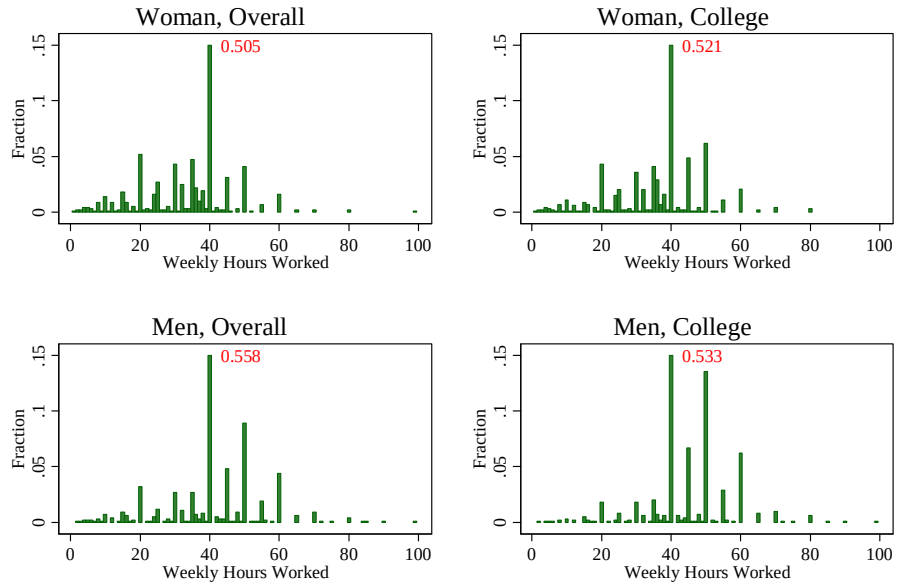


Figure 10a: Gender Gaps in Part Time Occupations

(Men vs. Women with Part Time Jobs)

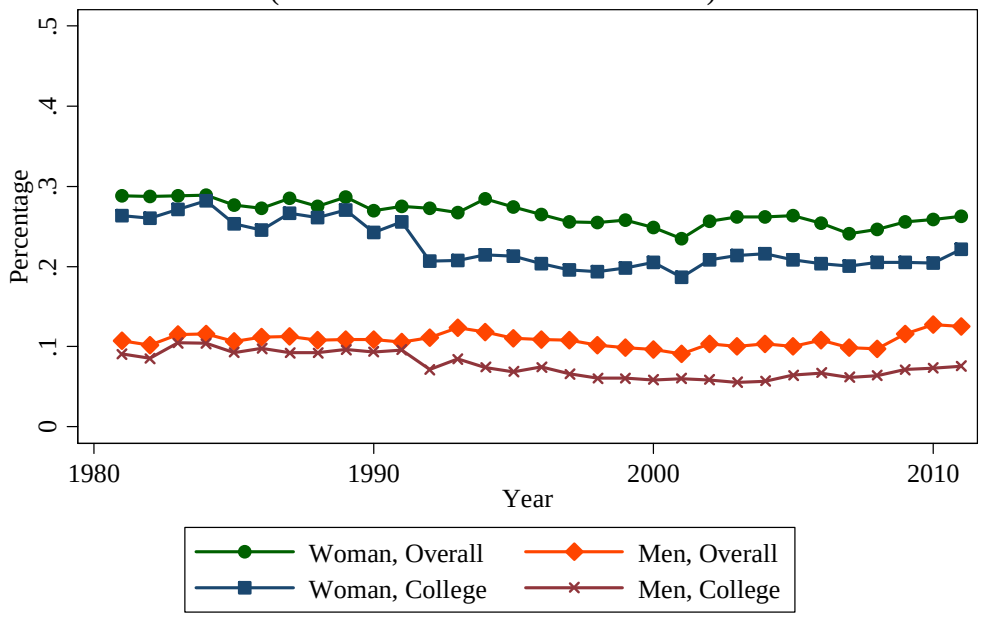


Figure 10b: Percentage of Women in Part Time Jobs by Age

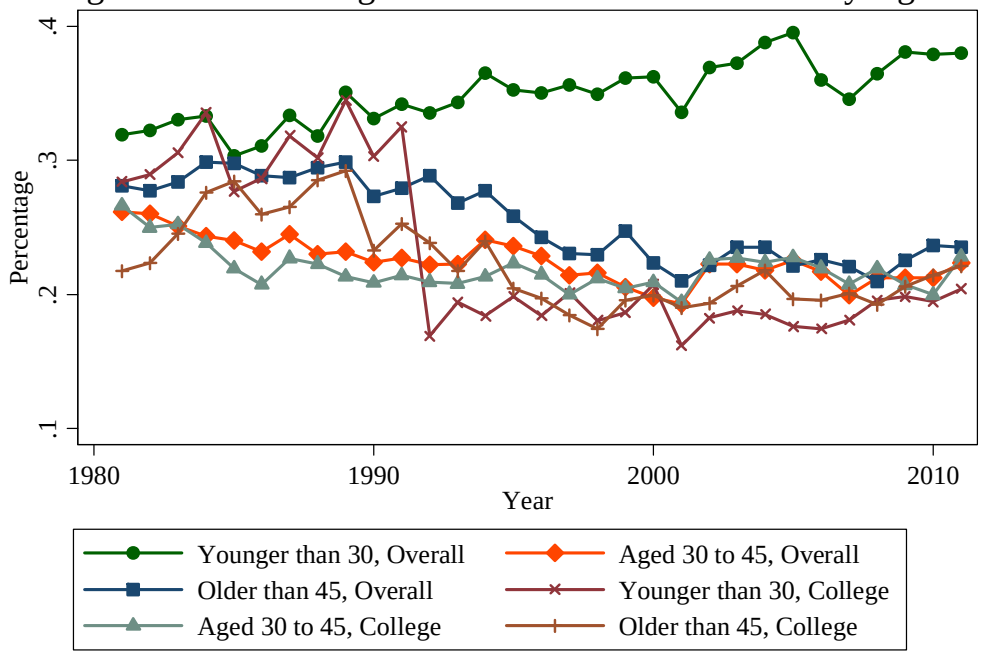
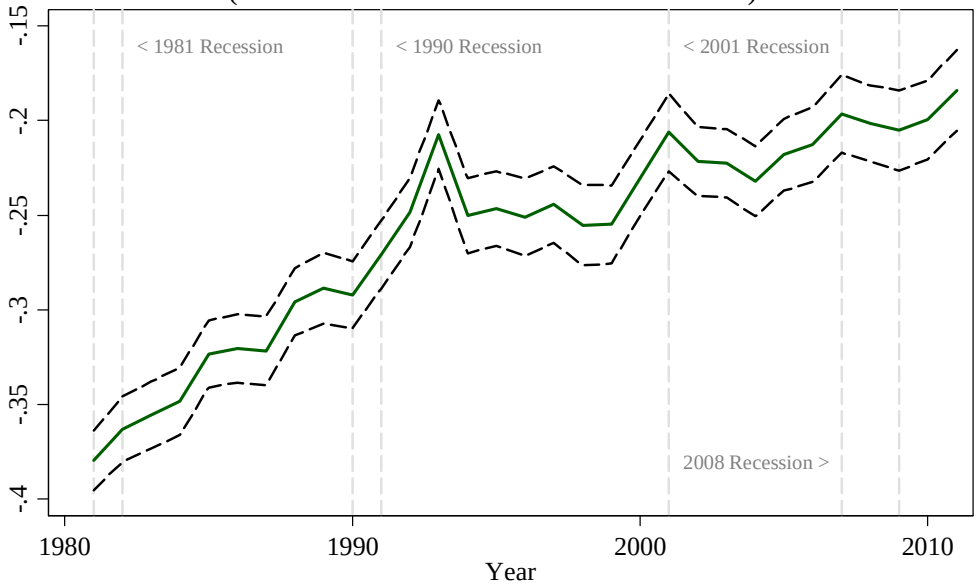


Figure 11a: Gender Earnings Differential Over Time, Unconditional Case

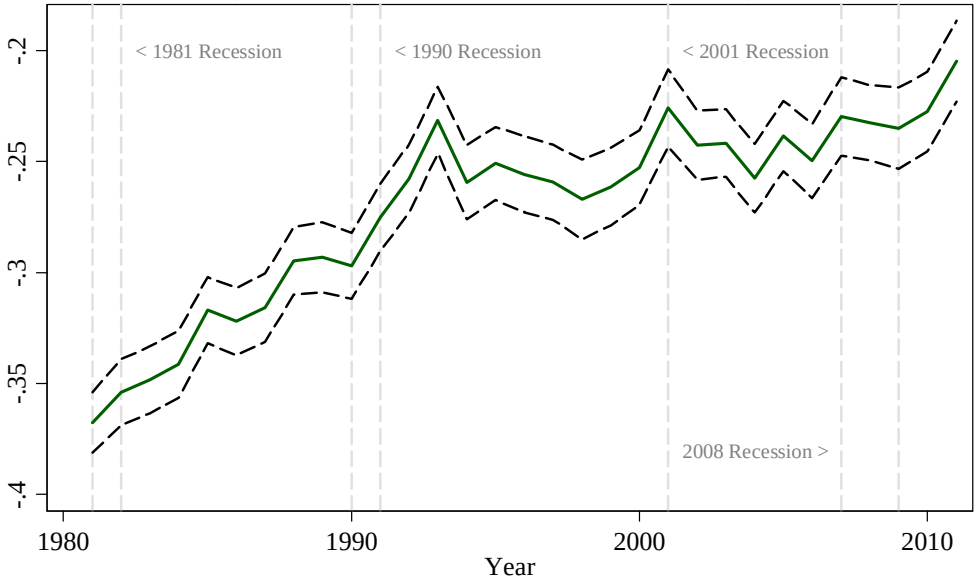
(Point estimates and 95% confidence interval)



Note: Dashed Lines represent 95% confidence interval.

Figure 11b: Gender Earnings Differential Over Time, Conditional Case

(Point estimates and 95% confidence interval)



Note: Dashed Lines represent 95% confidence interval.

Figure 11c: Gender Gap in Log Hourly Earnings by Percentile and Year, Unconditional Case

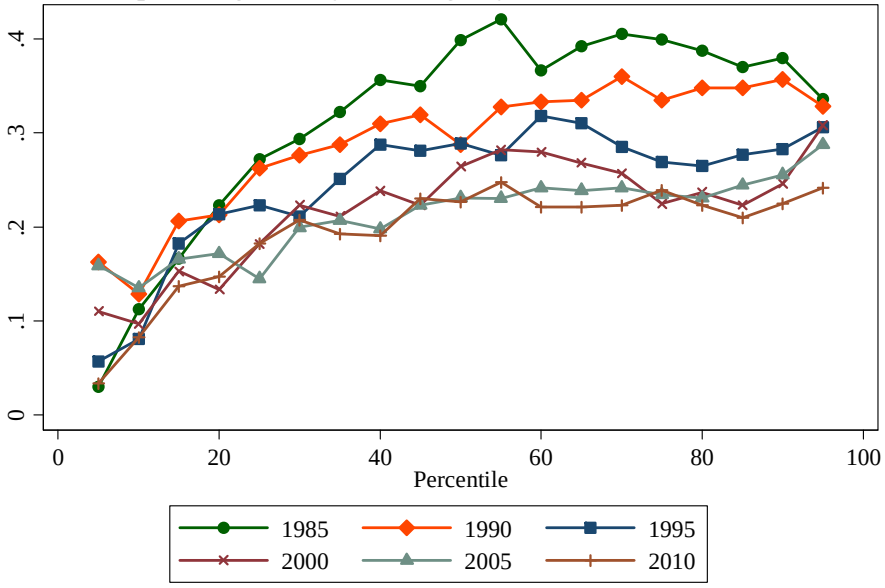


Figure 12a: Gender Composition of Managerial Occupations

(Proportion of Women)

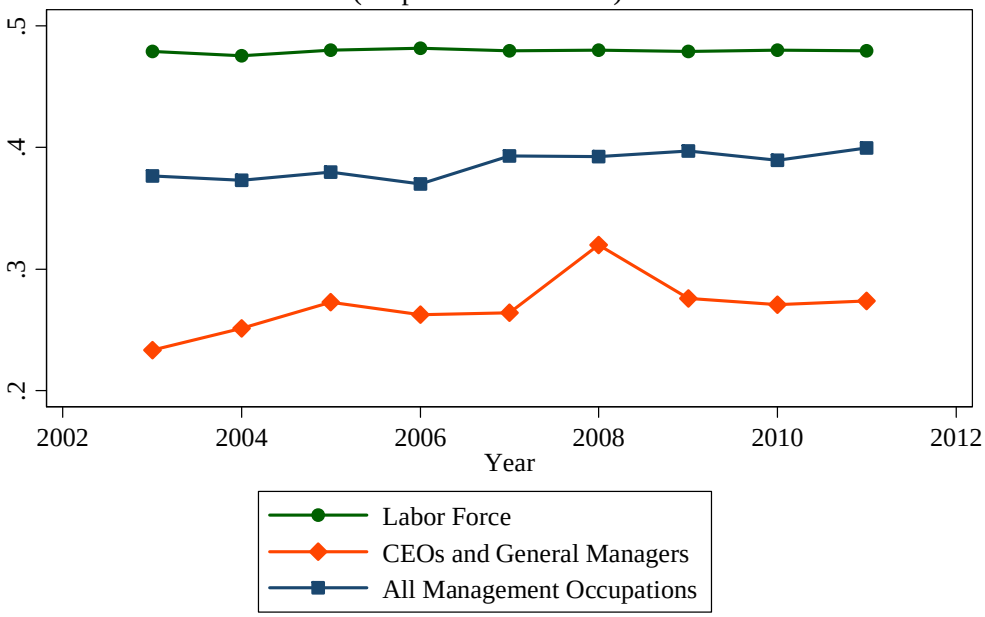
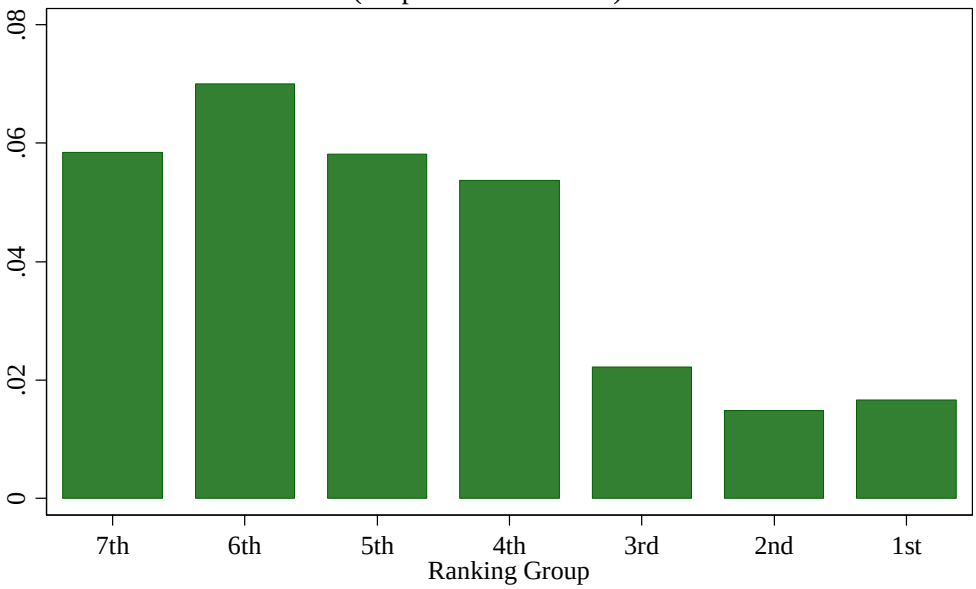


Figure 12b: Positions within CEOs and General Managers

(Proportion of Women)



Note: Gayl, Golan and Miller (2009), Table 1.

Table 1. Estimation Results - Predicted Productivity and Wages

	Master and PhD		College		High School	
	1995	2005	1995	2005	1995	2005
			Men			
Productivity						
Average	49.88	58.48	45.11	48.32	35.23	36.93
	(1.509)	(2.052)	(0.739)	(0.834)	(0.396)	(0.453)
Variance	911.33	1748.68	858.14	1271.58	602.23	862.33
	(113.0)	(241.0)	(56.4)	(86.6)	(26.9)	(41.4)
Average Wage Offers						
Overall	30.36	34.68	27.03	29.16	21.28	22.46
	(0.754)	(1.026)	(0.370)	(0.417)	(0.198)	(0.226)
Average Accepted Wages						
Overall	30.50	34.95	27.17	29.56	21.50	22.99
	(0.753)	(1.028)	(0.370)	(0.417)	(0.198)	(0.227)
			Woman			
Productivity						
Average	42.68	45.39	42.40	39.56	32.31	30.28
	(1.485)	(2.210)	(2.727)	(1.198)	(2.295)	(0.887)
Variance	367.94	567.13	385.36	629.37	337.18	488.36
	(50.5)	(113.3)	(29.6)	(42.9)	(14.6)	(25.3)
Average Wage Offers						
Overall	25.03	25.92	21.78	23.18	14.68	17.84
	(2.600)	(5.538)	(0.450)	(0.595)	(0.367)	(0.829)
Prejudiced Employers	20.63	19.19	20.18	21.75	14.68	16.81
	(15.127)	(24.569)	(0.359)	(0.404)	(0.367)	(0.493)
Unprejudiced Employers	25.81	27.73	24.56	23.68	18.99	18.39
	(0.742)	(1.105)	(1.363)	(0.599)	(1.147)	(0.444)
Average Accepted Wages						
Overall	24.53	26.44	20.85	22.83	15.86	18.07
	(1.711)	(3.453)	(0.337)	(0.292)	(0.305)	(0.187)
Prejudiced Employers	17.32	21.46	18.72	20.13	15.86	16.97
	(9.401)	(16.124)	(1.000)	(0.458)	(0.305)	(0.321)
Unprejudiced Employers	25.81	27.78	24.56	23.78	19.01	18.66
	(0.740)	(1.092)	(1.362)	(0.556)	(1.125)	(0.361)

Notes: The table reported predicted values based on the Maximum Likelihood estimated structural parameters. Estimated parameters are reported in Table A2 in the Appendix. Asymptotic standard errors by Delta method in parentheses.

Table 2. Estimation Results - Predicted Labor Market Dynamics

	Master and PhD		College		High School	
	1995	2005	1995	2005	1995	2005
			Men			
Hazard Rate out of Unemployment						
Overall	0.11107	0.27512	0.17459	0.18792	0.20459	0.22586
	(0.04535)	(0.07942)	(0.02796)	(0.02296)	(0.01395)	(0.01437)
To a Prejudiced Empl.	0.01673	0.05824	0.11086	0.04916	0.20458	0.07899
	(0.00859)	(0.03413)	(0.03831)	(0.02198)	(0.01396)	(0.03785)
To an Unprejudiced Empl.	0.09435	0.21688	0.06373	0.13876	0.00001	0.14687
	(0.03887)	(0.06929)	(0.03545)	(0.02710)	(0.00060)	(0.03866)
Hazard Rate out of Employment						
Overall	0.00169	0.00819	0.00443	0.00711	0.01174	0.01378
	(0.00098)	(0.00337)	(0.00101)	(0.00124)	(0.00115)	(0.00126)
Labor Market Status						
Emp. Rate at Prejudiced	0.14833	0.20557	0.61926	0.25205	0.94568	0.32960
	(0.04621)	(0.10484)	(0.18965)	(0.10843)	(0.00454)	(0.15654)
Emp. Rate at Unprejudiced	0.83667	0.76551	0.35598	0.71150	0.00005	0.61288
	(0.04649)	(0.10503)	(0.18964)	(0.10848)	(0.00277)	(0.15655)
Unemployment Rate	0.01500	0.02891	0.02476	0.03645	0.05427	0.05752
	(0.00608)	(0.00823)	(0.00392)	(0.00437)	(0.00360)	(0.00355)
			Woman			
Hazard Rate out of Unemployment						
Overall	0.24914	0.23926	0.26069	0.22785	0.25167	0.22910
	(0.11140)	(0.07975)	(0.03685)	(0.02871)	(0.01948)	(0.01581)
To a Prejudiced Empl.	0.00234	0.00313	0.16066	0.04805	0.25166	0.06727
	(0.02711)	(0.03145)	(0.05927)	(0.02514)	(0.01949)	(0.03905)
To an Unprejudiced Empl.	0.24681	0.23612	0.10003	0.17980	0.00001	0.16184
	(0.11358)	(0.08485)	(0.05656)	(0.03330)	(0.00082)	(0.04035)
Hazard Rate out of Employment						
Overall	0.00407	0.00493	0.00949	0.00774	0.01216	0.01230
	(0.00259)	(0.00233)	(0.00191)	(0.00139)	(0.00135)	(0.00122)
Labor Market Status						
Emp. Rate at Prejudiced	0.00923	0.01284	0.59465	0.20397	0.95385	0.27865
	(0.10697)	(0.12874)	(0.20269)	(0.10360)	(0.00466)	(0.16061)
Emp. Rate at Unprejudiced	0.97469	0.96698	0.37024	0.76318	0.00006	0.67040
	(0.10721)	(0.12936)	(0.20268)	(0.10382)	(0.00310)	(0.16066)
Unemployment Rate	0.01608	0.02018	0.03511	0.03284	0.04609	0.05094
	(0.00713)	(0.01229)	(0.00488)	(0.00882)	(0.00348)	(0.00626)

Notes: The table reported predicted values based on the Maximum Likelihood estimated structural parameters. Estimated parameters are reported in Table A2 in the Appendix. Asymptotic standard errors by Delta method in parentheses.

Table 3. Wage Gap Decomposition - Women/men ratio on average accepted wage

Gap	1995				2005			
Generated by:	Overall	Top 50%	Top 25%	Top 10%	Overall	Top 50%	Top 25%	Top 10%
				Master and PhD				
Productivity	0.835	0.883	0.865	0.857	0.699	1.628	1.344	1.130
Prejudice	0.935	0.958	0.971	0.982	0.881	0.923	0.950	0.974
Search Frictions	1.337	1.195	1.139	1.088	1.052	1.031	1.021	1.013
All Parameters	0.804	0.867	0.853	0.849	0.756	0.857	0.841	0.834
Sample	0.823	0.791	0.793	0.812	0.728	0.682	0.638	0.568
				College				
Productivity	0.922	0.949	0.918	0.895	0.769	0.845	0.847	0.856
Prejudice	0.812	0.872	0.906	0.937	0.936	0.958	0.970	0.982
Search Frictions	1.162	1.100	1.073	1.048	1.106	1.065	1.046	1.029
All Parameters	0.767	0.837	0.833	0.833	0.772	0.847	0.849	0.857
Sample	0.788	0.781	0.770	0.801	0.778	0.768	0.757	0.769
				High School				
Productivity	0.890	0.910	0.894	0.884	0.762	0.820	0.829	0.844
Prejudiced	0.747	0.817	0.862	0.904	0.930	0.951	0.965	0.978
Search Frictions	1.129	1.084	1.063	1.044	1.047	1.030	1.022	1.015
All Parameters	0.737	0.787	0.797	0.812	0.786	0.835	0.842	0.854
Sample	0.743	0.738	0.766	0.770	0.784	0.779	0.783	0.777

Notes: Women/men ratio on average accepted wage computed over the entire distribution and over the top 50%, the top 75% and the top 10% of the distribution. All counterfactuals are generated taking into account equilibrium effects.

Table 4. Welfare Returns to Schooling - Ratios of Average Welfare Measures

	1995			2005		
	MA and PhD / College	MA and PhD / High School	College / High School	MA and PhD / College	MA and PhD / High School	College / High School
Overall						
All Parameters	1.191	1.599	1.342	1.171	1.552	1.326
Without Prejudice	1.110	1.445	1.302	1.211	1.604	1.325
Men						
All Parameters	1.151	1.546	1.343	1.177	1.591	1.352
Women						
All Parameters	1.225	1.638	1.338	1.169	1.529	1.308
Without Prejudice	1.052	1.324	1.259	1.251	1.630	1.303

Notes: The table presents ratio of the average welfare measures of the corresponding education levels. Welfare measures are computed using the estimated structural parameters. See main text for the complete definition.

Table 5. Policy Experiments - Relative Average Welfare Measures

	1995			2005		
	Benchmark Model	Equal Pay	Affirmative Action	Benchmark Model	Equal Pay	Affirmative Action
			Master and PhD			
Men	1.0000	0.9459	0.9861	1.0000	0.9149	0.9799
Women	0.7857	0.8088	0.8067	0.7659	0.7657	0.7841
Overall	0.9063	0.8860	0.9076	0.8787	0.8376	0.8785
			College			
Men	1.0000	0.9159	0.9834	1.0000	0.9567	0.9825
Women	0.7386	0.8340	0.7578	0.7714	0.8060	0.7890
Overall	0.8758	0.8770	0.8762	0.8833	0.8797	0.8837
			High School			
Men	1.0000	0.8927	0.9827	1.0000	0.9567	0.9828
Women	0.7417	0.8599	0.7599	0.7971	0.8337	0.8149
Overall	0.8766	0.8771	0.8763	0.9006	0.8965	0.9005

Notes: The table reports average welfare normalized with respect to men in the Benchmark Model. *Benchmark Model* is the model at the estimated parameters. *Equal Pay* means each employer must pay one wage at same productivity. *Affirmative Action* means employers receive a flow subsidy equal to 5% of the men average accepted wage when hiring a woman and the subsidy is financed by a lump-sum tax on all workers.

Appendix to Chapter 2

Appendix A: Data Sources and Definitions

A.1 Data used in the Descriptive Evidence section.

The data used in the descriptive evidence section are extracted from the *Annual Social and Economics Supplement* (ASES or March Supplement) and the *School Enrollment Supplement* (October Supplement) of the *Current Population Survey* (CPS). The first supplement contains data on family characteristics, household composition, marital status, education attainment, earnings, labor market status, work experience, job characteristics. The second focuses on school enrollment, college attendance, fields of study, major choices. Both supplements are conducted annually. We use the March yearly supplement from 1981 to 2011 and the October supplement in 2002. We use only the 2002 supplement because it is the only one reporting Field of Study choice.

Individual characteristics

The individual characteristics are obtained from the CPS questions on gender, race, age, marital status, and presence of kids under 18 years in the household. The year of birth, for the analysis by cohort, was inferred from the year of the survey and the age.

Education Level and Fields of Study

The education levels and fields of study choices are obtained from the set of questions related to education attainment in the March Supplement and with school enrolment in the October Supplement. We classify Education level in three groups according to the highest degree obtained: (1) Master and Doctorate degree, (2) College degree and (3) High School degree. It is important to mention that this classification is used from 1992 onward, because in that year there was a major change in the coding of the CPS data to classify education attainment. For the survey years before 1992, we simply define college graduates as persons with 14 years or more of education.

The gap in Figures 1 and 2 is calculated as a percentage difference with respect to men, that is $(x_W - x_M) / x_M$ where x is the percentage of college graduates.
The Field of Choice variable is only collected in the October supplement of 2002 and this is why Figures 4 and 5 reports the distribution only on 2002.

Labor Market Status

The labor market status (employment, unemployment and nonparticipation) is obtained by a set of questions organized by the CPS team in the monthly Labor Force recode variable which directly assigns each individual in the sample to employment, unemployment or not-in-the labor force status. Excluded from the universe are kids and individuals in the armed forces.

Earnings and Hours Worked

Hourly earnings are obtained either by using the value directly reported in the CPS survey or by computing the value dividing weekly earnings by the usual hours worked per week. Earnings are measured in real terms. We express Earnings in 2005 US dollars by deflating them by the Consumer Price Index for All Urban Consumers. For hours worked we use, as before, the usual hours worked per week directly reported in the survey.

Unemployment Durations

Unemployment durations are measured in months and they are obtained by rescaling the original weekly unemployment durations reported in the CPS.

Job Characteristics

The job characteristics are obtained from the set of questions related to full/part time job and occupational classification. The codes in this last variable are the 2002 NAICS equivalent. It is important to mention that all the descriptive analysis related with occupations is done from 2002 onward because in that year there was a major change in the coding used in the CPS to classify occupations.

A.2 Data used in the Impact of Employer's Prejudice section

The data used in the structural estimation of the search-matching-bargaining model with employers taste discrimination are extracted from the March Supplement of the CPS for 1995 and 2005. These years were chosen because they satisfy two criteria. First, these are years neither boom nor recession years, and therefore they seem appropriate to describe a model under the steady state assumption. Second, they are equally spaced over-time and far away enough to potentially describe different steady-states.

An important assumption in the model is ex-ante agents homogeneity. To obtain the estimation sample, we extract individuals homogenous sample with respect to the following characteristics: race (white), age (30 to 55 years old) and education (MA and PhD; College; High School).

The variables used in the estimation are: real hourly wages, unemployment duration in month, gender, education level, and labor market status. Wages are available only for individuals currently employed and unemployed duration only for individuals currently unemployed. As a results unemployment durations are not complete spells but on-going spells.

Table A1 presents number of observations and descriptive statistics, by education level and year, of the sample used in the Maximum Likelihood Estimation procedure.

Appendix B: The Formal Model

B.1 The Search-Matching-Bargaining Model

B.1.1 Theory

Environment

The model is developed in continuous time and it is populated by four types of agents infinitely lived: two types of workers – Men (M) and Women (W) - and two types of employers - Prejudiced (P) and Unprejudiced (N). The workers' type is defined by gender. The employers' type is defined by a difference in preferences: prejudiced employers receive a disutility flow (d) from hiring women.

There is random matching and there is not on-the-job search. Workers meet employers following a Poisson process with an instantaneous rate of arrival λ. Once an employer and a worker meet, they observe a match-specific productivity value (x) , which is drawn from an exogenous distribution denoted by the cdf G. Once a match is formed, it can be terminated following a Poisson process at an instantaneous rate η.

Workers' utility functions are linear in wages and no disutility from working is assumed. While unemployed, workers receive an instantaneous utility flow b. Time is discounted by a constant and common rate ρ. All the model's parameters are common knowledge.

Markets are fully segmented along gender-education-year cells. We denote gender with g, employer's type with t, year with y and education with e.

Value Functions

The value of employment for a worker of type g working at an employer of type t, producing x, in year y, with an education e is:

(A.1) $$\left(\rho + \eta_{gye}\right)V_{gye}\left[w_{gtye}(x)\right] = w_{gtye}(x) + \eta_{gye}U_{gye}$$

where w(x) denotes the wage, which is determined by Nash-bargaining.

The value of unemployment conditioning on type, education and year is:

(A.2) $$\rho\, \mathrm{U}_{\mathrm{gye}} = b_{\mathrm{gye}} + \lambda_{\mathrm{gye}} \begin{Bmatrix} \mathrm{p}_{\mathrm{ye}} \int \max\left[\mathrm{V}_{\mathrm{gye}}\left[\mathrm{w}_{\mathrm{gPye}}(\mathrm{x})\right] - \mathrm{U}_{\mathrm{gye}}, 0\right] \mathrm{dG}_{\mathrm{gye}}(\mathrm{x}) + \\ (1-\mathrm{p}_{\mathrm{ye}}) \int \max\left[\mathrm{V}_{\mathrm{gye}}\left[\mathrm{w}_{\mathrm{gNye}}(\mathrm{x})\right] - \mathrm{U}_{\mathrm{gye}}, 0\right] \mathrm{dG}_{\mathrm{gye}}(\mathrm{x}) \end{Bmatrix}$$

Wages

We assume the axiomatic Nash-Bargaining solution to the bargaining problem faced by workers and employers bargaining over the wage, given the match-specific productivity x and their types. The solution corresponds to maximizing the product of the worker's and employer's surpluses, weighted by their bargaining power α:

(A.3) $$\mathrm{w}_{\mathrm{gtye}}(\mathrm{x}) = \mathrm{argmax}_{\mathrm{w}} \left\{\mathrm{V}_{\mathrm{gye}}[\mathrm{w}] - \mathrm{U}_{\mathrm{gye}}\right\}^{\alpha} \left\{\frac{\mathrm{x} - \mathrm{d}_{\mathrm{ye}} \mathbf{1}_{\{\mathrm{g=W}\}} - \mathrm{w}}{(\rho + \eta_{\mathrm{gye}})}\right\}^{(1-\alpha)}$$

The wages schedules reported in the main text as equations (1)-(3) are simply specializations of equation (A.3).

