OCCUPATIONAL GENDER SEGREGATION AND GENDER WAGE GAP IN SWITZERLAND

Ilaria Finzi

Submitted for the degree of Ph.D. in Economics at
Faculty of Economics
University of Lugano
Lugano, Switzerland

Thesis Committee:
Prof. Patrick Gagliardini, University of Lugano
Prof. Rico Maggi, University of Lugano
Prof. Yves Flückiger, University of Geneva

5 July 2007

To my family

Contents

List of Figures

List of Tables

Acknowledgments

During the course of my thesis work, there were many people who were instrumental in helping me. Without their guidance, help and patience, I would have never been able to accomplish the work of this thesis. I would like to take this opportunity to acknowledge some of them.

I would like to express my gratitude to my supervisor, Prof. Patrick Gagliardini, whose expertise, understanding, and patience, added considerably to my Ph.D. experience. I appreciate his vast knowledge and skill in many areas, and his assistance. I would like to thank the other members of my committee, Prof. Rico Maggi for the assistance he provided, and Prof. Yves Fluckigher from the Faculty of Economics in Geneva for taking time out from his busy schedule to serve as my external reviewer.

A very special thanks goes out to Prof. Pietro Balestra, he was the one professor/teacher who truly made a difference in my life. It was under his tutelage that I developed a focus and became interested in econometrics. He provided me with direction, technical support and became more of a mentor than a professor.

I would also like to thank my family for the support they provided me through my entire life, and in particular Siro deserves an award for his patience and understanding during the writing of this thesis.

I must acknowledge my friends Claudio, Ettore, Flaminio, Francesca, Maura, Miriam, Oscar and Sandro without whose love, encouragement and editing assistance, I would not have finished this thesis.

I must also acknowledge my friend Massimo, without whose motivation and encouragement I would not have considered a Ph.D. career in economic research.

Introduction

It is widely recognized in the economic literature that women suffer wage discrimination. The unexplained gender wage gap consists in the discrepancy between what females would earn with male characteristics and what women actually earn.

Gary Becker (1957) started the economic analysis of labour market discrimination. He developed the idea that employers and costumers may have prejudices against members of particular (minority) groups and introduced the concept of tastes for discrimination. These discriminatory tastes influence the behaviour of people and have an impact on earnings and employment chances of the discriminated group.

Common to the discrimination studies is the comparison between the income measures of employees in a protected class and those of employees not in the class, to detect potential discrimination. In our case, the protected group refers to males. Most of the existing studies are based on a sample of male and female workers and attempt to explain their wages by their observable characteristics, such as accumulated human capital (years of schooling and years of work experience, for instance) and other factors believed to influence wages, such as marital status and labour supply. The remaining portion of the wage gap is then often attributed to employer discrimination.

One natural candidate to explain at least part of the unobserved difference in wages across gender is occupational segregation. Researchers usually distinguish between labour supply and labour demand factors when explaining occupational segregation by gender. Factors related to labour supply generally focus on why women "prefer" certain types of occupation (for example, women may prefer works with flexible hours in order to allow time for child care). Explanations related to labour demand focus on why employers generally "prefer" to hire women or men for particular occupations, and why women and men have different opportunities for promotion and career development within firms.

In chapter 1, I present a detailed review of the literature on gender wage gap measurement and I highlight the contributions of my work. The main part of the thesis consists of Chapters 2 and 3.

In chapter 2, we are interested in the analysis of horizontal occupational segregation by gender in Switzerland. We try to answer the following questions: Does segregation arise because "female" occupations have financial advantages for women planning to spend some time as homemakers, as human-capital theory claims? Do "male" occupations have more onerous working conditions that explain their higher earnings, as the neoclassical notion of "compensating differentials" suggests? To answer these questions we use a cross-sectional survey in year 2000. We extend the linear model by introducing a new variable, FEM, which represents the percentage of the female labour force employed in an occupation. This new variable is included in order to explain the impact of the density of females in a certain occupation on individual wages (see England et al. 1988, Sorensen 1990 for similar studies).

In chapter 3 we try to give some answers to an interesting question concerning the vertical occupational segregation: Are women more segregated into lower level of the occupational hierarchy? To answer this question, we use a binomial Probit model with random effects to predict the career stages and the earnings of men and women with a panel data, for a period of 5 years (1999-2003), using data from the *Enquête Suisse sur la Popolation Active*.

Chapter 1

A discrimination measurement

1.1 Introduction

In the first part of this chapter I present an overview on the labour economics literature and on econometric models to detect wage discrimination. In particular, I first briefly review the literature from Becker (1957) to Oaxaca (1973), analysing the different economic models. I then present a different approach used by Polachek and Kim (1994) and developed by others economists, namely panel data methods. In the second part of this chapter, I expose the purpose of my analysis, and finally in the last part I describe the data used in my analysis.

1.2 Reviewing the literature

The most used approach in measuring the extent of gender wage discrimination is based on the human capital theory of wage determination. According to human capital theory, wages depend on the productivity of the individual. In a non-discriminatory environment, the observed male-female wage differentials should be due to differences in productivity between men and women only. Gender wage discrimination takes place when equally productive workers are paid different wages. In other words, when there is discrimination, male-female wage differentials cannot be explained only in terms of differences in productivity. In empirical analysis, productivity is not observed by researchers and thus measures of discrimination are usually adjusted for all observable characteristics that might be expected to affect productivity.

The findings of several empirical studies point out that part of wage differential is due to differences in objective characteristics such as education and work experience, while a part remains even when male-female differences in these traits are controlled for. However, the quantitative dimensions of the various causes of unequal wages are not well known.

Gary Becker (1957) was the first to analyse economically the labour market discrimination. He developed the idea that employers and customers may have prejudices against members of particular (minority) groups and introduced the term of tastes for discrimination. These discriminatory tastes influence the behaviour of people; they do have an impact on earnings and employment chances of the discriminated group. The discrimination coefficient captures costs associated with discriminatory tastes, and suggests that employers with discriminatory tastes are willing to trade reduced profits for fewer minority employees. The way that minority and nonminority workers enter into the production function, and the way they enter in to the utility function, will obviously have implications on the interpretation of the resulting wage differentials.

For simplicity, it is assumed that male labour and female labour are perfect substitutes. That is, men and women are equally productive and thus they deserve an equal wage in absence of discrimination. However discriminatory tastes may exist among employers, co-workers and/or customers.

Employers with tastes for discrimination against women will hire women workers only at a wage discount that is sufficiently large to compensate them for the disutility of employing women. Becker (1957) also showed that even if employers themselves have no tastes for discrimination against women, profit-maximizing behaviour by employers may result in gender discrimination if the employees or the customers have such discriminatory tastes. Indeed, male employees with tastes for discrimination against women will work with them only at a wage premium that is sufficient to compensate them for the disutility of female co-workers. Similarly, customers with tastes for discrimination against women will buy products and services produced or sold by women only at a lower price. Intuitively, we would expect this type of discrimination to be more important in sales or service occupation where face-to-face contact with the customer occurs. As a consequence of co-worker or customer discrimination, employers may, in certain circumstances, discount female wages to compensate for the higher costs or lower revenues they attend in employing women.

The basic idea stated by Becker (1957) is that market discrimination is defined by a comparison of the wage rates of two groups, f and m, as they are actually observed, and as they would be observed in the absence of discrimination. For example, if f and m are perfect substitutes in production, in the absence of discrimination f and m would have the same wage rates. In this case, the difference between the wage of the two groups f and m would be a measure of discrimination.

More generally, if the observed wages of groups f and m are π_f and π_m respectively, and if they would be π_f^0 and π_m^0 in the absence of discrimination, then the percentage market discrimination against group f is

$$D = \left[\frac{\pi_m}{\pi_f} - \frac{\pi_m^0}{\pi_f^0} \right] / \frac{\pi_m^0}{\pi_f^0} \tag{1.1}$$

D is the percentage due to discrimination.

Continuing discrimination can also be due to certain beliefs of employers regarding the productivity of particular groups. The theory of statistical discrimination argues that hiring or promoting decisions may not only be based on information about skills and qualification. Employers, who are not assumed to have a taste for discrimination, may also use easily observable characteristics like race or gender (Arrow, 1972). Such behaviour is clearly discriminatory but has even more serious consequences if discriminated people react in the expected way, i.e. do quit jobs more often and becomes less productive. Discrimination persists since employers are confirmed of their expectations and see no need to change then (Blau, Ferber, and Winkler, 1998; Becker 1992).

After Becker, Oaxaca (1973) and Blinder (1973) were among the first economists to treat the problem of wage discrimination from an econometric point of view. Over the last 30 years, the approach to explain wage differentials introduced into the literature by these two researchers has been intensively used and further developed.

Oaxaca (1973) and Blinder (1973) suggest that the average wages of two groups (which can be distinguished by gender, race, country of origin and others) can be decomposed into two parts. The first part is due to differences in the observable characteristics that measure both productivity (skill) and endowment. The second part is due to the different treatment of the two groups in the labour market (what is often referred to as the discrimination component). In this analysis I refer to the group with the greater mean wage (conditional on the observable individual characteristics) as advantaged (a) and the other group as disadvantaged (d).

The decomposition of wage differences was construct by using the measure of discrimination, introduced by Becker (1971), defined as the difference between the observed wage ratio and the wage ratio that would prevail in the absence of discrimination. From equation (1.1), we have

$$\left(\frac{\pi_a}{\pi_d}\right) = \left(\frac{\pi_a^0}{\pi_d^0}\right) \tag{1.2}$$

in the absence of discrimination $(D = 0)$[1]. Oaxaca (1973), and then Cotton (1988), expressed the formulation given in (1.1) in logarithmic form as

$$\ln \pi_a - \ln \pi_d = \ln \pi_a^0 - \ln \pi_d^0 + \ln(D+1) \tag{1.3}$$

where the first term on the right hand side (the difference in the logs of the marginal productivities) is due to differences in productivity of the two groups and the second term on the right hand side $\ln(D+1)$ is due to discrimination. Oaxaca (1973) showed that, since regression lines must pass through the variables' means, separate linear models of the log wage specification can be estimated for the disadvantaged group

$$\ln(\pi_d) = \bar{X}_d' \widehat{\beta}_d, \tag{1.4}$$

and for the advantaged group with

$$\ln(\pi_a) = \bar{X}_a' \widehat{\beta}_a. \tag{1.5}$$

The formulation given in (1.4) and (1.5) follows Neumark's (1988) notation where $\bar{X}_a'$ and $\bar{X}_d'$ are vectors containing the means of the variables which are presumed to impact productivity (and subsequently wages) and $\widehat{\beta}_a$ and $\widehat{\beta}_d$ are the estimated coefficients.

The estimates can then be combined in the following way:

$$\ln \pi_a - \ln \pi_d = \bar{X}_a' \widehat{\beta}_a - \bar{X}_d' \widehat{\beta}_d. \tag{1.6}$$

Empirical work using (1.6) has been done using two decompositions. If $\Delta \bar{X}' = \bar{X}_a' - \bar{X}_d'$ and $\Delta \widehat{\beta} = \widehat{\beta}_a - \widehat{\beta}_d$, then (1.6) becomes either,

$$\ln \pi_a - \ln \pi_d = \Delta \bar{X}' \widehat{\beta}_a + \bar{X}_d' \Delta \widehat{\beta} \tag{1.7}$$

[1]Equation (1.2) also follows from the usual cost minimization problem and $\frac{\pi_a}{\pi_d}$ to the ratio of the marginal productivities.

or

$$\ln \pi_a - \ln \pi_d = \Delta \bar{X}' \widehat{\beta}_d + \bar{X}_a' \Delta \widehat{\beta} \tag{1.8}$$

The Oaxaca model decomposes wage differentials into a part due to differences in average productivities (first terms in the right hand side of (1.7) and (1.8)) and a part due to different wage structures (second terms in the right hand side of (1.7) and (1.8)). Formulas (1.7) and (1.8) differ in the group taken as reference, which is a in formula (1.7) and d in (1.8). The β coefficients have this interpretation since they reflect the returns that individuals will get from their personal characteristics with respect to wages.

Recently, panel data methods have been used to control for individual wage effects. In fact the individual effect, which is taken to be constant over time and specific to the individual, will capture all disturbances that are specific for the individual. Consider, for instance, a simple wage model in which the log-wage is explained in terms of formation, experience and managerial ability. This last variable is typically non-observable. In a panel data context we are able to control for this latent variable by introducing into the equation an individual effect (assumed to be constant through time), which can be fixed or random.

Polachek and Kim (1994) use fixed effects to estimate the gender earnings gap with intercept and slope specific effects. Since gender is a time invariant variable in the panel data models, a two-stage procedure is employed to estimate the gender gap. Rosholm and Smith (1996) estimate separate wage equations for male and female workers in Denmark using panel data techniques in order to identify the sources of changes in the wage gap. Other refinements to measuring labour market discrimination incorporate the gender and ethnic compositions as determinants of wages in the different occupations: Hirsch and Macpherson (1994), Hirsch and Schumacher (1992), and Macpherson and Hirsch (1995). Panel data techniques are used to control for occupational characteristics and unobserved worker characteristics encompassing skill and tastes.

1.3 The purpose of my thesis

Many studies have been carried out in order to determine the wage differential between men and women in Switzerland. For instance: Kugler (1988) uses the traditional approach, i.e., estimation of selectivity-corrected wage functions and then apply an Oaxaca decomposition; Brüderl et al. (1993), use data for their standard analysis of the wage gap; Bonjour (1997), applies several different methodological approaches in her estimation of the wage functions and decomposition of the wage gap; Flückiger and Ramirez (2000), show with annual aggregate data the decline of the wage gap.

In Chapters 2 and 3, I will investigate this wage gap from a different point of view, namely the occupational segregation. More precisely, in Chapter 2, I will examine the horizontal occupational segregation, which occurs when women are gathered in "female" occupations. I will extend the analysis based on a cross-sectional Survey by the Swiss Federal Statistical Office in year 2000, by introducing a new variable, FEM, which represents the percentage of the female labour force employed in an occupation. This new variable is included in order to explain the impact of the density of females in a certain occupation on individual wages (see England 1986, Sorensen 1990). The results of the studies carried out by England and Sorensen suggest that there is a significant and negative relationship between the proportion of females in an occupation and their wages. I then aggregate the FEM-variable into three categories depending on the portion of women in the occupation. The resulting selection model is estimated by Heckman two-stage approach.

Therefore, in Chapter 2 I will apply the same methodology used by Sorensen and developed by Hansen and Wahlberg (2000), using Swiss data. The goal is find a relationship between segregation and the gender wage gap.

In Chapter 3 the occupational segregation is treated in a vertical way, namely how the different phases of the careers affect wage of employees and how males and females differ in the career opportunities. I will estimate a usual wage equation by introducing a new latent variable d which is defined equal to 1 if the employee is in the high level of the job hierarchy and 0 otherwise.

Therefore, in Chapter 3 I will apply a new model to estimated the chance of career and the wage gap between males and females, using Swiss panel data.