Equilibrium

Optimal decision rules are characterized by the reservation value property, i.e. the match is mutually accepted if the match-specific productivity value is higher than an appropriate reservation value. The reservation value is determined as the value at which the agents are indifferent between accepting or rejecting the match. The reservation values are reported in equation (4)-(6) in the main text. By adding the optimal decision rules to the value functions, we obtain an equation that implicitly defines the only necessary equilibrium object, the value of unemployment U:

(A.4) $$\rho\, \mathrm{U}_{\mathrm{gye}} = b_{\mathrm{gye}} + \lambda_{\mathrm{gye}} \begin{Bmatrix} \mathrm{p}_{\mathrm{ye}} \int_{\rho\, \mathrm{U}_{\mathrm{gye}} + \mathrm{d}\mathbf{1}_{\{\mathrm{g=W}\}}}^{\infty} \left[\mathrm{x} - \mathrm{d}_{\mathrm{ye}} \mathbf{1}_{\{\mathrm{g=W}\}} - \rho\, \mathrm{U}_{\mathrm{gye}}\right] \mathrm{dG}_{\mathrm{gye}}(\mathrm{x}) + \\ (1-\mathrm{p}_{\mathrm{ye}}) \int_{\rho\, \mathrm{U}_{\mathrm{gye}}}^{\infty} \left[\mathrm{x} - \rho\, \mathrm{U}_{\mathrm{gye}}\right] \mathrm{dG}_{\mathrm{gye}}(\mathrm{x}) \end{Bmatrix}$$

We are now ready to propose the following:

Definition 1: In each market defined by year and education group, given the vector of parameters $\{\lambda_{\mathrm{gye}}, \eta_{\mathrm{gye}}, \rho, b_{\mathrm{gye}}, \alpha, \mathrm{d}_{\mathrm{ye}}, \mathrm{p}_{\mathrm{ye}}\}$ and the cdf of match-specific productivity values $\mathrm{G}_{\mathrm{gye}}(\mathrm{x})$, the *equilibrium* is defined by the vectors of values of unemployment U^*_{gye} that solves

equation (A.4), which in turn determine the reservation values characterizing the optimal decision rules.

B.1.2 Estimation Results

The estimates of the structural parameters are reported in Table A2. The estimation is performed jointly for 1995 and 2005 but separately by education level. The joint estimation is done to constrain the relative prejudiced preferences to be the same over the 10 years period. Following Flabbi (2010b), we assume that the proportion of prejudiced employers is quicker to adjust than preferences, therefore we leave the first one free to change over time while we constrain the second to be the same over the two periods. In estimation, we reparametrize the model and we estimate the disutility of prejudiced employers relative to the average male productivity. This ratio is the parameter *k* reported in Table A2.

The estimates of the structural parameters are in line with previous literature (Flabbi (2010a,b); Flinn (2006); Bowlus and Eckstein (2002)): women usually have higher arrival rates of offers and lower average productivity. The proportion of prejudiced employers and the relative disutility of discrimination are consistent with the result of Flabbi (2010b): the labor market for college graduates see a decrease in the proportion of prejudiced employers and a disutility value equal to about 30% of average male productivity. If High School also experience a decrease in the proportion of prejudiced employers, this is not the case on the sample of Master and PhD. However, the estimates of the prejudiced parameters are much more imprecise on this sample, probably due to the smaller sample size.

B.2 Policy Implications and Policy Experiments.

We report here the wage scheduled used in performing the policy experiments. In the *Equal Pay* policy, we impose the requirement that each employer has to pay the same wage to workers with identical productivity by interpreting the Nash-Bargaining wage schedules defined before as reduced form sharing rules. Defining the proportion of men in the population with m, the new wage equations become:

(A.5) $w_N(x) = \rho U + \alpha [x - \rho U]$

(A.6) $w_P(x) = \rho U + \alpha [x - (1-m)d + \rho U]$

where:

(A.7) $\rho U = m\rho U_M + (1-m)\rho U_W$

Notice that by definition the wage equations are not gender-specific. However, they remain employer-specific, as indicated by the subscript $_P$ and $_N$.

The *Affirmative Action* policy is defined as a flow subsidy received by each employer for each woman hired. The subsidy is paid by a lump-sum tax on workers. Defining with γ the subsidy and with *t* the endogenous tax rate necessary to finance it, the new wage equations become:

(A.8) $w_M(x,\gamma) = \rho U_M(\gamma) + t(\gamma) + \alpha [x - t(\gamma) - \rho U_M(\gamma)]$

(A.9) $w_{WN}(x,\gamma) = \rho U_W(\gamma) + t(\gamma) + \alpha [x + \gamma - t(\gamma) - \rho U_W(\gamma)]$

(A.10) $w_{WP}(x,\gamma) = \rho U_W(\gamma) + t(\gamma) + \alpha [x + \gamma - d - t(\gamma) - \rho U_W(\gamma)]$

The first equation states that men receive a wage that should compensate for the tax they pay but takes into account the reduced surplus implied by the tax. The second equation states that women working at unprejudiced employers receive the same tax effects but at the same time see the surplus increased by the subsidy γ. Finally, the third equation states that women working at prejudiced employers receive similar impacts from the presence of the tax and the subsidy but still share the cost of the disutility implied by prejudice. We denote the tax rate and the value of unemployment as a function of γ to emphasize that they are endogenous objects changing with the subsidy.

Table A1. Descriptive Statistics - Estimation Sample

	Master and PhD		College		High School	
	1995	2005	1995	2005	1995	2005
Observations						
N	711	861	3000	3756	7585	8416
N (Wages,Women)	306	437	1375	1855	3456	3912
N (Duration,Women)	5	9	50	63	167	210
N (Wages,Men)	394	403	1536	1771	3747	4047
N (Duration,Men)	6	12	39	67	215	247
Hourly Earnings in Dollars						
Overall						
Average	27.79	31.26	24.06	26.24	18.71	20.61
Std. Dev.	12.89	36.74	12.19	22.42	10.54	17.05
Women						
Average	24.55	26.43	20.87	22.94	15.86	18.12
Std. Dev.	10.10	12.04	10.34	15.25	8.96	12.51
Men						
Average	30.30	36.49	26.91	29.70	21.35	23.02
Std. Dev.	14.20	51.06	12.99	27.62	11.18	20.22
Diff(%)						
Average	-18.97	-27.56	-22.42	-22.76	-25.72	-21.29
Monthly Unemployment Duration						
Overall						
Average	6.73	3.87	4.67	4.87	4.49	4.40
Std. Dev.	7.07	4.19	5.43	5.56	5.35	5.16
Women						
Average	4.02	4.18	3.84	4.39	3.97	4.36
Std. Dev.	4.14	4.40	4.32	5.02	4.40	5.34
Men						
Average	9.00	3.63	5.73	5.32	4.89	4.43
Std. Dev.	8.52	4.21	6.50	6.03	5.96	5.02
Diff(%)						
Average	-55.38	14.99	-32.96	-17.54	-18.72	-1.42

Note: Data extracted from the Annual Social and Economic Supplement (March Supplement) of the CPS for the years 1995 and 2005. In each education label the sample includes individuals who are white and 30 to 55 years old.

Table A2. Maximum Likelihood Estimation Results - Structural Parameters

	Master and PhD		College		High School	
	1995	2005	1995	2005	1995	2005
λ_M	0.1119	0.2781	0.1759	0.1917	0.2076	0.2336
	(0.0457)	(0.0803)	(0.0282)	(0.0234)	(0.0142)	(0.0149)
λ_W	0.2906	0.3003	0.2740	0.2450	0.2799	0.2543
	(0.1355)	(0.1249)	(0.0391)	(0.0313)	(0.0231)	(0.0192)
η_M	0.0017	0.0082	0.0044	0.0071	0.0117	0.0138
	(0.0010)	(0.0034)	(0.0010)	(0.0012)	(0.0011)	(0.0013)
η_W	0.0041	0.0049	0.0095	0.0077	0.0122	0.0123
	(0.0026)	(0.0023)	(0.0019)	(0.0014)	(0.0013)	(0.0012)
μ_M	3.7536	3.8622	3.6333	3.6603	3.3640	3.3639
	(0.0290)	(0.0335)	(0.0156)	(0.0170)	(0.0109)	(0.0126)
σ_M	0.5587	0.6426	0.5931	0.6594	0.6290	0.7000
	(0.0220)	(0.0258)	(0.0119)	(0.0133)	(0.0085)	(0.0099)
μ_W	3.6618	3.6937	3.6501	3.5088	3.3353	3.1968
	(0.0374)	(0.0518)	(0.0760)	(0.0407)	(0.0890)	(0.0419)
σ_W	0.4289	0.4931	0.4407	0.5814	0.5291	0.6535
	(0.0246)	(0.0378)	(0.0300)	(0.0232)	(0.0350)	(0.0224)
p	0.1506	0.2117	0.6350	0.2616	0.9999	0.3497
	(0.0469)	(0.1079)	(0.1945)	(0.1125)	(0.0029)	(0.1661)
k	1.3796		0.3058		0.2445	
	(3.8471)		(0.0541)		(0.0826)	
w_M^*	10.8382	10.8800	8.9373	10.0000	7.3413	8.0000
w_W^*	8.9373	10.0713	6.7189	7.8040	5.6744	6.5000
lnL	-6142		-26014		-58552	
N	1,572		6,756		16,001	

Note: Asymptotic standard errors in parentheses. Joint estimation on all years by education level.

Chapter 3: The Impact of Family Issues and Career Development on Gender Gap: Evidence from Spain

Sara de la Rica[*25](University of the Basque Country, FEDEA and IZA)
Ainara González de San Román (University of the Basque Country)

3.1 Introduction

Among the most important determinants of a country's competitiveness is its human capital talent – the skills, education and productivity of its workers. Over time, this competitiveness depends very much on how countries use their human resources. The implications of a shrinking working-age population and the high educational level that women are accomplishing over recent decades[26] make it essential to consider women as fundamental pillars of the workforce. Governments may play an important role in creating the correct legal framework for improving women's education and their participation in the economy. Moreover, it is also imperative that firms invest increasingly a suitable working environment where men and women can adequately combine work and family.

Needless to say, women have made huge progress in the workplace, especially in the more industrialized countries. Goldin (2004) describes very effectively the evolution of women in the US labor market in the 20th century. Until the 1920s, working women were basically young and single and worked in factories or as domestic servants. From the 1930s onwards many more went to school and got jobs in offices. In the 1950s many married women entered the labor market and got jobs as secretaries, teachers or nurses. By the 1970s, their daughters saw their mothers working and took it for granted that they would also work. And since the end of the 1980s, women are overtaking men in graduating from college. This revolutionary process, to a greater or lesser degree, has been observed in most industrialized economies.

[25] Financial support from the Spanish Ministry of Science and Innovation under the project BES-2007-14507, and ECON2009-10818 is gratefully acknowledged. Address: Departamento de Fundamentos del Análisis Económico II, Universidad del País Vasco, Lehendakari Aguirre 83, 48015 Bilbao, Spain.
[26] See Goldin, Katz and Kuziemko (2006) for an exploratory analysis of the catch-up and reversal in the gender gap in university graduation of American college women.

However, after the educational period, male and female work careers often start to diverge. In addition to gender differences in the educational fields, which clearly condition their future performance in the labor market, and will be analyzed in detail in Chapter 4 and 5 of this report, family issues play a crucial role in understanding these differences. Women must combine employment with home responsibilities to a much larger extent than their male partners and this affects their decisions with respect to their labor supply which directly affects their human capital accumulation, and hence their labor performance in the future. There are many studies exploring the trade-offs between family and career among similarly educated men and women. For instance, Wood, Corcoran, and Courant (1993) find that 40% of the gender gap among American lawyers is explained by children. More recent works by Goldin and Katz (2010), and Bertrand, Goldin and Katz (2011) point to the differences in the intensive margin as the main determinant of the gender gap in male and female careers through the negative effect of children on women's hours worked. Finally, Molina and Montuenga (2009) confirm the existence of a wage penalty for Spanish working-women with children.

Our goal in this paper is to focus on highly educated men and women and try to explore the trade-offs between family and work career in a country like Spain, where changes in female behavior in the labor market have been rapid in the very recent years. We compare male and female behavior with respect to labor supply and performance along the life cycle for different birth cohorts to explore the connection between family and work over time. Further, we compare the impact of children on the labor supply decisions and labor market outcomes in two different time periods – 1994 and 2008.

The chapter is organized as follows. In the next section, we provide descriptive evidence on gender gaps in pre-labor markets and labor market outcomes. In section 3 we briefly describe the policy measures implemented in Spain, since mid-nineties, in an attempt to reconcile family and work. Section 4 relates the observed gender gaps in the Spanish Labor Market with family issues. Gender gaps in employment rates, part-time employment rates and incidence of women in top positions are presented along the life cycle of men and women for different birth cohorts. Section 5 is devoted to estimating and quantifying the impact of children on the career

development of men and women (employment rates and part-time employment rates) as well as on their performance in the labor market. Finally, section 6 concludes with a summary of our findings and some policy implications of the results.

3.2 Descriptive Evidence: Pre-Market and Labor Market Gender Gaps

3.2.1 Pre-market gender gaps

The individual decision on the quantity and quality of education constitutes the first important human capital investment decision which undoubtedly affects the individual future attachment and performance in the labor market. In most developed countries, the participation of women in higher education has increased enormously in the last few decades reaching and surpassing equality with men. There is still a large difference in choice of college majors.

3.2.1.1 Gender Gaps in University rates

Figure 1 depicts the trends in university graduation rates in Spain by gender (2000-2011). These rates are computed as the total number of students who completed their university studies every year divided by the total population at the typical graduation age[27]. Overall graduation rates are higher for women and the gap widens significantly over time. In 2000, university graduation rates for women were 2.5 times higher than those for men, but by 2011, the ratio has increased to 10 times.

This picture is not homogeneous if we disaggregate by field of study. Figure 2 presents the share of women among university graduates by field. Two main features emerge. First the shares are stable over time. Second, women have a spectacularly lower presence in Technical Studies[28] (around 28%) relative to others – Health, Social and Experimental Studies.

[27] The typical graduation age is the age at the end of the last academic year of the corresponding level and program when the degree is obtained. The typical age is based on the assumption of full-time attendance in the regular education system without grade repetition. For University graduation in Spain the OECD (see Education at Glance; 2002) establish this collective as the population aged 25-27.

[28] Technical studies group college majors as Computer Science, Engineering or Architecture.

3.2.1.2 Gender Gaps in Post-Graduate Studies: PhD Programs

Figure 3 presents the PhD graduation rates in Spain by gender in the last 10 years. As before, these are ratios of the total number of students who completed their PhD studies every year by the total population at the typical graduation age, which in this case is commonly set from 30 to 34 years. Given the selected population of reference, the percentage of PhD graduates appears to be small (almost 3% in 2000), but it increases along time reaching 4.5% by 2010. Male PhD graduation rates are somewhat higher than womens', but the distance has been shrinking with barely no difference by gender in 2010.

With respect to the incidence of women in PhD Programs by field of study, the first thing to note is that over-representation of women is not as clear as in university graduation rates. If we consider all fields of study together (total), women were around 44% of all students who finished PhD programs, and by 2010 it almost reaches parity. However, as before, enormous differences emerge when comparing PhD graduates in Technical Studies with respect to other majors, in particular, experimental sciences, where in the last decade the relative presence of women in PhD programs has increased from 46% in 2000 to 55% of total graduates ten years later.

3.2.2 Gender gaps in Labor Supply

In this sub-section, we provide evidence of gender differences in employment rates[29], hours of work, and part-time work.

3.2.2.1 Gender Gaps in Employment rates

Figure 5 presents employment rates for all men, all women, as well as college educated men and women (as a percentage of the working age population – 16 to 65 years). The first thing to note is that employment rates have increased for all women, independently of their

[29] Given that unemployment in Spain has been above 10% in the last two decades, labor force participation rates and employment rates differ to some extent. However, the general features are quite similar. To keep similar indicators as the rest of the chapters, we have used employment rates, instead of labor force participation rates, as one of the indicators of labor supply.

educational level. Still, employment rates for college women are much higher than for the rest. This education premium in employment rates was much larger for women than for men.

Following the precise definitions of the OECD (Employment Outlook 2010) we compute gender gaps in employment rates as the difference between the rates of males and females – the use of this definition makes them comparable to those of other countries. Changes in gender gaps in employment can be better appreciated in figure 6, where we depict overall and college gender gaps in employment rates. There has been a clear decrease in gender gaps in employment rates over the last 15 years, and the speed of this decrease is similar considering all men and women than when restricted to individuals with a college education. However, the level of the gap is remarkably different: for college individuals, the gap decreases from around 12% (in 1996) to around 2% (by 2010), whereas if we consider all individuals, the employment rate gap ranges from 27% (1996) to 12% (2010). Therefore, as of 2010, there is an almost negligible gap in employment rates for college individuals whereas this is still of a substantial size for the rest of the population.

Increases in women's labor supply are often negatively associated with fertility rates. However, the sharp decline in fertility rates in Spain took place in the mid-seventies. From the late eighties to the present day fertility rates have remained stable in Spain and – at around 1.3 births per mother on average- during the period under consideration (1996-2010). Hence the fall in employment gender gaps is unlikely to be due to changes in fertility rates, but rather to the changes in the way women are able to reconcile family and work.

3.2.2.2 Gender Gaps in Part-time employment rates:

Figure 7 and 8 present the relative incidence of part-time employment for total and college women overall and by age group respectively. These values are relative to men since they are computed as the number of employed women in part-time work divided by the total number of part-time workers. From figure 7 we observe that the gaps are increasing over time and that the relative incidence is clearly smaller among college workers – although in the last years, both

gaps tend to converge, mainly as a result of a more extensive use of part-time employment among college women.

Figure 8 shows that total women aged 30-45 is the group with the highest relative presence in part-time jobs (around 85%), closely followed by total women older than 45. The youngest groups of women, both overall and college-educated, shows similar trends, lying almost all the time about 20 pp. below the relatively older ones, what indicates that the gap in part-time rates with respect to young men is of smaller size. More interestingly, from the figure we also observe a clearly distinct behavior of college and total women for the two oldest categories of age. In particular, the relative incidence of college women aged 30-45 in part-time jobs is close to the corresponding total for women in the same age category but it increases relatively more over time. Something similar occurs with the oldest group of college women, with an increasing presence in part-time jobs over the period under consideration, so that the convergence observed in the gaps for total and college workers already observed in figure 7 is mainly driven by the increase in the relative presence of college women older than 30.

3.2.3 Gender Gaps in Labor Market Performance

Next we provide descriptive evidence on gender differences in the labor market performance. A first obvious indicator of such performance is earnings. A second one is the access to top labor market positions – managerial jobs.

3.2.3.1 Gender Gaps in Annual Earnings

Unfortunately, we do not have time series information with average wages for representative men and women. Instead, information on wages, and therefore gender wage gaps, can be extracted from different waves of Spanish data of the European Earnings Structure Survey – 1995, 2002 and 2006. In addition, for the years with no survey during the period 2004-2009 the Spanish Institute of Statistics updates information[30] on annual earnings by gender and age for

[30] These estimates of gross annual earnings are obtained by combining information from the Social Security files and the IRPF reports of the State Agency of Tax Administration, together with the variables occupation and working hours from the Quarterly Labor Cost Survey conducted by INE.

all workers, although information is not reported by different educational levels. Figure 9 provides this information for all workers. For workers younger than 30, the gap is smallest, around 15%, and it is quite stable over time. Gender gaps are highest for oldest men and women – older than 45, where gender gaps reach 35% in 1995 but decrease by 10 pp. as of 2009. For intermediate ages, 30-45, the gender gap is around 24% and it remains quite stable over time.

For college workers, we only have information for three time periods: 1995, 2002 and 2006 – the three releases of the European Earnings Survey for Spain. Figure 10 presents the gender gaps in annual earnings for different age groups. All we can observe is that the picture is not very different from the one presented above: gender gaps are smallest among younger workers, but over time there does not seem to change significantly.

3.2.3.2 Gender Gaps in Access To Managerial Jobs

Previous studies have documented a significant gender gap in top managerial jobs. This is particularly apparent for the group of college men and women[31]. This may be, in part, the result of women's preferences for more flexible jobs allowing better reconciliation between work and family life. In part it may also be due to greater difficulty for women to access these top jobs. Figure 11 presents this trend for college male and female workers (as a percentage of employment). The first interesting issue to highlight is that very few workers access these jobs. Top managerial jobs constituted less than 1% of female jobs in 1996 and 1.5 % by 2010. In comparison 2.3% of males were in managerial jobs in 1996, and 3% by 2010.

When we look at the relative incidence of women in managerial jobs by age groups, we can see – figure 12 – that the relative presence of women is highest among the youngest group, and moreover, it increases with time – from 26% in 1996 to 50% by 2010. Note however that the overall incidence of managers in this group is tiny. With respect to the other age groups, the

[31] For instance, Mateos, Gimeno and Escot (2011) find a really scarce presence of women on the boards of directors from the top 1000 Spanish companies in 2008 of around 6%.

presence of women is relatively low, although it has increased from 8-12% in 1996 to 15-25% by 2010 (depending on the age group).

3.3 Policy measures undertaken in Spain to reconcile family and work

During the 1990s and 2000s, Spain implemented specific policies to help workers reconcile work and family. The first and most important one was the Law of November, 1998, which concerned a change in the regulation of part-time work. The second one was the Gender Equity Law of 2007. In this section, we present the main features of each of these policies.

3.3.1 November 1998 Law: Regulation of Part-Time Work

Part-time working in Spain was much less widespread than in most other European Union countries. To encourage its use, improve its quality and increase the social welfare protection of part-time workers, major reforms were introduced in November 1998. The reform introduced certain improvements in the regulation of part-time working, including clarification of some rights of workers with part-time contracts, better social welfare protection, the extension and stabilization of relief contracts and greater stability in part-time contracts. The most important changes that the reform introduced were the following:

- The principle of non-discrimination between full-time and part-time workers was affirmed.
- Greater scope was given for partial early retirement and relief contracts, which had been used very little.
- Social Security contributions for additional hours are now included in the calculation of contribution periods for entitlement to Social Security benefits, and in the calculation of the regulatory basic amounts for those benefits.
- Job stability in part-time contracts is reinforced with incentives for indefinite contracts and the changeover from short-term part-time contracts to indefinite ones.

3.3.2 Other policies related to family

In June 1996, the European Union Directive on parental leave (96/34/EC) came into force. The Directive provides a set of minimum requirements all member states must meet; in particular, each employee (male or female) should be entitled to at least 3 months of parental leave after the birth or adoption of a child. We next summarize the basic regulation of family policies in Spain since the beginning of the nineties.

(i) Maternity Leave

- New parents in Spain have 16 paid weeks of maternity leave, two paid weeks of paternity leave, up to three years of child-care leave for each parent, and the right to work part-time until their child is 6 years old. Of all of these types of leave, only six weeks of maternity leave is mandatory.
- Spain guarantees 16 weeks of maternity leave. New mothers must take six weeks immediately after childbirth. They may take the remaining 10 weeks before or after giving birth, on either a part-time or full-time basis.
- During maternity leave, mothers receive Social Security payments equal to 100 percent of their usual salary. To qualify for leave and benefit, women must be currently employed, self-employed, or receiving unemployment benefits, and must have made Social Security contributions for at least 180 days in the seven years prior to taking leave, or 360 days in their entire working life.

(ii) Paternity Leave

Until 2007, Spain basically followed the European guidelines with regards to Paternity Leave. However, by 2007, Paternity Leave was enhanced with the New Gender Equity Law, which introduced the following measures:

- New fathers receive two different types of paternity leave. First, they receive two employer-paid days for having a birth in their family, or four days if travel is necessary for the childbirth.

- Second, fathers are entitled to paternity leave of 13 calendar days, paid by Social Security, with benefits and eligibility requirements equal to those that apply to new mothers, described above.
- Finally, as with maternity leave, fathers have the additional guarantees of returning to the same position at their employer and having two 30-minute feeding breaks during the first nine months of their child's life (although if they opt to use these by shortening their workday, they may only reduce it by 30 minutes per day).

(iii) Childcare Leave:

Since the Gender Equity Law of 2007, after maternity and paternity leave, parents may both access child-care leave: an unpaid leave of absence that can last until the child's third birthday. Both parents can use this leave simultaneously, unless they are working for the same employer (in which case the employer may object). Some interesting insights regarding childcare leave are described below.

- Child-care leave must be taken in one uninterrupted period, but may be taken on a full-time or part-time basis. To access child-care leave, parents must have at least one year's service at their current employer. If parents take more than one year of child-care leave, employers must only guarantee a similar post upon return.
- If parents wish to extend their leave beyond child-care leave, they may take an additional "voluntary leave of absence", which can last between four months and five years.
- Finally, all parents have the right to request part-time schedules until their child's eighth birthday. In doing so, parents may lower their work schedule to between one-half and seven-eighths of their usual schedule, but the employer is permitted to lower their salary accordingly.

3.4 Gender Gaps in Labor Supply and Performance in the Labor Market along the life cycle

At this point we focus on the evolution of gender gaps in labor supply and wages over the life cycle of women and men. This allows, on the first hand, to compare gender gaps at different ages, in particular, at pre-maternal and post-maternal ages. If these indicators show that

gender gaps enlarge between 30 to 40 years of age, this could be preliminary evidence of the importance of family issues in the professional careers of women versus those of their male counterparts. Second, by comparing gender gaps along the life cycle for different birth cohorts it is possible to assess the changes with respect to family and work balance of the Spanish women. We will focus mainly on the behavior of college-educated women.

3.4.1 Employment Rates along the life cycle

Making use of the *Spanish Labor Force Surveys* from 1986 to 2011, we construct employment rates as the ratio of individuals who are working at the time of the survey relative to the total population. For our analysis, we are restricted to nationals aged between 25 and 60 years. Moreover, in order to assess how these rates evolve for different groups in our sample, we also construct indicators for gender, education – college/non-college – and for different age intervals. Figure 13 presents life cycle employment rates for college-educated females (y-axis) for seven age groups (x-axis). Each line corresponds to a different cohort. i.e., to women born in different years[32], and therefore, for each of them it is possible to look at average changes in labor market indicators throughout the life cycle.

The figure shows that the employment rate of college educated women increases rapidly with age with no significant "dip" during the child-rearing years[33]. Moreover the more recent cohorts have increased participation at all ages, especially during the parenting years.

Figure 14 presents the relative employment rate of college women with respect to men. Traditionally the family has not played any role in the male work decisions. Therefore, if family is basically what makes the difference between male and female decisions in the labor market, the gender gaps may help to see the extent to which family plays a role in female labor supply decisions.

[32] Given the data availability we can construct 4 different cohorts, five years apart from each other, and thus cover from those born between 1960 and 1975. Moreover, since we have constructed 7 age categories that group 3 years each, the cohort born in 1960 refers to those individuals born between 1959 and 1961 and that are between 25 to 27 years of age in 1986. The same applies for the rest of cohorts.

[33] The same figure for total instead of college women is available from the authors upon request. It reveals lower rates of employment along the life cycle and a similar concave profile as that of the college only for the two youngest cohorts.

The figure reveals that at very young ages, college female employment rates are higher than those of males. However, by the age of 30 – the beginning of maternal age – there is an employment gender gap of 15% for the oldest cohorts and 7% for the youngest one. The gender gap along the life cycle has a U-shape, and by the 40s, the employment gap is reduced, although parity in employment rates is never reached at any older age.

3.4.2 Part-Time versus Full-Time Jobs along the life cycle

Family issues are very likely to affect not only the extensive concept of labor supply – work or not – but rather, the intensive margin. Given that childcare is very time intensive, it may entail a reduction in hours worked. Part-time versus full-time work indicates, in absence of precise measures of actual hours worked, a usual approximation of time dedicated to work. Figure 15 describes life cycle part-time rates (as a % of employment for college women)[34]. Two issues deserve attention. The first one concerns comparison across birth cohorts: part-time has been increasingly used by more recent cohorts at all ages. Secondly, we can also see a slightly increasing pattern with age, especially for the two youngest cohorts.

The relative incidence of women in part-time jobs offers a very similar pattern to the absolute one since the incidence for men is almost negligible – Figure 16. For the oldest cohorts, women represent on average around 80% of the total labor force in part-time jobs and this share does not seem to change much along the life cycle. However, for the youngest cohorts – especially for those born in 1975, the relative presence of women is gradually increasing with age, starting with 60% at ages 25-27 and reaching the peak at ages 37-39 with a relative presence of nearly 90%. This again confirms the fact that the youngest generation of women uses part-time employment to a greater extent especially at maternity ages what could be associated with a need of reconciling work and family. Moreover, additional evidence[35] suggests that whereas in

[34] We do not provide evidence of part-time rates for total women since the patterns are basically the same. We neither depict part-time rates for males because they are negligible at all ages and for all birth cohorts.

[35] The *Labor Force Survey* includes a question about the main reason for working part-time, so that we can compute the frequency with which college women answer that the reason is the family for each wave and for the different age groups.

1996 only 15% of college females aged 35-39 declared to work on a part-time basis for family reasons, by 2004 that fraction had doubled and today it reaches 45%. This increase in the last decade could be associated with the 1998 Law of part-time described in the previous section.

3.4.3 Gender Gaps in the access to Managerial Jobs along the Life-Cycle

In order to complete the descriptive information provided in this section, we describe the incidence of college-educated men and women in managerial jobs and look at possible occupational changes along their life cycles[36]. Figure 17 presents the incidence of female workers in managerial jobs along the life cycle and for different birth cohorts. As we saw before, the average percentage of women who are top managers is very low. However, a clear increase with age - more pronounced for the youngest cohorts than for the oldest ones, is also observed.

Finally, when comparing men and women along the life cycle, we can see – figure 18 that there is a clear decrease in the relative incidence of women in managerial jobs with age, and the pattern is similar across cohorts. Therefore, there seems to be a strong decreasing correlation between relative incidence of women in managerial jobs and age, which somewhat suggests a connection between family and access to managerial jobs.

3.5 The impact of Children on Gender Gaps in Labor Supply and in Labor Performance

In this section we quantify the impact of children in each of the labor market indicators presented above – employment rate, part-time employment rate, wages and probability of getting a managerial job. To do so, we use two micro-data sets: the first one is the first wave (1994) of the Spanish data from the European Household Panel, and the second corresponds to the 2008 wave of the Spanish data of the European Survey of Living Conditions (EUSILC). One of the advantages of using these panels is that their design is the same which makes them highly

[36] Unfortunately, the scarce data on wages for Spain does not allow us to construct gender wage gaps along the life cycle for different cohorts. Figure 10 already described gender wage gaps at different ages, but we cannot go further in the description of gender wage gaps for Spain along the life cycle.

comparable. Moreover, these panels are the only datasets containing individual information on demographics, labor market indicators and family issues for individuals in Spain. We use these particular waves so as to compare evidence from the mid-nineties with the more recent years– 2008 is the last available wave of the EUSILC[37]. In addition, it is interesting to take into account that those college individuals included in the 1994 sample, whose age is between 25 and 45 years of age belong to the two oldest birth cohorts analyzed in section 4 – those born during the 1960s, whereas those included in the 2008 sample are individuals who belong to the two latest cohorts – born in the 1970s.

3.5.1 Average Gender Gaps: 1994 versus 2008

Before getting into the estimation of the impact of children on the observed gender gaps, it is interesting to look at the raw average gaps for the two periods under consideration. Since our main goal is to detect the extent to which family issues (children) affect these gaps for the highly educated men and women, we have restricted individuals to be between 25-45 years of age, when parenting affects labor choices more strongly[38] and to those who have achieved the highest level of education (ISCED 5-7). Table 1 presents the average employment rates, part-time employment rates and incidence in managerial positions (following the corresponding ISCO classification of occupations for each period) for different type of college individuals, in particular, for total men and women, men and women without children and men and women with children[39]. In addition average gender gaps are reported for all individuals and separately depending on the family situation. Finally, the family gap is also included which compares the rates of women with and without children.

Employment Rates

Table 1 shows a substantial increase in the employment rate of all college-educated women between 1994 and 2008. However, this increase is particularly large among women with

[37] At the same time, the period under study in this section corresponds basically to the one presented in the descriptive part of the chapter, which facilitates somehow the interpretation of the results and helps relating them to the stylized facts already shown in the previous sections.

[38] This is the standard age interval considered by other studies which have analyzed the impact of children on gender issues, see Harkness and Waldfogel (2003), and Moluenga and Molina (2009) among others.

[39] With children refers to having at least one, regardless of the total number of children or their age.

children – around 17 percentage points, whereas for women without children the increase amounts to 11 percentage points. The employment rate of men, to the contrary, is very stable. With respect to gender gaps, these are particularly high among men and women with children (31% in 1994), although by 2008 the gap decreases to less than half – to 14.4%. It is also interesting to observe that the gender gap in the employment rate among men and women with no children disappears by 2008. Finally, comparing women with and without children – family gap – we can observe that it is substantial in 1994 (11%), but decreases to a great extent (5%) by 2008.

Part-time employment Rates

Part-time employment is essentially a female choice. The incidence of part-time employment among college male workers – fathers or not – is negligible. However, for women with children the increase in the use of part-time employment between 1994 and 2008 is substantial (from 11% to 19%), whereas for women with no children the contrary can be observed – a decrease from 15% to 11%. This first raw evidence clearly suggests that the use of part-time employment and family issues are strongly positively correlated, at least along the 15 years under study. Hence, the gender gap in part-time work increases among men and women with children and decreases among their counterparts without children, but the change is basically driven by female workers. Finally, the family gap increases considerably, given the more extensive use of part-time work by women with children in 2008 as compared to 1994. This might be in part related to the political measures already described in the third section of the chapter.