1.4 The data

The data for the empirical analysis in Chapter 2 are drawn from a cross section of the *Enquete Suisse sur la structure des salaires* (ESS) in 2000, complemented with information on occupational segregation taken from the census in 2000. Both these data sources are supplied by Switzerland Statistics.

The survey (ESS) provides information on labour force market activities and incomes for a random sample of firms. It contains 10'500 firms which employ 510'000 workers. The firms are from the private and public sectors and are chosen randomly from the different regions of Switzerland. The questionnaires are sent to the firms, which are asked to answer some questions about the characteristics of the employees, work conditions, such as work place, education, field of activity, type of salaries and wages. In the case of small firms (1-50 employees), questions are placed to all employees, for medium firms (50-500) to 50% of the employees and for bigger firms (more than 500) to 10% of the employees. To construct the FEM-variable, which measures the proportion of workers who are women in a given occupation, I used information from the census.

The data for the empirical analysis in Chapter 3 are drawn from a rotating panel data, the Swiss Labour Force Survey (SLFS). A subset of individuals is replaced every year, and an individual can stay in the survey for a maximum of five years. These data sources are supplied by Switzerland Statistics. The purpose of this survey is to obtain data about the working environment and professional life in general. It also provides insight into the living conditions of the employment people, retired people, housewives as well as students. The SLFS is based on a representative sample of Swiss households. The interviews are conducted by phone call in the 2^{nd} quarter of each year (from April to June). To perform my analysis I selected individuals that are in the sample for a period of five years (from 1999 to 2003), therefore the data used are balanced panel data.

Chapter 2

Occupational gender segregation

2.1 Introduction

The term occupational segregation indicates a greater concentration of women or men in given levels of activity and occupation. Because of occupational segregation in the labour market, women are often confined in a more narrow range of occupations and in lower levels of activity compared to men (Robinson, 1998). We can distinguish two types of segregation: horizontal segregation, which occurs when there is a concentration of women and men in a determined fields and occupations, and produces disparity in terms of career, pension, benefits accessories, etc.; and vertical segregation, which occurs when there is a concentration of women and men in determined degrees and levels of responsibility or positions, and produces disparity on salaries. In this chapter I will focus my attention on the case of horizontal segregation.

The purpose of this Chapter is to quantify the degree of occupational segregation in Switzerland, and if it exists, to find the relationship between wages and the proportion of women in each occupation. The Chapter is organised as follows. First, I will give an overview of the literature and some definitions of market segregation. Secondly, I will measure the level of occupational segregation in Switzerland using two different indices. Thirdly, I estimate the effect of occupational segregation on male and female wages. Finally, using an ordered probit model I will eliminate the endogenity problem due to the correlation between the density of females in an occupation and the error term in the wage equation.

2.2 Theories of occupational segregation by gender

Most of the studies that try to estimate the unexplained portion of the gender wage gap are based on a sample of male and female workers and try to explain their wages by their observable characteristics, such as accumulated human capital (formation and years of work experience) and other factors expected to influence wages (such as marital status and labour supply). The remaining portion of the wage gap is then often attributed to employer discrimination.

One candidate for explaining part of the unobserved difference in wages across gender is occupational segregation, as argue Hansen and Wahlberg (2000). Researchers usually distinguish between labour supply and labour demand factors when explaining occupational segregation by gender. Factors related to labour supply generally focus on why women "prefer" certain types of occupation (for example, women may prefer works with flexible hours in order to have time for child care). Explanations related to labour demand focus on why employers generally "prefer" to hire women or men for particular occupations and why women and men have different opportunities for promotion and career development within firms in particular sectors.

Theories explaining the existence of occupational segregation by gender can be classified into three broad categories: neoclassical and human capital theories, institutional and labour market segmentation theories, and non-economic and feminist (or gender) theories.

2.2.1 The neoclassical human capital model

Neoclassical economics assumes that workers and employers are rational and that labour markets function efficiently (Anker, 1997). According to this theory, workers look for the best-paying jobs after taking into consideration their own personal endowments (as education and experience), constraints (e.g. young child to take care of), and preferences (as a pleasant work environment). Instead the goal of employers is to maximize profits by maximizing productivity and minimizing costs to the extent possible. And given the competition and efficient labour markets, employers pay workers their marginal product.

The neoclassical economic view explains occupational, or pay, differentials between individuals or groups by different human capital investment, or by different choices in the trade off between pecuniary and non pecuniary job rewards.

Zellner (1975) and Polachek (1979, 1982, 1985) pioneered the application of neoclassical theory in explaining occupational segregation. They suggest that women, who plan intermittent employment, will maximize lifetime earnings if they choose occupations with low rates of appreciation and depreciation of human capital. Appreciation of human capital refers to formal or informal on-the-job training that makes a worker more productive and leads to wage growth. Human capital theory asserts that those who plan more years of employment will choose jobs with the highest returns to experience. Because, on average, men anticipate more job experience than women, women will select occupations with higher starting wages but smaller returns to experience. This would lead to occupations with these characteristics to become disproportionately female. Briefly, these economists assert that average differences in pay between male and female jobs are all "compensated" for by other advantages of female jobs, such as fewer skill demands, more pleasant job requirement or working conditions, higher starting wages, or lower risk of depreciation.

But no empirical analysis has found the higher starting wages in female occupations that the theory predicts; to the contrary, starting wages are generally lower in female than male occupations requiring the same education (Greenberger and Steinberg 1983; England 1986).

2.2.2 Institutional and labour market segmentation theories

The starting point of the institutional and labour market segmentation theories is the assumption that institutions, such as unions and large enterprises, play an important role in determining who is hired, fired and promoted, and how much employees are paid. Institutional theories are also based on the assumption that labour markets are segmented in certain ways. And while each labour markets segment may function according to neo-classical theory, it is difficult for workers to pass from one segment to another, resulting in non-competitive phenomena in the aggregate.

The most well-known institutional theory is the *dual labour market theory*, which distinguishes between a "primary" and a "secondary" sector. Jobs in the primary sector are relatively good in terms of pay, security, opportunities for advancement and working conditions (Anker 1997). Secondary sector jobs tend to be relatively poor with respect to pay, chances for promotion and working conditions, and they tend to provide little protection or job security. The distinction between primary and secondary sectors has become less marked in recent years in both industrialized and developing countries.

It is a relatively short step to adapt the concept of dual labour markets to occupational segregation by gender, with one labour market segment consisting in "female" occupations and the other in "male" occupations. This segmentation implies relatively low wage rates in "female" occupations because many women workers are "overcrowded" into a small number of "female" occupations. "Male" occupations, on the other hand, benefit from reduced competition within a wider set of occupations and, consequently, tend to enjoy relatively high wage rates. If females, but not males, are crowded into low earnings jobs only due to discrimination, then the gender composition of a job becomes an index of labour quality for males and, to a small degree, for females (Hansen and Wahlberg 2000). Males who are relatively less productive accept low earnings work in primarily female occupations. Over time, low earnings occupations, crowded by females, would attract relatively less productive males and loose high productive females. Thus, over time we should observe workers with lower productivity and lower wages in these occupations.

Labour market segmentation theories are very useful to explaining gender inequality in the labour market, since they stress the existence of segregated labour market occupations. However, they do not explain why occupations are in fact segmented by gender.

2.2.3 Gender theories

The basic idea in the gender theories is that women's disadvantaged position in the labour market is caused by, and is a reflection of, patriarchy as well as women's subordinate position in the society and in the family.

In many societies, household work and child care are seen as women's primal responsibility, while being the breadwinner is perceived as men's main responsibility. As Anker (1997) explains, this division of responsibilities and the patriarchal ordering of society are instrumental in determining why women usually accumulate less human capital compared with men before entering the labour market. That is, why girls receive less education than boys, and are less likely to pursue fields of study such as sciences, but are more talented for literature or languages study. The same influences are also instrumental in explaining why women acquire less labour market experience, on average, because many of them withdraw from the labour force earlier, and many others have discontinuous labour experiences.

Gender theories explain occupational segregation by gender by showing how closely the characteristics of "female" occupations mirror the common stereotypes of women and their supposed abilities. One can typically consider three groups of stereotypes (positive, negative and other). The five "positive" stereotypes typically mentioned in the literature are: a caring nature, skill and experience in household-related work, greater manual dexterity, greater honesty, and attractive physical appearance. The five "negative" stereotypes are: disinclination to supervise others, less physical strength, less ability in science and mathematics, less willingness to travel, and less willingness to face physical danger. Three other "neutral" stereotypes are presented: greater willingness to take orders, greater willingness to accept lower wages and less need for income, and greater interest in working at home. These stereotypes, if true, would have a great influence on the general characteristics typifying "female" occupations. Masculine stereotypes also play a role in determining the occupations which become typically "male" (such as engineer, truck driver, police officer and construction worker). In order to break down the gender segregation of occupations, it is important to change both male and female stereotypes and to integrate men into "female" occupations as well as women into "male" occupations.

2.3 Measure of the occupational segregation

In the last years, some studies (Silber 1989; Hutchens 1991; Deutsch, Flückiger and Silber 1994) have suggested that several approaches applied in the income inequality literature can be used to measure occupational segregation.

The first index of occupational segregation was proposed by Duncan and Duncan (1955). In fact, they derived the concept of the Segregation Curve from the traditional Lorenz Curve which is popular in the income inequality literature. Trends in occupational segregation are commonly measured by the index of segregation (Duncan and Duncan Dissimilarity Index). The index is defined as follows:

$$S = \frac{1}{2}\sum_{i=1}^{n}\left|\frac{F_i}{F}-\frac{M_i}{M}\right|, \tag{2.1}$$

where M_i is the number of males employed in occupation i, and F_i is the number of female employed in occupation i, and M and F are the total number of male and female in the labour force. The index may take on a value between 0 and 1, where zero represents perfect integration and 1 represents complete segregation. This number gives us the proportion of women (or men) that would have to be redistributed among occupations in order for the occupational distribution to reach complete equality between the sexes[1].

Later on Williams (1979) introduced the Size-Standardized Dissimilarity Index (S_d), which is the absolute measure of the segregation that controls for the effects of the occupational structure, using all occupations as if they were of the same size, computed over a fixed number of comparable occupational categories. The standardized dissimilarity index is expressed as

$$S_d = \frac{1}{2}\sum_{i=1}^{n}\left|\left[\frac{\left(\frac{F_i}{T_i}\right)}{\sum_{i=1}^{n}\left(\frac{F_i}{T_i}\right)}\right]-\left[\frac{\left(\frac{M_i}{T_i}\right)}{\sum_{i=1}^{n}\left(\frac{M_i}{T_i}\right)}\right]\right|, \tag{2.2}$$

where T_i is the total number of men and women in occupation i, (F_i/T_i) and (M_i/T_i) denote the female and male proportions in occupation i, and the dominators adjust such values to the proportions in the other occupations.

[1]Unfortunately these indices of occupational segregation are sensitive to the degree of aggregation. For example, at an highly disaggregated level, a large number of jobs are occupied only virtually by males or only by females (see Groshen 1991).

Index S_d is not affected by the shape of the occupational distribution. More precisely, suppose the total number of workers in a sector doubles but the proportions of women and men in the sector stay the same. Index S_d does not change, whereas index S may change. Thus, changes in the absolutes size of the occupations in time do not affect the value of index S_d.

2.4 Models for empirical analysis

The purpose of this Chapter is to study whether any wage penalty exist for working in occupations which are characterized by a high concentration of female workers in Switzerland. In the following sections, after a briefly review of the literature, we introduce the model used for the empirical analysis. We start the analysis with a wage regression with controlling for occupational segregation (section 2.4.1) and in the sesond part we estimate a regression with Heckman correction for occupational segregation (section 2.4.2).

The interest among economists in occupational gender segregation stems from the fairly well established relationship between the gender differential in earnings and women's concentration in a small number of occupations. There is a general consensus that occupational segregation exists and that females are gathered disproportionally in occupations with lower earnings. There is no agreement on the causes of these outcomes and two alternative explanations have been given in the literature. The first one argues that females are gathered disproportional in occupations with low earnings due to market discrimination (The Crowding Hypothesis[2]), and the second one argues that it is due to a self-sorting mechanism.

Recently, there have been a large number of studies[3] devoted to empirically determine the impact that the density of females in a certain occupation has on individual wages. See for instance Bayard et al (1999), Macpherson and Hirsch (1995) and Sorensen (1990) for applications on U.S. data, Baker and Fortin (1999) using Canadian data, Miller (1987) using data from the U.K., and Hansen and Wahlberg (2000) and Le Grand (1991) using Swedish data. The results from these studies suggest that there is a significant and negative relationship between the proportion of females in an occupation and their wages. The finding that individual wages vary systematically with the gender composition of occupations is well established in cross-sectional empirical studies. For examples, Killingsworth (1990) provides two stylized facts regarding the "femaleness" of occupations: (1) both women and men earn less as the proportion of females in an occupation increases, and (2) the negative relationship between wages and percentage of female is stronger for men than for women.

2.4.1 Regression controlling for occupational segregation

In this section I am interested in testing the hypothesis that an individual earns less if she/he is employed in a job held predominantly by women in Switzerland. One method of testing this hypothesis is to estimate an earnings equation which includes

[2] According to the crowding model, employers discriminate against women by excluding them from occupations considered "men's work". Since these jobs are reserved for men, relatively few women are hired into these positions. Given that the demand for women in these jobs is limited, they are crowded into other occupations, typically referred to as "women's work". The supply of women accordingly increases for "women's work", which in turn reduces their wage.

[3] The models, used to explain the existence of occupational segregation and lower relative earnings for women, assume that women and men have equal abilities and thus without discrimination, would deserve equal remunerations.

an independent variable measuring the proportion of women workers in an occupation, FEM, in addition to other traditional explanatory factors.

Following the model introduced by Hirsch and Macpherson (1995), and then developed by Hansen and Wahlberg (2000), I can define the relationship between wages and gender composition in two equations as:

$$\ln w_i = X_i'\beta_f + \theta_f FEM_i + \varepsilon_{if} \quad \forall i \in f \tag{2.3}$$

$$\ln w_i = X_i'\beta_m + \theta_m FEM_i + \varepsilon_{im} \quad \forall i \in m \tag{2.4}$$

where f denotes the set of females in the sample and m denotes the set of males in the sample. Sub index i denotes individuals ($i = 1, .., N_f$ for women and $i = 1, ..., N_m$ for men). In addition, $\ln w_i$ is the natural log of the salary for individual i working in occupation k, X_i includes a constant, information on the experience (defined as: $age - education\ level - 6$)[4], occupational seniority, marital status, domiciled and job characteristics; FEM_i is the percentage of the female labour force employed in occupation k[5]. The coefficients β and θ for characteristics X and variable FEM are allowed to be different for males and females. The last term in equation (2.3) and (2.4) concern the error structure[6] of the model. Unobserved individual-specific disturbance term is assumed to be captured in ε_i, effects of unobservable variables that vary across individuals. It is further assumed that the sequence $\{\varepsilon_i\}$ consists of normal i.i.d. random variables with mean zero and a constant variance σ^2_ε.

---

[4]We subtract 6 because in Switzerland children start the school when they are six years old.

[5]In Table 2.5 are reported the proportions of women in each occupation specification. Note that the number of occupations k is 24.

[6]Hansen and Wahlberg (2000) used a different error structure. They consider in addition to the individual specific effect ϵ_i an unobserved occupational specific effects v_k. They assume that a part of the error term v_k is correlated across workers within occupations. In our model we neglect this occupation-specific effect as did Bayard et all (1990), Sorensen (1990) and le Grand (1991).

The interpretation of θ_f and θ_m depends on the causes of the occupational segregation and on the ways FEM and wage rates are related. Some explanations of the occupational segregation are: human capitals differentials, employer discrimination, restrictions to labour mobility and premarket differences for family and educational background and the socialization process.

Hirsch and Macpherson (1995) in their analysis find out a negative coefficient for θ of magnitude -0.090 and -0.139 for males and females respectively. The crowding model is useful to explain why $\theta_f < 0$. Women can be concentrated in particular occupations, due to their preferences or to past and current barriers to alternative occupations. This crowding compress the wages to a level below the one of similarly qualified workers in other occupations and the interoccupational mobility is insufficient to equalize wages. For men, if $\theta_m < 0$, predominantly female occupations attract less skilled men. In other words, if women face barriers to better remunerated occupations, low-wage occupations would attract a disproportionally large number of women and a small proportion of men, so that there would be a negative correlation between FEM and male and female wages.

Summarising, a value of $\theta_f < 0$ and $\theta_m < 0$ implies that wages decrease with respect to the proportion of females. Over time, low-paying occupations crowded by women would attract relatively lower-quality males and lose many high-quality females; thus, we observe workers in predominantly female jobs with lower average productivities and wages. FEM serves as a quality index since the probability of a woman being hired into predominantly male job is an increasing function of productivity. In short, workers are sorted into occupations based in part on expected productivity, and this productivity may be correlated with the gender composition (Macpherson and Hirsch, 1995).

We can see that a potential problem with most of the previous studies[7] in this area is that they assume that occupational attainment can be treated as exogenous, i.e. there is no correlation between the density of females in an occupation and the error term in the wage equation. As argued by Macpherson and Hirsch (1995), there exist at least two reasons for why the exogeneity assumption may be not valid. First, if men and women with higher unmeasured skills (captured by the error term in the wage equation) are more likely to be sorted into male jobs and those with lower skills into female jobs, then the exogeneity assumption will obviously be violated. Second, the error term may also capture unobserved taste differences among workers. For example, female workers may foresee future work interruptions due to childbearing and thus prefer part-time jobs or jobs where the wage "penalty" for absence from work is low. Therefore, we would observe a concentration of female workers in these types of jobs, which may also pay lower wages. It is again clear that the assumption of no correlation between the density of females in an occupation and the error term can be violated. In other words a problem with the specification (2.4) and (2.5) is that the variable FEM is endogenous. Indeed, the job characteristics regarding the percentage of female labour force employed are chosen (to some extent) by the individual, and the factors determining this choice may be correlated with the error forms in equations (2.4) and (2.5). To avoid the potential problem with endogeneity, Macpherson and Hirsch (1995) use longitudinal data covering the period from 1983 to 1993 and appling a fixed-effect estimator. In my dissertation, to overcome this difficulty, I introduce a selection model and I estimate it with a version of Heckman two step approach as in Hansen and Wahlberg (2000). In the next section the model is exposed.

[7]In Bayard et al (1999), Sorensen (1990) and le Grand (1991)

2.4.2 Regression with Heckman correction for occupational segregation

I transform the variable FEM in a discrete variable, representing different types of occupation, e.g. male dominated, female dominated or an intermediate occupation. An ordered Probit model is specified for the selection of the type of occupation by the individuals. The selection equation can be specified as follows:

$$\begin{aligned} FEM_i &= k \quad if \quad \mu_{k-1} < FEM_i^* < \mu_k, \quad for \quad k = 0, 1, 2 && (2.5) \\ FEM_i^* &= \gamma_j Z_i + \eta_i, \quad j = f, m && (2.6) \end{aligned}$$

where FEM_i^* is a continuous latent variable, coefficients μ are unknown parameters to be estimated, and reflect threshold values for moving through the occupational choice decision. In my analysis, the vector Z contains the explanatory variables in X and some additional variables, which will be described in the next section.

Then, separate wage regressions are specified for males and females and for each value of the FEM variable.

$$\begin{aligned} \ln w_i &= X_i' \beta_{fk} + \varepsilon_{if} \quad if \quad FEM_i = k \quad and \quad \forall i \in f && (2.7) \\ \ln w_i &= X_i' \beta_{mk} + \varepsilon_{im} \quad if \quad FEM_i = k \quad and \quad \forall i \in m && (2.8) \end{aligned}$$

The errors ε_i and η_i are assumed jointly normal,

$$\begin{pmatrix} \varepsilon_i \\ \eta_i \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 & \sigma_{\varepsilon\eta} \\ & \sigma_\eta^2 \end{pmatrix} \right). \qquad (2.9)$$

The covariance term $\sigma_{\varepsilon\eta}$ accounts for correlation between variable FEM and individual error ε_i in the regression.

The model (2.5)-(2.8) is estimated by a Heckeman-type two-stage procedure. More precisely, it is possible to show that (ref. Maddala (1999))

$$E(\varepsilon_i \mid FEM_i = k) = \frac{\sigma_{\varepsilon\eta}}{\sigma_\eta^2} \frac{\phi(\mu_{k-1} - \gamma_j Z_i) - \phi(\mu_k - \gamma_j Z_i)}{\Phi(\mu_k - \gamma_j Z_i) - \Phi(\mu_{k-1} - \gamma_j Z_i)} = \frac{\sigma_{\varepsilon\eta}}{\sigma_\eta^2} \lambda_{ijk}, \tag{2.10}$$

where ϕ and Φ are the p.d.f and the c.d.f. of standard Gaussian. A consistent estimation $\widehat{\gamma}$ of γ and $\widehat{\mu}_k$ of μ_k, $k = 0, 1, 2$, are obtained by apply Maximum Likelihood on the Probit model (2.5), (2.6). Then, I construct an estimator of the selection term.

$$\widehat{\lambda}_{ijk} = \frac{\phi(\widehat{\mu}_{k-1} - \widehat{\gamma}_j Z_i) - \phi(\widehat{\mu}_k - \widehat{\gamma}_j Z_i)}{\Phi(\widehat{\mu}_k - \widehat{\gamma}_j Z_i) - \Phi(\widehat{\mu}_{k-1} - \widehat{\gamma}_j Z_i)}. \tag{2.11}$$

In the second step I estimate the following equation by OLS

$$\ln w_i = X_i' \beta_{jk} + \delta_{jk} \widehat{\lambda}_{ijk} + errors \quad j = f, m \quad and \quad k = 1, 2, 3 \tag{2.12}$$

where the coefficients[8] of variables X_i is β_{jk} if individual i is of gender j in occupation k.

2.5 The results

In the next subsection I show the results obtained from a general analysis of the gender composition in the workplace. Results from the linear regression with control for occupational segregation and concluding with the model decomposed in three occupational categories.

[8] The standard errors of the estimated coefficients are biased because we do not consider that λ is estimated.

2.5.1 The gender stratification in Switzerland

There is a general perception that in Switzerland many women are becoming increasingly oriented toward non-traditional family roles and non-traditional jobs in the workplace. However, if I take a look at the statistics, I note that in Switzerland the choice between working or studying is still related to gender.

To better understand the nature of the gender segregation in Switzerland, it is useful to consider a few things about the Swiss educational system. The Swiss educational system is highly differentiated and vocationally-oriented. After completing eight or nine years of compulsory school, young people can either join the labour force as unskilled workers, go to the secondary school (university-track), or enrol in vocational training (either in a full-time vocational school or in a part-time vocational school combined with an apprenticeship). The choice between these options is largely determined by the earlier educational career of the individual. Moreover, access to nearly all medium and high status occupations in Switzerland is strictly limited to those who have completed the corresponding vocational education.

In Table 2.1 we can observe the highest educational level vested from women and men in Switzerland for the years 1970 until 2000. It can be noted that the percentage of women who decide to stop the school career after the compulsory school is clearly bigger than for men, for all periods. However, this percentage is decreasing over time, for both men and women, and the gender gap is becoming smaller.

Through the analysis of 6 classes of age (Table 2.2), one can see the generational differences in the educational choices of employed people. It can be noted that the percentage of women between 25 and 34 years old, that have achieved a university degree, is bigger than for older women. It seems that the young generation, of women, is more motivated to continue its studies compared to the older generation. In the period from 1980-2000, the participation of women in universities (Table 2.3) increases (to 46.5%), and women tend not be segregated in "female" faculties but they choose also "male" faculties (as medicine, pharmacies and technical sciences), even if the percentage of women remains lower than that of men. Concerning the vocational school (Table 2.4), women are more oriented towars professions in the sanitary sector, textile and clothes industry, as well as commerce, organisation and administration. Instead men have a fan of choices more rich.

In general, differences between men and women in overall educational attainment remain substantial in Switzerland. Women complete fewer years of formal education, are less likely to pursue vocational education and tend to graduate from universities at a lower rate with respect to men. When women do complete vocational training, it is generally of shorter duration than men. This fact is consistent with the arguments put forward by neoclassical economists. They argue that women may prefer not to make large educational investments. This may partially account for the large gender differences in the educational environment (see Table 2.1 and 2.2).

Another characteristic of the Swiss society, that may be relevant while studying occupational gender segregation, is the extreme structural incompatibility of work and family roles. Although nearly all Swiss women are professionally active prior the marriage, labour force participation of wives is dramatically lower (Charles, 1994). The majority of women either give up or greatly curtail their labour force activity with the birth of the first child.

A number of contextual factors limit Swiss mothers' labour attachment. First, public and private childcare arrangements are extremely scarce in Switzerland. Second, school schedules are highly incompatible with the demands of two-earner families: children return home daily for a two-hour lunch break, and school hours vary widely from day to day and from child to child. Third, the ideology of "motherhood" and the sexual division of labour within the family are firmly entrenched in the Swiss culture.

And fourth, male wages are traditionally quite high in Switzerland; this makes it generally possible to comfortably raise a family on a single income. However, during the last few years life in Switzerland has become more expensive and most families now need two salaries. This determines an increase in the demand of public or private childcare arrangements (Banfi, 2006) and grandparents are forced to help growing up children.

2.5.2 Indices of occupational segregation

Table 2.5 reports some measures of occupational segregation in Switzerland for the years 1980-2000. Note that the number of occupations is different in the years 1970-1980 and in the years 1990-2000. The Dissimilarity Index decreases during the years from 0.63 to 0.44. This means that to reach perfect integration, 44% of women must be redistributed among occupations in 2000. The Size-Standardized Dissimilarity Index features the same trend, from 0.63 to 0.56. In both cases we can see that gender segregation in Switzerland is declining over time.

In Figure 2.1 are reported the segregation curves for the years 1970-2000 in Switzerland. We can see that all of the segregation curves moved to the left over time, hence the observed segregation has decreased.

A comparison of standardized segregation-index scores in 25 industrial countries shows that Switzerland is second only to Luxembourg in the overall level of occupational gender segregation (Charles, 1992). As Charles (1994) argues, this distinction is due in large part to Swiss women's overrepresentation in sales occupations and their underrepresentation in management.

A general examination of Swiss women's and men's occupational distributions reveals that gender segregation tends to occur very early in the individual life and persists for a long time in the professional career.

A research by Charles and Buchmann (1994) helps to identify some macro-level factors that also contribute to explain Switzerland's extraordinarily high level of gender segregation. These factors include a relatively conservative ideological climate, a low rate of female labour force participation and the relatively large proportion of the Swiss labour force with employee (as opposed to self-employed) status.

2.5.3 Wage regressions controlling for occupational segregation

In this section I present the results on estimation of the wage regressions, equations (2.4)-(2.5). The explanatory variables are reported in Table 2.1. They include marital status, dummies for the level of education, dummies related to the foreigner status, experience, occupational seniority, and dummies for job characteristics, namely the variable *position*. This variable gives an indication on the job hierarchy level of the employee, and it is subdivided in 5 levels. From *position* 1 to *position* 5, where *position* 1 is the low level and *position* 5 is the high level (top management category).

Estimation is performed separately for males and females, without introducing the FEM variable (Table 2.8), and by including the FEM variable (Table 2.9). The first and second columns of the Table 2.8 represent the estimates for males, the third and fourth columns for females. The variable *married* is significant for both males and females, but while for women the impact is negative, for men it is positive. For women, being married reduces the monthly earnings by 3%, whereas earnings are increased for men by 5%. The dummy variables corresponding to the different levels of formation are all significant. As expected, the estimated coefficients are positive and increasing as the level of educations increases, for both males and females. In particular for women having a high formation increases the monthly salary of 56%, instead for men this effect is about 50%. Let us consider the variables *experience* and *occupational seniority*. For men and women, they are both significant, with a quadratic specification implying a positive effect, and the *experience* has an higher impact (0.025 and 0.022 for males and females respectively) than the *occupational seniority* (0.008 and 0.009 for male and females respectively). As expected the *position* variable is significant and positive and increasing as the level of position increases, for both males and females. But while the dummy *position* 5 has a big impact on the wage for males, in the case of females this impact is relatively smaller.

Let us now consider the results in Table 2.9. First overall note that the introducing of the variable FEM do not improve the estimation of the wage equation. In fact the estimate of the coefficients are similar to those reported in Table 2.8. The effect of the variable FEM is negative and significant for both female and male wages, with a value of -0.0531 for women, and -0.0658 for men. The meaning of the negative values of θ_w and θ_m imply lower female and male wages in jobs with higher densities of female workers. Note that the coefficient for women is smaller than the coefficient for men. Macpherson and Hirsh (1995) found a larger negative effect of FEM on wages for females than males.

2.5.4 Two step Heckman estimation

In Table 2.10 and 2.11 are reported some descriptive statistics for males and females and by occupational type (male dominated, intermediate and female dominated occupation). In Table 2.12 are report the results for the ordered probit model. The first two columns refer to males, while the last two columns display results for females. Note that in these regressions new explanatory variables are added, such as the dummy $part-time$, dummies for skill, dummies for the firm size and the variable age. The variable *skill* is a 4 levels variable, which indicate the characteristics of the job, from the lowest skill namely a job with repetitive tasks ($skill1$) until job that involve high qualify tasks ($skill4$). For the size of the firm we consider 4 categories, from the smallest with 1-10 employees ($Firm\ Small$) to the extra-large with more than 500 employees ($Firm\ Extra-large$).