Access to Managerial Jobs

The first feature to highlight from the last two columns of Table 1 is that the incidence of college men and women aged 25-45 years working in managerial jobs is very low. For women, particularly if they have children, is null, and for women with no children amounts from 1 to 2 percent. For men, it reached 7% in 1994 but by 2008 it has decreased to half – 3%. Its incidence is higher for men with children than for childless men. Contrary to what we observed with the part-time employment indicator, in this case the gender gaps reflect the evolution of male incidence in these jobs, given that women are barely represented in these occupations,

particularly if they have children. This is consistent with the descriptive figures we depicted in Section 3.2, although the percentages there were somewhat higher for both groups, which could be associated with the selected age range of our sample.

In addition to these three indicators, Table 2 presents the descriptive evidence of the most obvious indicator of labor performance in the Job Market – *(Log) Hourly Wages*. When comparing log hourly wages of men without and with children, the first issue to highlight is that fathers earn more than childless men in the two periods, whereas the contrary happens for women. With respect to the gender wage gap, it is slightly higher in 2008 than in 1994 – it rises from 3.2% to 4.5%. In addition, comparing men and women without children, it can be seen that childless women earn more than their male counterparts in 1994 (6% more), but by 2008 the gender gap reverts sign, favoring men in around 5%. When comparing men and women with children, men earn between 5 and 6 percent more in the two periods under consideration. Finally, note that these gender gaps are much lower than those observed for the whole population of men and women – for all educational and age levels.

3.5.2 The impact of children on each of the Labor Market Indicators

After presenting the raw gender gaps in the labor market indicators, we proceed to isolate the impact of children from other potential determinants. Tables 3, 4 and 5 present the estimates from discrete choice (Probit) estimations of three outcomes: (*i*) the probability of working versus non-working, (*ii*) the probability of working part-time versus full-time, and (*iii*) probability of having a managerial job. In all estimations we use the pooled sample of college men and women of 25-45 years of age. The tables display the marginal effects of the variables related with having children, as well as the marginal effect of children for different groups: women with children versus women without children – marginal contribution to the family gap – and women versus men – marginal contribution to the gender gap.

Depending on the specification, controls for age, presence of a working couple, other family income and regional fixed effects are also included. Our final preferred family specification is captured by the variable *Any child*. Estimations are done separately for 1994 and for 2008 to

compare the impact of having children for the three indicators in the two periods. In the first column of each of the two periods – column 1 – the raw marginal effect of children is displayed – no additional controls included. Columns 2 and 3 report adjusted marginal effects of having children, therefore including the additional set of controls in the estimation. In column 2 we assume exogeneity of the children variable. However, given the potential endogeneity of the decision to have any child for any decision related to the labor market (work, work part-time or work in managerial jobs versus others) we must include a specification where we instrument the children variable, so that we avoid spurious causal interpretations of the impact of interest due to reverse causation in the estimation equation. Conceptually, it is not an easy task in this setting to think of an obvious valid instrument. Following Harkness and Waldfogel (2003) we use as an instrument the fraction of women who have any child by region, age and marital status (married or not). The correlation between this variable and the three labor market indicators is very significant (i.e. with working amounts to 0.65 in 1994 and 0.73 in 2008).

The computational method that we use to calculate the IV estimates for each column 3 is two-stage least-squares (2SLS). In the first stage, *Any Child* is regressed on all of the exogenous control variables by using a probit model, including also the instrument in the equation of interest. After obtaining the predicted values of that equation, in the second stage, the regression of interest is estimated as usual, except that the endogenous variable is replaced with the predicted values from the first stage. From the results of the first stage estimation, we can assure that the instrument is a good predictor of the endogenous variable[40] – *Any Child*. Unfortunately, the additional requirement for the instrument to be valid (it should be uncorrelated with the error term in each equation of interest) is not testable in exactly identified models as we have here. However, the fact that its correlation with each of the labor market indicators is close to null seems to suggest that the orthogonality condition is not violated.

[40] To verify that the chosen instrumental variable is not weak we compute the F-statistic against the null that the instrument is irrelevant in the first-stage regression which turns out to be much larger than 10, what ensures that the instrument is a good predictor of the endogenous children variable.

3.5.2.1 The Impact of Children on the Employment Decision

In 1994, having any child decreases the probability of working for women with respect to men by 22% – coefficient of *Female*Any child* in the third raw of the table. By 2008, this differential increases to 26%. However, the total gender gap, which is composed of the specific gender effect plus the children differential effect for women, decreases from 28% in 1994 to 24% by 2008. This is because the specific gender effect – coefficient of *Female*, decreases significantly from 1994 to 2008. In addition, when comparing mothers with childless women (family gap), having children decreases the probability of working by 10% in 1994, but interestingly, the impact of children by 2008 is even higher – reduces the probability of working by 17%.

In summary, the results suggest that in the mid-nineties employment levels of college men and women differ mainly as a result of the specific gender gap, rather than as a result of family reasons. This gap may arise as a result of social norms, which prevent that even college women have a high attachment to the labor market. This is reversed clearly by 2008, where the there is no specific gender gap in labor supply, but instead, family issues generate a clear gap in employment between college men and women.

3.5.2.2 The Impact of children on the Decision of working Part-Time

Table 4 presents the estimations on the impact of children on the part-time versus full-time choice for those who have decided to remain in the labor force. The most interesting result is that the impact of children in the use of part-time employment for women with respect to men (coefficient of female*any child) is not significant in 1994, but increases substantially by 2008 – to 13%. In addition and consistent with this (given the small incidence of part-time work among men, either with or without children), having children does not increase the use of part-time employment for women in 1994 (which is consistent with a lower part-time employment rates for women with children with respect with women without children observed in Table 1) but it does by 2008. If we consider the whole adjusted gender gap in part-time rates i.e., the specific gender plus the family effect, this amounts to 7% in 1994 (which comes basically from the

specific gender effect) but rises to 22% by 2008, where the specific female effect is added to the significant family effect.

In summary, it looks as if by 2008 some women with children who have decided not to quit the labor market decide to use the possibility of part-time work to reconcile family and work[41].

3.5.2.3 The impact of children on the access to Managerial Jobs

Table 5 shows that having children decreases the probability of accessing managerial jobs for women with respect to men. However, given the almost negligible incidence of females in these jobs, this is a result of men increasing their access to managerial jobs when they have children, and not as a result of a decrease in female access to them. In addition, and consistently with the observed decrease in the incidence of males in managerial jobs by 2008, the adjusted gender gap in the access to managerial jobs decreases from 10% to 3% when 1994 is compared with 2008. Again, this is because the specific gender gap decreases, but not because of a specific family effect. In summary, we do not find much with respect to the family impact in the access to managerial jobs, partly because the incidence of men, and particularly of women, with and without children in our sample in managerial jobs is negligible.

3.5.2.4 The impact of children on Gender Gaps in Hourly Wages

Table 6 presents log hourly wage estimations of college male and female workers to estimate the extent to which children affect the observed gaps. We present the impact of children in the average wage gap – column 1 – by including the gender indicator, the family variable and the interaction between the two. In column 2 we add the standard controls (age and its square, tenure, part-time employment, industry dummies and regional dummies) but without conditioning on occupation. Column 3 adds controls for occupation[42]. Finally, column 4

[41] We have not taken into account the possible selection of women into the labor market in this estimation because our aim is not to infer the decision of part-time for all college women, but rather the determinants of this decision for women who have already decided not to quit from the labor market. Our aim is to measure to what extent women with children who have decided not to quit from the labor market decide to go on part-time to facilitate family and work life. The same applies to the estimation of the incidence in managerial jobs, as well as the wage estimations.

[42] The aim for presenting these two specifications separately is to measure the extent to which part of the gender gap in wages is due to the fact that women segregate into low paid occupations.

presents the IV estimation of the latter (more general) specification. Estimations for 1994 and 2008 are presented separately.

The most interesting issue to highlight is that the impact of the family differs to a great extent when comparing 1994 with respect to 2008. In 1994, there is no significant gender wage gap between childless women and men (coefficient of *Female*). However, children decrease female wages by around 13% when compared to their male counterparts. By 2008, results are just the opposite: the adjusted gender wage gap between childless women and men is significant (9% without conditioning for occupation and 6% conditioning for it) but the impact of children is negligible. This result is very consistent with what we saw in the estimations of the employment probability: as we argued before, in 1994 there is high selection in labor market entrance of college women, which results in average higher wages for them and hence in no gender wage gaps. Hence the whole wage gap is due to the impact of children. By 2008, he wage gap is basically a gender specific gap, and it is not driven by family issues, but to other potential explanations. Among these, we might tentatively mention as possible segregation of women in low paid jobs (low paid firms or even lower paid categories within the firm), discrimination, differences in specific skills, etc.

It is also interesting to see the different effect of part-time on wages when 1994 is compared with 2008. In 1994, part-time employment's impact is either positive or null, depending on whether we condition for occupation or not. However, by 2008, its impact is clearly negative, which is a more usual result. Given the low incidence of part-time work in 1994, it is likely that workers who made use of it were a very particular sample of workers. Finally, when comparing the impact of children on wages with and without controlling for occupation (2-digit disaggregation of the corresponding ISCO classification for each year) – columns 2 and 3 – we find very small differences. This means that gender gaps in wages of college individuals are not driven by segregation of women with children in low paid (broad) occupations to reconcile family and work. Finally, estimations from IV methods are qualitatively identical to those of OLS, although in quantitative terms they are somewhat bigger.

3.5.3 Contribution of the Family to Average Gender Gaps

To finish the section of the impact of the family on gender gaps, we aim to quantify the absolute and relative contribution of children to each of the raw average gaps, based on the estimations presented above. This way of quantifying the impact of interest allows us to measure by how much differences in the labor market indicators between men and women can be explained by differences in their response associated with having children. In the case of wages the contributions are computed by simply multiplying the estimated coefficients of having children by the mean frequency of having children in the sample for men and women respectively. However, given that for employment rates, part-time employment rates and incidence of managerial jobs the estimated models are non-linear, the absolute contribution of children to the average gap is obtained following Yun (2004), who uses an extension of the Oaxaca Decomposition Method to account for non-linear estimations[43]. In practice, to compute the absolute contributions we just need to multiply the estimated coefficient of interest by a proper weight, which in this context is the standard normal probability density function evaluated at the mean predicted characteristics. The relative contribution is the ratio between the absolute contribution and the raw corresponding average gap. The computed contributions are presented in the last raw of Tables 3-6 both for the most general specification that includes controls and for the IV estimates when it proceeds.

The first thing we observe is that the relative contribution of the family to the employment gender gap was very small in 1994 but increases to 56% by 2008. However, note that average gender gaps in employment are much higher in 1994 than by 2008. Second, for those women who decide to stay in the labor market, having children contributes to explaining 4% of the gender gap in part-time work use in 1994, but this increases to 9% by 2008. Although the contribution of the family is not very remarkable in the two years, it is interesting to observe that in 15 years it has doubled, whereas the gap has remained quite stable. This means that

[43] There are alternative ways of decomposing probit functions, see Even and Macpherson (1990), Nielsen (1998), but the one proposed by Yun (2004) is the most suitable in our setting since the emphasis of our approach relies on decomposing the differences into coefficients effects. However, for comparability purposes, we have also computed the contribution of the family to these three indicators assuming linearity by taking the coefficients from the estimation of a Linear Probability Model instead of a probit.

college women with children are increasingly using part-time jobs to remain in the labor market instead of quitting for childcare issues. Third, family does not contribute to the differential impact of male and female workers in managerial jobs[44].

Finally, whereas average gender wage gaps have doubled in these 15 years, the contribution of the family to it has decreased substantially: having children contributed to explaining 85-100% of the average gender gap in wages in 1994, whereas by 2008, the family does not explain it at all. Therefore, and as we already mentioned, the gender gap in average wages in the mid nineties is mainly a result of family issues, whereas by 2008, the average gap, which is double, has its roots on an specific gender gap not related with family issues, but with other potential explanations, such as discrimination, segregation into different jobs, differences in the acquired skills, etc.

3.6 Summary, conclusions and Policy Implications

In this paper, we explore the trade-offs between family issues and work career development for highly educated men and women in Spain. Although women reach a higher educational attainment in terms of graduation rates and on average grades with respect to men, male and female work careers often diverge substantially. Part of it may be due to a different choice in the fields of study, as analysed in chapters 4 and 5 of the present book. However, family issues are undoubtedly an important additional potential determinant of this gender divergence. Women must combine employment with home responsibilities to a much larger extent than their male partners and this affects their decisions with regards to their labor supply and hence to their labor performance in the future.

We present evidence of gender gaps in labor supply – employment rates and part-time employment rates – and in labor performance – wages and incidence in managerial jobs along

[44] By comparing these contributions with those coming from the Linear Probability Model estimations the results with the latter are qualitatively unchanged but the magnitude is always much higher (about 3 to 5 times more) what suggests that the contributions following Yun's methodology represent a lower bound, which was expected given that they are Taylor approximations and hence an error is associated to them.

the life cycle and for different cohorts – those born in 1960, 1965, 1970 and 1975. This allows, on the first hand, to compare gender gaps at different ages, in particular, at pre-maternal and post-maternal ages and observe whether these indicators suggest that gender gaps emerge during parenting years – from 30 to 40. Second, by comparing gender gaps along the life cycle for different birth cohorts we can assess changes in the Spanish women (in particular, college women) with regards of their way to combine family and work. In addition to this descriptive evidence we use two micro-datasets to account for the impact of the family, more precisely, we estimate the effect of having any child on the observed gender gaps. The first one is the first wave of the European Household Panel for Spain – 1994, and the other one is the 2008 wave of the European Survey of Living Conditions. Both of them share the design and hence are highly comparable. The use of these two datasets is very convenient because, given that we consider college men and women of 25-45 years of age, the 1994 sample captures men and women born between 1950 and 1970, which on average reflects the behavior exhibited by our two oldest cohorts – 1960s and 1965s. On the other hand, the sample of 2008 captures college men and women who were born between mid-sixties and early eighties and hence basically reflects the behavior of the two youngest ones.

Our findings are consistent with the following insights: first, with respect to gender gaps in the extensive margin of labor supply – employment rates – we observe a very different pattern when comparing the mid-nineties (1994) to recent years (2008). In the former, gender gaps in employment rates are quite substantial even among childless men and women – 11%, which increase to 31% when comparing mothers with fathers. Children account for one tenth of this substantial gap, which tells us that besides children there were several other causes of inequality. However, by 2008, the pattern of gender gaps in employment rates had changed substantially: there was basically no gap between childless men and women, although this rose to 14% between fathers and mothers. Furthermore, children accounted for more than half of this gap. In summary, in the mid nineties employment levels of college men and women differ mainly as a result of the specific gender gap, rather than as a result of family reasons. This gap may arise as a result of social norms, which prevent that even college women have a high

attachment to the labor market. This is reversed clearly by 2008, where the there is no specific gender gap in labor supply, but instead, family issues generate a clear gap in employment between college men and women.

Second, for those men and women who stay in the labor market – workers – we look at gender gaps in labor supply in the intensive margin – number of hours – proxied by part-time employment rates. The data reveals, consistently with what it has been documented for other countries, that the use of part-time employment is basically a female choice, the use of part-time employment among men is negligible. At this stage, we see interesting differences in the behavior of women with respect to the use of part-time work in the mid-nineties as compared with the more recent years: on the one hand, the use of part-time employment was higher among childless women as compared to mothers (15% versus 11%), which suggests a very small, if any, connection between family issues and part-time work at that time. This changed by 2008, when the use of part-time employment of women with no children was 10% but it was 18% for mothers. Additionally, children accounted for 4% of the gap in 1994 and 9% of it by 2008. The fact that children accounted only for 9% of the gap even in 2008 suggests that although children is an important determinant for the choice of part-time employment versus full time work, there may be other determinants, among others demand restrictions, which lead some women work on part-time basis. Indeed, the percentage of women who work part-time and report that the main reason for it is children amounts to 40% in 2008, which suggests that under the decision of part-time employment, although children is clearly an important determinant, there may be others, which may come from demand side restrictions.

Third, with respect to the performance in the labor market, we compare hourly wages of those college men and women who stay in the labor market. The first interesting issue we observe is that in the mid-nineties, childless women earned on average 6% more than men, which suggests that these women were a highly selected sample of college females. However, when we compare fathers with mothers, the gap is 7% in favor of men. Furthermore, children account for almost all the gap. By 2008, the pattern is quite different: the gender gap between

childless men and women is of the same magnitude than that among fathers and mothers (5% in favor of men) and children do not contribute to explain the gap because the whole of it is explained before family issues play a role. In addition, selection of women into low paid occupations does not help to explain the source of the gap. In sum, the gender gap in average wages in the mid nineties is mainly a result of family issues, whereas by 2008, the average gap, which is double, has its roots on a specific gender gap not related with family issues, but with other potential explanations, such as discrimination, segregation into different jobs, differences in the acquired skills, etc.

Finally, although the descriptive evidence of the life cycle profiles of men and women suggest that female decrease her participation in top managerial jobs with age relative to men, the samples of 1994 and 2008 do not help us to account for the impact of family in this performance gap, given that the incidence of our samples of men and particularly of women into these jobs is negligible.

What are the policy implications of these findings?

As our results reveal, by 2008 childless college women behave very similarly to men in the labor market in terms of their labor supply decisions. Still, men on entrance earn approximately 5% more than their female counterparts, which might be due to differences in the chosen fields of study undertaken by men and women. If this different behavior is due to differences in preferences, there is not much to be done from a policy perspective. Rather, if females do not engage in these studies because they perceive future discrimination in those occupations as some studies suggest (see The National Academy Press (2007)), then there is scope for improvement with respect to trying to encourage women to engage in more technical studies at university degree and PhD levels. We analyzed some of these issues in the policy experiment of Chapter 2.

A second policy implication concerns the use of part-time work for mothers. Although some of them use it to combine family and work, it looks like the use of part-time employment for many mothers is not voluntary. It would be very desirable that the decision of using part-time work

would be voluntary for both men and women. However, part-time work should not be the only way to combine family and work for mothers. Going a little bit further, it would be desirable that in addition to the possibility of part-time for those women for whom working part-time may be a first best choice, firms would develop other possible mechanisms to combine family and work and hence allow mothers to participate in the labor market as much as fathers in terms of their intensive margin. This can be achieved by developing flexible timing work schedules and *home-work (tele-trabajo)* not only for mothers, but also for fathers. Only if these mechanisms are widely available for mothers and fathers, we will achieve gender equality and hence allow women with children to achieve work careers similar to their male counterparts if they wish to do so. As we already said in the introduction, a shrinking working-age population and the high educational level that women are accomplishing in the last few decades, make it essential to consider women as fundamental pillars of the workforce. If they are provided with adequate mechanisms to combine family and work society will surely take more advantage of the full application of these highly educated females in the workforce throughout their whole working lives.

References for Chapter 3

Azmat, Ghazala, Güell, Maia, Manning, Alan (2006), "Gender Gaps in Unemployment Rates in OECD countries", *Journal of Labor Economics,* vol 24 no.1: 1-37.

Bertrand, Marianne, Goldin, Claudia, Katz, Lawrence F. (2010), "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors", *American Economic Journal: Applied Economics,* 228–255.

De la Rica, Sara, Anghel, Brindusa, Alloza, Mario (2011), "Diferencias de Género en la Formación Científica Evidencia Empírica", *Situación de las Mujeres en la Ciencia Española* – Libro Blanco (Unidad de Mujeres y Ciencia) Capítulo 2.

División General de Renovación Pedagógica (1992), "Las Desigualdades en la Educación en España", *Ministerio de Educación y Ciencia*.

Escot, Lorenzo, Gimeno, Ricardo, Mateos de Cabo, Ruth (2011), "Disentangling Discrimination on Spanish Boards of Directors", *Corporate Governance: An International Review*, 2011, 19(1): 77–95.

Even, William E., Macpherson, David A. (1990), "Plant Size and the Decline of Unionism", *Economics Letter*, 32, 393-398.

Fuchs, Victor (1988), "Women's quest for economic equality", *Cambridge: Harvard University Press.*

Goldin, Claudia (2005), "From the Valley to the Summit: The Quiet Revolution that Transformed Women's Work", *Regional Review*, 14:3, pp. 5-12.

Goldin, Claudia, Katz, Lawrence F., Kuziemko, Ilyana (2006), "The Homecoming of American College Women: The Reversal of the College Gender Gap", *Journal of Economic Perspectives,* Volume 20, No. 4 pp. 133-156.

Goldin, Claudia (2006), "The 'Quiet Revolution' that Transformed Women's Employment, Education and Family", *American Economic Review*, 96 (2): 1-21.

Goldin, Claudia, Katz, Lawrence F. (2008), "Transitions: Career and Family Life Cycles of the Educational Elite", *American Economic Review*, 98:2, 363–369.

Harkness, Susan, Waldfogel, Jane (2003), "The Family Gap in Pay: Evidence from Seven Industrialized Countries", *Research in Labor Economics,* 22, 369-414.

Herr, Jane, Wolfram, Catherine **(**2009), "Opt-Out Rates at Motherhood across High-Education Career Paths: Selection versus Work Environment", *National Bureau of Economic Research,* Working Paper 14717.

Instituto de la Mujer (2009), "Mujeres en Cifras", *Ministerio de Igualdad*.

Joshi, Heather (1991), "Sex and motherhood as handicaps in the labor market", *Women's Issues in Social Policy*, London: Routledge.

Joshi, Heather, Newell, Marie-Louise (1989), “Pay Differentials and Parenthood: Analysis of Men and Women Born in 1946”, Coventry: *University of Warwick Institute for Employment Research.*

Molina, José Alberto, Montuenga, Víctor M. (2009), “The Motherhood Wage Penalty in Spain”., *Journal of Family Economic Issues,* 30:237–251.

Neumark, David, Korenman, Sanders (1994), “Sources of bias in women’s wage equations: Results using sibling data”, *Journal of Human Resources*, *29*, 379–405.

Nielsen, Helena S. (1998), “Discrimination and detailed decomposition in a logit model”, *Economics Letters*, 61, 115-120.

OECD *Employment Outlook* (2002) Chapter 2 – ISBN 92-64-19778-8.

The National Academy Press (2007), “Beyond Bias and Barriers: Fulfilling the potential of women in academic science and engineering”, *Washington, D.C.*

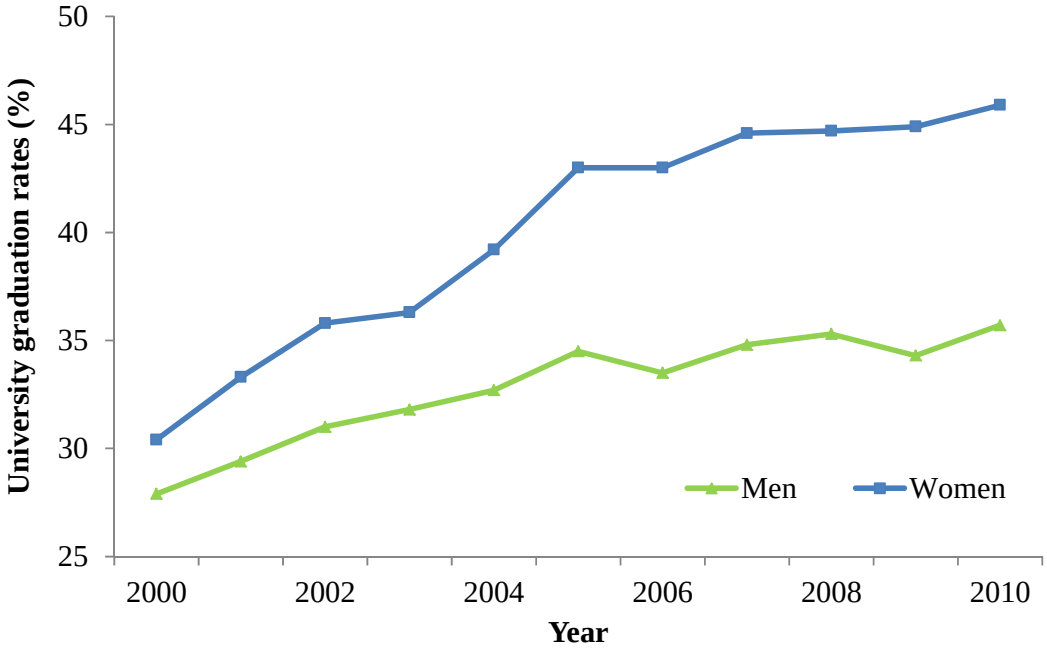
Waldfogel, Jane (1998), “Understanding the ‘Family Gap’ in Pay for Women with Children”, *Journal of Economic Perspectives*, vol. 12: 137-156.

Wood, Robert G., Corcoran, Mary E., Courant, Paul N. **(**1993), “Pay Differences among the Highly Paid: The Male-Female Earnings Gap in Lawyers’ Salaries”, *Journal of Labor Economics*, 11(3): 417–41.

Yun, Myeon-Su (2004), “Decomposing Differences in the First Moment”, *Economic Letters*, Vol. 82: 275-280.

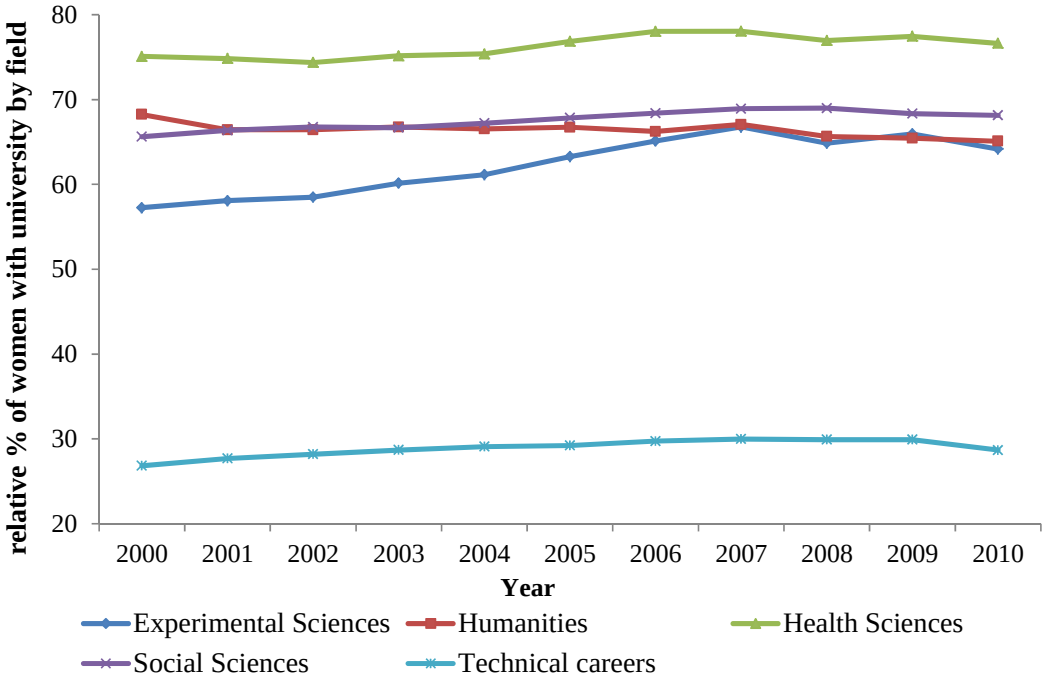
Yun, Myeon-Su (2000), “Decomposing Analysis for a Binary Choice Model”, *Working Paper, No. 2000-01*, Department of Economics, Rutgers University.

Figure 1 – University Graduation Rates in Spain by Gender (2000-2011)



Source: Eurostat

Figure 2 – Relative Presence of Women by College Major

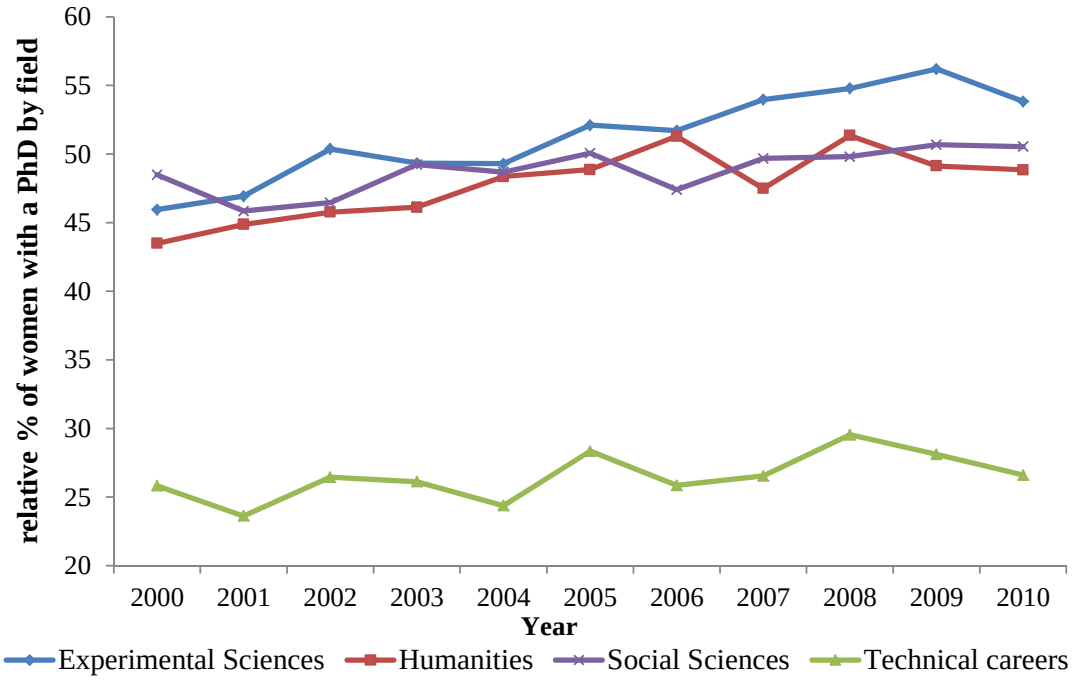


Source: Estadística de la Enseñanza Univeristaria en España (INE)

Figure 3 – PhD Graduation Rates in Spain by Gender (2000-2010)

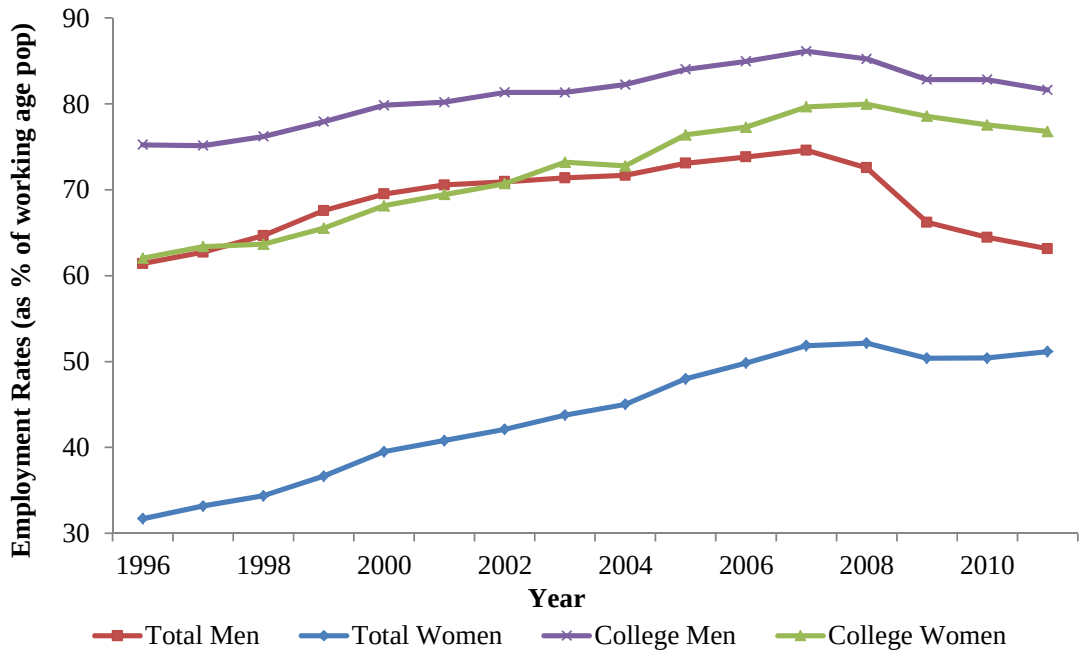
Source: Estadística de la Enseñanza Superior (INE)