For both males and females, the results show that the education and occupational choice are strongly correlated and that the probability of working in a male dominated occupation is higher for individuals with higher education. For females, the estimated coefficients are decreasing with respect to the education level. For males, the estimated coefficient is smaller for $formation4$ than for $formation2$, as expected, the coefficient for $formation3$ is not significant. Women without education are more likely to work in female dominated occupations. Work experience has the same effect as education, the higher the experience, the higher the probability of working in a male dominated occupation. These results are in accordance with economic theory (see Section 2.2). Also the variable *married* has a significant impact on both male of female occupational choice. The impact is negative, which means that married individuals are more likely to work in a male occupation. As far as women are concerned, this result is rather difficult to reconcile with economic theory, which argues that married women are more likely to work in female occupation where the wage penalty for absence from work is low. The negative coefficient for the variable foreign implies that foreigners are more likely to work in a male dominated occupation.

Looking the effects of the additional variables on occupational choice, I see that *age* has a positive estimate coefficient for males, but not for females. This means that men from older cohorts are more likely to have jobs in segregated occupations. Therefore, the assumption that occupational segregation is more pronounced among older cohorts than among younger ones is confirmed only for the male case. For the $part-time$ dummy, the estimate for females is positive, which implies that, females who are working part-time are more likely to hold a job in a female dominated occupation. For males, the same reasoning applies. These results are not surprising. Indeed, assuming that the age-earnings profile for women is flatter in female dominated occupations, the wage penalty of work absence is lower in these jobs. I would then expect women with many children (and therefore with more work absence) to prefer working part-time rather than in jobs where working full-time is mandatory (as in most male dominated jobs). For both males and females, the results show that the probability of working in a male dominated occupation is higher for the individuals with higher skill. For the dummies firm size only the variable $Firm\ Extra-large$ has a positive effect, and only for males. This means that the probability for males of working in a female dominated occupation is higher for men employed in an extra-large firm.

Table 2.13 contains the estimates for the wage equation (2.12) including the Heckman's correction term. The results in the first two columns refer to male dominated occupations, columns three and four show the results for integrated occupations and the last two columns display estimates for females working in female dominated occupations. The results in the three occupations can be compared with the estimates reported in Table 2.8. For all regressors, the estimated coefficients in Table 2.13 are rather similar in the male dominated, female dominated and intermediary sector, and similar to the estimates in Table 2.8 (see section 2.5.3 for a discussion of the estimates).

However, the constant differs across sectors and appears to be larger in the male dominate sector. Thus, in accordance with economic theory (see Section 2.2), females employed in the male dominated sector have an higher wage than in the other two sectors[9]. Finally, the selection correction variable, λ, is significant supporting the hypothesis that women self-select themselves into different occupations. Hence, I can reject the hypothesis that women are randomly allocated into different occupations.

Table 2.14 contains the wage estimates for males. Most of the results are in accordance with prior expectations. The results regarding the effect of human capital imply higher return to education in all the different occupations and also an higher return in work experience. Further, there is a significant and positive wage effect on being

[9]We note that the estimated constant for the female dominated sector is larger than for the intermediary sector. This fact is rather unintuitive. However the difference is not statistically significant, and is likely due to sample variability.

married. Also in this case, the selection correction variable is significant, suggesting that men self-select themselves into different occupations.

Table 2.15 contains the estimates of the observed wage differentials in the OLS model and in three occupational groups. The observed gender wage gap is of 13.3%in the model without control for occupational segregation. In the model with the variable FEM the gap is, as expected, slightly smaller. Concerning the three occupational groups, I note that the intermediate group has the biggest wage differential (24.4%). The wage gap is smallest (2.1%) in male dominated occupations. This suggests that for women it would be better to work in male dominated occupations instead of in the female dominated occupations, where the wage differential is of 7.7%.

2.6 Conclusions

In the first part of this chapter I have gathered some basic evidence regarding the individual-level determinants of occupational gender segregation in the Swiss labour market. A general examination of Swiss women's and men's occupational distributions reveals that gender segregation tends to occur very early in the individual life and persists in the labour market career. There is some evidence that Swiss mothers are more likely than their childless counterparts to work in female-dominated occupations. Two factors, which are formal education and organization of everyday life, might influence the individual-level processes of human capital development and occupational allocation.

The purpose of the second part of this chapter was to study the existence of any wage penalty for working in occupations are characterized by high concentration of female workers in Switzerland. I found a negative effect on male and female estimated wages, confirming the hypothesis of a wage penalty. We can conclude that occupations dominated by female pay lower wages to women and men. Considering the two-step Heckman estimation, we can notice that the inclusion of correction terms for self-selection show to have a significant impact on the results, supporting the hypothesis that women and men self-select themselves into different occupations. Moreover occupational segregation appears to influence the gender wage gap. In fact in different occupational categories we find different wage gap. The lowest wage gap, that we found, is in the male dominated occupation.

We must remember that in Switzerland legislation which assures equal opportunities and equal wages between men and women was only effectively passed in 1990s. And issue such maternity benefit and child-care facility are more often present in political debate. These measures will most probably, in the next years, induce a decrease in the index of occupational segregation and also in the wage gap.

Variable Name	Definition
	Dependent Variables
Log Earnings	The natural log of monthly earnings
FEM	0 if he/she is in male dominated occupation, 1 if he/she is in intermediate occupation and 2 if he/she is in female dominated occupation
	Independent Variables
Age	The individual age
FEM	The percentage of female in an occupation k
Married	1 if he/she is married, 0 otherwise
Formation 1	1 if the highest level of study is compulsory schooling, 0 otherwise
Formation 2	1 if the highest level of study is apprenticeship, 0 otherwise
Formation 3	1 if the highest level of study is high school, 0 otherwise
Formation 4	1 if the highest level of study is university degree, 0 otherwise
Foreign	1 if he/she is foreign, 0 otherwise
Permit C	1 if he/she has a permit C, 0 otherwise
Experience	Number of years of prior work experience
Experience squared	Work experience squared
Occupational seniority	Number of years worked for current employers
Occupational seniority squared	Occupational seniority squared
Position 1	1 if he/she is no one management category, 0 otherwise
Position 2	1 if he/she has a supervisor position, 0 otherwise
Position 3	1 if he/she is in the low management category, 0 otherwise
Position 4	1 if he/she is in the middle management category, 0 otherwise
Position 5	1 if he/she is in the top management category, 0 otherwise
Skill 1	1 if the job involves simple or repetitive activities, 0 otherwise
Skill 2	1 if the job involves specialized professional acquaintances, 0 otherwise
Skill 3	1 if the job involves qualify activities, 0 otherwise
Skill 4	1 if the job involves the most qualify activities and the most difficult tasks, 0 otherwise
Firm Small	1 if in the firm are employed 1-10 individuals, 0 otherwise
Firm Medium	1 if in the firm are employed 11-50 individuals, 0 otherwise
Firm Large	1 if in the firm are employed 51-500 individuals, 0 otherwise
Firm Extra-Large	1 if in the firm are employed 500 and more individuals, 0 otherwise

Table 2.1: **Variables Definitions**

	1970		1980		1990		2000	
	Males (%)	Females (%)	Males (%)	Females (%)	Males (%)	Females (%)	Males (%)	Females (%)
Without Formation	0.33	0.37	3.40	3.34	2.55	2.43	2.33	2.60
Compulsory School	34.02	50.99	29.89	39.98	20.55	27.29	19.05	25.55
Vocational School	39.59	20.37	46.94	40.71	52.57	53.16	44.07	44.71
High School and Teaching	5.80	8.58	2.77	6.44	2.56	6.40	5.98	11.76
Superior Vocational School	10.73	10.95	9.12	3.59	13.34	5.62	16.38	7.61
University	4.66	1.25	6.84	2.86	7.81	3.97	12.18	7.77
Other	4.87	7.48	1.04	3.08	0.62	1.13	0.00	0.00
TOTAL	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table 2.2: **Percentage of the formation degree for employed persons in Switzerland for the years 1970-2000**

In the table are reported the percentage of employed male and female in Switzerland for the years 1970-2000.

Males in Switzerland in year 2000						
Formation Degree	15-24 (%)	25-34 (%)	35-44 (%)	45-54 (%)	55-64 (%)	65-74 (%)
Without Formation	2.71	1.84	2.25	2.68	2.54	1.49
Compulsory School	45.06	12.95	15.00	16.44	18.12	17.28
Vocational School	39.85	46.72	42.91	44.91	45.30	37.04
High School and Teaching	9.54	7.13	5.58	4.79	3.74	4.59
Superior Vocational School	1.94	18.23	19.72	17.47	17.69	19.31
University	0.90	13.13	14.54	13.72	12.60	20.29
Other	0.00	0.00	0.00	0.00	0.00	0.00
TOTAL	100.00	100.00	100.00	100.00	100.00	100.00

Females in Switzerland in year 2000						
Formation Degree	15-24 (%)	25-34 (%)	35-44 (%)	45-54 (%)	55-64 (%)	65-74 (%)
Without Formation	2.46	1.71	2.72	3.08	3.41	3.47
Compulsory School	40.48	15.26	21.85	27.26	32.41	38.92
Vocational School	39.40	49.02	45.38	44.27	42.73	35.41
High School and Teaching	14.76	12.93	11.28	10.60	9.01	9.34
Superior Vocational School	1.83	10.23	9.08	7.30	6.61	6.35
University	1.06	10.85	9.70	7.49	5.84	6.50
Other	0.00	0.00	0.00	0.00	0.00	0.00
TOTAL	100.00	100.00	100.00	100.00	100.00	100.00

Table 2.3: **Percentage of the formation degree for 6 classes of age**

In the table are reported the percentage of male and female in the different degree of formation for 6 classes of age in Switzerland for the year 2000.

	1980/81		1992/93		2001/02	
Univesity Students	M%	F%	M%	F%	M%	F%
Social and Human Sciences	48.0	52.0	41.1	58.9	36.9	63.1
Medicine and Pharmacy	65.8	34.2	52.6	47.4	43.8	56.2
Law	73.1	26.9	58.3	41.7	51.2	48.8
Natural Sciences	78.3	21.7	73.6	26.4	67.3	32.7
Technical Sciences	91.9	8.1	82.7	17.3	77.0	23.0
Economic Sciences	84.6	15.4	76.1	23.9	71.6	28.4
Interdisciplinary, other	68.5	31.5	62.3	37.7	60.2	39.8
Total	67.7	32.3	59.9	40.1	53.5	46.5

Table 2.4: **University students in the different faculty**

Percentage of the participation of students in the different university faculty for the years 1980, 1992 and 2001.

Professional Sector 2001/02	M%	F%
Textile Industry and clothing	11.1	88.9
Hygiene and Treatment	11.2	88.8
Organisation, Administration and Commerce	37.0	63.0
Artistic Profession	58.3	41.7
Hotel Industry and Domestic Science	67.4	32.6
Agriculture and Breeding	86.3	13.7
Interdisciplinary, other	68.5	31.5
Leather and Scalp-Working	50.0	50.0
Communications and Transports	72.7	27.3
Graphic art	83.1	16.9
Design and Technical profession	81.2	18.8
Jewels and Watchmaker Industry	78.3	21.7
Land, Stone and Glass-Working	100	0
Painting	97.6	2.4
Wood and Cork-Working	99.4	0.6
Building Industry	100	0
Chemistry Industry	100	0
Metallurgist Industry and Machine	98.7	1.3
Cleaning	59.4	40.6
Others Professions	97.1	2.9
Total	69.6	30.4

Table 2.5: **Students in the different vocational study**

In the table are reported the percentage of women and men in different vocational study in Switzerland for the year 2001/02.

Occupation	Prop. of women
Construction	0.016
Manufacturing: Precision-tool	0.099
Administration: Government and Business	0.101
Workshop and Construction	0.139
Transport and Communication	0.161
Analysis and Development	0.166
Security service (police)	0.181
Manufacturing	0.187
Research and development	0.220
Sales and Purchase (business services)	0.227
Logistics, General Staff	0.251
Agriculture and Breeding	0.280
Law	0.344
Others Activities	0.363
Sport and culture	0.421
Hotel Industry and Domestic Science	0.533
Accounting, personnel management, finance management	0.545
Administration: Other	0.574
Education	0.618
Public cleaning	0.627
Retail Sales (goods and services)	0.715
Secretariat and back office	0.782
Medicine, Health and Nursing	0.809
Hygiene and Treatment	0.836

Table 2.6: **Occupational Specification and Proportion of Women in each Occupation**

Values used for the construction of the variable FEM.

Census year	1970	1980	1990	2000
Number of occupations	321	321	378	378
Females	1’019’896	1’107’960	1’402’009	1’654’996
Males	1’969’904	1’959’273	2’178’904	2’134’420
Total	2’989’800	3’067’233	3’580’913	3’789’416
Weighted average, F/M	0.51774	0.56549	0.64344	0.77538
Dissimilarity Index	0.62849	0.61193	0.54662	0.44098
Size-Standardized Dissimilarity Index	0.65629	0.63299	0.57706	0.55801

Table 2.7: **Measures of the occupational distribution in Switzerland 1970-2000**

Summary measures of the occupational distribution of the gender ratio in Switzerland 1970-2000.

Variables	Males Est.	Males Std. Err.	Females Est.	Females Std. Err.
Constant term	8.092676	0.003095	7.993337	0.003202
Married	0.050017	0.001477	-0.037031	0.001607
Formation2	0.135631	0.001971	0.172148	0.001977
Formation3	0.232371	0.002369	0.255005	0.002568
Formation4	0.511448	0.002659	0.563092	0.003631
Foreign	-0.127841	0.00173	-0.102204	0.002293
Permit C	-0.097293	0.001702	-0.0779	0.002139
Experience	0.024965	0.000257	0.021843	0.000273
Experience squared	-0.042047	0.000508	-0.040457	0.000568
Occupational seniority	0.007852	0.000216	0.00934	0.000299
Occupational seniority squared	-0.01028	0.000619	-0.009436	0.001026
Position 2	0.083919	0.002109	0.131335	0.00257
Position 3	0.21796	0.002037	0.216754	0.002852
Position 4	0.359339	0.002394	0.320205	0.004135
Position 5	0.509075	0.002717	0.363924	0.006251
Observations	220158		143435	
R-squared	0.524801		0.366437	
F-statistic	17299.72		5922.334	

All values are significant at 1 percent level

Table 2.8: **Regression analysis of the salaries of active employees in 2000**

In this table I have made two separeted regressions on the gender, without control for occupational segregation.