Figure 4 – Relative Incidence of Women with PhD by Field of Study



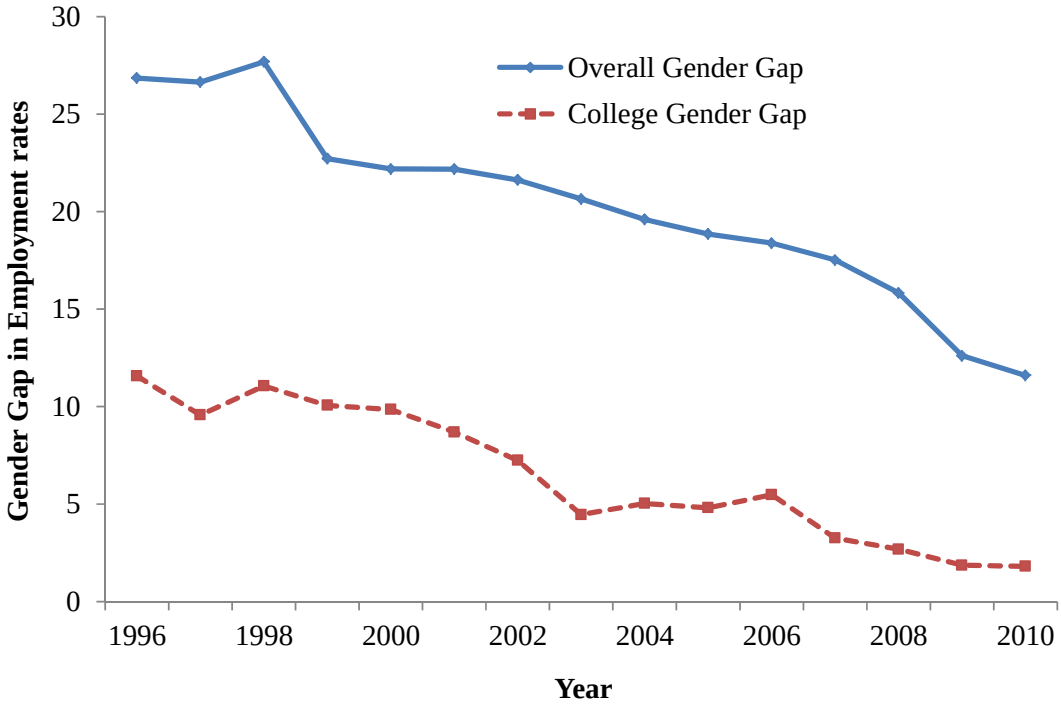
Source: Estadística de la Enseñanza Superior (INE)

Figure 5 – Employment Rates (%) – Overall and College Men and Women



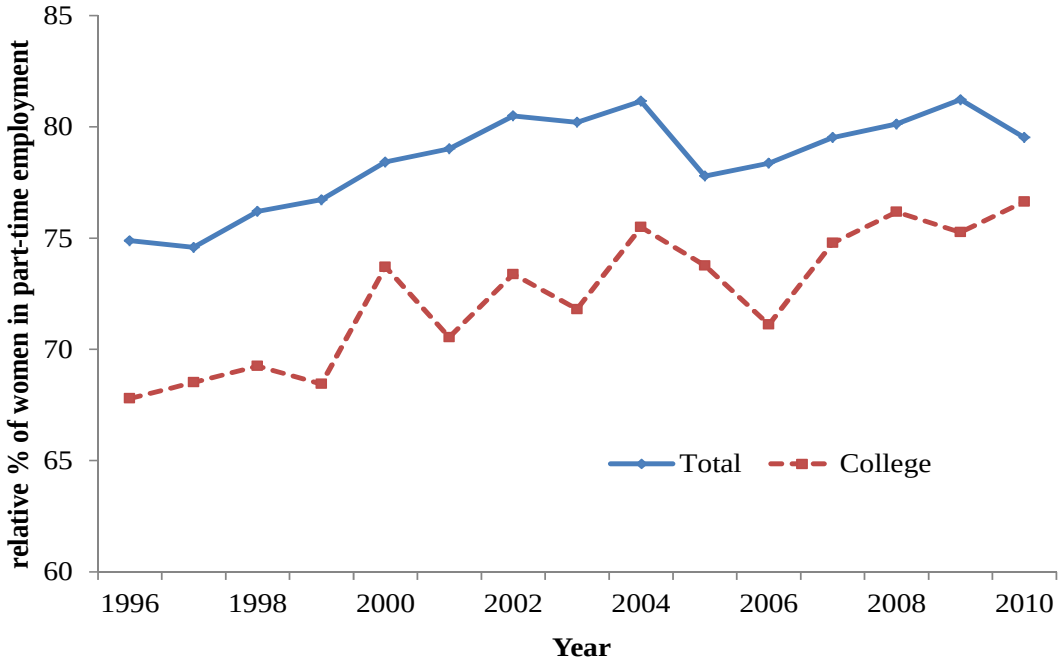
Source: Encuesta de Población Activa (INE) 1996-2011

Figure 6 – Gender Gaps in Employment Rates (1996-2011) – Overall and College Men and Women



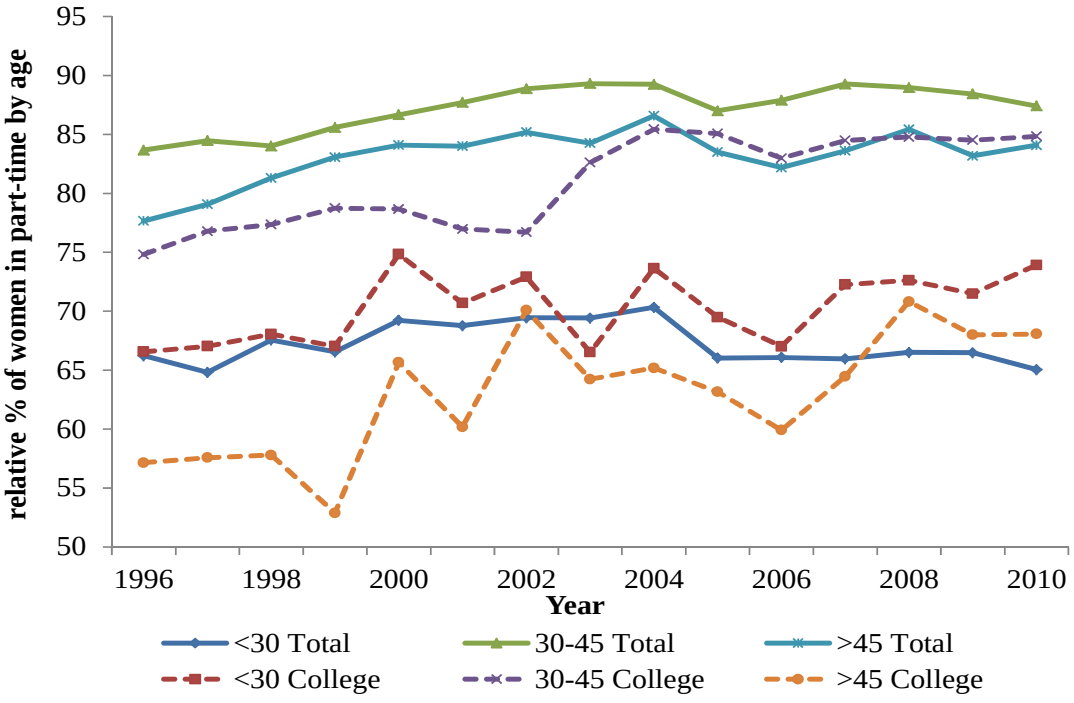
Source: Encuesta de Población Activa (INE) 1996-2011

Figure 7 – Relative Incidence of Women in Part-Time Employment (1996-2011)
Overall and College Females



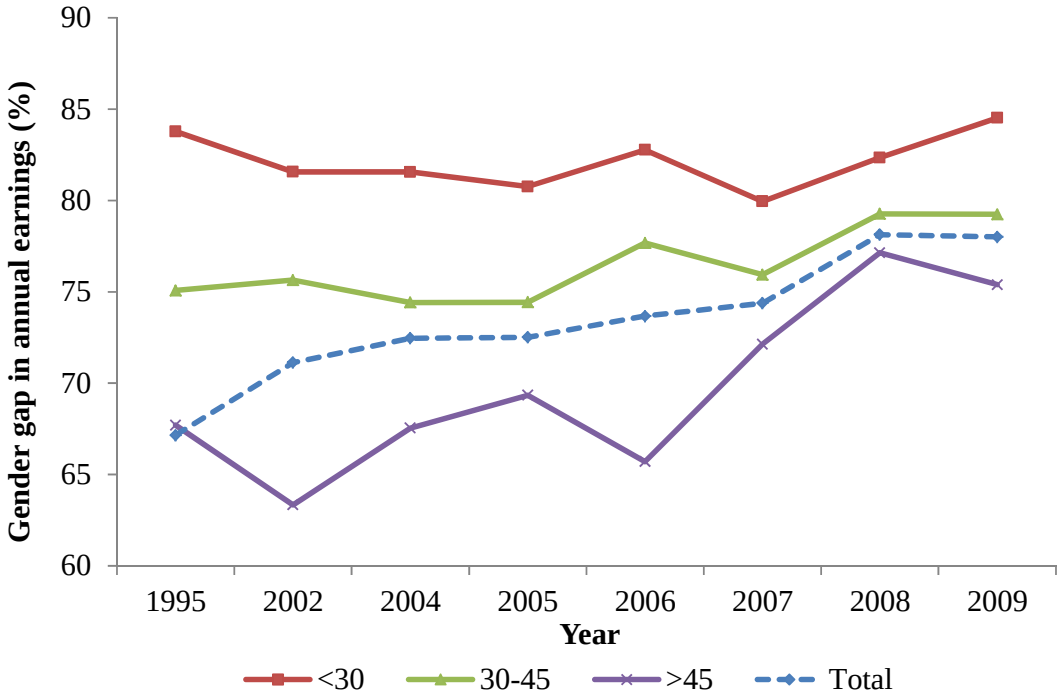
Source: Encuesta de Población Activa (INE) 1996-2011

Figure 8 – Relative Incidence of Females in Part-Time Jobs by Age



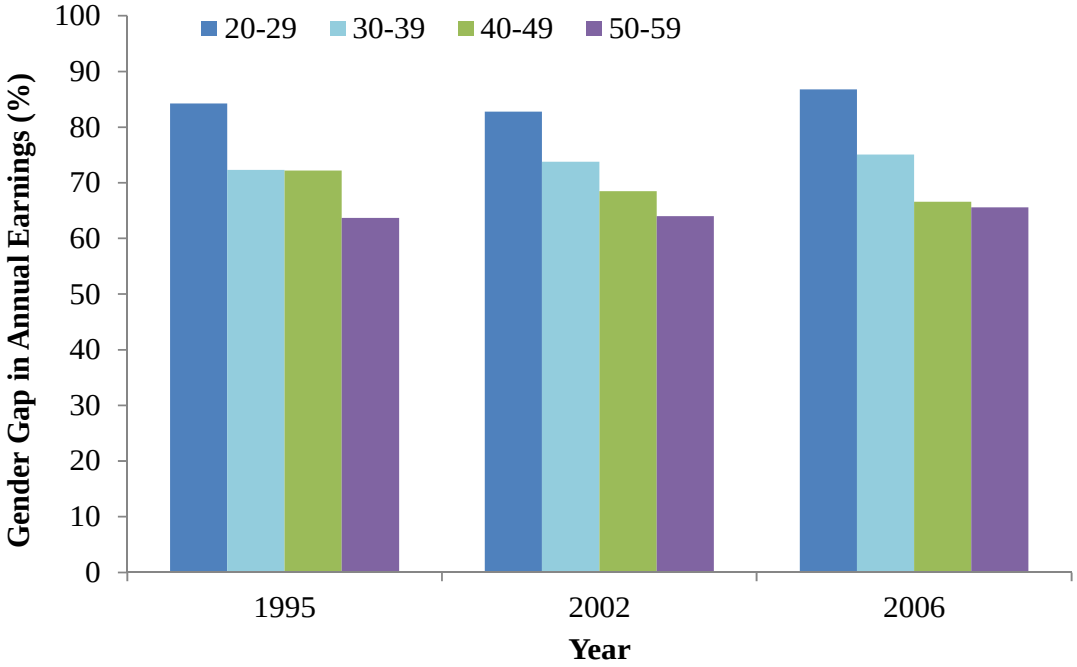
Source: Encuesta de Población Activa (INE) 1996-2011

Figure 9 – Gender Gap in Annual Earnings Total Women/Men (1995-2009)



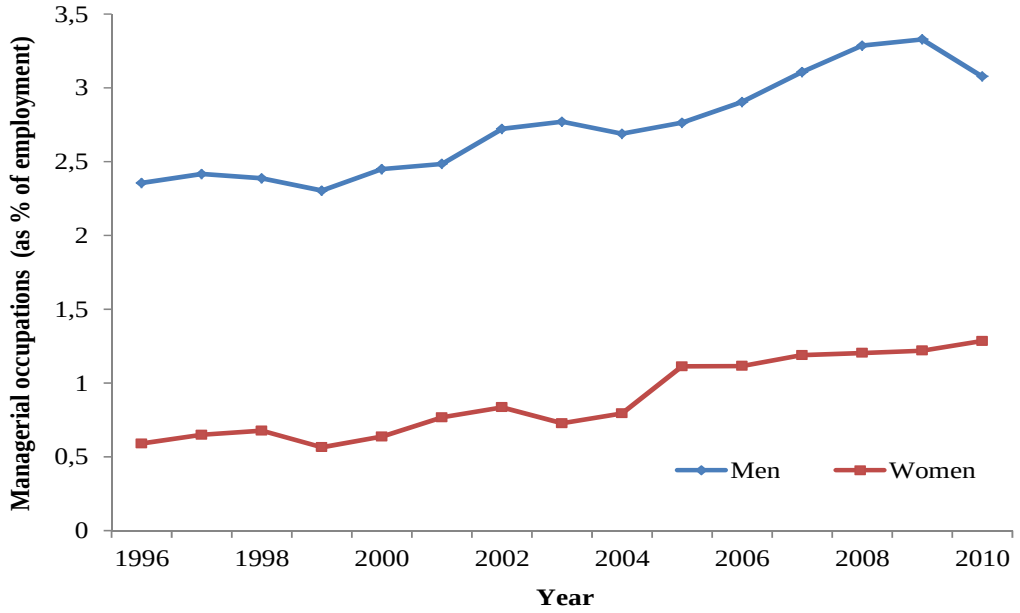
Source: Encuesta de Estructura Salarial (EES) 1995-2009

Figure 10 – Gender Gap in Annual Earnings by Age – College Workers



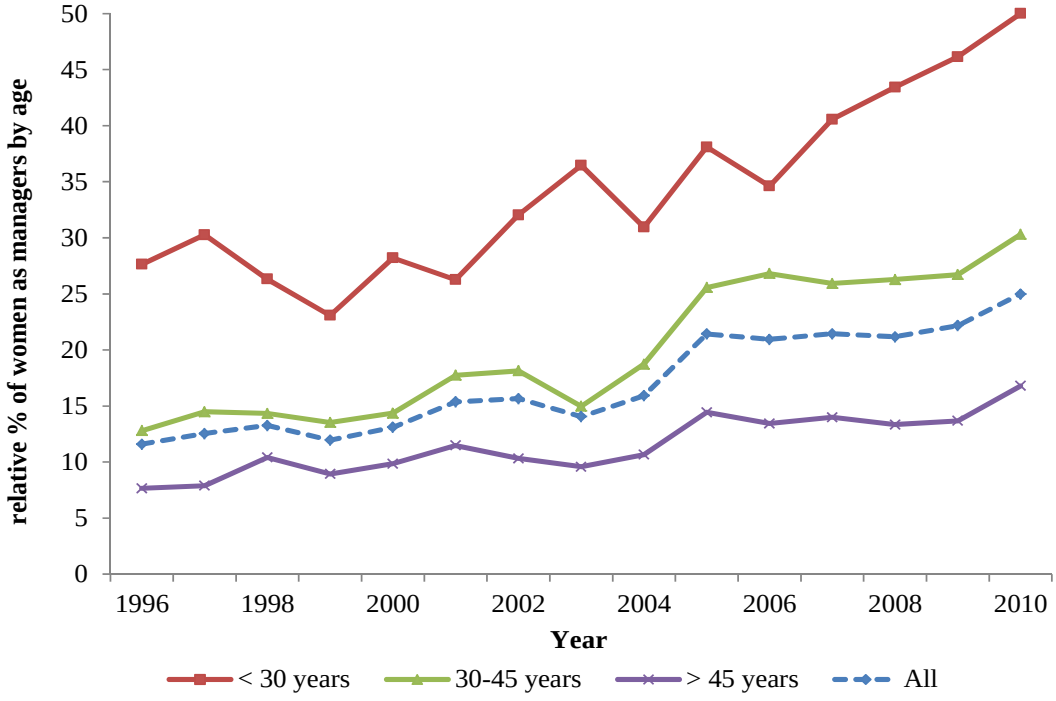
Source: Encuesta de Estructura Salarial (EES) 1995-2006

Figure 11 – Incidence of Workers in Managerial Occupations (1996-2011) – College Male and Female Workers



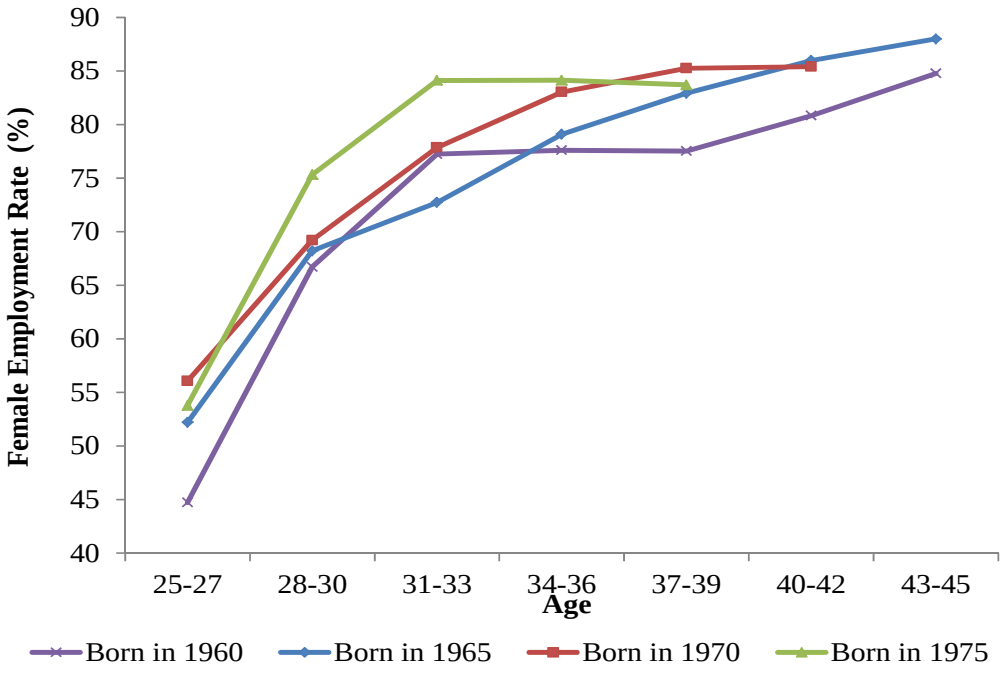
Source: Encuesta de Población Activa (INE) 1996-2011

Figure 12 – Relative Incidence of Women in Managerial Occupations by Age - College Workers



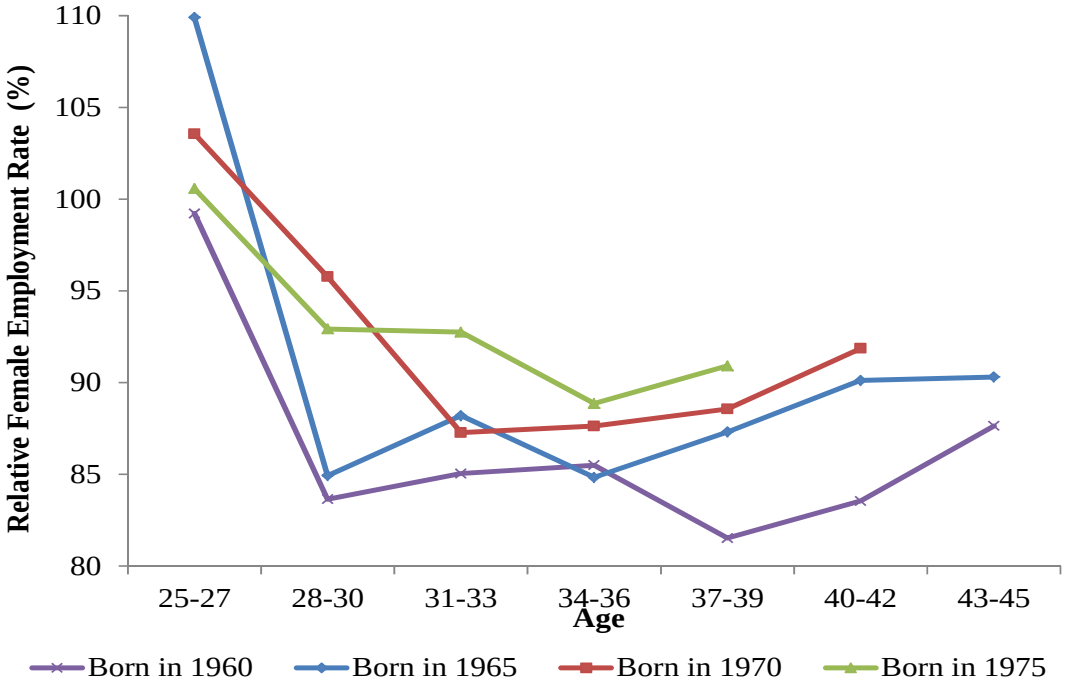
Source: Encuesta de Población Activa (INE)

Figure 13 – Life Cycle Employment Rates (%) – College Women



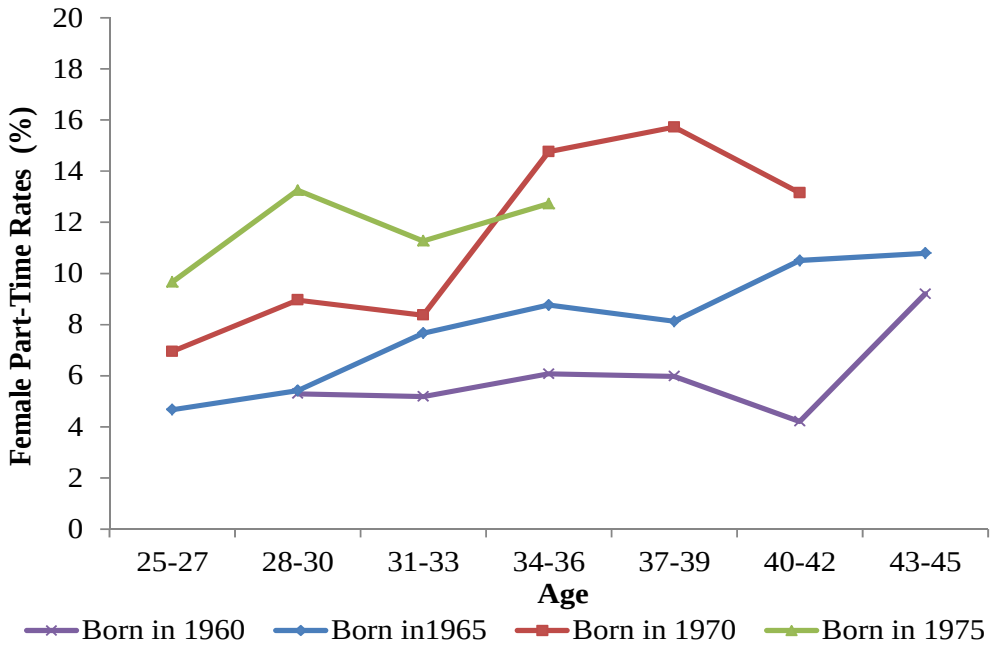
Source: Encuesta de Población Activa (INE) 1986-2011

Figure 14 – Life Cycle Relative Female Employment Rates (%) – College Women/Men



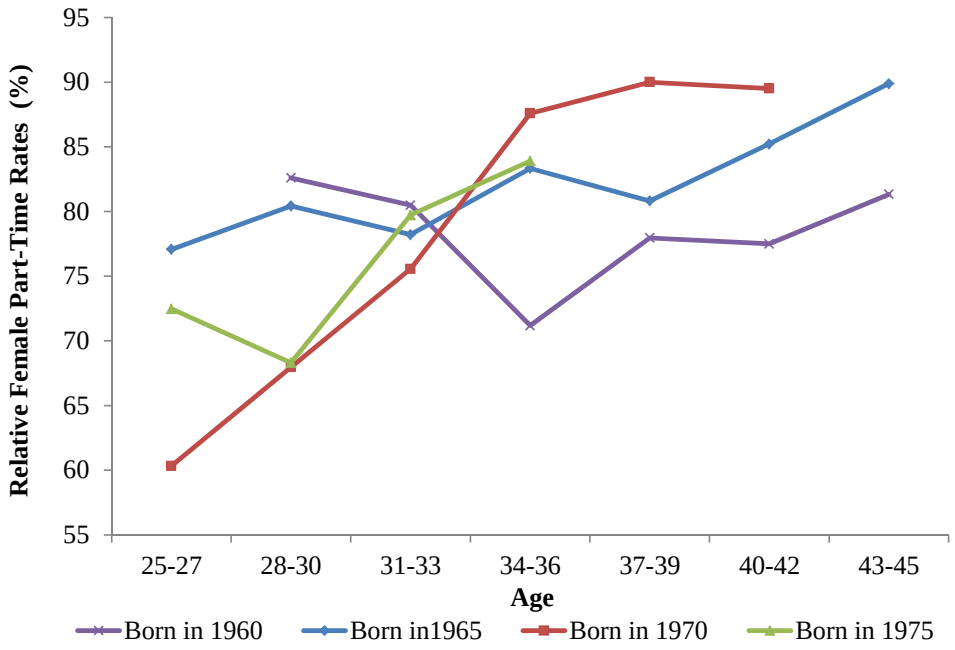
Source: Encuesta de Población Activa (INE) 1986-2011

Figure 15 – Life Cycle Part-Time Rates (as % of employment) – College Women



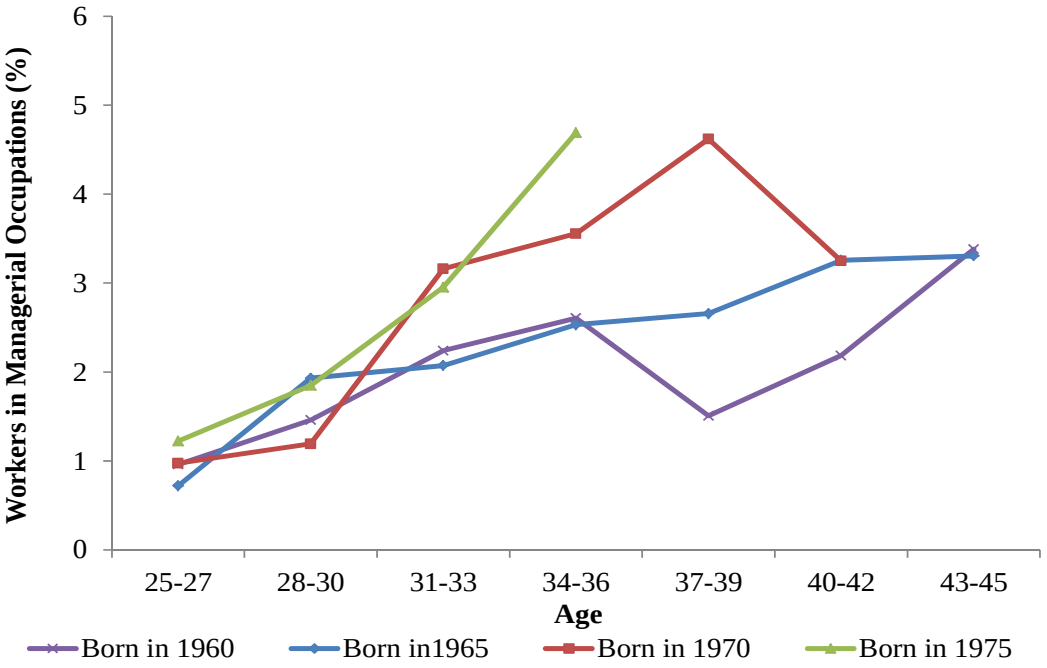
Source: Encuesta de Población Activa (INE) 1989-2011

Figure 16 – Life Cycle Relative Female Part-Time Employment Rates (%) – College Women/Men



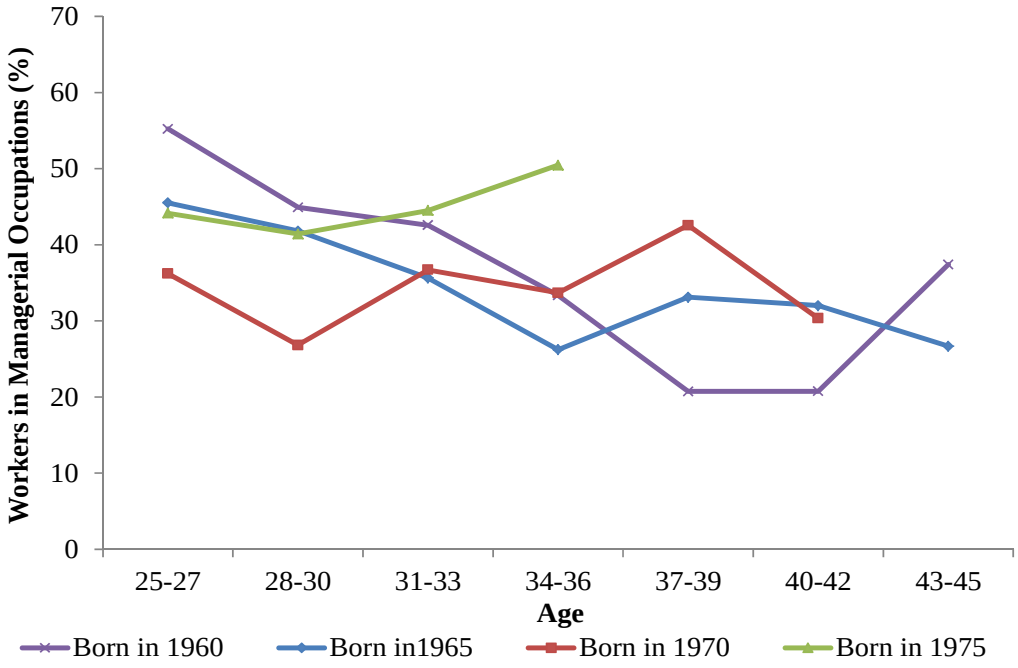
Source: Encuesta de Población Activa (INE) 1989-2011

Figure 17– Life Cycle Incidence of Women Workers in Managerial Occupations – College Females



Source: Encuesta de Población Activa (INE) 1986-2011

Figure 18 – Life Cycle Relative Incidence of Females Working as Managers – College Women/Men



Source: Encuesta de Población Activa (INE) 1986-2011

Table 1. Labor Market Statistics of College Men and Women Age 25-45 in Spain 1994-2008

	Employment Rates		Part-Time Rates		Managers incidence	
	1994	**2008**	**1994**	**2008**	**1994**	**2008**
1. All Men	93.12	91.82	2.31	1.84	6.85	3.14
	(25.34)	(27.41)	(15.03)	(13.47)	(25.30)	(17.45)
2. All Women	68.52	85.20	12.33	14.45	0.98	0.80
	(46.49)	(35.52)	(32.93)	(35.18)	(9.91)	(8.93)
3. Men without children	87.19	88.93	3.04	2.58	4.71	1,39
	(33.54)	(31.38)	(17.24)	(15.86)	(21.27)	(11.74)
4. Men with children	95.73	96.72	2.01	0.71	7.76	5.84
	(20.24)	(17.88)	(14.06)	(8.38)	(26.80)	(23.48)
5. Women without children	76.07	87.65	15.03	11.06	1.61	1.18
	(42.81)	(32.91)	(35.89)	(31.38)	(0.66)	(10.84)
6. Women with children	65.01	82.29	10.86	18.73	0.66	0.50
	(47.76)	(38.19)	(31.19)	(39.04)	(8.12)	(7.05)
Gender Gap All (2-1)	-24.60	-6.62	10.02	12.61	-5.87	-2.34
Gender Gap no child (5-3)	-11.12	-1.28	11.99	8.48	-3.10	-0.21
Gender Gap children (6-4)	-30.72	-14.43	8.85	18.02	-7.10	-5.34
Family Gap Women (6-5)	-11.06	-5.36	-4.17	7.67	-0.95	-0.68
No. observations	1033	3452	842	3027	842	3027

Notes: Standard deviations in parenthesis. Rates are computed using individuals' weights.
Source: Household Panels Phogue 1994 and EU-Silk 2008.