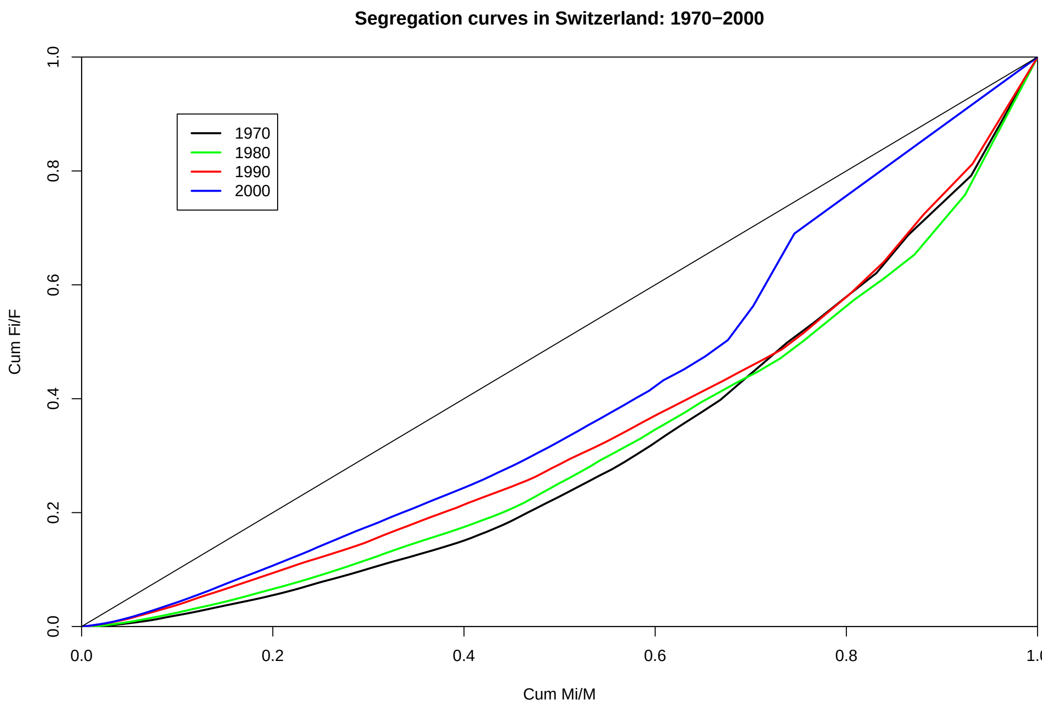

Figure 2.1: **Segregation curves: 1970-2000**

The figure reports the segregation curves, in Switzerland, for the years: 1970-2000. The curves are obtained by using the cumulative of the ratio F_i/F and M_i/M.

Variables	Males Est.	Males Std. Err.	Females Est.	Females Std. Err.
Constant term	8.117092	0.003286	8.021562	0.003762
FEM	-0.065828	0.002939	-0.05307	0.003682
Married	0.048836	0.001483	-0.036342	0.001612
Formation2	0.134677	0.001974	0.176709	0.002003
Formation3	0.233442	0.002374	0.258086	0.002583
Formation4	0.512722	0.002673	0.565456	0.003695
Foreign	-0.130661	0.001737	-0.10538	0.002304
Permit C	-0.098775	0.001706	-0.079398	0.002144
Experience	0.024741	0.000258	0.021755	0.000274
Experience squared	-0.041637	0.000509	-0.040168	0.000569
Occupational seniority	0.007669	0.000216	0.009065	0.0003
Occupational seniority squared	-0.010027	0.000621	-0.00889	0.001029
Position 2	0.086725	0.002118	0.131665	0.002581
Position 3	0.221519	0.002047	0.215104	0.002862
Position 4	0.363247	0.002412	0.317365	0.004153
Position 5	0.50654	0.002724	0.354881	0.006293
Observations	218720		142724	
R-squared	0.525746		0.365	
F-statistic	16115.85		5463.79	

All values are significant at 1 percent level

Table 2.9: **Regression analysis of the salaries of active employees in 2000**

In this table I have made two separate regressions on the gender, introducing also the new variable FEM which capture the density of female in the labour force.

WOMEN						
	Male Dominated Occupation		Intermediate Occupation		Female Dominated Occupation	
Characteristics	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Wage monthly	4'902.20	2'443.20	5'441.50	2'309.70	4'414.80	1'526.20
Married	0.55	-	0.47	-	0.53	-
Formation 2	0.40	-	0.58	-	0.60	-
Formation 3	0.13	-	0.18	-	0.12	-
Formation 4	0.07	-	0.10	-	0.03	-
Foreign	0.20	-	0.12	-	0.11	-
Permit C	0.22	-	0.13	-	0.15	-
Experience	22.34	11.40	20.00	11.53	21.42	12.16
Occupational seniority	9.09	7.97	7.61	7.51	7.43	6.99
Position 2	0.06	-	0.11	-	0.08	-
Position 3	0.05	-	0.11	-	0.05	-
Position 4	0.03	-	0.05	-	0.02	-
Position 5	0.03	-	0.02	-	0.00	-
Part time	0.29	-	0.38	-	0.52	-
Skill 2	0.35	-	0.51	-	0.48	-
Skill 3	0.10	-	0.17	-	0.08	-
Skill 4	0.03	-	0.02	-	0.01	-
Firm Medium	0.13	-	0.16	-	0.10	-
Firm Large	0.28	-	0.22	-	0.16	-
Firm Extra-large	0.52	-	0.50	-	0.63	-
Age	39.34	10.91	38.04	11.22	38.62	12.00
Number of observations	37'069		69'912		98'220	

Table 2.10: **Descriptive statistics by occupational type, for females**

MEN						
	Male Dominated Occupation		Intermediate Occupation		Female Dominated Occupation	
Characteristics	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Wage monthly	6'658.70	4'015.40	7'627.90	4'512.10	5'547.10	2'527.10
Married	0.67	-	0.59	-	0.56	-
Formation 2	0.57	-	0.50	-	0.59	-
Formation 3	0.15	-	0.22	-	0.16	-
Formation 4	0.12	-	0.20	-	0.07	-
Foreign	0.19	-	0.14	-	0.14	-
Permit C	0.20	-	0.13	-	0.20	-
Experience	23.41	11.21	20.56	11.17	21.07	11.93
Occupational seniority	11.40	9.91	9.75	9.37	8.96	8.73
Position 2	0.09	-	0.10	-	0.10	-
Position 3	0.09	-	0.16	-	0.10	-
Position 4	0.06	-	0.12	-	0.06	-
Position 5	0.06	-	0.07	-	0.03	-
Part time	0.05	-	0.08	-	0.17	-
Skill 2	0.48	-	0.46	-	0.49	-
Skill 3	0.19	-	0.29	-	0.18	-
Skill 4	0.09	-	0.10	-	0.04	-
Firm Medium	0.15	-	0.13	-	0.10	-
Firm Large	0.29	-	0.22	-	0.18	-
Firm Extra-large	0.45	-	0.54	-	0.63	-
Age	41.40	11.12	39.54	10.89	38.49	12.05
Number of observations	172'418		75'561		35'612	

Table 2.11: **Descriptive statistics by occupational type, for males**

	Males		Females	
Variables	**Est.**	**Std. err.**	**Est.**	**Std.err.**
Married	-0.08896	0.006257	-0.03381	0.006745
Formation 2	-0.03312	0.013404	0.329821	0.014813
Formation 3	0.028105	0.017027	0.20387	0.019072
Formation 4	-0.29956	0.027374	-0.0395	0.035314
Foreign	-0.13553	0.007779	-0.18335	0.009462
Permit C	-0.06741	0.007572	-0.07707	0.008886
Experience	-0.04971	0.003832	-0.00913	0.004792
Experience squared	0.024758	0.002165	0.042763	0.00232
Occupational seniority	-0.02618	0.000943	-0.01563	0.00122
Occupational seniority squared	0.033901	0.002723	0.007951	0.004142
Position 2	0.227925	0.009068	0.022753	0.010609
Position 3	0.34972	0.008996	-0.05589	0.012102
Position 4	0.430752	0.010921	-0.13524	0.017629
Position 5	0.194338	0.014593	-0.46832	0.028317
Part-time	0.57928	0.01011	0.429596	0.006765
Skill 2	-0.03683	0.008418	0.165741	0.008042
Skill 3	0.068513	0.010598	0.108467	0.012259
Skill 4	-0.11466	0.015143	-0.09712	0.026125
Firm Medium	-0.05127	0.009808	-0.20953	0.010973
Firm Large	-0.07825	0.009196	-0.32989	0.010352
Firm Extra-large	0.221783	0.008855	-0.15272	0.009673
Age	0.036229	0.003736	-0.00984	0.004708
$\mu 1$	0.740181	0.059132	-1.08817	0.073572
$\mu 2$	1.701346	0.059193	0.005097	0.073553
N	218104		142509	

All values are significant at 1 percent level

Table 2.12: **Ordered probit estimates**

Estimation of the probit model for male and female, where the dependent variable admits three values: 0 if male dominated occupation, 1 if intermediate and 2 if female dominated.

Variable	Male dominated Est.	Male dominated Std.err.	Intermadiate Est.	Intermadiate Std.err.	Female dominated Est.	Female dominated Std.err.
Constant term	8.025139	0.00674	7.99943	0.003082	8.01249	0.004785
Married	-0.035233	0.00168	-0.035172	0.001679	-0.035272	0.001681
Formation 2	0.178901	0.00219	0.179336	0.002198	0.178505	0.002175
Formation 3	0.259513	0.00284	0.259873	0.002847	0.259167	0.002829
Formation 4	0.565687	0.004151	0.565554	0.00415	0.565838	0.004151
Foreign	-0.108573	0.002663	-0.108927	0.002667	-0.108238	0.002657
Permit C	-0.080923	0.002223	-0.081129	0.002227	-0.080731	0.002218
Experience	0.021571	0.000269	0.021553	0.000269	0.021588	0.000269
Experience squared	-0.039721	0.000584	-0.039673	0.000585	-0.039763	0.000584
Occupational seniority	0.008977	0.000326	0.008958	0.000326	0.008991	0.000326
Occupational seniority squared	-0.008891	0.001132	-0.008895	0.001132	-0.008879	0.001132
Position 2	0.132205	0.002471	0.13226	0.002471	0.132153	0.00247
Position 3	0.214771	0.003031	0.214715	0.003031	0.214811	0.003031
Position 4	0.315682	0.005117	0.315387	0.005117	0.315975	0.005115
Position 5	0.352383	0.01001	0.350934	0.010017	0.35388	0.009996
λ	0.022178	0.004238	0.02029	0.003687	-0.028299	0.005618
R-squared	0.362721		0.362708		0.362741	
N	142564		142564		142564	

All values are significant at 1 percent level

Table 2.13: **Wage equation estimates for females by occupation**

In this table I have made three separated regressions on the occupation for the female sample. The dependent variable is the logarithm of wage.

Variable	Male dominated Est.	Male dominated Std.err.	Intermadiate Est.	Intermadiate Std.err.	Female dominated Est.	Female dominated Std.err.
Constant term	7.931463	0.005474	8.05603	0.003301	8.000197	0.00427
Married	0.061083	0.001523	0.06298	0.001524	0.058618	0.001519
Formation 2	0.134983	0.001624	0.133663	0.001623	0.135998	0.001626
Formation 3	0.213733	0.002357	0.210011	0.002362	0.218263	0.002348
Formation 4	0.519358	0.002798	0.519243	0.002794	0.519051	0.002801
Foreign	-0.106422	0.001896	-0.101749	0.001913	-0.112043	0.001874
Permit C	-0.083989	0.001764	-0.081186	0.00177	-0.087379	0.001756
Experience	0.027041	0.000266	0.027246	0.000265	0.026687	0.000266
Experience squared	-0.046038	0.000535	-0.046413	0.000534	-0.045379	0.000536
Occupational seniority	0.010814	0.000234	0.011377	0.000236	0.010126	0.000232
Occupational seniority squared	-0.014197	0.000636	-0.014819	0.000636	-0.01336	0.000636
Position 2	0.060423	0.001958	0.055414	0.001974	0.066483	0.001934
Position 3	0.176074	0.002371	0.16898	0.002389	0.185502	0.002327
Position 4	0.313793	0.003214	0.305959	0.003231	0.324137	0.003175
Position 5	0.50151	0.004385	0.499722	0.004388	0.503341	0.004382
λ	-0.200986	0.005992	-0.141263	0.00368	0.297932	0.010769
R-squared	0.52381		0.524391		0.523168	
N	218283		218283		218283	

All values are significant at 1 percent level

Table 2.14: **Wage equation estimates for males by occupation**

Estimation for male. The dependent variable is the logarithm of wage.

	OLS	OLS with FEM	Male-Female Wage Gaps: Male Dominated	Intermediate	Female Dominated
Observed Gap	0.1332	0.1207	0.0211	0.2424	0.0770
	(0.0010)	(0.0010)	(0.0066)	(0.0032)	(0.0045)

Table 2.15: **Observed wage gaps**

OLS with FEM: estimation with control for occupational segregation.

Chapter 3

Determinants of vertical occupational segregation and the wage gap

3.1 Introduction

Traditionally, the gender wage gap has been explained by gender differences in the level of human capital and by discriminatory forces in the labour market that reduce the earning capacity of women. In the previous chapter I have analysed the gender wage gap from the point of view of horizontal occupational segregation. Now, I am interested in studying the characteristics of vertical occupational segregation, namely how the different phases of the careers affect the wage of employees and how male and female differ in career opportunities.

Income growth is the typical marker for a successful career. However, careers also include episodes of joblessness, which may inhibit career development. Since careers unfold over time, analyzing job rewards in cross-sectional samples overlooks an important sorting process that creates and maintains wage disparities among workers. The purpose of this chapter is to give new insights into the dynamic nature of the career development by studying the impact of some specific variables on the work mobility through job hierarchy.

The chapter proceeds as follows. In the second section, theoretical bases are exposed. The data are described in the third section. In the fourth section, empirical models are discussed to identify the main variables that affect wages and level of the job hierarchy. The methods of estimation are presented in the fifth section, and in the sixth section is illustrated the computation of the gender wage gap. Estimation results are reported in the seventh section. In the last section, the main results are summarized.

3.2 Theoretical bases

Much of the research on occupational segregation (Acker 1990, Baron and Newman 1990, Reskin 1988) has examined its effect on reward. However the findings from these studies imply also that segregation affects promotion chances. That is, unequal career opportunities exist for men and women, with women often confined to dead-end jobs. Despite the increasing presence of women in corporations, they fail to attain top managerial positions in the firm. Numerous studies indicate that women and men with similar abilities are allocated to different positions when entering in a firm (Bielby and Baron, 1986; Kanter, 1977; Reskin, 1988; Tomaskovic-Devey, 1993).

Three explanations for the phenomenon of vertical segregation are dominant in the sociological literature on organizational discrimination against women.

First, Kanter (1977) documented how women are excluded from informal work groups when they reach a management position, thereby contributing to women's failure as managers. Kanter (1977) argues that managers are often unsure about their job performance and are dependent on colleagues in the firm to accomplish their goals. In particular, managers must build a network to do an effective job, and relationships are easier to establish when managers share the same background and experiences. Consequently, firms create segregated job ladders for women, which results in their slower progress within the firm.

A second explanation suggests that men resist to attempts at gender integration in order to preserve their privileged position within the firm (Reskin, 1988). By segregating women into specific occupations with short job ladders, men are free to compete among themselves for higher paying jobs that offer better career opportunities. This process is referred to as "social closure", in which employers discriminate in hiring and offering promotions, often under the influence of their male employees.

The third way in which women are disadvantaged with respect to men is the process of "status composition". That is, jobs with a large number of female incumbents are devalued in the eyes of the organization. Female-typed jobs are viewed as having lower skill requirements, which explains their low pay and minimal promotion opportunities.

Although these studies undoubtedly offer a very valuable contribution to the literature, their reliance on cross-sectional samples precludes an examination of career changes. By definition, careers take place over time as individuals move between jobs, and movements from non supervisory to supervisory position is often associated with wage growth and career development. For this reason, in this chapter, I use a panel data for my empirical analysis. Another advantage of panel data is the possibility to account for individual heterogeneity by means of individual specific effects. Cross-sectional studies not controlling for this heterogeneity may obtain biased results.

3.3 Data and descriptive statistics

The sample used in this chapter is a sub-sample of the Swiss Labour Force Survey consisting of individuals who are employed in each year during the period 1999-2003. I have excluded apprentices because they are not representative for my analysis, and also self-employees for reasons which will be explained in the next paragraph. We get a balanced panel data with 927 individuals, of which 484 are men and 443 women, followed over 5 years.

The list of variables with their definitions is reported in Table 3.1. Table 3.2 reports some descriptive statistics of the data. The two dependent variables used in this chapter are the natural log of monthly earnings and the level of occupational hierarchy (denominated as *career*). The *career* variable is constructed from a variable which has originally 5 levels: (1) self-employees , (2) directors, (3) employees with supervisory position, (4) employees without supervisory position and (5) family members (namely members of the family who work in the family company, i.e. the wife of a farmer who helps her husband in the farm, or the son who works with his father). To illustrate the data and to estimate a binary probit model (see section 3.4), this variable is aggregated into two levels: high (individuals with supervisory position and directors) and low (including individuals in a non supervisory position and family members). Therefore, we consider in the high level respectively position (2) and (3), and in the low level position (4) and (5). The category self-employees (1) is not considered in our *career* variable. In fact, self-employees do not always have a managing role as directors or employees with supervisor position, and also do not always have a low hierarchical position. So it is difficult to identify in which category self-employee belong. Therefore, we decide to drop out of the sample all individuals with the characteristics of self-employee in at least one year.

Considering the independent variables, the marital status variable is measured by an indicator variable: *married* takes the value 1 if the person is legally married, and 0 otherwise. If the person is living with a partner but not legally married is considered a single person. The sample contains information on the type of education. Therefore we can construct 4 dummy variables for the level of the formation, from the lowest to the highest level. The 4 categories are: compulsory school, apprenticeship, high school and university degree. We consider as well the variable $foreign$, which is equal to 1 if the individual is not Swiss and equal to 0 otherwise. We know, also, the number of years an employee is working foe the current employer. We define this variable as *occupational seniority*. However, we do not have the exact number of years of actual job experience, and a proxy variable is used instead. Following the literature, potential work experience is defined as $experience_{it} = age_{it} - Edu_{it} - 6$, where age_{it} is corrected for the number of years of formal schooling Edu_{it}, and the age at which children start school, which is equal to 6. A problem of using this proxy variable is that, unless individuals work full-time and continuously, this proxy measures actual work experience with error and hence application of OLS may give inconsistent estimates. This is because working life cycles for both males and female may be characterised by frequent interruptions, especially in the case of female. For the discussion the results of the estimated model, we have to remember the limits of this variable. In addition, in the analysis is included the variable $part - time$, which is equal to 1 if the individual works for a percentage less than 90%, and 0 otherwise. The *children* dummy variable is defined equal to 1 if the individual has children of five years old or less, and 0 otherwise.

Figure 3.1 displays the evolution of the females and males average wages by two different career levels. We can notice that although the female wages increases during the years[1], it is always behind the low level of male average wages. In fact the male average wage in the case of a low level of the career is slightly higher than the female average wage in the high level.

Figure 3.2 shows the evolution of the average years of occupational seniority in the two levels of the career (high and low). What is remarkable is that for males and females the occupational seniority is typically larger in the case of high level of the job hierarchy. And it is not surprising that the occupational seniority for males, in both states of career, is always greater than for females. We can also notice that the occupational seniority for females, in the case of high level of the job hierarchy, has a large increase during the period 1999-2003.

In Figure 3.3 we display the distribution of the different formation levels in the sample, by distinguishing between males and females in the two career levels. What is interesting is that from all women working in a high career level only 15% have a university degree. The majority, 55%, have an apprenticeship degree.

In Figure 3.4, the ratio of female wages to male wages is reported for the different educational levels during the five years. If there is perfect equality between female wages and male wages the ratio will be equal to 1. In our case, in all formation groups, this ratio is below 1 in all years. The figure shows that among employees with only compulsory degree, women's wages are about the 70% of those of men (for the years 2001-2003). This means that the wage differential is about 30%, namely women earn 30% less of men. In the other groups the ratio is around 80-85%. We can conclude that the higher the formation of an individual, the smaller the wage differential.

Figure 3.5 displays the ratio of female wages to males wages for the two classes of career (low and high). Also in this case, this ratio is below 1 in all years. It appears that the highest overall difference between male and female wages is among employees

[1]Remember that wages are nominal wages. The wage growth rate in real terms can be obtained by taking into account inflation.

who have a high career level. In this group women's wages are about 75% of those of men. This means that wage gap between males and females with high career level is about 25%. Instead, for the low level we have a 15% wage gap.

Table 3.3 reports the shares of employees in the sample by educational level. In the third column it is seen that during the years 1999-2003 about half of the individuals in the sample have an apprenticeship degree and about a third have a university degree. Groups corresponding to low levels of education (compulsory school and apprenticeship) consists of approximately equal shares of men and women (fourth column). Women dominate the group of employees who have the high school degree as highest education level; 65 per cent of the employees in this group are women. Probably this is do to the fact that in this category are also included individuals with teacher's diploma. The group of individuals with university degree is dominated by men.

In Table 3.4 are reported the shares of employees by career level. It can be seen that about 60% of the individuals are in the low hierarchy level and about 50% in the high level. It appears that about 65 per cent of the employees who have a high career level are men. The low level of the work hierarchy consist for 57.5% of women.

3.4 The model

In this section I am interested in the impact of individual characteristics on the chance of receiving a job promotion, and in the evaluation of the gender wage gap at the different levels of the job career. To conduct this analysis we start by considering a simple panel model for a wage equation (section 3.4.1). Thereafter, we introduce in the wage equation a dummy variable which admit two values which are, 1 if the employee is in a high level of the job hierarchy, and 0 if he/she is in the low level (section 3.4.2). Finally, we specify the probability for an employee to be in the high level of the job hierarchy as a Probit model with random effects (section 3.4.3).

To introduce the models used for the empirical analysis, consider a sample of n individuals indexed by i ($i = 1, ..., n$), and let T denote the number of observations for individual (in our case $n = 927$ and $T = 5$). Note that T does not vary across individuals, therefore the sample is a balanced panel.

3.4.1 Wage equation

A simple panel model of wage determination, which has often been considered in the empirical literature[2], is

$$y_{it} = x'_{it}\beta + \alpha_i + w_{it} \qquad i = 1, .., n \quad \text{and} \quad t = 1, .., T \tag{3.1}$$

where the dependent variable y_{it} is the logarithmic wage. The $k \times 1$ vector of explanatory variables x_{it}, which are assumed exogenous, includes measures for observed human capital characteristics, which can be time varying or time invariant. The unobservable error term contains an individual specific component, α_i, which is constant over time, and an idiosyncratic error term, w_{it}. The individual specific component α_i captures unobserved individual specific skills, such as motivation or ability, while the idiosincratic error component w_{it} accounts for transitory shocks or luck. We assume further that

$$\begin{aligned}
E[\alpha_i] &= E[w_{it}] = 0, \\
E\left[w_{it}^2\right] &= \sigma_w^2, \\
E\left[\alpha_i^2\right] &= \sigma_\alpha^2, \\
E[w_{it}\alpha_j] &= 0 \quad \text{for all} \quad i, t \quad \text{and} \quad j, \\
E[w_{it}w_{js}] &= 0 \quad \text{if} \quad t \neq s \quad \text{or} \quad i \neq j, \\
E[\alpha_i\alpha_j] &= 0 \quad \text{if} \quad i \neq j.
\end{aligned} \tag{3.2}$$

[2] Mincer and Polanchek (1987), Dolton and Makepeace (1986).

For the T observations of individual i let us now define

$$\varpi_{it} = \alpha_i + w_{it} \qquad t = 1, \ldots, T, \tag{3.3}$$

and

$$\varpi_i = \begin{bmatrix} \varpi_{i1} & \varpi_{i2} & \ldots & \varpi_{iT} \end{bmatrix}'. \tag{3.4}$$

This error structure for ϖ_{it}, is often called an error component model[3] with random effects. Then we have,

$$Var(\varpi_{it}) = \sigma_\alpha^2 + \sigma_w^2, \tag{3.5}$$

and

$$Cov(\varpi_{it}, \varpi_{is}) = \sigma_\alpha^2 \quad \text{for} \quad t \neq s. \tag{3.6}$$

The individual specific effect α_i introduce a correlation between the observation of the same individual. For the T observations of individual i, let $\Omega = E[\varpi_i \varpi_i']$. Then

$$\Omega = \begin{bmatrix} \sigma_\alpha^2 + \sigma_w^2 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 + \sigma_w^2 & \sigma_\alpha^2 & \vdots \\ \vdots & & \ddots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 & \sigma_\alpha^2 + \sigma_w^2 \end{bmatrix}. \tag{3.7}$$

[3]Balestra and Nerlove (1966), Matyas and Sevestre (1996).

Since individuals i and j are independent, the disturbance covariance matrix for the full sample of nT observations is

$$V = \begin{bmatrix} \Omega & 0 & \cdots & 0 \\ 0 & \Omega & 0 & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \Omega \end{bmatrix} = I_n \otimes \Omega. \tag{3.8}$$

3.4.2 Wage equation with career effect

We now extend the panel model (3.1) to account for the effect on wage of the different levels of career. Consider the dummy variable d_{it}, which can take two values, 1 if the employee i is in the high level of the job hierarchy at t and 0 if he/she is in the low level. We write

$$d_{it} = j, \quad j = 0, 1 \tag{3.9}$$

where j corresponds to the two admissible states in job hierarchy. The wage equation is given by,

$$y_{it} \;=\; x_{it}'\beta_j + \alpha_i + w_{it} \qquad \text{if} \quad d_{it} = j, \quad j = 0, 1, \quad i = 1, .., n \quad \text{and} \quad t = 1, .., T, \tag{3.10}$$

where the vector of the coefficients β_j is different in the two states ($j = 0$ and $j = 1$), the same assumptions (3.2) as before apply to the error term in equation (3.10). Thus, we allow for an effect of career level on the coefficients of the wage regressions. Since x contains a constant and explanatory variables, this effect may correspond in general to a shift in the level or in the slope of the wage function. Note that when $\beta_0 = \beta_1$ we are in the special case of the equation (3.1). In the next section we introduce the process that determines whether the individual is in the high level or in the low level of career at different dates.

3.4.3 Probit model for career effect

The model for d_{it} is deduced from a latent regression. More specifically, we consider the following panel model

$$d^*_{it} = z'_{it}\gamma + \epsilon_{it} \qquad i = 1, .., n \quad \text{and} \quad t = 1, .., T \tag{3.11}$$

where d^*_{it} denotes a latent continuous variable, z_{it} is an observable vector of strictly exogenous characteristics, γ is a vector of unknown coefficients and ϵ_{it} is an unobservable error. The observable endogenous variables d_{it} are given by:

$$d_{it} = 1 \quad \text{if} \quad d^*_{it} > 0 \qquad \text{and} \qquad = 0 \quad \text{otherwise.} \tag{3.12}$$

The error terms ϵ satisfy a random effects specification:

$$\epsilon_{it} = u_i + v_{it} \tag{3.13}$$

where u_i is an individual specific effect and v_{it} is a idiosyncratic component. In the Probit random effects model, the errors u_i and v_{it} ($i = 1, \ldots, n,\ t = 1, \ldots, T$) are assumed mutually independent and normally distributed with $u_i \sim N(0, \sigma^2_u)$ and $v_{it} \sim N(0, \sigma^2_v)$. We allow, however, for correlation between u_i and α_i, the individual specific component defined in the wage equation (3.1). This is natural since both u_i and α_i are specific effects for individual i.

From the unobservability of d^*_{it}, it follows that parameters β, σ^2_u and σ^2_v are identified only up to a multiplicative constant. As identification condition we set $\sigma^2_v = 1$. Thus, the errors ϵ_{it} are normally distibuted independent across individuals and such that:

$$Var(\epsilon_{it}) = \sigma^2_v + \sigma^2_u = 1 + \sigma^2_u \tag{3.14}$$

and

$$Corr(\epsilon_{it}, \epsilon_{is}) = \rho = \frac{\sigma^2_u}{1+\sigma^2_u}, \quad t \neq s. \tag{3.15}$$

The equicorrelation between the errors of individual i at different dates is originated by the individual effect u_i.

In the panel probit model the probability of an individual i to be in a high level of job hierarchy at time t is:

$$P[d_{it} = 1 \mid z_{it}] = P\left[d^*_{it} > 0 \mid z_{it}\right] = P[\epsilon_{it} > -z'_{it}\beta] = \Phi\left(\frac{z'_{it}\gamma}{\sqrt{1+\sigma^2_u}}\right), \tag{3.16}$$

where Φ denotes the c.d.f. of the standard normal distribution. The probability of promotion at time t is an increasing transformation of the score $\frac{z'_{it}\gamma}{\sqrt{1+\sigma^2_u}}$. Thus, at each date t the variable d_{it} follows a standard Probit specification. Note however that the individual effect u_i introduces dependence across observations of the same individual at different dates (cf. (3.15)).

The wage equation (3.10) and the Probit model for career effect (3.11)-(3.12) define a panel switching regression model. For the sake of the future reference, we repeat here the equations,

$$y_{it} = x'_{it}\beta_j + \alpha_i + w_{it} \quad if \quad d_{it} = j, \quad j = 0, 1 \tag{3.17}$$

$$d_{it} = 1 \quad if \quad d^*_{it} > 0 \quad and \quad = 0 \quad otherwise \tag{3.18}$$

$$d^*_{it} = z'_{it}\gamma + u_i + v_{it} \tag{3.19}$$

$$i = 1, .., n \quad and \quad t = 1, ..., T$$

3.5 Estimation

In this section we discuss the estimation methods for the models introduced in section 3.4.

3.5.1 Random effects estimator

The panel model specified in equation (3.1) does not satisfy the conditions of the classical regression. In fact, the matrix of variance-covariance is not a diagonal matrix (cf. (3.7)-(3.8)). In this case the BLUE estimator is the Generalized Least Squares (GLS) estimator

$$\widetilde{\beta} = \left(\sum_{i=1}^{n} x_i' \Omega^{-1} x_i \right)^{-1} \sum_{i=1}^{n} x_i' \Omega^{-1} y_i, \tag{3.20}$$

where x_i is a $T \times k$ matrix of explanatory variables for individual i and y_i is the vector of T observations y_{it} for individual i. The $T \times T$ matrix Ω is the variance-covariance matrix defined in equation (3.8).