Table 2. Log Mean Hourly Wages of College Men and Women Age 25-45 in Spain 1994-2008

	1994			2008		
	All	Without children	With children	All	Without children	With children
1. Men	1.987	1.819	2.056	2.445	2.353	2.591
	(0.678)	(0.746)	(0.636)	(0.448)	(0.418)	(0.457)
2. Women	1.923	1.926	1.921	2.334	2.237	2.459
	(0.696)	(0.719)	(0.685)	(0.452)	(0.444)	(0.432)
3. Relative Women/Men	96.77%	105.88%	93.43%	95.46%	95.07%	94.91%
Gender Gap (2-1)	-0.064	0.107	-0.135	-0.111	-0.116	-0.132
Gender Gap (3-100%)	-3.22	5.88	-6.57	-4.54	-4.92	-5.09

Source: Household Panels Phogue 1994 and EU-Silk 2008.
Note: Standard deviations in parenthesis.

Table 3. Probability of Employment – Probit Estimation College Men/Women

	1994			2008		
	(1)	(2)	(3)	(1)	(2)	(3)
1. Female	-0.101**	-0.088*	-0.065	-0.012	-0.015	0.019
	(0.047)	(0.047)	(0.064)	(0.017)	(0.017)	(0.019)
2. Any child	0.138***	0.098*	0.116	0.117***	0.073**	0.081**
	(0.046)	(0.050)	(0.082)	(0.027)	(0.028)	(0.045)
3. Female*Any child	-0.214***	-0.192***	-0.221**	-0.161***	-0.162***	-0.257***
	(0.058)	(0.059)	(0.088)	(0.032)	(0.032)	(0.043)
Adjusted Family Gap						
For Women (2+3)			-0.106*			-0.176***
			(0.062)			(0.034)
For Men (2)			0.116			0.081**
			(0.082)			(0.045)
Adjusted Gender Gaps						
Women with any child versus men any child (1+3)			**-0.173*** (0.026)**			**-0.237*** (0.033)**
Childless Women versus Childless Men (1)			**-0.012 (0.017)**			**0.019 (0.019)**
Control variables	No	Yes	Yes	No	Yes	Yes
Observations	1033	1033	1033	3452	3452	3452
Pseudo R^2	0.1262	0.1484	0.1518	0.0339	0.1023	0.0795
Family contribution to the gender gap (%)		10.08 [-0.025]	11.24 [-0.028]		30.19 [-0.019]	56.20 [-0.035]

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The set of control variables include age, presence of a working spouse/partner, other family income in the household and regional fixed effects. Individual sampling weights are used in the estimations. The presented coefficients are marginal effects from probit models. The computation of the family contribution to the overall gender gap follows the methodology of probit decomposition into coefficient effects developed by Yun (2003). Absolute contributions in brackets.

Table 4. Probability of Working Part-Time vs Full-Time – College Men/Women

	1994			2008		
	(1)	(2)	(3)	(1)	(2)	(3)
1. Female	0.097***	0.106***	0.139**	0.096***	0.089***	0.088***
	(0.035)	(0.036)	(0.060)	(0.018)	(0.017)	(0.020)
2. Any child	-0.020	-0.008	0.074	-0.067**	-0.045	-0.086**
	(0.034)	(0.043)	(0.096)	(0.032)	(0.029)	(0.042)
3. Female*Any child	-0.003	-0.012	-0.060	0.112***	0.111***	0.129***
	(0.042)	(0.045)	(0.080)	(0.035)	(0.034)	(0.047)
Adjusted Family Gap						
For Women (2+3)			**0.014**			**0.042***
			(0.065)			**(0.027)**
For Men (2)			**0.074**			**-0.086****
			(0.096)			**(0.042)**
Adjusted Gender Gaps						
Women with any child versus men any child			**0.079****			**0.217*****
(1+3)			**(0.033)**			**(0.036)**
Childless Women versus childless Men			**0.139****			**0.088*****
(1)			**(0.060)**			**(0.020)**
Control variables	No	Yes	Yes	No	Yes	Yes
Observations	841	750	750	3027	3027	3027
Pseudo R^2	0.0898	0.1079	0.1087	0.1220	0.1502	0.1386
Family contribution to the gender gap		-5.43 [-0.005]	4.10 [0.004]		10.79 [0.014]	8.58 [0.011]

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The set of control variables include age, presence of a working spouse/partner, other family income in the household and regional fixed effects. Individual sampling weights are used in the estimations. The presented coefficients are marginal effects from probit models. Absolute contributions in the last raw in brackets.

Table 5. Probability of Working in Managerial Occupations – College Men/Women

	1994			2008		
	(1)	(2)	(3)	(1)	(2)	(3)
1. Female	-0.041	-0.038	-0.007	-0.016*	-0.014	-0.007
	(0.035)	(0.034)	(0.041)	(0.010)	(0.008)	(0.011)
2. Any child	0.022	0.022	0.044	0.028***	0.018*	0.045**
	(0.019)	(0.021)	(0.040)	(0.009)	(0.011)	(0.017)
3. Female*Any child	-0.051	-0.052	-0.098	-0.014	-0.017	-0.028
	(0.044)	(0.044)	(0.058)	(0.013)	(0.012)	(0.017)
Adjusted Family Gap						
For women (2+3)			**-0.054**			**0.016**
			(0.044)			**(0.017)**
For men (2)			**0.044**			**0.045****
			(0.040)			**(0.017)**
Adjusted Gender Gaps						
Women with any child versus men any child			**-0.105*****			**-0.035*****
(1+3)			**(0.033)**			**(0.012)**
Childless Women versus childless Men			**-0.007**			**-0.007**
(1)			**(0.041)**			**(0.011)**
Control variables	No	Yes	Yes	No	Yes	Yes
Observations	855	855	855	3020	3020	3020
Pseudo R^2	0.0740	0.0918	0.0940	0.0817	0.1458	0.1555
Family contribution to the gender gap		13.93 (-0.008)	25.3 (-0.015)		1,92 (0.000)	-6.71 (0.002)

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The set of control variables include age, presence of a working spouse/partner, other family income in the household and regional fixed effects. Individual sampling weights are used in the estimations. The presented coefficients are marginal effects from probit models. Absolute contributions in the last raw in brackets.

Table 6. Log Wage Regressions College Men/Women - Dependent Variable: Log Hourly Wages

	1994				2008			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Female	0.085	0.066	0.039	0.118	**-0.115*****	**-0.092*****	**-0.066*****	**-0.064*****
	(0.063)	(0.061)	(0.052)	(0.084)	**(0.022)**	**(0.021)**	**(0.019)**	**(0.022)**
Any Child	0.189***	0.010	0.055	0.083	0.116***	0.099***	0.056**	0.228***
	(0.050)	(0.052)	(0.044)	(0.084)	(0.029)	(0.026)	(0.025)	(0.033)
Female*Any Child	**-0.209*****	**-0,138****	**-0.136*****	**-0.232*****	-0.016	-0.004	-0.005	-0.042
	(0.076)	**(0.069)**	**(0.062)**	**(0.097)**	(0.034)	(0.031)	(0.028)	(0.038)
Age		0.089**	0.049	0.058		0.054**	0.056***	0.044***
		(0.043)	(0.040)	(0.043)		(0.020)	(0.017)	(0.017)
Age squared		-0.001	-0.001	-0.001		-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)
Tenure		0.009**	0.013***	0.013***		0.002	0.006***	0.006***
		(0.004)	(0.003)	(0.003)		(0.002)	(0.002)	(0.002)
Part-Time		0.145***	0.045*	0.035		-0.151***	-0.117***	-0.126***
		(0.068)	(0.063)	(0.063)		(0.030)	(0.026)	(0.026)
Industry dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Regional dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Occupations (4dig)	No	No	Yes	Yes	No	No	Yes	Yes
Observations	828	802	796	793	2602	2537	2533	2533
R^2	0.021	0.1955	0.3917	0.3960	0.0795	0.3132	0.4533	0.4539
Family contribution to the gender gap			**85.50 [-0.055]**	**100 [-0.101]**			**-16.51 [0.018]**	**0 [0.065]**

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The set of control variables include age, age squared, years of tenure in the actual job, industry dummies, occupations and regional fixed effects. Individual sampling weights are used in the estimations. Absolute contributions in brackets.

Chapter 4: Gender Gap in the choice of College-Major and in labor market outcomes: evidence from Italy

Massimo Anelli (UC Davis)
Giovanni Peri (UC Davis)

4.1 Introduction

In this chapter we quantify the impact of the choice of major for college-educated men and women on some important labor market outcomes such as their employment rate, occupational choice and income. We are interested, in particular, in how much the choice of college-major depends on specific skills and on general abilities or on preferences that may differ between men and women and how much it affects the gender wage and employment differential once in the labor market. There is abundant evidence from the literature that women are significantly under-represented in college-majors such as Engineering, Mathematics, Computer Science and Economics while they are over-represented in Humanities and Education. This seems to be a feature prevalent in several countries (as we documented in Chapter 1). In the US, Turner and Bowen (1999) show that women are significantly less likely to choose Math and Physical Sciences as college major, relative to men. Flabbi (2011) shows that in OECD countries, overall, women tend to choose Education and Humanities with much higher probability and Engineering and Mathematical Sciences with lower probability than men. For the UK, Chevalier (2011) also shows very different propensity of women to choose some college majors and he documents how relevant this is for some labor market outcomes. This difference in the choice of college major could be responsible for a significant portion of women under-performance in terms of wage, as the college major is a very important determinant of the access to high-paying occupations. In this chapter we first characterize the gender gap in the choice of majors in Italy and we also characterize how this choice is related to the ability of men and women as measured by several academic indicators. We focus on Italy because we have collected a unique database on the high school career, university career and labor market performance for young Italians who have graduated from high school between 1985 and 2005. While most of the analysis on college major and performance on the labor market has used data from the US and the UK, a country like Italy provides a case which should be even more

interesting. The choice of college major in Italy is potentially more relevant in determining the job that a person will get, relative to the US or the UK. In the Italian educational system, in fact, College majors correspond quite closely to the choice of a career/occupation. Each student, in fact, satisfies the general education requirements during high school, and then he/she applies and gets admitted to a specific major of a specific university instead of applying to a university and declare major only afterwards as it happens in the US. Once he/she gets admitted to a major he/she focuses, since the beginning of college, on the specific fields that define the major. Moreover, changing major after some time implies the large cost of losing most of the college units obtained. For instance a student majoring in Economics and Business will begin in her first year studying Mathematics for economists, Econometrics, Microeconomics, Marketing and Economic Law and will continue studying those topics for four years. A student in Italian Literature will begin in her first year studying Italian Composition, Italian Literature and Poetry and will continue to do so for four years. Changing major would imply the need to start again with the new major. Also, Law and Medicine in Italy are college majors rather than post-graduate degrees; hence becoming a doctor or a lawyer is a choice made right out of high school. The labor market opportunities of the two students, therefore, will be very different. As post-graduate education is still very rare in Italy (less than 0.5% of the population achieves a graduate degree), the choice of college will essentially direct the choice of the occupation within a very specific set. Its effect on the income of a person will be very prominent and long-lasting.

The main goal of this chapter is to quantify how much the gender-gap in the choice of major affects the men-women income gap as measured during their early years of their working career (five to fifteen years after their entry in the labor market). The years of high school and college, however, are extremely relevant not only in shaping the working career of individuals but also for their family choices (the choice of a partner). Hence decisions about major may also have an important indirect effect on family choices that should be considered when analyzing the choice of men and women. As men and women may have different preferences for monetary (wage) and non-monetary (partner, family) outcomes after college we should analyze

the effect of college choice on both of those outcomes. While in this chapter we focus on monetary and labor market outcomes in the choice of college majors, in the next we also consider the relative importance of career perspectives and quality of marriage on the choice of college-major of women and men.

The sample of our analysis is a group of highly educated Italians who graduated from high school between 1985 and 2005 and hence are currently between 25 and 45 years of age. We collected a novel database that includes all the graduates from college-preparatory high schools in the city of Milan (Italy) over the period 1985-2005. The sample includes around 30,000 individuals. The database includes information about their high school exit test, the type of high school attended, their class in high school and the location of their home during high school. Then we have gathered information about their university career, their college major, time to graduation and degree obtained and the final grade in college. Finally, we have information (for all of them) on total income in year 2005, as revealed by administrative sources (the Italian internal revenue service) and for a representative 10% sub-sample we also have information on occupation, family background, education and job of the mother and the father and current labor, location and family conditions, obtained with an interview in Summer of 2011. The next section and Appendix 4.1 at the end of the chapter describe the database, the original data, the method of collection, the variables included in the database and our merging and cleaning procedures.

We present in this chapter descriptive statistics and a regression approach to quantify the gender gap in the income of these college-educated individuals. We quantify how much the gender-difference in choice of college major contributes to the overall measured gender gap in income for the individuals in our sample. Our findings indicate that the choice of college major is the most relevant aspect of the school career of individuals in predicting their future income and in explaining the gender gap in income. Importantly we also consider how gender interacts with the academic ability of men and women, as measured by their high school exit test scores and by other measure of performance in college, such as their exit-grade and the time to

graduation. We find that the gender-gap in choosing highly-paid majors becomes particularly strong among highly performing students. This is a sign that academic quality, even in Math and Science, is not likely to be the determinant of the gender-differences in choice of field. Relative preferences for subjects or preferences for post-college outcomes must determine the differences.

One interesting finding of our analysis is that the gender gap within our sample (of college educated people in a city with relatively high income) is mainly in income and wages. We find that there is not much gender gap in employment, as most of the women in the sample work, at least part-time[45]. Unfortunately we do not have data on hours worked and hence we cannot quantify the gender gap in that variable. We can however restrict our sample to women and men with no children for which several studies show that the difference in hours worked is not significant, to isolate the differences in wages.

4.2 The data and some descriptive statistics: high school and college

In this section we describe our data, and their main characteristics.

4.2.1 Description of the data base

The database that we use for the analysis is a new database collected by the authors over the course of the year 2010-2011 and including information on the high-school and the college career of all the graduates of college-preparatory public high schools (Liceo Classico and Liceo Scientifico) in the city of Milan (for a total of 13 schools). We consider all individuals who graduated between the years 1985 and 2005[46]. Due to the lack of any database including information on the high school and college career of Italian students we collected this information manually visiting all the College Prep High schools in Milan and accessing the hard-copies of the records directly from the schools. While some missing and destroyed records existed we were able to cover more than 90% of all records for students who graduated

[45] The Italian Labor Force survey confirms that the difference in employment rate between men and women with a college education is smaller than 7%.

[46] There are few private high schools in Italy and they tend to select those students that, for some particular reasons, do not make it in public high schools. We did not include them in our sample.

between 1985 and 2005. This includes about 30,000 individuals. Milan is a large service-oriented metropolitan area in the richest part of Italy. The college educated individuals from this city often become part of the northern Italian elites in business, finance, Academia and in the professions. Hence they are an interesting group that allows us to analyze the gender gap at the top of the income and educational distribution in Italy. For these individuals we have information on the year of graduation, the grade in the high school exit exam, the school attended, the location where they lived during high school and the identity of their parents. Then we were able to link (using names and date of birth) these data with the student records from all universities in Milan (there are five of them, two private universities, Università' Cattolica and Bocconi, and three public universities, Politecnico, University of Milano, University of Milano-Bicocca). We can, therefore, reconstruct the university career of the individual who graduated from a college-preparatory high school in Milan, including the following information: whether they graduated, in what year, in what field of study, in what university and their exit-grade.

In a further step we link these records to data on the personal income of the individuals in year 2005, as revealed to the internal revenue service. This is information on their total income, reported on which individual pay taxes. There are some advantages and one main disadvantage in using these income data. The advantage is that the administrative file of reported income includes all individual in the national territory as it is mandatory to report any income. Hence if a person does not appear is because he/she has no income. Self-employed are included in the sample. If there is a higher tendency of a gender to be self-employed this does not affect our estimates of the gap. The disadvantage is that vis-à-vis the precise measure of income we do not have a measure of hours worked. Indeed differences in income are due for a large part to differences in hours worked. We will, however, identify some measures of hourly wages (or at least of wages at fixed labor supply) and we will discuss in the results how the gender gap differs between the two.

Data on income or wages are very rare in Italian datasets and hence this data provides a very important step in the analysis of income differential and gender gap and one of the very few

sources to link school performance and income. We use these income data only for people who graduated from high school before 1995. Considering, in fact, an average college attendance of five years, people in our sample would have been on the labor market between 5 and 15 years in 2005. Hence this provides a good assessment of the long-run consequences of the school quality/school choice on labor market outcomes during the early career. The total number of individuals for which we have data on the high school and university career is around 30,000. For 14,000 of them, those graduated between 1985 and 1995, we match the information on income in year 2005.

For a stratified[47] 10% random subsample of the initial universe (equal to 3,069 individuals) we collected much more detailed information from telephone interviews conducted in June 2011 by the professional company "Carlo Erminero & Co.". The additional information covers several variables regarding the family background, parental income, job and education, current employment and current family situation of the individual.

This database provides a unique window on the schooling career of young Italians with information on their schooling achievements in high school and college, and then we also link this information to their following labor market performance, between five and fifteen years from their entry in the labor market. We first describe some features of the pre-labor market abilities and characteristics of men and women, focusing on their high-school and college performance and how these and preferences differ between men and women and how they are related to their choice of major. Then we consider how the choice of major affects income, wage and labor supply once those individuals are in the market and how this choice differs between men and women. As we find a stark discrepancy between the high academic ability of women and their choice of less selective (and less rewarding in terms of future wage) college majors, in the next Chapter (Chapter 5) we explore alternative possible explanations for this. First we focus on how less "academic" characteristics, revealed by the choice of doing sport, or of participating in charity and volunteering has any bearing on the choice of college-major. Second, we consider how the economic success of the partner (husband) varies by choice of major, to see whether economic consideration in the partner choice may rationalize, for

[47] The stratification has been done along the following four dimensions: type of high school, year of graduation in 4-years bins, final test score in high school, in 5-point bins, gender. There is a total of 80 cells.

women, the preference for certain majors. Finally we assess the impact of peers in high school on the choice of majors, to see if women or are affected by their peers (and possibly teachers) in high school and if such peer effect is different between women and men, especially in their choice of major.

4.2.2 Description of College Prep high schools and college major for Men and Women

The Italian system of education encourages those students who are planning to go to college to select a different track already in high school. The very large majority of people who enroll to college, about 80% of the total, attended a type of college-prep high school called "Liceo". This can be "Liceo Classico" with an emphasis on classic studies such as Literature, Philosophy, Latin and Greek or "Liceo Scientifico" with an emphasis on Mathematics and Sciences[48]. Either of these preparatory high schools allows access to any type of college major. Students finish their general education in high school and, beginning with the first year of college, they focus only on the topics of their major. Among the students who attended the high schools included in our sample almost 80% enrolled to college. Of those enrolled, about 70% graduated from college. This is a very high graduation rate for Italian standards. This is due to the fact that college preparatory schools in Milan are, generally, of very high quality, they select highly achieving individuals and these individuals, as we describe below, have parents with high education and relatively high income.

The universe of students who attend college-preparatory high school is an interesting one, to study gender differences in the scholastic career and then in the labor market of young Italian individuals. First these are the upper part of the distribution (in terms of ability and possibly income of the family) and hence they are the more relevant group to analyze gender gap for highly educated and high performing individuals on the labor market, and "glass-ceiling" type of phenomena. Second, their family background is relatively similar, especially in our case, as they come from middle class families, living in the same city and valuing college education. Finally, the high school experience that lasts 5 years and often generates a very close-knit group of

[48] There is also very small number of "Liceo Artistico" with an emphasis on Art.

class-mates is considered as a crucial educational experience for Italian youth. Importantly the group of class-mates is the same over the 5 years, time devoted to each subject of study is the same across classes within type of school (and is centrally decided by the Ministry of Education) and there are no elective classes. This would be the environment in which peer effects and the effect of teachers are strongest, especially in relation to the choice of college major and careers.

As for university education, the Italian system essentially pushes the choice of a career at the end of high school, as the college major (field of study) is highly focused on topics that are related to the specialty of the major. So students essentially choose to become Engineers, Doctors, Accountants, Lawyers by choosing at the end of high school, the corresponding Field of study (Engineering, Business, Accounting, Law). The correlation between university studies and the type of occupation is much stronger than in Anglo-Saxon systems.

Let us characterize, then, some features of the gender differences in performance and choice in high school and college beginning with their choice of major and their high school performance. The first difference between men and women (gender gap) apparent in our data is that women are much more largely represented in the “Liceo Classico”, the high school more focused on Humanities and Classics. While overall in our universe of graduates from college-preparatory schools women are about 52% of the total, in “Liceo Classico” they are 66% of the population while in “Liceo Scientifico” they are only 40% of the population. As that choice of high school is made when students are 13 years old it may reveal the preferences of the student for field of study, or a forward looking preference of the family for a job outcome. Figure 1 shows how the percentage of women overall, and in each type of school, has changed over time.

While overall the share of women among high school graduates is close to 50%, with some fluctuations over time, the share of women in “Liceo Classico” fluctuates around 65% (with not much of a trend) while for the “Liceo Scientifico” that share fluctuates around 40% with an initial upward sloping trend that flattens in the 1990’s and 2000’s.

This preference of women for Humanities and Arts and of men for Scientific and Mathematical subjects becomes even clearer when we analyze the choice of college major for students graduating from the high school in our sample. Figure 2 shows the percentage of women among those who enrolled and graduated in 11 different college majors that span all the possible choices available to the Italian students[49].

Three things emerge very clearly from Figure 2. First the three majors of Engineering, Business and Economics and Math/Computer-Science have a very low percentage of women (lower than 50%). They are also large majors and, as we will see, they are associated with relatively high income. Hence this compositional difference will contribute significantly to the gender gap. Second, because of the significant attrition rate in college we report both the gender ratio among enrolled (dark bar) and graduated (light bar). It is clear that, for each major, the graduation rate of women is larger than the graduation rate of men as their percentage increases among graduated. Third, Humanities and Education, but also Medicine and Natural Sciences have a very large share of women enrolled and graduated (higher than 60%). Some of those are large majors and, except for Medicine, they are not associated with large income.

Are some of these college majors "harder" to get in? Do some major attract/select better students? In our sample and during the considered years some of the degrees in Engineering and Business and Economics, as well as Medicine, had entry tests while others (typically humanities and social sciences) did not. It is likely, therefore, that some of those majors were more selective than others. Figure 3 shows the average grade in the graduating high school exam, standardized to be between 0 (barely passing) and 1 (top grade), for each of the considered college major. The average grade of men in the major is represented by the light bar, while the average grade of women by the dark one.

The figure reveals significant differences in "average quality" of the students admitted in different majors and, interestingly, several of the most selective majors, in particular Engineering, Economics/Business and Mathematics/Statistics are those that Figure 2 revealed

[49] Each of the 11 groups merges several specific majors, the exact definition of majors in each of our groups is reported in Appendix 4.1.

as “male dominated”. The female-dominated majors, instead, are spread through the distribution (Humanities has a rather high average quality of admitted student while Education has low average quality). Two further interesting things emerge from Figure 3. First, women admitted to each major exhibit higher quality than men, as measured by their high school exit-test grade[50]. Second, in some majors, such as Social Sciences and Architecture and Design the gender gap in high school grade is larger than 0.1, which is an extremely large value. This may be the sign that several academically gifted women direct themselves toward these two majors (rather than towards more demanding ones) while those majors attract relatively less talented men.

The previous graphs reveal a striking gender difference in major choice. But is the choice of college major driven mostly by the previous choice of high school type? Namely, is it the case that graduates from “Liceo Scientifico” choose Engineering and the Business-Economics (independently of their gender) while graduates of the “Liceo Classico” mainly chose Humanities and Law (also independently of their gender)? Or are women graduating from “Liceo Scientifico” still less likely to choose engineering and Economics/Business than men from the same type of school? Figure 4 (part a and b) shows that the second scenario is the relevant one. Even among the “Liceo Scientifico” graduates, women are much less likely to choose Engineering and Economics and Business. In fact more than 55% of male graduates from that “Liceo” choose one of these two major, while only about 20% of women does. Interestingly Natural Sciences and Architecture and Design (besides Humanities) are among the very top choices for women in the “Liceo Scientifico”. The more “math-science” oriented women, that is those attending “Liceo Scientifico”, seem to prefer Natural Science and Design to Engineering and Business/Economics, relative to comparable men[51]. The distribution of choices in the “Liceo Classico” is more similar between men and women. While there is the usual difference in

[50] The high school graduation test in Italy is identical for all schools in the nations and is graded on the same scale. However as it is not multiple-choice, but involves essays and problems as well as an oral exam there may be differences in the grading. We will account for these potential differences in using the high school test as measure of academic quality, in regressions.

[51] Dolado, Felgueroso and Almunia (2010) show a comparable finding among women within the field of economics: A larger share of them chooses fields such as labor and development (with more “human content”) and fewer choose finance and theory (with more abstract content).

Engineering and Business-Economics (chosen by 25% of men and less than 10% of women), Law and Humanities, together, absorb more than 50% of both men and women graduates (with a bias in favor of Law for men and in favor of Humanities for women).

Are these choices likely to be correlated with the labor market performance of men and women, especially with their wage income? While we will devote the later sections to a more formal analysis, the answer is a resounding yes and Figure 5 already shows the main reason why.

The figure shows the average gross (pre-tax) yearly income in Euros as of 2005, reported to the internal revenue service by the individuals in our sample, divided by college major. Among the four top-paying college majors two are among those strongly men-dominated. Conversely the two bottom-paying majors are those strongly women dominated. Law, another high paying major, is also among those with lowest percentage of women, while only Medicine among the high-paid majors is an exception and has a large percentage of women. Notice that the three bottom majors, which include about 35% of the total of college-graduate women earn an average yearly income of less than 13,000 Euros (about 17,000 $)! Those majors include only 10% of the men. A huge 65% of college-educated men graduated from one of the four top-earning majors (whose average income is above 30,000 euros), while only 35% of college educated women did.

Not only this gender difference in the choice of major is large and correlated with earnings, it has also not changed much over time. Considering the four top-paid majors (Engineering, Medicine, Economics & Business and Law) and the four least-paid majors (Humanities, Art-Design, Agriculture and Education) Figure 6 shows the evolution of the share of women in each group between 1985 and 2005.

The share of women in high paid majors has been below 40% for most of the period and possibly declining in the last seven years of the sample. The share of women in low paying majors has been above 65% and stable. Whatever the reasons for the large difference in major choices they do not seem to have changed much in the twenty years from 1985 to 2005.

4.3 Gender gap in Academic Quality in high school

The previous section has shown three interesting tendencies, somewhat hard to reconcile with each other. The first is that Engineering and Economics/Business, two of the most math-intensive majors are among those with lowest share of women among enrolled and graduated students. At the opposite end of the spectrum Education and Humanities, more language intensive, are the majors with highest share of women among enrolled and graduated students. The second fact emerging is that Engineering and Economics/Business are also among the most selective majors (when measured in terms of the high school exit-exam grade of people in the major) and are among the best paid. The third fact is that in any major, including the math-intensive ones, the quality of women, as captured by the high school exit-test grade is higher than the quality of men. Hence, the idea that men have stronger math abilities, and those are highly rewarded on the labor markets and because of this they can access better paid majors and lucrative jobs could be reconciled with the first two facts but not with the third. A possibility is that women have an absolute advantage academically, both in math-science and humanities-art, but a comparative advantage in humanities-art. They outperform males, that is, in those subjects by more than they do in science-math subjects. However, at least as revealed by the test score in high-school, there seems to be only a small relative advantage of women in Humanities, versus Math-Science, certainly not comparable with the very large difference in Major choice. Moreover as math-science majors are associated to much higher wages than humanities majors, if the motivation to choose humanities is higher productivity (skills) on the labor market in order to obtain higher wage, such differential should further attenuate the incentive to choose those majors. Something besides wage must motivate women.

To inquire into how academic quality interacts with gender and choice of major we begin by analyzing in greater detail the academic quality of women and men as revealed by their high school final test grades. In particular we first check, by exploiting the difference in the scores in the exit-test at the “Liceo classico” and “Liceo Scientifico”, whether women have a disadvantage in math-scientific fields, as revealed by those scores.

Figure 7 shows the distribution of exit-test grades, smoothed with a kernel, for men and women in the “Liceo Scientifico” (panel (a)) and in the Liceo Classico (panel (b)). It is apparent the similarity of the two and in both cases the women’s distribution is significantly shifted to the right relative to men’s distribution. Hence even in the “Liceo Scientifico” whose exit test is much more intensive in Math and Science, women outperform men. The distribution of test scores is not centered and reflects the tradition of Italian schools to give very few high grades and concentrate most of the grades in the lower part of the distribution.
A formal test of the difference in average test scores between men and women reveals that for both “Liceo Classico” and Liceo Scientifico” the standardized average exit-test score for women is around 0.43 while for men it is 0.39 with a difference around 0.04 points, significant at any level of confidence.

While the high school exit exam is common to all high schools of each type in the country, the committee grading the exit tests is different across high schools and years. This may generate some lack of comparability for test scores across schools and years. If this is correlated with the presence of women it may generate some spurious correlation. Moreover as we have information about where, in the city of Milan, each student lived at the time of high school graduation, we can control for the real estate values in that area, which is a crude control for the wealth of the family of origin. Hence, we perform a simple regression of the high school graduation test on year-school dummies, on the house value in the area where the student lived and on a gender dummy equal to one for females. The coefficient on this dummy would reveal the quality difference between men and women, as revealed by their high school exit test score, once we account for differences in schools, years and in neighborhoods. The coefficient on the female dummy is 0.045 with a standard error of 0.004. Hence, confirming the raw data, there is a very significant and quite large average difference in academic performance between men and women, independently of the high school type, of the graduation year and of the part of the town where students lived. Women outperformed men in high school, on average, by a significant amount equal to one seventh of the overall standard deviation of high

school grades. A rapid comparison with the available PISA-test scores relative to the region of Italy where Milan is (Lombardia), for year 2000 (within our sample) reveals that men outperform women in math (530 vs. 513), but do worse in both reading and sciences (496 vs. 537 in reading and 541 vs. 545 in sciences). If we look at an equally weighted average of the scores in these tree skill categories we see that girls outperform boys overall (532 vs. 522). If we consider the structure of the two type of high schools the PISA results are not surprising: indeed at Liceo Scientifico (that is the high school track relatively more oriented to sciences) Language and Science matter more than Math while in the Liceo Classico, Language matters much more than Science and Math together. This is consistent with the traditional bias of the Italian education system towards humanities.

In spite of this better performance a large part of women did not choose the most demanding math-intensive and highly selective college majors that would also be associated with higher pay later on. Understanding why this is the case, or at least where and how this gender bias arises and what effect it has on the labor market outcomes of women is the goal of the rest of this chapter.

4.4 College Majors as determinant of gender Gap: A regression approach

4.4.1 From High school to College: the gender gap in university performance

We have considered so far high school exit exam-scores as a measure of academic quality of the students. But are they good predictors of their performance in college? Is it possible that men, who were under-performing relative to women in high school, then out-perform them in college? Is this true at least for those who choose the math-intensive highly paid major? Or at least do men fill the academic gap with women, so that by the end of college their performance is comparable? Table 1 analyzes these issues. Conditional on their choice of college, we consider three measures of academic performance in College. The first is the graduation rate. The second is, conditional on graduation, the time to graduation. The third is the final grade at college graduation. Specification 1 in Table 1 shows the marginal effects of a probit regression

in which the probability of graduating from college is regressed on the high school exit-test grade, a female dummy and the interaction between female and high school grade.

Three results emerge from specification 1. First the high school grade is a very significant and strong predictor of the probability of graduation. Going from the lowest to the highest high school grade increases on average the probability of graduating by 42%! Second, even controlling for high school grade (that was higher for women), being a woman carries a significant advantage in graduating probability (6% at the average probability) on top of that. Third, such gender advantage is particularly strong for students who performed at the bottom of the high school test score distribution, while at the top it is close to 0. This can be seen by using the coefficient on the interaction between gender and grade for lowest grades (0) and highest grade (1) and adding it to the main gender effect. Each regression controls for dummies absorbing the year by high school effects. Specifications 2 and 3 consider Least Square regressions using as dependent variables two alternative measures of college performance. One is the time to graduation expressed in months, which in the Italian System has a very large variance, and the other is the grade at college graduation. That grade is also standardized to be between 0 (barely passing) and 1 (top). Those regressions are restricted to those who graduated from college only and confirm the first two findings described above. First, the high school exit-test score is a very strong predictor of time to graduation. The magnitude of that coefficient implies that a student graduating with top grades will, on average, take almost one year and a half less, (17 months) to graduate relative to one graduating at the bottom of the high school grade distribution. Similarly each 1% increase in high school grade corresponds to a 0.2% increase in the college exit-grade. Second, women take on average 3 months less to graduate than men and get a final grade higher by 6.5% of the total range, even after controlling for their high school grade. Is this due to the fact that women choose in larger percentage “easier” college majors? Columns 4 and 5 dispel that thought. First, by including major dummies women still outperform men, by a very significant average of 0.038 points. This value is lower than before, hence the distribution of grades across majors does affect the gap. Even focusing only on the high-paid majors (Engineering, Economics-Business, Medicine and

Law) three of which are also largely dominated by men in the enrollment, the advantage of women is large and extremely significant (estimates in column 5). In this case the advantage appears to be uniform over the grade distribution (as the interaction female X grade is small and not significant).