In our case Ω contains unknown parameters that must be estimated and the GLS is not feasible. To make GLS estimation feasible, we shall use $\widehat{\Omega}$ instead of Ω, where $\widehat{\Omega}$ is obtained using some estimators $\widehat{\sigma}_\alpha$ and $\widehat{\sigma}_w$. The Feasible Generalized Least Squares[4] (FGLS) becomes

$$\widehat{\beta} = \left(\sum_{i=1}^{n} x_i' \widehat{\Omega}^{-1} x_i \right)^{-1} \sum_{i=1}^{n} x_i' \widehat{\Omega}^{-1} y_i. \tag{3.21}$$

[4]Matyas and Sevestre (1996).

The FGLS estimator $\widehat{\beta}$ is consistent under the hypothesis of strict exogeneity, that is

$$E\left[\alpha_i + w_{it} \mid x_i\right] = 0 \qquad \forall i, t \tag{3.22}$$

which we have assumed in section 3.4.1. If the condition of strict exogeneity is not satisfied, we must proceed with an IV approach. We will not consider this approach in this work.

3.5.2 Two-stage estimation of the panel switching regression

The wage equation model with career effect (3.17)-(3.19) can be rewritten as,

$$\begin{aligned} y_{it} &= d_{it}x'_{it}\beta_1 + (1-d_{it})x'_{it}\beta_0 + \alpha_i + w_{it} \\ &= x'_{it1}\beta_1 + x'_{it0}\beta_0 + \alpha_i + w_{it}, \end{aligned} \tag{3.23}$$

where $x_{it1} = x_{it}d_{it}$, $x_{it0} = x_{it}(1-d_{it})$ and d_{it} is defined in (3.17)-(3.19). This is a switching regression model for panel data.

The natural estimation approach for the specification in equation (3.20) is a Random Effects estimator as in section 3.5.1. However, the orthogonality condition between the explanatory variables (x_{it1}, x_{it0}) and the individual error α_i is violated, that is

$$E[\alpha_i \mid x_{it1}, x_{it0}] \neq 0. \tag{3.24}$$

In fact, the explanatory variables x_{it1} and x_{it0} contain the selection variable d_{it}, which by definition includes the individual error term u_i, which is likely correlated with α_i. This endogeneity problem causes in general inconsistency of the random effects estimator.

To overcome the problem of endogeneity, a two-stage estimation method can be used. We opt for a the two-stage method, because it gives us estimates that are consistent and easy to compute[5].

A two-stage method was suggested by Heckman (1976) for his labour-supply model. It was further extended to a wide class of models[6] and is widely used in the literature. The basic idea of the two-stage approach is to first estimate the selection equation for d_{it} by a Panel Probit method, and then to use these estimates to correct the wage equation for y_{it} including the expectation of the error term. These two stages are presented in the following subsections.

Probit model with Random Effects

As shown in Heckman (1981), the parameters of model (3.11)-(3.13) can be estimated by noting that, conditionally on individual effect u_i, the variables d_{it}, $t = 1, \ldots, T$, for individual i are independent. Their distribution is such that:

$$P[d_{it} = 1 \mid z_{it}, u_i] = P\left[v_{it} > -z'_{it}\gamma - u_i\right] = \Phi\left(z'_{it}\gamma + u_i\right), \tag{3.25}$$

(recal that $\sigma_v^2 = 1$). The likelihood of the observations for individual i is obtained by integrating out factor u_i:

$$l_i(\theta) = \int \left\{ \prod_{t=1}^{T} \Phi(z'_{it}\gamma + u_i)^{y_{it}} [1 - \Phi(z'_{it}\gamma + u_i)]^{1-y_{it}} \right\} \frac{1}{\sigma_u} \phi\left(\frac{u_i}{\sigma_u}\right) du_i \tag{3.26}$$

[5] Maximum likelihood estimation (MLE) could also be used (see Vella and Verbeek (1999) and Ridder (1990)), but this approach is computationally more complicated.

[6] Lee(1976), Amemya (1978, 1979).

where ϕ is the p.d.f. of the standard normal distribution, and $\theta = (\gamma, \sigma_u^2)$. The maximum likelihood estimator is obtained by maximizing the log-likelihood

$$L(\theta) = \sum_{i=1}^{n} \log l_i(\theta). \tag{3.27}$$

Pooled OLS with correction term

To derive the correction term, we take the expectation of the wage equation (3.17), conditional on explanatory variables and get

$$E[y_{it} \mid x_{it}, z_{it}, d_{it}] = x'_{it1}\beta_1 + x'_{it0}\beta_0 + E[\alpha_i \mid x_{it}, z_{it}, d_{it}], \tag{3.28}$$

where we used that $E[w_{it} \mid x_{it}, z_{it}, d_{it}] = 0$. To compute $E[\alpha_i \mid x_{it}, z_{it}, d_{it}]$ we distinguish two cases, when $d_{it} = 1$ (high level of the job hierarchy) and $d_{it} = 0$ (low level of the job hierarchy). Thus, we need to evaluate the expectation for the individual error term in the two states, namely,

$$\begin{aligned} E[\alpha_i \mid x_{it}, z_{it}, d_{it} = 1] &= E[\alpha_i \mid \epsilon_{it} > -z'_{it}\gamma] \\ &= E\left[E[\alpha_i \mid \epsilon_{it}] \mid \epsilon_{it} > -z'_{it}\gamma\right], \end{aligned} \tag{3.29}$$

and

$$\begin{aligned} E[\alpha_i \mid x_{it}, z_{it}, d_{it} = 0] &= E[\alpha_i \mid u_i + v_{it} < -z'_{it}\gamma] \\ &= E\left[E[\alpha_i \mid u_i + v_{it}] \mid u_i + v_{it} < -z'_{it}\gamma\right], \end{aligned} \tag{3.30}$$

where conditioning on z_{it}, x_{it} in the right hand side of (3.29) and (3.30) is omitted for the purposes of exposition. Using $\epsilon_{it} = u_i + v_{it}$, the joint distribution of the error terms is

$$\begin{pmatrix} \alpha_i \\ \epsilon_{it} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \sigma_{\alpha u} \\ \sigma_{\alpha u} & 1 + \sigma_u^2 \end{pmatrix}\right), \tag{3.31}$$

where $\sigma_{\alpha u}$ is the correlation between α_i and u_i. Now, we can evaluate the two expectations in (3.29) and (3.30) knowing that, from joint normality,

$$E[\alpha_i \mid \epsilon_{it}] = \frac{\sigma_{\alpha u}}{1+\sigma_u^2}\epsilon_{it}. \tag{3.32}$$

Moreover from the standard results for the Probit model (see Maddala, 1999) and $V(\epsilon_{it}) = 1+\sigma_u^2$ we have

$$E[\epsilon_{it} \mid \epsilon_{it} > -z_{it}'\gamma] = \sqrt{1+\sigma_u^2}\frac{\phi\left(\frac{z_{it}'\gamma}{\sqrt{1+\sigma_u^2}}\right)}{\Phi\left(\frac{z_{it}'\gamma}{\sqrt{1+\sigma_u^2}}\right)}, \tag{3.33}$$

and

$$E[\epsilon_{it} \mid \epsilon_{it} < -z_{it}'\gamma] = -\sqrt{1+\sigma_u^2}\frac{\phi\left(\frac{z_{it}'\gamma}{\sqrt{1+\sigma_u^2}}\right)}{1-\Phi\left(\frac{z_{it}'\gamma}{\sqrt{1+\sigma_u^2}}\right)}, \tag{3.34}$$

where ϕ and Φ are the density function and the cumulative distribution function of the standard normal, respectively. From equations (3.29) and (3.30) follows that

$$\begin{aligned} E\left(\alpha_i \mid x_{it}, z_{it}, d_{it}=1\right) &= \frac{\sigma_{\alpha u}}{1+\sigma_u^2}\sqrt{1+\sigma_u^2}\frac{\phi\left(\frac{z_{it}'\gamma}{\sqrt{1+\sigma_u^2}}\right)}{\Phi\left(\frac{z_{it}'\gamma}{\sqrt{1+\sigma_u^2}}\right)} \\ &= \frac{\sigma_{\alpha u}}{\sqrt{1+\sigma_u^2}}\frac{\phi(\Theta_{it})}{\Phi(\Theta_{it})} = \delta\frac{\phi(\Theta_{it})}{\Phi(\Theta_{it})} = \delta\zeta_{it1}, \end{aligned} \tag{3.35}$$

and

$$E\left(\alpha_i \mid x_{it}, z_{it}, d_{it}=0\right) = -\frac{\sigma_{\alpha u}}{1+\sigma_u^2}\sqrt{1+\sigma_u^2}\frac{\phi\left(\frac{z_{it}'\gamma}{\sqrt{1+\sigma_u^2}}\right)}{1-\Phi\left(\frac{z_{it}'\gamma}{\sqrt{1+\sigma_u^2}}\right)}$$

$$= \quad -\frac{\sigma_{\alpha u}}{\sqrt{1+\sigma_u^2}} \frac{\phi(\Theta_{it})}{1-\Phi(\Theta_{it})} = -\delta \frac{\phi(\Theta_{it})}{1-\Phi(\Theta_{it})} = \delta \zeta_{it0}, \tag{3.36}$$

If γ and σ_u^2 were known, the variables ζ_{it0} and ζ_{it1} could be constructed. Then, from (3.35) we have

$$E[y_{it} \mid x_{it}, z_{it}, d_{it}] = x'_{it1}\beta_1 + x'_{it0}\beta_0 + \delta[\zeta_{it1} d_{it} + \zeta_{it0}(1-d_{it})]. \tag{3.37}$$

In practice γ and σ_u^2 are unknown, but can be consistently estimated by the panel Probit regression. Thus, the second step of the two-stage procedure is to estimate the pooled OLS regression

$$y_{it} = x'_{it1}\beta_1 + x'_{it0}\beta_0 + \delta[\widehat{\zeta}_{it1} d_{it} + \widehat{\zeta}_{it0}(1-d_{it})] + error, \tag{3.38}$$

where $\widehat{\zeta}_{it1}$ and $\widehat{\zeta}_{it0}$ are obtained using the panel Probit estimates of the first step. This gives us consistent estimates of β_1, β_0 and δ. From the estimation of δ we get also consistent estimates of the covariance term $\sigma_{\alpha u}$.

3.6 Gender wag gap

The estimates from the panel switching regression model are used to evaluate how the gender wage differential changes over the different career phases. To evaluate this, we use the decomposition of the wage gap suggested by Oaxaca (1973). We distinguish the wage gap for the two states of career, high level and low level. First we define the expected wage for males and females at different dates and for the two states, and then we construct the wage gap taking the difference between male and female expected wages.

For males the expected average wage in the high career level is defined as follows,

$$\begin{aligned} E[y_{it} \mid x_{it}, d_{it}=1] &= x_{it}'\beta_m + E[\alpha_i \mid x_{it}, z_{it}, d_{it}=1] \\ &= x_{it}'\beta_m + \delta_m \frac{\phi\left(\frac{z_{it}'\gamma_m}{\sqrt{1+\sigma_{um}^2}}\right)}{\Phi\left(\frac{z_{it}'\gamma_m}{\sqrt{1+\sigma_{um}^2}}\right)}, \end{aligned} \tag{3.39}$$

and for females as

$$\begin{aligned} E[y_{it} \mid x_{it}, d_{it}=1] &= x_{it}'\beta_f + E[\alpha_i \mid x_{it}, z_{it}, d_{it}=1] \\ &= x_{it}'\beta_f + \delta_f \frac{\phi\left(\frac{z_{it}'\gamma_f}{\sqrt{1+\sigma_{uf}^2}}\right)}{\Phi\left(\frac{z_{it}'\gamma_f}{\sqrt{1+\sigma_{uf}^2}}\right)}, \end{aligned} \tag{3.40}$$

where the indices m and f denote parameters for the subsample of males and females respectively. In the case of low career level, the equations are,

$$E[y_{it} \mid x_{it}, d_{it}=0] = x_{it}'\beta_m - \delta_m \frac{\phi\left(\frac{z_{it}'\gamma_m}{\sqrt{1+\sigma_{um}^2}}\right)}{1-\Phi\left(\frac{z_{it}'\gamma_m}{\sqrt{1+\sigma_{um}^2}}\right)}, \tag{3.41}$$

$$E[y_{it} \mid x_{it}, d_{it}=0] = x_{it}'\beta_f - \delta_f \frac{\phi\left(\frac{z_{it}'\gamma_f}{\sqrt{1+\sigma_{uf}^2}}\right)}{1-\Phi\left(\frac{z_{it}'\gamma_f}{\sqrt{1+\sigma_{uf}^2}}\right)}. \tag{3.42}$$

We define the (unexplained) wage gap for the two states of career $j = 0, 1$ at each period t as:

$$\Delta_{1,t} = \overline{X}'_{tf}\left(\widehat{\beta}_m - \widehat{\beta}_f\right) + \widehat{\delta}_m \frac{\phi\left(\frac{\overline{Z}'_{tf}\widehat{\gamma}_m}{\sqrt{1+\widehat{\sigma}^2_{um}}}\right)}{\Phi\left(\frac{\overline{Z}'_{tf}\widehat{\gamma}_m}{\sqrt{1+\widehat{\sigma}^2_{um}}}\right)} - \widehat{\delta}_f \frac{\phi\left(\frac{\overline{Z}'_{tf}\widehat{\gamma}_f}{\sqrt{1+\widehat{\sigma}^2_{uf}}}\right)}{\Phi\left(\frac{\overline{Z}'_{itf}\widehat{\gamma}_f}{\sqrt{1+\widehat{\sigma}^2_{uf}}}\right)}, \tag{3.43}$$

$$\Delta_{0,t} = \overline{X}'_{tf}\left(\widehat{\beta}_m - \widehat{\beta}_f\right) - \widehat{\delta}_m \frac{\phi\left(\frac{\overline{Z}'_{tf}\widehat{\gamma}_m}{\sqrt{1+\widehat{\sigma}^2_{um}}}\right)}{1 - \Phi\left(\frac{\overline{Z}'_{tf}\widehat{\gamma}_m}{\sqrt{1+\widehat{\sigma}^2_{um}}}\right)} + \widehat{\delta}_f \frac{\phi\left(\frac{\overline{Z}'_{tf}\widehat{\gamma}_f}{\sqrt{1+\widehat{\sigma}^2_{uf}}}\right)}{1 - \Phi\left(\frac{\overline{Z}'_{tf}\widehat{\gamma}_f}{\sqrt{1+\widehat{\sigma}^2_{uf}}}\right)}, \tag{3.44}$$

where $\overline{X}_{tf}$ and $\overline{Z}_{tf}$ are the vectors of sample means of explanatory variables at period t for females, $\widehat{\beta}_m$ and $\widehat{\beta}_f$ are estimated coefficients from the male and female panel switching regression, respectively.