Hence women outperform men academically in high school, both in the "Liceo Scientifico" and "Classico". Controlling for the quality in high school, they outperform men in terms of probability to graduate, time to graduation and exit grade in college. This is true overall and also in the math-intensive, men-dominated high paying majors. Hence, this evidence point to higher academic quality of women, emerging in high school and deepening in college. Conditional on the choice of math-intensive and male-dominated fields women still significantly out-perform men. If academic quality is associated with productivity, and if skills developed in high paying majors are in high demand women should have an advantage and a disadvantage in terms of productivity on the labor market, relative to men. On one hand, they exhibit higher academic quality and hence should perform better. On the other hand they choose majors associated with lower wages at a much higher rate than men. It seems also the case that their choice of major is dictated by taste preferences and by consideration beyond their potential wages (possibly family and flexibility) and not to their abilities: certainly not their absolute abilities. Even the difference in relative abilities is small enough, as women outperform men of a similar margin in science-math as they do in language to make a relative productivity based story unlikely to explain the large difference in specialization.

4.4.2 Choice of Majors and Gender Income Gap

After having analyzed the academic performance of this group of college-prep graduates it is interesting to see how this relates to their labor market performance. We have three types of outcome measures for these individuals and each of them has some limits. One is their personal gross income earned in year 2005, as recorded by the internal revenue service. This income does not only include labor income, although for most of the individuals wage earnings would be the overwhelming part of the income. It includes employees and self-employed

workers. We consider only the group that graduated from high school between 1985 and 1995 and that has a college degree. Hence, as of year 2005, most of those individuals have been in the labor market between five and fifteen years. The other two outcomes that we can consider are the employment status (whether a person is working or not) and the occupation/sector of activity. These last two outcomes are available only for the subsample on which we have conducted some telephone interviews (10% of the total, around 3,000 individuals). We know whether the individual works, but not the number of hours worked, and in which occupation and sector within a list of nine possible occupations and twelve sectors.

First we illustrate with a regression approach how large is the wage gap between men and women in the sample and then we analyze how the inclusion of different controls explains away that gap. Our presumption is that the choice of college majors will contribute a significant amount to the explanation of the gender gap of income (wage), while other characteristics of the schooling career of men and women (grades, time to graduate, type of high school) will explain relatively little as women performed well relative to men and the other characteristics in terms of family background and location in the city are relatively similar between men and women. Specification 1 of Table 2 shows a regression of the logarithm of income on year of graduation dummies, age of the individual and a female dummy. The regression includes all the high school graduates over the period 1985 and 1995. These individuals were mostly between 29 and 39 years of age in 2005, and hence well into their working career.

In the first specification of Table 2 the coefficient on the female dummy is an estimate of the average gap in (logarithmic points) of gross income between women and men (controlling only for age and type of high school attended). Remarkably, this difference equals 0.46 logarithmic points (about 37%) which is by all means a very large average difference. This is yearly income and it combines the difference in average wage including bonuses, overtime and other incentives to production and in average hours worked which can compound the effect. On the other hand this is the gap in a group of college graduates, all from a rich city in Northern Italy, with relatively homogeneous family backgrounds. Hence this should be a group for which prejudice and discrimination on the work-place is as low as possible within Italy and for which access to education and to opportunity is high and similar for men and women. Recent

estimates of the gender earning gap in Italy in recent years reveal values that are actually comparable with those estimated in column 1 of table 3 (equal to 37%). For instance the Global Gender Gap report of 2011, by the Global Economic Forum reports that the earning of women were 46% below those of men in Italy overall. The gender gap in yearly wages estimated from the EU-SILC data, and limited to college educated over 25 was 34% in 2009, while for all workers it was 40%. Our figures of the gender wage differential therefore are larger than those measured in the EU-SILC data.

Interestingly, when restricting the sample to individuals who actually enrolled to college (specification 2) and to those who graduated from college (specification 3) the gender income gap does not change much and in specification 3 it is still equal to 0.46 logarithmic points. Hence selection into college and then into the group that graduated from college does is not particularly affect the gender gap. Then in specification 4 we include some controls for the academic quality of the individual, namely the grade in her/his high school exit test, the grade in his college exit test and a control for the value of the real estate in the part of town where the student lived at the time of high school as a coarse proxy of the wealth of his family of origin[52]. This specification reveals that the high school grade is highly significant in affecting the logarithmic wage of an individual. A student graduating at the top of the class (score of 1) would on average earn 0.37 logarithmic points (44%) more than a student graduating at the bottom (score of 0). Academic merit counts on the labor market. However, not surprisingly, the inclusion of these controls does not affect much the gender gap. While, on average, academic performance pays on the labor market it is not at all the cause of the gender gap, in fact its inclusion does not even significantly change the point estimate of the gap. Column 5 includes also more detailed controls for the quality of high schools attended by students in the form of high-school by year effects. This does not change the estimated gender gap either. After all this is not surprising: women are better than men academically, and higher quality is associated to

[52] The value of the house in the area was calculated based on the exact address where the student was living at the time of high school, merged with a map of the real estate values of homes in Milan in year 2008, by block. The location in the city of Milan is the most important determinant of value of the house and highly correlated with average income.

higher income so controlling for high school grades and quality does not affect the estimates of the gap.

In column 6 we simply add two dummies to the previous controls (notice that the previous controls are 3 variables plus 130 year by school dummies and they have essentially no effect on the female coefficient). These are a dummy for graduating from a highly-paid major (Engineering, Business/Economics, Law or Medicine) and one for graduating from a low-paying major (Humanities, Art and Design, Education or Agriculture). This, by itself, reduces the gender gap by full 0.1 log points, one fourth of the whole difference. If we include all the 11 major dummies the female gap is reduced by 0.14 logarithmic points to -0.29. This corresponds to a wage gap of 25% vis-à-vis the 37% estimated in specification 3. Hence, the choice of major is responsible for one third of the gender gap between college educated young Italians in this sample. Notice also, interestingly, that when we control for the College-Major dummies the grades in high school have no predictive power on logarithmic income, while the grades in college, now within major, are very relevant. This is consistent with the interpretation that the high school grade affects the working career mainly by affecting the choice of major in college, and then the specific academic performance within the major still affects the income. This fact certainly penalizes women very much, relative to men. In fact women with high grades in high school do not choose as frequently as men to enter highly paid major and hence they dissipate large part of their academic quality advantage in economic terms by channeling their higher quality in lower paid majors. Moreover this penalty is likely to be larger for particularly talented women who would be fully qualified and competitive in majors such as Engineering and Economics, but choose Architecture and Design or Humanities, instead.

The fact that one third of the gender gap in observed income is explained by the choice of broad major and that most of that economic effect is simply captured by the choice of a high-paid or a low-paid major, pushes us to analyze more closely the gender differences in this choice.

In Table 3 we show the estimated gender bias in the probability of choosing a high-paid major. Column 1 reports the marginal effects of a probit in which we include the high school grade of

the individual (that increases very significantly the probability of choosing a high paid major) a female dummy (that exhibits a very strong negative effect) and the interaction between female and grade. Column 2 reports the estimates of the same coefficients with a linear probability model, and show very similar values. The meaning of those estimates is that, for given quality (grade in high school) a woman has a 20% lower probability of choosing a high paid major relative to men. As on average 65% of men choose a high paid major, then the difference is very large. Even more surprising is the interaction effect. This means that among high performing students (say those at the very top of the distribution with a grade of 1) there is a further 10% lower probability for women to choose a low-paid major. Hence women at the top of their class are fully 30% less likely than men to choose a high paid major.

In column 3 we interact the female dummy with dummies for the range of high school grade, instead of linearly with the grade. The dummy for the lowest quartile of the distribution is omitted, then the female dummy is interacted with the second lowest quartile (low) the third lowest (intermediate) and the top. The estimates confirm that the bias against high paid majors is particularly strong among women of high academic quality. Column 4 to 6 repeat the same specifications, limited to the sub-sample of individuals for which we have more detailed information on family background and on labor market outcomes (the one we will use in the next section). The estimates are similar, although with higher standard errors, and they confirm that the subsample that we use for some specific part of the analysis appears to be representative of and similar to the whole sample.

Table 4 shows similar estimates when we consider the probability of choosing a "low-paid" major. Again column 1 shows the marginal effects of a probit estimate, column 2 shows the effects estimated in a linear probability model. Column 3 shows the effect when the interaction between the female dummy and the high school grade is broken into four (of which one is omitted) grade intervals. Columns 4 to 6 show the same estimates on the random sub-sample used for the telephone interviews. Interestingly, Table 4 shows the exact symmetric effect in terms of gender bias and high school grades in affecting the probability of accessing low-paid majors relative to what we found in Table 3. Using the estimates of column 2 students at the

top of the grade distribution in high school were 16% less likely to go to a low-paid major, however women were, in general 15% more likely and those at the top of the grade distribution an additional 5% more likely. So a woman at the very top of the distribution of high school grade had actually still a higher probability than a man at the very bottom of the distribution of high school grade, to choose a low-paid major (namely the difference would be 0.149+0.051-0.167=0.033 logarithmic point). The estimates for the restricted sample show a smaller main effect of the gender dummy but a stronger interaction effect so that the effect of being women at the top of the grade distribution is that of being 16 to 20% more likely to choose a low-paid major than a man.

The results shown in Table 3 and 4 are robust to changing the definition of high and low paid majors and including only three majors in each group. Those results are very useful but they do not provide a clear measure of how much of the wage differential across genders is explained by this choice. To do this we produce, in Table 5, a synthetic measure of how much of the income gender gap is purely determined by the different choices in major, combining all majors. We construct as dependent variable the "major-imputed" wage, that is, we associate with each individual the average income of male individuals with the same college degree from the 2005 income sample. This translates the major dummies into wages, at the value of the average male wage earned by individual with that college major as of 2005. Then we regress this "major-imputed" wage on the female dummy, the high school grade and the interaction.

Column 1 and 2 of Table 5 show the results when we include all the sample of graduates (from 1985 to 2005). As the wage is imputed only based on the major we can use all observations as long as we have a college major for them and this is what we do in the first two columns. Column 3 and 4 of Table 4 show those estimates including in the sample only the individuals who graduated between 1985 and 1995, which are those for which we also had actual reported income that we used in Table 2. Column 5 and 6, finally, show the results obtained on the restricted sample on which we have conducted telephone interview. In all three cases we have that the major-imputed wage is associated with a gender gap between 0.13 and 0.15 log points,

which is almost exactly the difference in the point estimate of the effect of the gender gap explained in Table 1 by the inclusion of the major dummies. Hence, the difference in the choice of major is a crucial channel through which gender affects wages, producing an income differential of about 15%. Another measure of how significant is the wage effect of gender gap through choice of major is given by the effect of high school grade on the major-imputed wage. The difference between men and women in major-imputed wage is as large as or somewhat larger than the difference between the bottom and the top high school performer. Hence the women graduating from high school with the top grades will have a major-imputed wage that is lower than the men graduating with lowest grades. The income effect purely channeled through the choice of major implies that the women graduating at the top of their class will be at a disadvantage relative to a man graduating at the bottom of his. As we have seen in Table 2 the most important effect of high school grades on income operates through the effect on the choice of major (as once we included the major dummy its effect on wages disappeared, see column 5 and 6 in Table 2). Hence this gender gap in the choice of major has consequence for women wages much larger than the positive effect of their higher grades in high school. Whatever the reason for this difference in the choice of major, women are giving up substantial income opportunities and the Italian economy is giving up a large share of its productivity (captured by wages) through this channel. Incentives in the admission process to math-intensive majors, more information during high school, a woman-friendly environment in math-intensive majors or other measures could have a very high return in terms of aggregate productivity.

Before including the controls for family background we consider the gender gap in employment rate. We only know the employment status of our individuals for the interviewed sub-sample.Table 6 confirms the fact that the gender gap in income is explained by college-major choice for one fourth of the total (difference between column 1 and 2). Moreover we see no difference in employment rate (Column 3 and 4), at least in the restricted sample, between men and women. This is reasonable as most of these college educated women work (at least part-time) and, as we will see below, the fertility rate of these women is very low. While higher

grades are associated with higher income and higher employment neither effect is significant, due to large standard errors.

4.4.3 Adding wealth and family background controls

Are the results shown above, in particular the existence of a very substantial gender-gap in income and the fact that difference in the choice of college major accounts for one quarter to one third of it, robust to the inclusion of other controls? It is reasonable to think that most of the family background variables are distributed rather similarly between men and women and hence not responsible for spurious effects. The students considered in the sample, after all, are from a relatively homogenous socio-economic background. Table 7 shows the same regression as Table 6, including several family background controls. We control for the imputed wage of the father at the time the child was in high school (obtained by assigning to him the average wage for the occupation/sector of activity that he revealed in the phone interview), his education, the mother's education, and the average value of house in the area where the family lived at the time of high school. The only control that has an effect on the income of the individual in 2005 is the imputed wage of the father. Education of the father and of the mother do not matter (remember that all individual we are considering are college educated), and also the value of the family home does not have a significant effect. The gender gap in income is unchanged when we include the controls, as shown in specification 1 and 2. At the same time we do not find any gender gap in employment, even when we including controls.

Finally we check whether the gender gap in the choice of major, captured by the major-imputed wage, is affected by the family background variables that we can control for. Table 8 shows the effects of home value, parental education, father's income besides the high school grade and the female dummy, in affecting the major-imputed wage of an individual. Column 1 to 4 use the whole sample, during the 1985-2005 period or during the 1985-1995 period, while column 5 and 6 considers the restricted sample of individuals who were also interviewed.

The regressions continue to find a gender gap in major-imputed wage of 0.12 to 0.15 logarithmic points when we consider the whole sample and around 9% when considering the interviewed sample. The high school grade continues to have a positive and significant effect on

the major-imputed wage. Among the controls the most significant are the father's wage and, in some specification, the dummy indicating that the student lived in an area ranking in the top 10% of housing values. These results confirm those of Table 5 above and emphasize that the different gender preferences for the choice of major is very pervasive across families and controlling for different background, location and schooling level of parents does not affect that.

4.4.4 College-Major Choice of Siblings

We finally perform the most demanding and perhaps cleanest test to see that the disproportionate choice of high paying major by men (and of low-paying major by women) does not depend on family-specific factors such as income, location of the house, education and job of parents. While the career choice of parents certainly affects the choice of children, we demonstrate that the overwhelming tendency to prefer different college majors is as strong within a family as in the whole sample. We begin by selecting families that have more than one sibling in our sample. Then we perform probit and OLS estimates of the gender bias in choosing high and low-paid major on this sub-sample, and we introduce also family-dummies in it. Table 9 below reports the coefficients. Specifications 1 to 3 show the coefficient estimates obtained for high-paid major as outcome. Specifications 4 to 6 consider low-paid major as outcome. The total sample includes 1525 siblings. Of those there are 751 siblings making opposite choices of major within a family (between high-paid and non-high-paid or low-paid and non-low-paid). The gender gap identified only on the families with siblings making different choices (and using the OLS estimates) implies that men from the same family are 47% more likely to choose a high paid major (than their sisters) and 38% less likely than their sisters to choose a low-paid major. If we include all siblings (also those making the same choice of major in a family) then brothers are 23% more likely to choose a high paid major and 12% less likely to choose a low-paid major than their sisters. The gap in the overall sample (estimated with OLS and same controls) from table 3 and 4 is, respectively, -20% for woman in high-paid majors and 15% in low paid majors. Hence while family characteristics certainly affect the choice of major, the differences across them do not change at all the extremely strong and significant bias of men for high-paid college major. That bias is as strong within the average family as it is in the overall sample.

4.5 Gender-gap in Occupational Choice

The third outcome that we can analyze with our data is the occupation of people revealed by the interviews in year 2011. We have grouped occupations into 12 groups and the sectors of activity into 8. Then we construct a grid by occupation and sector of activity and we assign to each individual the average hourly wage in Italy for that group, in 2011. This is a way of summarizing with one variable the occupation-sector choice of an individual, and to capture what wage differential is associated with differences in occupation-sectors. As we are measuring hourly wages, this variable will not include any difference in hours and weeks worked. Hence we expect the variation of this variable and, possibly, of the estimated gender gap, to be much smaller than for the actual income. Table 10 shows the gender gap in (the log of) occupation/sector imputed wage. We notice some interesting features, all consistent with the previous findings. First, the occupational wage gender gap is around 6.3% when we do not control for major dummies, and decreases by about a fifth to 5.4% when we control for those. Hence it is much smaller than the gender gap in logarithmic income. Variation of specific jobs within an occupational category, differences in hours worked across jobs, differences in non-wage income across people all will contribute to generate the larger variance for income. Those characteristics seem also to increase substantially the gender gap. Second, quite significant and also affecting the gender gap is the inclusion of the wage of the father (also imputed through the occupation and sector of activity). The occupation of the father is significantly correlated with the occupational choice of the children and including this variable explains a fifth of the gender gap in occupational wage. A reasonable explanation for this would be that in several professions (lawyer, doctor, accountant, notary), where the intergenerational transmission within the family is common in Italy (see Pellizzari and Orsini 2012), this takes place more through the transmission to the son than to the daughter. In fact if we also include an interaction of the occupational wage of the father with a female dummy (not reported) the effect is negative (-0.03) and the gender gap loses significance.

The other variable that is significant in determining the occupational income in the sample is the value of the real estate in the area where the person lived. We only include a dummy for living in the part of town with the 10% highest value of houses, which has a positive effect on occupational wage. We also performed some other regressions (not reported) including a few further controls, such as a dummy for having both working parents, and an interaction of that dummy with the female dummy. These controls do not change the gender gap, nor are they very significant themselves.

4.6 Conclusions

We have documented in this chapter three very interesting features in the gender-gap for college educated Italians, currently in their 30's and 40's, who belong to the richest part of the country and to families with high education and medium to high income. The first is that there is a very significant gender gap, in favor of women in all measures of academic performance, beginning in high school and continuing with college. Women show significantly higher exit test scores, higher percentage of graduation from college, lower time to graduation and higher college grades. Even more interestingly and, possibly in contrast with previous findings, we clearly see that their better performance is not limited to the humanities/classic oriented topics. Women outperform men also within the "Liceo Scientifico" and, once in college they outperform men in each major, including the math/science-intensive ones.

The second fact is that their choice of college major is very different between women and men. In particular women, somewhat in contrast with their higher academic performance, tend to shun the highly selective and high paying major of Engineering and Business/economics and this tendency is particularly strong, relative to men when their academic quality is very high. Women reveal a preference for the less selective, low paying majors of humanities. In particular women in the "Liceo Scientifico" who earn high grades tend to prefer Natural Science and Art and Design to Engineering and Economics/Business.

The third interesting fact is that this remarkable difference in the choice of major has not really decreased over time and it more than offsets the advantage that higher academic quality would afford to women in terms of income. We also find that the choice of college major explains one quarter to one third of the gender gap. While there are obviously other important factors that contribute to the significant gender gap which may include the continuity of career, discrimination and differential non-academic skills, the choice of major is an important component of the gender gap and, by far, the largest factor among those related to school performance, at least in high school and college. These results are robust to the inclusion of

family background controls. Finally, among college educated women in our sample we do not find much evidence of a gender gap in employment.

Figure 1

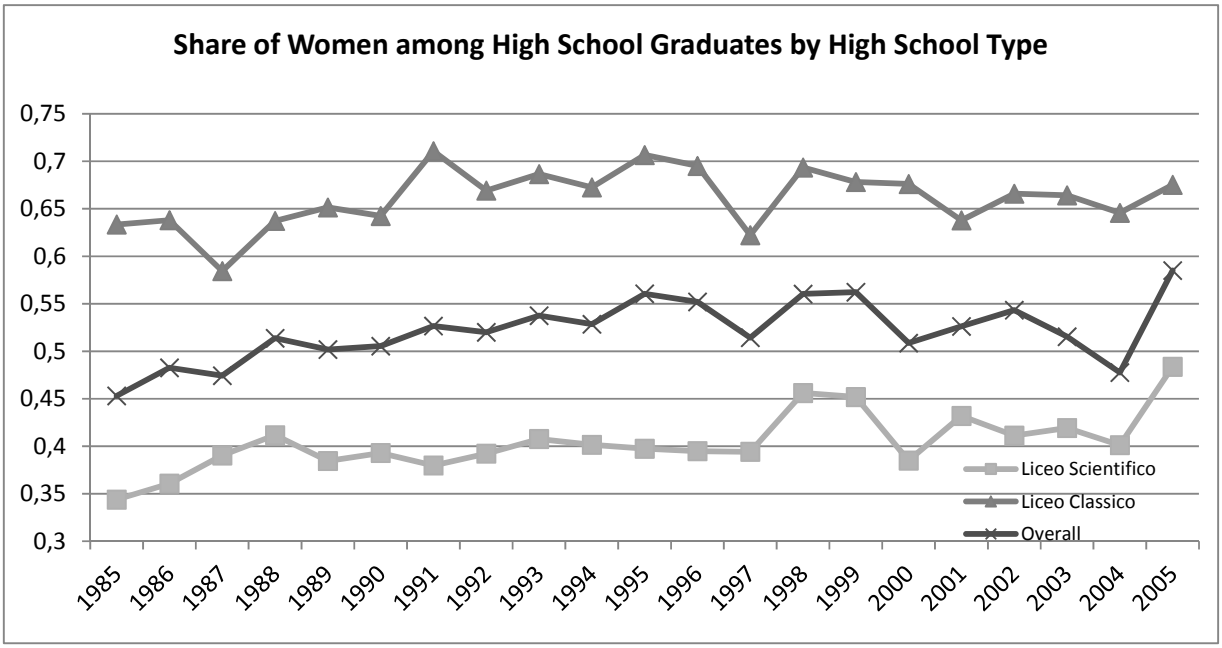
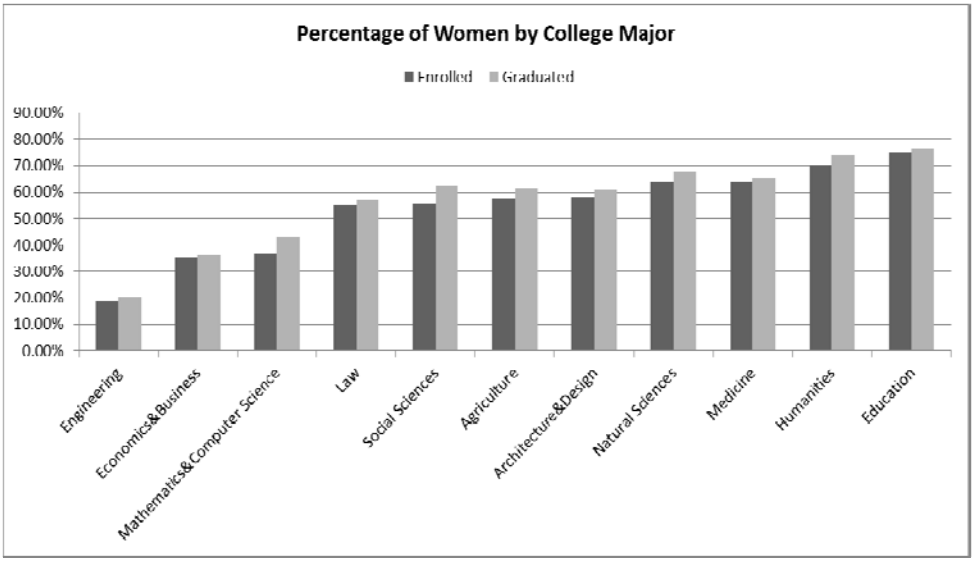
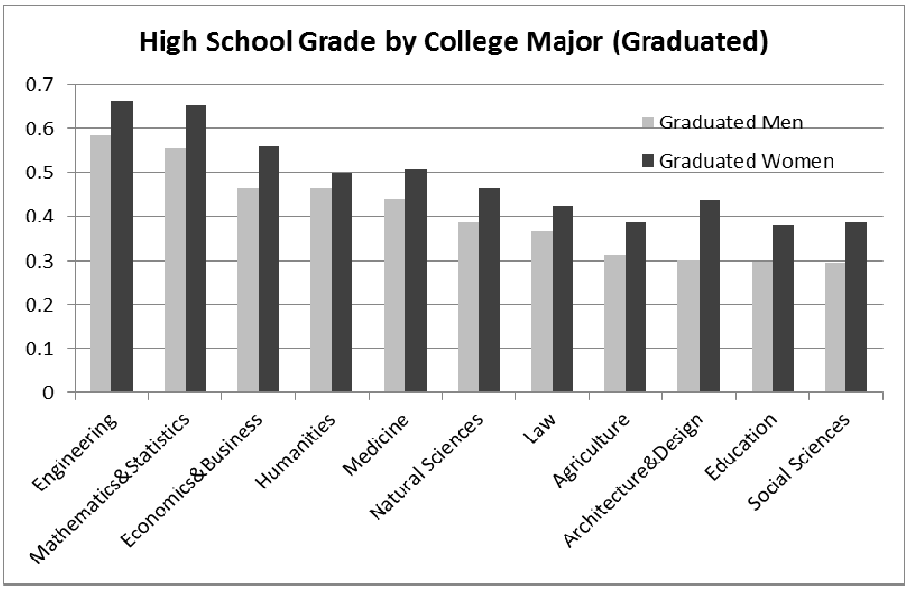


Figure 2



Note: The universe is represented by all college-prep students graduated between 1985 and 2005 in Milano, Italy.

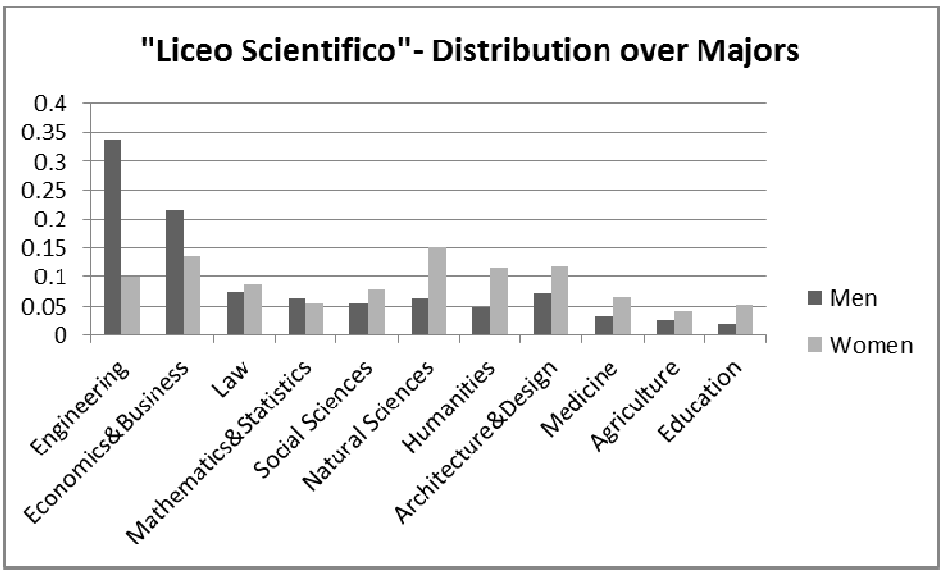
Figure 3



Note: The universe is represented by all college-prep students graduated between 1985 and 2005 in Milano, Italy.

Figure 4

(a)


"Liceo Scientifico"- Distribution over Majors
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Engineering
Economics&Business
Law
Mathematics&Statistics
Social Sciences
Natural Sciences
Humanities
Architecture&Design
Medicine
Agriculture
Education
Men
Women

(b)

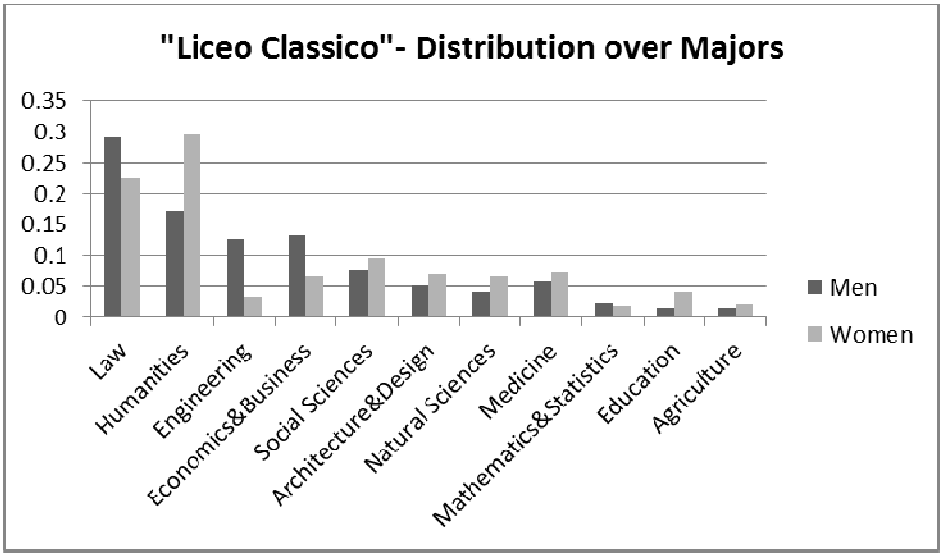
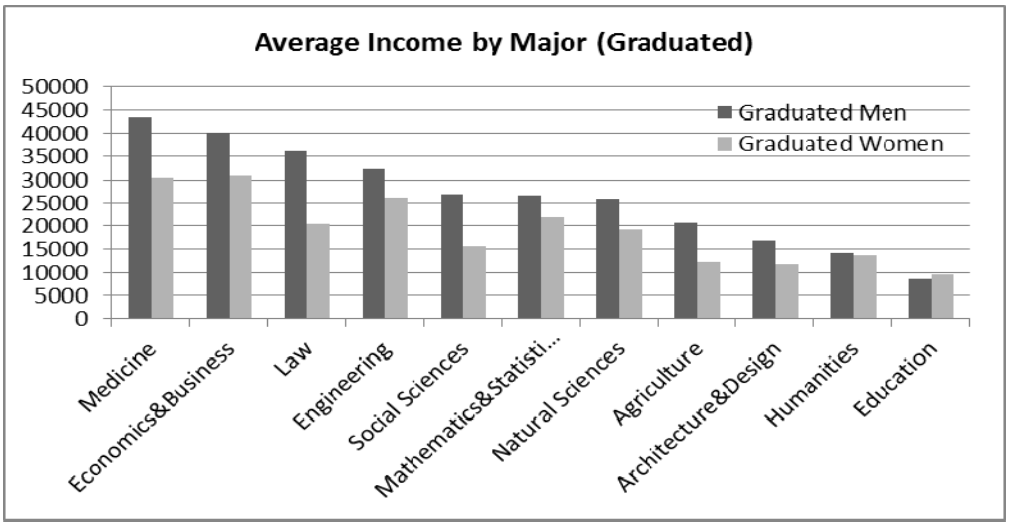
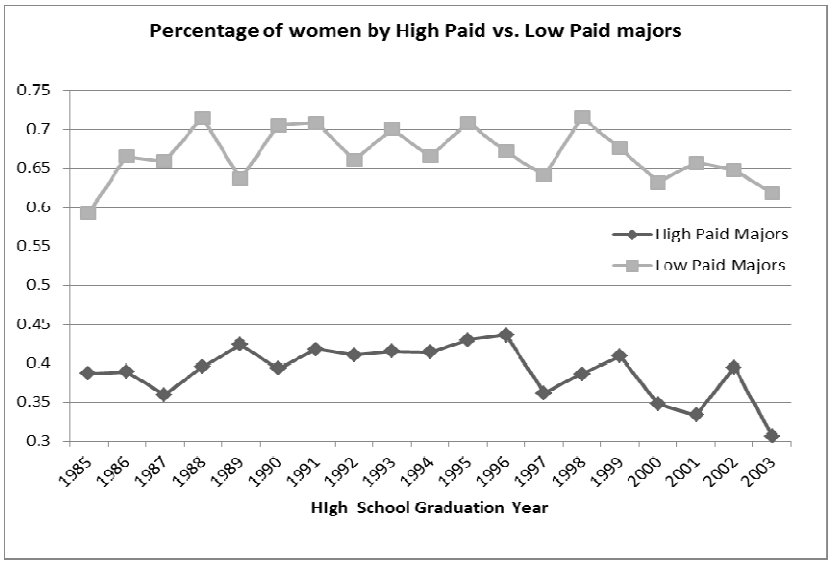

"Liceo Classico"- Distribution over Majors
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Law
Humanities
Engineering
Economics&Business
Social Sciences
Architecture&Design
Natural Sciences
Medicine
Mathematics&Statistics
Education
Agriculture
Men
Women

Figure 5



Note: The income is the average for the whole sample of yearly income in year 2005 in Euros, as reported to the internal revenue service.

Figure 6



Note: high paid majors are Engineering, Medicine, Economics and Business and Law and the low-paid majors are Humanities, Art-Design, Agriculture and Education.

Figure 7

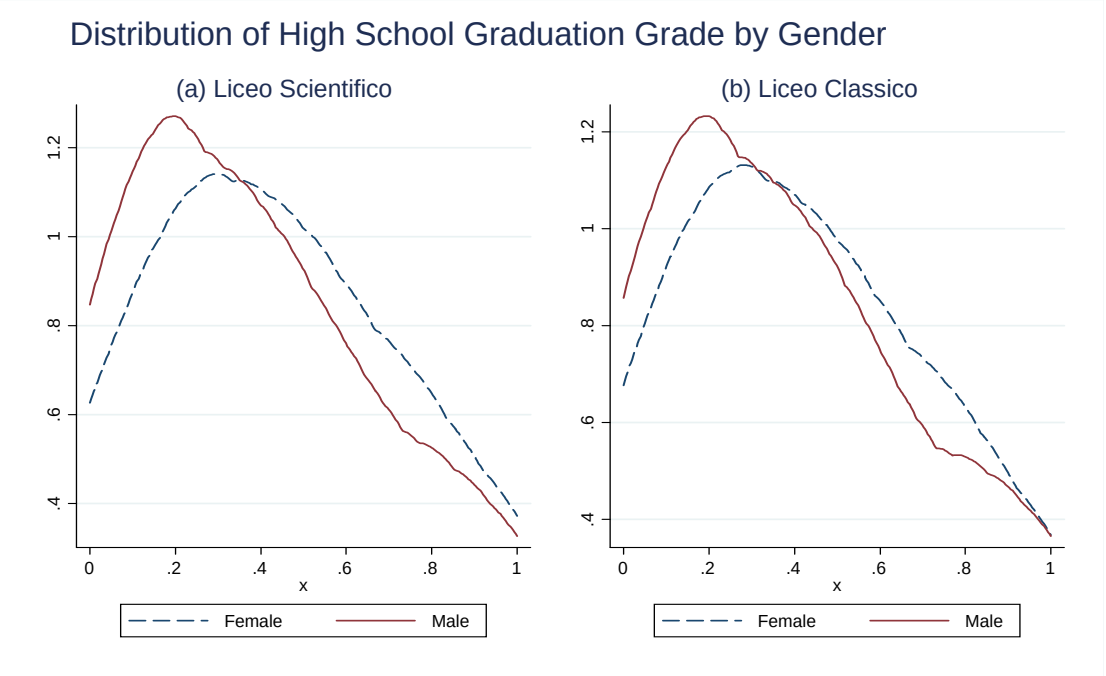

Distribution of High School Graduation Grade by Gender
(a) Liceo Scientifico
(b) Liceo Classico
.4
.6
.8
1
1.2
0
.2
.4
.6
.8
1
x
Female
Male

Table 1: Measures of Performance in College and Gender-Ability

Dependent Variable	Graduated=1	Time to Graduation	Final Grade in College	Final Grade in College	Final Grade in College
Specification	(1)	(2)	(3)	(4)	(5)
VARIABLES					
Female	0.063***	-2.972***	0.065***	0.038***	0.046***
	(0.01)	(0.769)	(0.005)	(0.004)	(0.008)
(Female)X(grade)	-0.062***	-0.711	-0.020**	-0.054***	-0.003
	(0.022)	(1.281)	(0.008)	(0.007)	(0.012)
HS exit test grade	0.424***	-17.031***	0.207***	0.263***	0.247***
	(0.015)	(1.072)	(0.006)	(0.006)	(0.007)
School X Year Dummies	X	X	X	X	X
Major Dummies				X	
Constant		62.768***	0.712***	0.748***	0.650***
		(0.174)	(0.003)	(0.007)	(0.003)
Sample	enrolled	graduated	graduated	graduated	graduated in high paid majors
Observations	22,873	16,035	16,561	16,561	8,406
R-squared		0.165	0.221	0.392	0.265

Note: Probit for specification 1, OLS for all other specifications, standard errors clustered at School/Year level, *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Gender gap in Log income and the importance of College Majors

Dependent variable: Log of Actual Wage							
SPECIFICATIONS	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES							
Female	-0.466***	-0.436***	-0.462***	-0.435***	-0.432***	-0.329***	-0.292***
	(0.021)	(0.022)	(0.025)	(0.03)	(0.031)	(0.028)	(0.029)
HS exit-test grade				0.375***	0.395***	0.125**	0.032
				(0.047)	(0.047)	(0.051)	(0.054)
Value of real estate in Area				0.011	0.025	0.021	0.032
				(0.038)	(0.042)	(0.042)	(0.041)
College exit grade				-0.025	-0.039	0.720***	0.817***
				(0.092)	(0.09)	(0.096)	(0.11)
Year Dummies	X	X	X	X			
School by Year Dummies					X	X	X
High Paid Majors Dummy						X	
Low Paid Majors Dummy						X	
Majors Dummies							X
Constant	10.409***	10.412***	10.540***	10.274***	10.042***	9.779***	9.467***
	(0.043)	(0.052)	(0.046)	(0.316)	(0.341)	(0.346)	(0.358)
Sample	all sample	enrolled	graduated	graduated	graduated	graduated	graduated
Observations	13,469	11,033	8,322	6,871	6,871	6,871	6,871
R-squared	0.091	0.088	0.099	0.109	0.124	0.186	0.2

Note: OLS, Standard errors clustered at School/Year level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1.

Table 3: gender bias in the choice of highly paid majors

Probability High Paid Majors=1						
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Probit	OLS	OLS	Probit	OLS	OLS
Female	-0.191***	-0.204***	-0.218***	-0.139***	-0.139***	-0.145***
	(0.014)	(0.015)	(0.014)	(0.040)	(0.041)	(0.040)
Female X grade	-0.105***	-0.096***		-0.091	-0.100	
	(0.027)	(0.027)		(0.075)	(0.076)	
Fem X (grade low range)			-0.028*			-0.036
			(0.015)			(0.042)
Fem X (grade intermediate range)			-0.029*			-0.059
			(0.017)			(0.048)
Female X (grade high range)			-0.070***			-0.067
			(0.022)			(0.061)
HS exit test grade	0.262***	0.258***	0.246***	0.169***	0.177***	0.166***
	(0.021)	(0.021)	(0.020)	(0.055)	(0.058)	(0.055)
School Dummies				X	X	X
Year Dummies				X	X	X
School X Year Dummies	X	X	X			
Constant		0.523***	0.526***		0.371***	0.372***
		(0.010)	(0.010)		(0.114)	(0.115)
Sample	graduated	graduated	graduated	graduated & interviewed	graduated & interviewed	graduated & interviewed
Observations	16,560	16,561	16,561	1,965	1,965	1,965
R-squared		0.110	0.110		0.070	0.070

Note: Standard errors clustered at School/Year level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Gender bias in the choice of low-paid majors

Probability High Paid Majors=1						
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Probit	OLS	OLS	Probit	OLS	OLS
Female	0.125***	0.149***	0.145***	0.033	0.039	0.047
	(0.013)	(0.014)	(0.014)	(0.037)	(0.038)	(0.040)
Female X grade	0.106***	0.051**		0.210***	0.197***	
	(0.025)	(0.024)		(0.070)	(0.069)	
Fem X (grade low range)			0.023*			0.051
			(0.014)			(0.043)
Fem X (grade intermediate range)			0.039**			0.134***
			(0.017)			(0.049)
Female X(grade high range)			0.050**			0.163***
			(0.020)			(0.061)
HS exit test grade	-0.210***	-0.167***	-0.171***	-0.173***	-0.154***	-0.158***
	(0.021)	(0.017)	(0.017)	(0.056)	(0.048)	(0.047)
School Dummies				X	X	X
Year Dummies				X	X	X
SchoolXYear Dummies	X	X	X			
Constant		0.198***	0.199***		0.440***	0.451***
		(0.009)	(0.009)		(0.081)	(0.084)
Sample	graduated	graduated	graduated	graduated & interviewed	graduated & interviewed	graduated & interviewed
Observations	16,560	16,561	16,561	1,965	1,965	1,965
R-squared		0.096	0.096		0.105	0.106

Note: Standard errors clustered at School/Year level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Gender Bias in College Major-Related Wage

Dependent Variable: Imputed log Wage by Major of graduation						
Specification	(1)	(2)	(3)	(4)	(5)	(6)
Reference Year:	1985-2005	1985-2005	1985-1995	1985-1995	1985-2005	1985-2005
VARIABLES						
Female	-0.157***	-0.154***	-0.135***	-0.128***	-0.091***	-0.083***
	(0.011)	(0.010)	(0.013)	(0.013)	(0.024)	(0.025)
Female X grade	0.003		-0.004		-0.064	
	(0.019)		(0.023)		(0.045)	
Fem X (grade low range)		-0.005		-0.016		-0.040
		(0.011)		(0.014)		(0.029)
Fem X (grade intermediate range)		0.005		-0.006		-0.052
		(0.012)		(0.016)		(0.033)
Female X(grade high range)		-0.003		-0.012		-0.055
		(0.015)		(0.020)		(0.039)
HS exit test grade	0.119***	0.120***	0.099***	0.101***	0.093***	0.092***
	(0.012)	(0.012)	(0.015)	(0.015)	(0.033)	(0.032)
School Dummies					X	X
Year Dummies					X	X
SchoolXYear Dummies	X	X	X	X		
Constant	10.366***	10.365***	10.426***	10.417***	10.229***	10.226***
	(0.006)	(0.006)	(0.011)	(0.014)	(0.069)	(0.070)
Sample	graduated	graduated	graduated	graduated	graduated & interviewed	graduated & interviewed
Observations	16,561	16,561	8,615	8,615	1,965	1,965
R-squared	0.109	0.109	0.089	0.089	0.098	0.099

Note: Standard errors clustered at School/Year level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Gender Gap in Income and employment

Dependent Variable:	Actual log income	Actual log income	Employment status	Employment status
Specification	(1)	(2)	(5)	(6)
VARIABLES				
Female	-0.399***	-0.313***	-0.007	0.002
	(0.117)	(0.115)	(0.024)	(0.024)
Female X grade	0.139	0.191	-0.003	0.012
	(0.220)	(0.210)	(0.044)	(0.045)
HS exit test score	0.202	0.131	0.027	0.017
	(0.157)	(0.148)	(0.027)	(0.030)
School Dummies	X	X	X	X
Major Dummies		X		X
Year Dummies	X	X	X	X
Constant	10.128***	9.850***	0.939***	0.943***
	(0.203)	(0.257)	(0.048)	(0.053)
Sample	graduated & interviewed	graduated & interviewed	graduated & interviewed	graduated & interviewed
Observations	1,297	1,297	1,965	1,965
R-squared	0.274	0.335	0.176	0.192

Table 7: Gender gap in income and employment, including family controls

Dependent Variable:	Actual log income	Actual log income	Employed==1	Employed==1
	(1)	(2)	(5)	(6)
VARIABLES				
Female	-0.395***	-0.326**	0.018	0.025
	(0.129)	(0.127)	(0.023)	(0.024)
(Female) X (grade)	0.078	0.157	-0.048	-0.033
	(0.234)	(0.222)	(0.043)	(0.045)
HS exit-test grade	0.159	0.075	0.041	0.029
	(0.164)	(0.157)	(0.028)	(0.030)
Imputed wage of father	0.187*	0.147	0.016	0.016
	(0.113)	(0.109)	(0.022)	(0.021)
Father education	-0.040	-0.067	0.003	0.002
	(0.053)	(0.051)	(0.011)	(0.010)
Mother education	0.004	-0.006	0.005	0.008
	(0.051)	(0.048)	(0.008)	(0.008)
House value in top decile	0.195	0.123	0.028	0.026
	(0.138)	(0.131)	(0.033)	(0.032)
House value in bottom decile	0.075	0.074	0.029*	0.030**
	(0.107)	(0.102)	(0.015)	(0.014)
School Dummies	X	X	X	X
Year Dummies	X	X	X	X
Major Dummies		X		X
Constant	8.669***	8.927***	0.811***	0.831***
	(0.850)	(0.841)	(0.162)	(0.159)
Sample	graduated & interviewed	graduated & interviewed	graduated & interviewed	graduated & interviewed
Observations	1,132	1,132	1,713	1,713
R-squared	0.263	0.323	0.188	0.202

Note: OLS, Standard errors clustered at School/Year level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1

Table 8: Gender gap in Major-Imputed Income, including controls

Dependent Variable: Imputed log Wage by Major of graduation						
Specification:	(1)	(2)	(3)	(4)	(5)	(6)
Reference Year:	1985-2005	1985-2005	1985-1995	1985-1995	1985-2005	1985-2005
VARIABLES						
Female	-0.156***	-0.154***	-0.134***	-0.127***	-0.088***	-0.087***
	(0.011)	(0.010)	(0.013)	(0.013)	(0.026)	(0.028)
Female X Grade	0.003		-0.004		-0.073	
	(0.019)		(0.023)		(0.047)	
Fem X Low grade range		-0.006		-0.016		-0.033
		(0.011)		(0.014)		(0.033)
Fem X intermediate grade range		0.005		-0.006		-0.050
		(0.012)		(0.016)		(0.033)
Fem X high grade range		-0.002		-0.012		-0.061
		(0.015)		(0.020)		(0.041)
HS exit test grade	0.119***	0.120***	0.099***	0.101***	0.107***	0.105***
	(0.012)	(0.012)	(0.015)	(0.015)	(0.033)	(0.033)
Home value in the top 10%	0.020**	0.020**	0.011	0.011	0.050	0.051
	(0.010)	(0.010)	(0.013)	(0.013)	(0.040)	(0.040)
Home value in the bottom 10%	-0.001	-0.001	-0.010	-0.010	-0.020	-0.019
	(0.009)	(0.009)	(0.011)	(0.011)	(0.025)	(0.025)
Father wage					0.062**	0.062**
					(0.024)	(0.024)
Father education					0.021*	0.020*
					(0.012)	(0.012)
Mother education					-0.001	-0.001
					(0.010)	(0.010)
School Dummies					X	X
Year Dummies					X	X
SchoolXYear Dummies	X	X	X	X		
Sample	graduated	graduated	graduated	graduated	graduated & interviewed	graduated & interviewed
Observations	16,561	16,561	8,615	8,615	1,713	1,713
R-squared	0.109	0.109	0.089	0.090	0.118	0.119

Note: OLS, Standard errors clustered at School/Year level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1

Table 9: College-Major Choice of Siblings

Dependent Variable:	Probability Highly Paid Major=1			Probability Low Paid Major=1		
Specification	(1)	(2)	(3)	(4)	(5)	(6)
Method	Probit	OLS	OLS	Probit	OLS	OLS
Female	-0.363***	-0.232***	-0.470***	0.322***	0.121***	0.383***
	(0.080)	(0.059)	(0.130)	(0.079)	(0.046)	(0.135)
Fem X HS Grade	-0.175	-0.055	-0.113	0.124	0.063	0.120
	(0.159)	(0.103)	(0.234)	(0.149)	(0.087)	(0.245)
HS Grade	0.463***	0.292***	0.439***	-0.473***	-0.174**	-0.473**
	(0.114)	(0.083)	(0.166)	(0.124)	(0.072)	(0.182)
Year Dummies	X	X	X	X	X	X
School Dummies	X	X	X	X	X	X
Family Dummies	X	X	X	X	X	X
Constant		-0.361	1.728***		0.964	0.573
		(0.281)	(0.440)		(0.595)	(1.281)
Sample	Mixed gender siblings choosing alternative outcomes	Mixed gender siblings	Mixed gender siblings choosing alternative outcomes	Mixed gender siblings choosing alternative outcomes	Mixed gender siblings	Mixed gender siblings choosing alternative outcomes
Observations	751	1,525	751	513	1,525	513
R-squared		0.585	0.337		0.556	0.345

Note: Standard errors clustered at School/Year level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Gender gap in occupation-sector imputed wage

Dependent Variable:	Imputed log wage by sector & occupation	Imputed log wage by sector & occupation	Imputed log wage by sector & occupation	Imputed log wage by sector & occupation
	(1)	(2)	(3)	(4)
VARIABLES				
Female	-0.063***	-0.054***	-0.052***	-0.045**
	(0.019)	(0.019)	(0.020)	(0.020)
Female X grade	0.018	0.028	0.001	0.013
	(0.037)	(0.036)	(0.040)	(0.040)
HS exit test score	0.042	0.020	0.047	0.028
	(0.030)	(0.030)	(0.033)	(0.033)
Imputed wage of father			0.064***	0.059***
			(0.020)	(0.020)
Father education			-0.005	-0.009
			(0.009)	(0.009)
Mother education			0.007	0.007
			(0.009)	(0.009)
House value in top decile			0.046*	0.034
			(0.028)	(0.027)
House value in bottom decile			0.022	0.022
			(0.021)	(0.020)
School Dummies	X	X	X	X
Year Dummies	X	X	X	X
Major Dummies		X		X
Constant	7.448***	7.447***	6.942***	7.006***
	(0.048)	(0.055)	(0.147)	(0.151)
Sample	graduated & interviewed	graduated & interviewed	graduated & interviewed	graduated & interviewed
Observations	1,727	1,727	1,515	1,515
R-squared	0.102	0.145	0.121	0.162

Note: OLS, Standard errors clustered at School/Year level in parenthesis,
*** p<0.01, ** p<0.05, * p<0.1.

Chapter 4, Appendix 4.1 – Description of the Data

In this work we present for the first time a unique database collecting information on the high school career, university career and labor market performance of young Italians who have graduated from high school between 1985 and 2005. This dataset involves different sources that have been carefully matched in a complicated merging process. The collection of the database has involved the collaboration of many parties. The help from the following persons and institutions made the collection possible: the directors of the high schools in Milan, the company "Ambroscuole", the Provincia di Milan, Daniele Checchi (for Universta' Statale), Carlo Lucifora (for Universita' Cattolica), Francesco Peri (For Universita' di Milan, Bicocca), Augusto Sarti and Mauro Santomauro (for Politecnico di Milan). Davide Malacrino and Francesca Barbiero provided excellent assistance in collecting and organizing the data.

We summarize the sources of the single datasets merged, the information contained in each dataset and the merging process in figure A4.1. The diagram representing the merging process must be read from left to right.

The core dataset include the universe of all high school graduates attending college-prep schools in the city of Milan between 1985 and 2005 (around 30550 individuals). Data have been collected manually by inputting the information contained in hard copies of the school records.

The list of the 13 college-prep high schools in the city of Milan involved in the data collection process by type of school (Classical Studies vs. Scientific Studies) is included in the following table:

Table A4.1

Liceo classico (focus on classical studies, humanities)	Liceo scientifico (focus on math and sciences)
• **ISTITUTO LSLR PITAGORA**	• DA VINCI
• **BECCARIA**	• DONATELLI-PASCAL
• **BERCHET**	• EINSTEIN
• **MANZONI**	• GALILEI
• **OMERO E TITO LIVIO**	• MARCONI
• **PARINI**	• SEVERI
	• VOLTA

Among the five major universities of Milan involved in the collection of our data three are public universities ("Universita' degli Studi di Milan", "Universita' degli Studi di Milan – Bicocca", "Politecnico di Milan") and two are private ("Universita' Bocconi", "Universita' Cattolica di Milan"). The first two public universities mentioned ("Universita' degli Studi di Milan", "Universita' degli Studi di Milan – Bicocca") have a very broad offer of majors while "Politecnico di Milan" offers degrees in Engineering, Architecture and Design only. Among the private universities "Universita' Cattolica di Milan" has a broad offer of majors comparable with "Universita' degli Studi di Milan" while "Universita' Bocconi" is a school of Business, Economics and Law.

The list of majors offered overall by universities located in Milan is very large. For our analysis we thus aggregated all these majors in 11 broader "Fields of Study" as follows:

Table A4.2

Fields of Study	Majors Aggregated
Agriculture	Agriculture, Nutrition, Veterinary
Architecture and Design	Architecture, Design
Economics and Business	Economics, Business
Education	Communication Studies, Education, Nursing, Exercise Science
Engineering	Engineering
Humanities	Literature, Languages, Philosophy, Cultural Heritage, Art History, Music, History, Archeology
Law	Law studies
Mathematics and Statistics	Mathematics, Statistics, Physics ,Computer Science.
Medicine	Medicine
Natural Sciences	Biology, Bio Technological Sciences, Pharmacy, Environmental Studies, Geology, Natural Science, Chemistry
Social Sciences	Political Sciences, Social Sciences, Sociology, Psychology

Figure A4.1 - Original data and merging process

References for Chapter 4

Chevalier Arnaud (2011) “Subject Choice and Earning of UK graduates” IZA Discussion paper N. 5652, April 2011, Bonn, Germany.

Dolado Juan,Florentino Felgueroso and Miguel Almunia (2010) “Are men and women economists evenly distributed across fields? Some New Empirical evidence” Discussion paper, Universidad Carlos III de Madrid.

Flabbi, Luca (2011). “Gender Differences in Education, Career Choices and Labor Market Outcomes on a Sample of OECD Countries.” Organization for Economic Co-operation and Development Background Paper for the WDR 2012.

Sarah E. Turner and William G. Bowen (1999) The Changing (Unchanging) Gender Gap” Industrial and Labor Relations Review, Vol. 52, No. 2 (Jan., 1999), pp. 289-313

Chapter 5: Non Academic Skills, Choice of Partner and Peer/Teacher effects as determinant of the Gender Gap in the choice of College Major

Massimo Anelli (UC Davis)
Giovanni Peri (UC Davis and NBER)

5.1 Introduction

In the previous chapter we have established the existence of a very significant gender gap in the choice of major, directing women disproportionately into Humanities, Education, Social Sciences and Design and away from Engineering, Economics & Business and Math. We have also established that this tendency is the determinant of a large portion of the income gender differential, up to 30% of it in some specifications and of the occupation-driven wage gap, (up to 15% of it).

In this chapter we focus on the determinants of the gender gap in the choice of major. We use the more detailed information that is available to us for the 10% random sample that was interviewed, in order to shed light on some of the possible motivations for the choice of major. In particular we explore three interesting channels that have been indicated by recent literature as possible important components of the women's decision and of the gender gap itself.

First, there is a large recent literature that emphasizes the importance of non-academic skills (psychological factors) that are not easily measured in standard academic testing. They differ significantly between men and women and may be the reason for their different choice of major and labor market performance. Bertrand (2010) reviews this literature and emphasizes two interesting and important "psychological attributes" that may differ systematically between men and women. The first is the attitude towards competition and the second is the attitude towards others or altruism. While we do not have information on the psychological attitudes of individuals, directly, we have information on some of their choices that may indirectly inform us about those attitudes. In particular, we have information on sport activities and on volunteering in charities while in high school. We will use that information and the variation within gender and across genders to infer their correlation with choice of major.

The second avenue that we will follow pursues the idea that the choice of major not only affects the job and career in the labor market, but may also affect the probability of marrying a

partner with certain characteristics and income potential. The college years are those in which people form relationship that may evolve in marriage or life partnership. Some majors may signal that an individual (especially a woman) is committed to a more flexible job and hence has a stronger family orientation. In a society as Italy with a traditional view of the family (in which the woman stays at home) this may render her more valuable to the partner. Hence we will analyze whether there is a difference across majors in the expectation of the average wage that a partner may bring. Based on the expected income of the partner, conditional on choosing a college major, we evaluate whether men or women choose majors associated with higher or lower income of the partner.

Finally we investigate the role of peers and of teachers in affecting the choice of major by women and men. Our database is very rich in that it allows us to identify the peer effects in the same class and school for each individual and also the group of students who had the same set of professors across cohorts and their choice of major. Several papers have studied the effect of academic quality of peers on the performance and choices of students (e.g. Sacerdote 2001, Carrell et al 2010, De Giorgi and Pellizzari 2010). We focus here on the effect of peers in the choice of college majors and we analyze, in particular, the gender component of it. At the same time there is an interesting literature on the effect of gender of teachers on the choice of students, specifically on the likelihood that female students choose a curriculum that is more intensive in Math and Sciences (Carrell et al 2010, Hoffmann and Oreopulos 2007 Neumark and Gardecki 1998). Given the structure of class-formation in Italian schools we are able to identify a “class” effect in the choice of major, controlling for all other school characteristics and for some individual characteristics. Moreover, by virtue of the way in which class are determined we can rely on random composition of classes within high schools and random assignment to the existing professors. Due to the fact that a high school class is always together and it shares all courses and professors, this is an ideal setting for intense and prolonged interactions between the students and between students and teachers, so that the teacher and peer effects are strongest. Our analysis studies the effects of peers on men and on women choice of college major and we can also analyze the effect of common teachers on that choice.

Our analysis can be particularly important as it concerns women in role of leadership and high professional level, hence the segment of the labor market where gender equality should be easier to achieve and most valuable. As Goldin and Katz (2010) do for the U.S. we are analyzing the outcomes of the Italian "educational elite" and hence of a very important group, from the economic point of view, and a group for which the talent of men and women should be maximally valued by the Italian society and economy. The rest of the chapter is organized as follows. Section 5.2 analyzes the effect of non-academic and psychological characteristics, as they differ between men and women, on the choice of major. Section 5.3 looks at how the expected partner quality may affects the choice of major for men and women. Section 5.4 analyzes the importance of peer/teacher effects in determining the choice of major of men and women. Section 5.5 concludes the chapter.

5.2 Sport and Volunteering and their impact on the gender gap in choice of College Major

While the literature on gender gap has traditionally focused on measures of human capital and of academic quality related to activity in the classroom (as we did in the previous chapter), recent literature has emphasized some complementary aspects of skills that may be important and also explain part of the gender gap in terms of choice of major and labor market performance. One is participation to sport activities while in high school. A recent paper (Stevenson 2010) estimates the causal effect of sport activity in high school on some measures of performance in college and participation to the labor market, focusing particularly on women. Lacking a clear identification strategy to isolate the effect of sport on subsequent outcomes, however, we take the choice of sport activity in high school as more likely to reveal some characteristics of the individual that may be hard to observe in class-room performance. Those characteristics may be correlated to subsequent performance and choices and hence we consider whether they are relevant for the choice of college major. For instance people doing sport on a regular basis are likely to be more competitive than people not doing sport, or at least to handle competitive pressure better than those not involved in sport activity. In this

respect, practicing sport in high school could reveal for men and women an easier attitude towards competition and if that characteristic is important in some majors (and jobs) it may be correlated with those choices.

Sport activity such as tournaments and official league games expose the young individuals to a competitive environment, hence choosing to practice sport on a regular/competitive basis may reveal that those individuals can better handle competition. Gneezy et al. (2003), Niederle and Vestlund (2007) and (2008) show in experimental set-up that women tend to under-perform when they are in a competitive setting, and tend to shy away from competition while men tend to perform better in competition and to seek competitive arrangements. Some studies (Gneezy et al. (2003)) find that women exhibit a performance gap especially when competing against men, while others (Gneezy and Rustichini 2004) find that the gender gap in performance is highest when children compete with same gender players. Sport is a situation in which individuals will compete with others of the same gender.

We would like to explore the possibility that participation to competitive sport settings reveals that individuals are comfortable with competition (both men and women) and we would like to check whether this revealed preference, in high school, is associated with the choice of college major. Are students doing sport in high school more likely to choose highly competitive (and paid) college majors, and is this effect different for men and women? In Italy sport activity at the time of high school is not associated with the school but performed in clubs, and private gyms. Moreover there are no sport fellowships in college and hence the motivation to do sport is completely separate from the educational motivation. Our database includes information on whether the individual did sport, on a regular and agonistic way, during high school. We will use that information and check whether it affects choice of major and in particular if it affects it differentially between men and women.

Similarly, there is a (mainly experimental) literature that considers the different "social" preferences between men and women, namely the fact that women are more "socially minded" and more altruistic than men, as another possible source of different choices and different performance on the labor markets. In particular college majors such as Education,

Humanities and Social Sciences may be associated with higher concern with social goods, while majors like Economics/Business and Engineering may be associated with a more competitive and selfish view of the world. Several papers (e.g. Andreoni and Verstlund 2001, Eckel and Grossman, 1998) find that in experimental settings women are more willing to give than men. Our data include information about the participation of the student, during high school, to volunteer activities for charities (social and religious). We consider this as a choice revealing the altruistic attitudes of individuals and we see if it is correlated with the choice of major in college, and whether the correlation is different for men and women.

Table 1 shows the results of including the sport and the volunteer dummies in our regressions about college major choice. In particular the first two columns report the average marginal effects of a probit regression for the probability of choosing a high-paid major (as defined in the previous chapter and hence including Medicine, Engineering, Economics/Business and Law). Columns three and four show the correlates with the choice of a low-paid major (Education, Humanities, Art and Design). Columns 5 and 6 summarize the effects on the correlates of major-imputed wage. For each dependent variable we include one regression that omits the sport and volunteering variables but includes the other individual characteristics, including the female dummy. Then Column 7 controls for the choice of major and analyzes any additional effect (not through the choice of major) of sport and volunteering on income in 2005. Finally column 8 considers the correlation between the explanatory variables, including the sport and volunteering dummies, on the probability of employment.

The finding in Table 1 are very interesting and in large part consistent with the interpretation that high-paying majors are chosen with higher probability by individuals who are more competitive (practice sport regularly) and less altruistic (do not participate to volunteering in charities). In all specifications we control for high school grade, a gender dummy and the interaction of gender and grade, plus we also include the controls for father's wage (imputed from his occupation and sector of activity), father's education and mother's education. The coefficient on the sport dummy is positive and significant in column 2, implying a correlation of sport activity with the probability of entering a high paid major. We also find a negative (and

not significant) effect of the sport dummy on the probability of choosing a low paid major. Conversely, being involved in volunteering activities has a positive and significant correlation with the choice of low paid majors and a negative correlation with the choice of low paid majors. The combination of these two effects on major-imputed wage is shown in column 6 and implies a positive but not significant effect of doing "sport" and a negative and significant effect of volunteering on that wage.

Interestingly those effects on the choice of major are attenuated for women. The estimates on the interaction between the women dummy and the sport-dummy reveals that the effect of participating to sport activities has less of an effect on choosing high paid major for women. Even stronger is the attenuation of the effect of volunteering for women. Women who did volunteering in high school were not more likely to choose a low-paid major than women who did not. Adding the main effect and the interaction with women gives an effect that is not significantly different from 0 of volunteering on the probability of choosing a highly paid or a low paid major. Hence while sport and volunteering seem to be significantly associated with (respectively) higher and lower probability of choosing a high-paid major and lower (higher) probability of choosing a low-paid major, this correlation is stronger for men than for women.
Hence we observe an effect of sport-volunteering on the choice of major (stronger for men) and the participation to sport and to volunteering is different between men and women. In our sample 27% of men practice sport regularly, while only 14% of women does. Moreover 36% of women was involved in volunteering while only 33% of men was. In principle, therefore, the regression effects and the differences between men and women in participating to these activities could reveal preferences that help to explain the gender gap. In practice the estimates of the gender gap is statistically the same with and without these controls for all of the outcomes analyzed in columns 1-6, hence quantitatively these factors do not seem to contribute much to explain the gap.

It is also interesting to consider the correlation of sport and volunteering with the income of individuals, after we control for college major. Column 7 shows the effects of sport and

volunteering on income in 2005 once we control for college major. Hence these estimates capture the additional positive (or negative) effect on wages associated with doing sport and volunteering, not channeled through the choice of college major. The only significant additional effect is a positive effect of sport on income, even controlling for college major. In fact the point estimate of this effect is quite large (equal to a difference of 27% between those who did and those who did not do sport). Moreover this effect is only due to men, as for women the effect (sum of the main effect and of the interaction) is not significantly different from 0. Finally, in column (8) we see whether sport and volunteering affect the employment probability of individuals in our sample. Given the very high participation rate of this sample there is no discernible effect of volunteering and sport on being employed. The idea that doing sport in high school reveals, for men, some psychological traits (desire of competing, ability to handle high pressure) that are positively correlated with choosing high paid majors in college and also to other productive characteristics seem consistent with these results. On the other hand volunteering in charities reveals a preference for altruism, this seems negatively correlated, in men, to the choice of high paid majors. Both mechanisms could help explain the gender gap, however quantitatively they do not matter much.

5.3 Expected partner income as determinant of college major

The choice of major of women in our sample is hard to rationalize if one considers only the income motivation and a measure of academic quality based on high school grades. In fact women reveal a systematic strong bias against the choice of high paid, math-intensive majors for a given academic quality. One explanation is that women have different skills, in other dimensions that make them less productive within those majors. This seems contradicted, however, by the fact that also in the high paid major, conditional on choosing them, women out-perform men academically. Nevertheless in the previous section we analyzed whether non-academic skills or Psychological attitudes may give a competitive advantage in highly paid majors and hence be associated with probability of attending such majors. We focused on propensity to competition (as captured by doing a sport activity agonistically) and altruism (revealed by doing a volunteering activity).

Another explanation is that women value different outcomes than men when choosing a major, and possibly give lower weight to the monetary income that can be accessed after graduation. In Chapter 1 we have emphasized that women value flexibility in their occupation and US data show that women are much more likely to choose majors that give access to occupation where a large fraction of individuals is employed part-time. This is likely due to the fact that women, especially while in their 30's and 40's attribute high value to spending time with the family (see results in Chapter 3). Here we consider a family-related motivation that, however, is perfectly in line with women maximizing future income of their household. One possibility for the choice of major is that, while not paying high wages, enrolling and graduating from a low-paying major could increase the probability of women to find a partner (husband) and specifically it may increase the expected income of the future partner/husband. Given the preference of female for low-paid major this story would be true only if there is an "assortative matching" in terms of majors on the marriage market. If majors characterized as "typically female" (and low paid) produce higher probability to marry high earning men as partners then a rational choice of the woman could be to trade off her wage, for the wage of the potential husband. This could happen if men consider the choice of a low-paid major by a woman, especially if she has high academic quality, as a signal of her commitment to the family. Moreover if men with higher potential income prefer marrying women with stronger commitment to the family (guaranteeing higher quality of the children), then there could actually be assortative matching. In order for this to be the case, we need to observe three facts that would be consequences of this type of matching. First, women in low-paying majors should have a higher probability of being married in 2011 (controlling for age, school and individual characteristics). Second, conditional on being married, women in low-paying major should marry men (or have domestic partners) earning higher income. Third, there should be a positive gender gap, controlling for other features, in the partner-wage predicted by the major; namely, in order to offset the negative gender gap in major-imputed wages, women should have a positive gender gap in major-imputed partner wage. If these effects exist, women may be optimizing in their choice of major accounting for the total expected household income (including the one of their partner).

Column 1 and 3 of Table 2 show the average marginal effects on the probability of being married in 2011 (Probit regression) on individual controls (including parents' education, father's wage and the high school grade) and on the high and low-paid major dummies. Column 1 is limited to women and column 3 to men. Then column 2 and 4 show the correlation, separately by gender, between the same explanatory variables and the income of the partner (imputed from his/her occupation/sector of activity).

The results do not support a theory of assortative matching in marriage, nor the idea that women would like to increase their probability of marrying high-paid men, by choosing low-paid majors. In fact the probability of marriage and the average income of the partner are not correlated to the choice of a low paid major. Mother education and father education seem both correlated with higher income of the spouse, and interestingly, the academic quality of women (as measured by high school grade) has a positive effect on the expected income of the husband. For men the academic quality in high school is negatively correlated with the income of the wife, instead, but not significantly.

Table 3 is a further test that assortative matching in the marriage (in terms of income) is not a good explanation for the preference of women in the choice of low-paid majors. In that table we consider, for each major, the expected income of the partner. Then we regress this expected income of the partner on the academic quality of the individuals and on his/her background and family characteristics. If the assortative matching is at work then we should observe a positive gender gap, namely we should observe women choosing majors that increase the probability of marrying a partner with higher income (relative to the choice of men). The four specifications of Table 3, that include progressively more background controls, do not support that idea. In fact women tend to choose majors with a lower predicted average income of the partner, relative to men, although the effect is small (0.5 logarithmic points). The effect is not significantly different from 0 when controlling for family background.

Hence the choice of low paid majors with high density of women does not seem to increase the probability of women to get married, nor to find a high income partner. The negative gender

gap on women income produced by the choice of major is not made up by a positive gender gap in the expected income of partners (predicted by the choice of major). While concerns for the family can certainly enter the choice of major, improving the probability of getting married or increasing the expected income of the spouse does not seem to be what drives that choice[53].

5.4 Peer effects and Teacher effects in the College Major's choice and Gender Gap

A third important component in the choice of college major, potentially affecting men and women differentially, is the effect of their peers. If women are more socially oriented (or insecure) than men they may attach a larger weight to the opinion and choices of their peer, so that their choice of major may affect them more deeply than men. Several studies have shown relevant peer effects on several student outcomes. Sacerdote (2001) finds effects of randomly matched roommates (in Dartmouth) on the grade average and the probability of joining fraternities. He does not find any evidence of peer effect on choice of major. De Giorgi et al (2010) find a significant effect of randomly assigned class-mates in first-year college courses on the choice of major within Bocconi University (one of the universities in our sample of Milan). Carrell et al (2011) find a causal effect of peers on the level of physical "fitness" of US Air force academy students and on the school performance of those. Neither of those studies focuses, specifically on the gender dimension of peer effects.

Our data and our setting are ideal to identify peer effects. First of all we can identify individuals belonging to the same class in the high school. Classes are formed randomly in public high schools, pooling all students and drawing a certain number in each class. In part following a criterion of fairness, in part to avoid parent's pressures, this practice has been implemented since a long time in most Italian high school. The "Licei" in Milan followed such a random rule. Second, a class in an Italian high school is a set of students who share the same identical choice of courses, the same professors and does not change for 5 years (except for some attrition). Hence students in the same class are exposed to a long period of interactions at an age (12 to

[53] Let us also notice that the fertility rate of women in our sample is very low. In general women in northern Italy have had among the lowest fertility rates in the whole world. The average number of children per woman, in 2011, covering an age range between 29 and 45 was 0.71!, the median number of children was one.

17) in which peers are very important. Friendship shaped during this period, often, last for the whole life of individuals. As we know for each person in our sample the class he/she was in during high school and, therefore his/her group of peer we will analyze whether the average major choice of people in the same class (and also in the same school and cohort) affected the choice of college major. Moreover we will also analyze whether peer of the same sex had a stronger effect on the choice, and whether the peer effect was stronger on men or women. While the gender difference in peer effects in the choice of major is unlikely to explain much of the overall difference in the choice between men and women, it may however provide very useful information for policy considerations. For instance if women are more affected by peers (independently of their gender) in the choice of their major, then it is likely that a high degree of segregation between women and men in classes would generate a lower probability of women to choose highly paid majors, as they find themselves in classes with high percentage of other women who will choose low-paid majors. On the other hand if men are not much affected by the choice of peers, especially those of the opposite gender, they will not reduce the probability of going to a high-paid major when they are in class with more women.

One important issue when estimating peer effects on choices is the fact that they are hard to distinguish from common fixed effect that affect the class. In our analysis we can control for fixed effects at the school-year level and exploit the random distribution of students across classes, in order to differentiate peer effects from common school-year shocks. Moreover the most relevant common fixed effect in a class is the set of teachers that the class gets. However, as teachers are common to all cohorts within a class in a school, we can control for a teacher-school fixed effects and see if what affects the choice of major are the peers or the set of teachers.

In order to check that classes are generated randomly from the distribution of students attending a certain school, and not by grouping students of similar background and quality together, we check that the dispersion of students' characteristics within classes in a school is similar to the dispersion of that characteristic in the school overall. Figure 1 shows for each

school and year, the standard deviation of the students' house value per square meter (a proxy of their family income) in a school (horizontal axis) and on average in all classes of the school (vertical axis). If the dispersion is the same this means that classes are assembled randomly. If the average class dispersion is systematically smaller than the school's (points below the 45% line) then we are in presence of sorting across classes by income, larger the larger is the distance. Except for three points, the observations are all very close to the 45 degree line, revealing essentially no sorting across classes within a high school on the family income dimension.

In order to test whether the college-choice of peers in the class or in the school helps explaining the choice of high paid and low paid college major we include in table 4 two variables that captures the share of peers in the class and in the whole school (same grade) that chose high paid or low paid majors. We also include the same shares for the peers in the whole school-grade. We control for the gender dummy, the high school grade, their interaction and the values of real estate in the area where the individual lived. Columns 1-3 of Table 4 report the basic results, showing first the estimate of the class-peer variable. In column 1 the dependent variable is the major-imputed wage, in column 2 is the dummy for choosing a high-paid major, in specification 3 it is a dummy for choosing a low-paid major. Columns 2 and 3 show the average marginal effect from a probit estimation. We focus on the coefficients on the "share of class-peers in high paid majors" and "share of class-peers in low paid majors". First of all notice that the effect of peers within the same school are not significant (except once and only marginally) and often have the wrong sign. To the contrary the effect of class-peer choice is almost always significant, it has always the expected sign and it is quantitatively large. For instance, the effect of having all class-peers choose a high-paid major, relative to having none, is to increase the probability of choosing a high paid major by 21%, decreases the probability of choosing a low paid major by 7.7% and increases the major-imputed wage by 11.7%. Conversely increasing the share of people in the same class who chose a low paid major from 0 to 1, decreases the major-imputed wage by 6.8%, decreases the probability of attending a high-paid major by 3.8% and increases the probability of attending a low paid major by 15.6%.

To provide a reference for the magnitude of this effect, we can compare it with the effect of the exit-test score on the choice of major (or of the gender gap). Having all class-mates attending the high paid majors (relative to having no class-mates) increases the probability of attending high paid majors of an individual almost by the same difference implied by graduating at the top or at the bottom of the class. The effect of peers on major-predicted wage is also similar (+12%) to the effect of graduating at the top versus the bottom of the class. A woman in a class of individuals who all choose high paying majors has somewhat higher probability of choosing a high paid major than a man graduating from a class with no students going to high paid major (for identical academic quality and other characteristics).

Specifications 4 to 6 include common effects at the school/year level in order to control for potential factors at which all individuals in a school are exposed over time, such as changes in type of students admitted, change in information about a major due to a school campaign and so on. The class-peer effects become somewhat smaller, but they are still quite significant with the largest being the effect of peers who go to high paid college major on the probability to go to a high paid college major.

Table 5 includes the control for the most relevant common effect at the class level that could be confused with a peer effect generating some omitted variable bias. This is the set of professors that a class has. As professors have a rather long tenure in these schools, because teaching in a high school within the city of Milan is considered as a very good and stable job, they do not vary much between years in a school. Moreover they are assigned to one specific class every year and we have information on which classes shared the same set of professors year after year. Hence we can include “school/teachers” fixed effects controlling for the same group of professors by school. Table 5 shows the same regression as in Table 4, including the teacher-school dummies. While the class peer-effects on the choice of college major are robust to this controls (column 1 to 3) when we also include the school by year effects, the class-peer effects become smaller and not very significantly different from 0. Hence the “class” effect may contain an important component that is related to teachers, rather than to peers.

While identifying peer effects and the effect of teachers in high school on the choice of college we are interested in the gender dimension of it. In particular: are women more sensitive to the influence of peers than men, and are individuals, in general, more sensitive to the influence of peer of the same gender? We explore these questions in Table 6, 7 and 8.

In Table 6 we estimate separately the effect of the share of class-mates choosing high paid and low paid majors on the choice of women (columns 1-3) and men (columns 4-6). While both genders seem sensitive to the peer effects, women seem to be more sensitive (although not significantly so) to peers. The share of classmates choosing a high-paid major and the share of those choosing low paid major have larger effects on the choice of women than of men. This can be seen considering the estimates in the first row of Table 6. Increasing the percentage of classmates choosing a high-paid major by 10% increases the probability of choosing a high paid major by 2.3% for women, but only by 1.8% for men. Similarly an increase of classmates choosing a low-paid major, by 10% would increase by 1.3% the probability that a men chooses a low paid major, but it will increase by 1.6% the probability that a woman chooses it.

Table 7 and 8 show, respectively, for women and for men the effect of the share of the female/ male classmates choosing high-paid and low-paid major, so that we can differentiate between the peer effects from same gender and opposite gender classmates. In columns 1-3 of table 7 women show to be highly sensitive to the choice of major made by both their female and male peers. The probability of choosing a high paid major increases by 15% going from no female choosing high paid majors to having only female classmates all choosing a high paid majors and +6.4% for the same measure of the share of male classmates choosing high paid majors. The same is not true for men: indeed the share of female classmates choosing high paid majors does not have any impact on the choice of major made by male students while the peer effect of male classmates is still significant. Hence results suggest that women are more affected by their peers (of both sexes) relative to men. Results in columns 1 and 3 for major-imputed log wage and probability of choosing a low paid major reproduce similar patterns.

Interestingly, when introducing school/year fixed effects (columns 4-6) the magnitude of the impact of the share of female peers choosing high paid majors on female is drastically reduced and it is not statistically significant, while the effect of the share of male peers choosing high paid majors is still present and significant. For men the same specifications (columns 4-6) show that men are not affected at all by other peers (no matter the gender) in terms of the probability of choosing a high paid major and in terms of major-predicted wage.

Overall, the peer effect analysis shows that women are more sensitive to the choice of their class-mates, and even more interestingly, they are very sensitive to the major choice of their male class mates. To the contrary men are not. One interesting implication of this is that, simply by adding more women in classes with a majority of men (for instance by encouraging women to attend the Liceo Scientifico) would have a positive impact on the probability of those women of choosing a math-intensive high paid major, while it would have no negative effect on men.

5.5 Conclusions

In this chapter we have taken as given the fact that the choice of college major has an important effect on the income gender gap on the labor market and we have focused on some of the potential determinants of this gap. We have identified and analyzed three factors that the previous literature has singled out as possibly relevant in determining the gender gap and we have studied them in the context of the choice of major for highly educated young Italians. The first is the fact that women and man have different psychological attitudes towards competition and altruism. Being less competitive and more altruistic may generate a preference in women driving them to more “socially oriented” majors (Education, Social Sciences) and less towards profit-oriented ones (Engineering, Business). The regression analysis uses participation to sports and to volunteer activity during high school as revealing whether a person is more altruistic or more competitive. We find that doing sport and not doing volunteering, controlling for academic quality, are predictors of a higher probability of enrolling in a high paid major. This

relationship is much stronger for men than for women. Hence the psychological explanation goes part way to explain the different choice of majors.

Then we consider the possibility that women may choose a major accounting for its effect on the probability of marriage and the average income of the spouse. If this is the case, and high-income men are more likely to marry women from low-paid major, then maximizing the income of the household could imply, for a woman, the choice of a low-paid major. Our results do not support this explanation, because we find that choosing low paid major neither increases the probability of marriage for women (given other characteristics) nor affects the expected income of the future spouse.

Finally we explore the possibility that women are more affected by their peers (or by their same-gender peers) in the choice of college relative to men. We find some evidence that this is true. In particular the share of their high school classmates going to high-paid majors affects significantly their probability of choosing a high-paid major, independently from the gender of their peers. To the contrary men are subject to less significant peer effects and those seem to be driven mainly by the influence of their male peers.

References for Chapter 5

Andreoni, James, Vesterlund, Lise, 2001. Which is the fair sex? Gender differences in altruism. Quarterly Journal of Economics 116 (1), 293–312.

Carrell Scott, Marianne Page and james West (2010) Sex and Science: How Professor Gender Perpetuates the Gender Gap," *Quarterly Journal of Economics*, Volume 125, Issue 3, August 2010.

Carrell Scott E. & Mark Hoekstra & James E. West, 2010. "Is Poor Fitness Contagious? Evidence from Randomly Assigned Friends," NBER Working Papers 16518, National Bureau of Economic Research, Inc.

De Giorgi, Giacomo, Michele Pellizzari, and Silvia Redaelli. 2010. "Identification of Social Interactions through Partially Overlapping Peer Groups." *American Economic Journal: Applied Economics*, 2(2): 241–75.

Eckel, Catherine C., Grossman, Philip J., 1998. Are women less selfish than men? Evidence from dictator experiments. Economic Journal 108 (448), 726–735.

Gneezy, Uri, Rustichini, Aldo, 2004. Gender and competition at a young age. American Economic Review 94 (2), 377–381.

Gneezy, Uri, Niederle, Muriel, Rustichini, Aldo, 2003. Performance in competitive environments: gender differences. Quarterly Journal of Economics. 118, 1049–1074.

Goldin, Claudia and Lawrence F. Katz (2010) “Transitions: Career and Family Life Cycles of the Educational Elite”. *American Economic Review*: Papers & Proceedings 2008, 98:2, 363–369

Hoffmann, Florian and Philip Oreopoulos, “A Professor Like Me: The Inuence of Instructor Gender

on College Achievement," Working Paper 13182, National Bureau of Economic Research, 2007.

Neumark, David and Rosella Gardecki, (1998) “Women Helping Women? Role Model and Mentoring Effects on Female Ph.D. Students in Economics," Journal of Human Resources, 33 (1998), 220,46.

Niederle, Muriel, Vesterlund, Lise, 2007. Do women shy away from competition? Do men compete too much? Quarterly Journal of Economics 122 (3), 1067–1101.

Niederle, Muriel, Vesterlund, Lise, 2010. Explaining the gender gap in math test scores: the role of competition. Journal of Economic Perspectives 24 (2), 129–144.

Sacerdote, Bruce 2001. "Peer Effects With Random Assignment: Results For Dartmouth Roommates,"The Quarterly Journal of Economics, MIT Press, vol. 116(2), pages 681-704, May.

Stevenson Betsey (2010) “Beyond the Classroom: Using Title IX to measure the returns to high school sports” The Review of Economics and Statistics, May 2010, 92(2): 284–301.

Weinberger, Catherine J.(1998) “Race and Gender Wage Gaps in the Market for Recent College Graduates," Industrial Relations, 37, 67-84

Weinberger, Catherine J.(1999) “Mathematical College Majors and the Gender Gap in Wages," Industrial Relations, 38, 407-413.

Figure 1

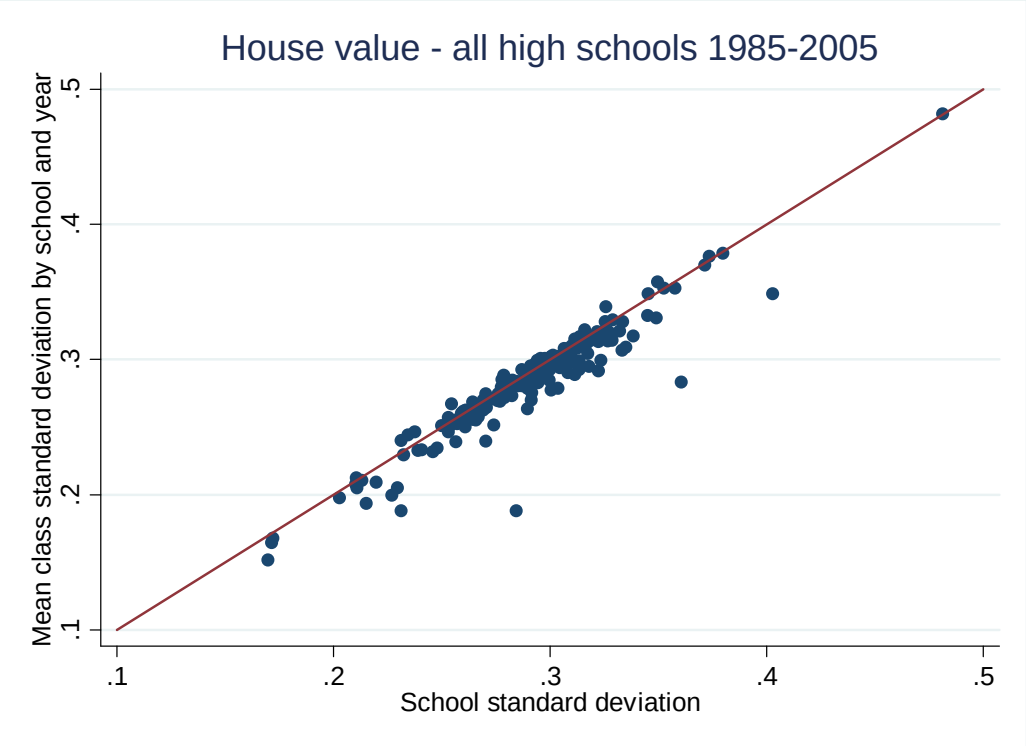

House value - all high schools 1985-2005
Mean class standard deviation by school and year
.1
.2
.3
.4
.5
School standard deviation
.1
.2
.3
.4
.5

Table 1 - Sport and Volunteering

Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Method	Probit	Probit	Probit	Probit	OLS	OLS	OLS	OLS
Dependent Variable	HP=1	HP=1	LP=1	LP=1	log of wage imputed by major	log of wage imputed by major	Log of actual wage	Employed=1
Female	-0.146***	-0.158***	0.084**	0.106***	-0.111***	-0.128***	-0.172	-0.003
	(0.033)	(0.036)	(0.033)	(0.037)	(0.020)	(0.023)	(0.123)	(0.023)
Fem X HS Grade	-0.111*	-0.124*	0.159**	0.168**	-0.054	-0.062	-0.214	-0.039
	(0.065)	(0.065)	(0.069)	(0.070)	(0.041)	(0.041)	(0.190)	(0.040)
HS Grade	0.219***	0.232***	-0.174***	-0.184***	0.114***	0.122***	0.275**	0.043*
	(0.049)	(0.049)	(0.055)	(0.056)	(0.029)	(0.030)	(0.131)	(0.024)
Sport		0.055*		-0.017		0.024	0.274***	-0.008
		(0.031)		(0.035)		(0.020)	(0.087)	(0.018)
Sport X Fem		-0.020		-0.029		0.010	-0.228	0.006
		(0.052)		(0.046)		(0.031)	(0.153)	(0.033)
Volunteering		-0.066**		0.062**		-0.058***	-0.121	0.013
		(0.029)		(0.026)		(0.018)	(0.084)	(0.017)
Volunt X Fem		0.076*		-0.065*		0.060**	-0.060	-0.012
		(0.042)		(0.034)		(0.027)	(0.119)	(0.022)
Family controls	X	X	X	X	X	X	X	X
Year Dummies	X	X	X	X	X	X	X	X
School Dummies	X	X	X	X	X	X	X	X
Major Dummies							X	
Constant					9.913***	9.936***	9.796***	0.790***
					(0.167)	(0.166)	(0.787)	(0.152)
Sample	Enrolled & Interviewed	Enrolled & Interviewed	Enrolled & Interviewed	Enrolled & Interviewed	Enrolled & Interviewed	Enrolled & Interviewed	Enrolled & Interviewed	Enrolled & Interviewed
Observations	2,261	2,261	2,261	2,261	2,261	2,261	1,234	2,261
R-squared					0.105	0.109	0.162	0.184

Note: Standard errors clustered at School/Year level in parenthesis. Family controls: father wage, father education, mother education.
*** p<0.01, ** p<0.05, * p<0.1.

Table 2 - Marriage and Partner of Income

	Women		Men	
Specification	(1)	(2)	(3)	(4)
Method	Probit	OLS	Probit	OLS
Dependent Variable	Married=1	Log wage of partner	Married=1	Log wage of partner
HS Grade	-0.044	0.049*	0.031	-0.022
	(0.047)	(0.027)	(0.052)	(0.030)
Father wage	0.082*	0.111***	0.038	0.054
	(0.048)	(0.025)	(0.050)	(0.034)
Father education	-0.042*	0.023*	0.005	0.035*
	(0.023)	(0.013)	(0.023)	(0.018)
Mother education	0.032	0.130***	0.014	0.119***
	(0.020)	(0.014)	(0.021)	(0.016)
High paid major=1	0.011	-0.008	0.046	0.027
	(0.037)	(0.020)	(0.032)	(0.020)
Low paid major=1	0.001	-0.029	-0.003	0.016
	(0.037)	(0.019)	(0.041)	(0.023)
Year Dummies	X	X	X	X
School Dummies	X	X	X	X
Constant		6.008***		6.465***
		(0.190)		(0.234)
Sample	Enrolled & Interviewed	Enrolled & Interviewed	Enrolled & Interviewed	Enrolled & Interviewed
Observations	1,153	703	1,041	643
R-squared		0.365		0.263

Note: Standard errors clustered at School/Year level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1.

Table 3 - Testing "assortative" matching in marriage

Specification	(1)	(2)	(3)	(4)
Dependent Variable	Log wage of partner imputed by major			
Female	-0.005***	-0.005**	-0.004	-0.006
	(0.001)	(0.002)	(0.003)	(0.007)
Fem X HS Grade		-0.001	-0.001	-0.002
		(0.004)	(0.005)	(0.005)
HS Grade		0.008***	0.007*	0.007*
		(0.003)	(0.003)	(0.003)
Father wage			-0.001	-0.001
			(0.002)	(0.002)
Father education			0.001	0.002
			(0.001)	(0.001)
Mother education			0.004***	0.003**
			(0.001)	(0.002)
Mother educ X Fem				0.001
				(0.002)
Father educ X Fem				-0.000
				(0.002)
Year Dummies	X	X	X	X
School Dummies	X	X	X	X
Constant	7.335***	7.331***	7.321***	7.321***
	(0.005)	(0.005)	(0.017)	(0.017)
Sample	Enrolled & Interviewed	Enrolled & Interviewed	Enrolled & Interviewed	Enrolled & Interviewed
	1,514	1,514	1,346	1,346
	0.047	0.055	0.078	0.078

Note: OLS, Standard errors clustered at School/Year level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1.

Table 4 - Class- and School-Peer Effects

Specification	(1)	(2)	(3)	(4)	(5)	(6)
Method	OLS	Probit	Probit	OLS	Probit	Probit
Dependent Variable	log of wage imputed by major	HP=1	LP=1	log of wage imputed by major	HP=1	LP=1
Classmates in HP	0.117***	0.213***	-0.077***	0.072***	0.107***	-0.053*
	(0.020)	(0.034)	(0.027)	(0.022)	(0.037)	(0.029)
Classmates in LP	-0.068**	-0.038	0.156***	0.006	0.009	0.022
	(0.031)	(0.040)	(0.044)	(0.035)	(0.045)	(0.051)
Female	-0.138***	-0.161***	0.128***	-0.139***	-0.165***	0.128***
	(0.008)	(0.011)	(0.010)	(0.008)	(0.011)	(0.010)
Fem X HS Grade	-0.017	-0.134***	0.093***	-0.013	-0.126***	0.091***
	(0.015)	(0.023)	(0.021)	(0.015)	(0.023)	(0.022)
HS Grade	0.126***	0.289***	-0.171***	0.127***	0.291***	-0.174***
	(0.010)	(0.018)	(0.018)	(0.010)	(0.018)	(0.018)
Schoolmates in HP	0.015	0.093	0.088			
	(0.046)	(0.095)	(0.057)			
Schoolmates in LP	0.125*	0.161	-0.197*			
	(0.072)	(0.103)	(0.119)			
House value controls	X	X	X	X	X	X
Year Dummy	X	X	X			
School Dummy	X	X	X			
School X Year Dummies				X	X	X
Constant	10.407***			10.438***		
	(0.023)			(0.012)		
Sample	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled
Observations	22,873	22,873	22,864	22,873	22,873	22,864
R-squared	0.086			0.093		

Note: Standard errors clustered at School/Year level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1.

Table 5 - Teacher vs. Peer Effects

Specification	(1)	(2)	(3)	(4)	(5)	(6)
Method	OLS	Probit	Probit	OLS	Probit	Probit
Dependent Variable	log of wage imputed by major	HP=1	LP=1	log of wage imputed by major	HP=1	LP=1
Classmates in HP	0.090***	0.166***	-0.052*	0.042	0.048	-0.031
	(0.024)	(0.038)	(0.030)	(0.026)	(0.043)	(0.034)
Classmates in LP	-0.055*	-0.027	0.126***	0.027	0.024	-0.022
	(0.033)	(0.043)	(0.044)	(0.038)	(0.048)	(0.052)
Female	-0.140***	-0.167***	0.130***	-0.141***	-0.169***	0.129***
	(0.008)	(0.012)	(0.011)	(0.008)	(0.012)	(0.011)
Fem X HS Grade	-0.011	-0.125***	0.089***	-0.009	-0.119***	0.088***
	(0.015)	(0.023)	(0.022)	(0.015)	(0.023)	(0.022)
HS Grade	0.123***	0.286***	-0.167***	0.124***	0.288***	-0.171***
	(0.010)	(0.018)	(0.018)	(0.010)	(0.018)	(0.019)
Schoolmates in HP	0.021	0.160*	0.103*			
	(0.047)	(0.088)	(0.061)			
Schoolmates in LP	0.157**	0.235**	-0.197			
	(0.076)	(0.101)	(0.124)			
House value controls	X	X	X	X	X	X
Year Dummy	X	X	X			
School Dummy	X	X	X			
School X Year Dummies				X	X	X
School X Teachers D.	X	X	X	X	X	X
Constant	10.277***			10.214***		
	(0.038)			(0.029)		
Sample	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled
Observations	22,298	22,298	22,298	22,298	22,298	22,298
R-squared	0.087			0.095		

Note: Standard errors clustered at School/Year level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1.

Table 6 - Peer Effects of Classmates and Schoolmates by Gender - Women

	Women			Men		
Specification	(1)	(2)	(3)	(4)	(5)	(6)
Method	OLS	Probit	Probit	OLS	Probit	Probit
Dependent Variable	log of wage imputed by major	HP=1	LP=1	log of wage imputed by major	HP=1	LP=1
Classmates in HP	0.139***	0.238***	-0.088**	0.092***	0.184***	-0.070*
	(0.032)	(0.050)	(0.042)	(0.028)	(0.049)	(0.037)
Classmates in LP	-0.067	-0.019	0.162**	-0.062	-0.049	0.137***
	(0.043)	(0.053)	(0.063)	(0.038)	(0.057)	(0.045)
Schoolmates in HP	-0.002	0.108	0.122	0.004	0.057	0.060
	(0.069)	(0.131)	(0.086)	(0.061)	(0.117)	(0.077)
Schoolmates in LP	0.057	0.020	-0.128	0.184*	0.291*	-0.258**
	(0.107)	(0.145)	(0.160)	(0.094)	(0.154)	(0.131)
House value controls	X	X	X	X	X	X
HS grade control	X	X	X	X	X	X
Year Dummy	X	X	X	X	X	X
School Dummy	X	X	X	X	X	X
Constant	10.271***			10.216***		
	(0.037)			(0.045)		
Sample	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled
Observations	11,747	11,747	11,744	11,126	11,126	11,120
R-squared	0.038			0.039		

Note: Standard errors clustered at School/Year level in parenthesis, *** p<0.01, ** p<0.05, * <0.1.

Table 7 - Peer effects from same gender vs. opposite gender – Women

Specification	(1)	(2)	(3)	(4)	(5)	(6)
Method	OLS	Probit	Probit	OLS	Probit	Probit
Dependent Variable	log of wage imputed by major	HP=1	LP=1	log of wage imputed by major	HP=1	LP=1
Fem classmates in HP	0.103***	0.156***		0.020	0.019	
	(0.027)	(0.038)		(0.029)	(0.043)	
Male classmates in HP	0.052***	0.064**		0.056***	0.069**	
	(0.018)	(0.026)		(0.021)	(0.029)	
Fem schoolmates in HP	0.035	0.119				
	(0.070)	(0.107)				
Male schoolmates in HP	-0.052	0.026				
	(0.053)	(0.077)				
Fem classmates in LP			0.108***			-0.039
			(0.042)			(0.047)
Male classmates in LP			0.104***			0.113***
			(0.032)			(0.036)
Fem schoolmates in LP			-0.087			
			(0.107)			
Male schoolmates in LP			-0.205**			
			(0.102)			
House value controls	X	X	X	X	X	X
HS grade control	X	X	X	X	X	X
Year Dummy	X	X	X			
School Dummy	X	X	X			
School X Year Dummies				X	X	X
Constant	10.302***			10.302***		
	(0.034)			(0.016)		
Sample	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled
Observations	11,584	11,584	11,581	11,584	11,584	11,581
R-squared	0.039			0.057		

Note: Standard errors clustered at School/Year level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1.

Table 8 - Peer effects from same gender vs opposite gender - Men

	(1)	(2)	(3)	(4)	(5)	(6)
Method	OLS	Probit	Probit	OLS	Probit	Probit
Dependent Variable	log of wage imputed by major	HP=1	LP=1	log of wage imputed by major	HP=1	LP=1
Fem classmates in HP	0.015	0.003		0.025	0.018	
	(0.017)	(0.028)		(0.019)	(0.031)	
Mal classmates in HP	0.072***	0.135***		0.006	0.019	
	(0.020)	(0.032)		(0.022)	(0.035)	
Fem schoolmates in HP	0.030	0.087				
	(0.052)	(0.082)				
Mal schoolmates in HP	-0.053	-0.031				
	(0.055)	(0.096)				
Fem classmates in LP			0.041			0.051
			(0.027)			(0.032)
Mal classmates in LP			0.047			-0.099**
			(0.041)			(0.046)
Fem schoolmates in LP			-0.044			
			(0.079)			
Mal schoolmates in LP			-0.214			
			(0.133)			
House value controls	X	X	X	X	X	X
HS grade control	X	X	X	X	X	X
Year Dummy	X	X	X			
School Dummy	X	X	X			
School X Year Dummies				X	X	X
Constant	10.250***			10.313***		
	(0.040)			(0.005)		
Sample	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled	Enrolled
Observations	10,952	10,952	10,946	10,952	11,584	10,946
R-squared	0.037			0.054		

Note: OLS, Standard errors clustered at School/Year level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1.