Similarly $\widehat{\gamma}_m$, $\widehat{\sigma}^2_{um}$, $\widehat{\gamma}_f$, $\widehat{\sigma}^2_{uf}$ are the estimate from panel Probit model for males and females, respectively. Note that we have used females as reference group.

The difference $\Delta_{j,t}$ is divided into two parts. The first part, $\overline{X}'_{tf}\left(\widehat{\beta}_m - \widehat{\beta}_f\right)$, is the wage gap which cannot be explained by differences in productivity. The second part is due to the correction for the career selection.

3.7 Estimation results

In this section I present the estimation results.

In Table 3.5 are reported the results for the estimation of the panel model (3.1) with random effects, where the dependent variable is the natural log of monthly earnings, and the independent variables are marital status, level of education, the foreign dummy, work experience, occupational seniority, part-time dummy and small children dummy.

The first and second columns of the table report the estimates for males, the third and fourth columns for females.

The variable *married* is significant for both males and females, but while for women the impact is negative, for men it is positive. For women, being married reduces the monthly earnings by 20%, whereas earnings are increased for men by 5%. The dummy variables corresponding to the different levels of formation are all significant. As expected, the estimated coefficients are positive and increasing as the level of educations increases, for both males and females. In particular, for women having a high formation increases the monthly salary of 41%, instead for men this effect is less strong, about 30%. Let us consider the variables *experience* and *occupational seniority*. For women, they are both significant, with a quadratic specification implying a positive effect, and the *experience* has an higher impact (0.037) than the *occupational seniority* (0.015). For males, the *occupational seniority* is not significant, while the coefficient of *experience* (0.048) is larger than for women. As expected, the $part - time$ variable is significant and negative for both males and females, with a value of -0.34556 and -0.48019, respectively. The variable *children* is significant only for women and has also a negative impact, with a value of -0.306642. Being married, having children and working part-time are all characteristics that have a negative impact on the women's wage.

Table 3.6 shows the estimation results for a panel model with random effects, where we control for vertical segregation by inserting the variable *career* among the regressors. The introduction of the *career* dummy variable doesn't modify the estimation results. In fact we have no significant change with respect to Table 3.5. The *career* dummy variable is significant and positive for both males and females, with a value of 0.029269 and 0.058271, respectively.

The estimates presented in Table 3.6 likely suffer from an endogeneity problem, since the *career* variable is correlated with the individual specific effect (see discussion in section 3.5.2). To correct for endogeneity we consider the dummy variable *career* as an endogenous variable determined by a Probit selection equation, and adopt the two-stage estimation approach.

In Table 3.7 are presented the estimation results for the Probit model with random effect. The first and second columns report the estimates for males, the third and fourth columns for females. We have also computed the probability for an individual, with average characteristics, to be in the high level of the job hierarchy for the years 1999-2003. In Table 3.8 are reported the probabilities for males and females during the 5 years. We have found that for men it is more likely to be in the high career level than for women. In fact this probability is about 55% for the 5 years, instead for women is about 30%.

In Tables 3.9 and 3.10 we present the results for the marginal effects[7] and their standard errors[8]. In Table 3.9 the marginal effects are computed for both males and females at the average value of characteristics in the corresponding sample. To distinguish the impact of different coefficients and average characteristics for males and females, in Table 3.10 the marginal effects for females are computed at the male average, and vice versa.

Considering the estimated coefficients of dummy variables of formation (table 3.7), they are significant and increasing as the levels of formation increases. This implies that, the higher is the level of education, the larger is the probability to be in the high

[7]The marginal effects for the variables $formation_i$, $i = 2, \ldots, 4$, are computed considering that the variables are discrete and exclusive of each others. For example, consider the variable $formation2$. Its marginal effect is defined as follow:

$$\alpha = P[d = 1 \mid form2 = 1, form3 = 0, form4 = 0, \bar{Z}] - P[d = 1 \mid form2 = 0, form3 = 0, form4 = 0, \bar{Z}]$$

where $\bar{Z}$ denotes the average in the sample for all other characteristics excluding formation. This marginal effect is the difference in the probability of being in the high career level between one individual with apprenticeship formation and one with no scholar formation, all other variables being at the average value in the sample.

[8]Detailed explanations on the computation of the standard errors for the marginal effects are given in the Appendix.

level of the job hierarchy. To be more accurate in the evaluation of the impact of the $formation$ variables on the chance of being in the high career level, we can look at the marginal effects (Table 3.9). For example the probability to be in the high career level increases of about 67% if a men has a university degree. For women this probability is of 20%. Note that in general the estimated coefficients and the marginal effects of formation for women are significantly smaller than for men, meaning that the impact of education levels on the chances of promotion is less important for women.

Let us now consider the impact of occupational seniority. The estimated coefficients (Table 3.7) and the marginal effects (Table 3.8) are similar for males and females and imply a positive effect. Thus, occupational seniority seems to be important to get job promotions for males and females to a similar extent. At the contrary, *experience* is not significant for women.

We also find that the variable $foreign$ is not significant, for both men and women. The $part - time$ variable is negative and significant for both men and women, even if its marginal effect doubles for men (Table 3.9). The variable *children* is not significant for females, but significant for males. I expected a negative and significant effect for women, explaining the smaller chance for women to be promoted, but this is not the case. Probably this has to do with the fact that only working women are considered in the sample.

Considering now the marginal effects for women computed using the mean characteristics of males (Table 3.10), we note that the values are not very different from the ones computed with female mean, except for the variables $formation$. In particular this means that the difference in the marginal effects between females and males have to be imputed mostly to difference in the coefficients.

In Table 3.11 are reported the results for the panel switching regression. In the wage equation we introduce as independent variable, besides the explanatory factors considered above, a correction term. The variable *correction* is our δ in equations (3.35) and (3.36). This variable is significant for both males and females, with a negative and positive sign respectively[9].

Note that only the constant term has a coefficient which is different for the two states of career, high and low. Instead we constrained the model (3.10) to have the same coefficients in $j = 0$ and $j = 1$ for all regressors except the constant. We tried also to estimate the wage equation allowing for different values of all coefficients in $j = 0$ and $j = 1$, but the estimates do not appear different.

Considering the constant term, it is significant and positive for both high and low career levels, and for males and females. As expected, we can notice that for males the value of the constant term is higher in the case of high level of career than for the low level. This means that men in the high level of the job hierarchy will earn a higher wage compared men in the low level, all other characteristics being equal. For females the result is different, namely the constant term is higher in the case of low career. We can argue that in the case of low career level we have a higher "starting" wage than in the high career level. More precisely, the constant in the wage equation can be interpreted as the wage rate of an individual with the lowest level of education, zero experience and zero seniority (as well as zero for the other dummies). Thus, the data suggest that a woman with low qualifications will earn a lower wage if, for some reasons, she is employed in a position in high levels of the occupational hierarchy. Of corse, this will not imply that a women in the high career level has always a lower wage than in the case of low career level. In fact the individual characteristics (formation, experience, seniority) will affect the wage allowing for higher wages.

The variable *married* has a positive impact for men (about 5%) and a negative impact for women (about -12%). Instead, the variable $formation$ is increasing for the three levels of education, for both males and females. The estimated coefficient of the

---

[9]Note that the standard errors, displayed in the second and forth column of the table, are not correct. This is because they do not account for the fact that the correction variable has been estimated.

experience variable is smaller for males than for females. In particular, for women an additional year of experience increases the monthly earnings by 3.5%, while for men only by 2.5%. For the *occupational seniority* variable we have the same trend. For females, an additional year of seniority increases the wage by 4.5%. For men the value is by 4%. As we expected, the variable $part - time$ is negative for males and females. The variable *children* has a negative impact in the females case and a positive impact if the individual is a man, but in this case the variable is not significant.

Making a comparison between the panel model with random effects and career variable (Table 3.6) and the panel switching regression (Table 3.11) we can observe that the estimation results are rather different in size, but not in sign. For example, for the $formation$ dummy variables, for the three dummy, the impact is larger in the panel switching regression. For instance, for females having a compulsory school degree increases the wage by about 20% in the first estimation, and by about 50% in the switching regression. We can notice that the values doubled from the first to the second estimation. This is true for both males and female. Only for the variable *experience* the estimated coefficients are almost of the same magnitude in the two estimations and for both males and females.

In Table 3.12 are reported the wage gap in the case of high or low career level, for the period 1999-2003. We can notice that the wage gap is always bigger in the case of low job hierarchy than in the high level. The gap in the case of high career level is between 21% to 25%, in the case of low level between 32% to 35%. Note that this result is opposite than the finding illustrated in Figure 3.5, where the female-to-male wage ratio is lower in the high level of occupational hierarchy. The difference between the two results is explained by the fact that the wage gaps computed in Table 3.12 accounts for individual characteristics, that is, they are the "unexplained“ gaps. Thus, although in the high level of career the numerical difference between male and female wages is higher, individual characteristics explain part of this gap, and the unexplained wage gap is smaller than in the low level of career.

3.8 Conclusions

In this chapter, I have investigated the gender wage differential and the gender promotion chance using data of the Swiss Labour Force Survey. The data are collected by the Swiss federal statistical office, covering the period 1999-2003.

The wage equation and the chance to be in the high level of the job hierarchy are analysed separately for males and females, to determine the impact of individuals' characteristics on the monthly earnings and to investigate whether same characteristics provide similar chances of career.

Unequal career advancement is an important part of women's unequal treatment in the labour market. Theoretical models frequently attribute this phenomenon to females shorter horizon in market work, because of comparative advantages in outside opportunities. For example, Lazear and Rose (1990) suggest that to be promoted, a woman must be somewhat better than a man in order to compensate her higher ex-ante probability of departure and the social investment loss. The more likely she is to leave (relative to male), the larger must be the compensating ability differential. Our empirical results show, however, that only a part of the unequal gender distribution in job position is explained by discontinuous labour market experience. In fact the endowment characteristics, such as the schooling formation, have an important effect on the chance to be in the high level of the job hierarchy.

From the descriptive analysis it is seen that men have best chances of career. In fact, the percentage of women in the high level of the job hierarchy is relatively small (30%). Furthermore observing the evolution of the average wage, it can be noticed that the gap between males and females is very high. While men reach almost an average level of 8'000 SFr, in the case of high career level, women achieve only a level of 6'000 SFr.

From the estimation of the wage equation we have found that wages in both gender categories increase with formation, with stronger education effects for females. Also the *occupational seniority* variable plays a more important role in the case women. In fact, this variable double in female estimates. For the *experience*, the opposite holds; it is for men that this variable has a much stronger effect. The *married* and *children* variables play also an important role. In fact, for women being married and having children, has a significant and negative impact on the wage. While for men, only the *married* variable is significant and the impact is positive.

In the probit estimation, $formation$ and *occupational seniority* appear to be the most important single factors to explain the different chances in the career. An others variable that have also an important influence is $part-time$. In both gender cases the impact is negative and, especially for men, $part-time$ has a large negative effect on the chance to be in the high job hierarchy. What is also surprising is the non significance of the *children* and *married* variables in the case of females.

The fact that occupational seniority and working part-time have such strong impact on the promotion chance suggests that for women will be difficult to overcome the fact to be often confined to dead-end jobs.

Three substantive conclusions can be drawn from this analysis. First, the social characteristics, such as the marital status and the number of children, have a different impact on men and women. For women being married and having children is disadvantageous, instead for men is almost favourable.

Second, our research suggests that for women it is more important to have a high occupational seniority than to have a lot of years of experience, to increase the probability to be in high level of the career or to observe raise in the wage. Instead for men the opposite holds. This can be explained by several empirical works[10], which generally associate a raise in the wage, with a job change within firm, but not always with a promotion.

Finally, also the *formation* dummy variable plays an important role especially for females. A raise in the formation is always associated with an increase in the wage and in the probability to be in the high level of the job hierarchy. This shows how much is important, for women, to have a high formation level.

[10] For example, Carmichael (1983)

Variable Name	Definition
	Dependent Variables
Log Earnings	The natural log of monthly earnings
Career	1 if he/she is in the high level of occupational hierarchy, 0 otherwise
	Independent Variables
Married	1 if he/she is married, 0 otherwise
Formation 1	1 if the highest achieved level of study is compulsory schooling, 0 otherwise
Formation 2	1 if the highest achieved level of study is apprenticeship, 0 otherwise
Formation 3	1 if the highest achieved level of study is high school, 0 otherwise
Formation 4	1 if the highest achieved level of study is university degree, 0 otherwise
Foreign	1 if he/she is a foreigner, 0 otherwise
Experience	Number of years of prior work experience
Experience squared	Work experience squared
Occupational seniority	Number of years worked for the current employers
Occupational seniority squared	Occupational seniority squared
Part-time	1 if he/she workes less than 90%, 0 otherwise
Child $\leq$5 years	1 if he/she has children of 5 years old or less, 0 otherwise

Table 3.1: **Variables Definitions**

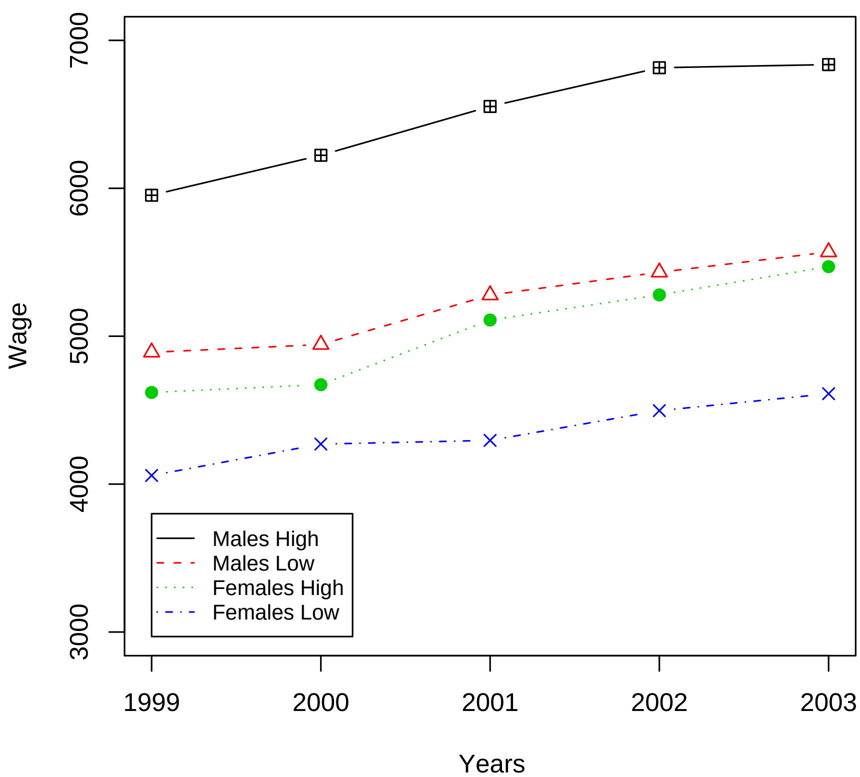

Figure 3.1: **Evolution of females and males average wages by career level**

The figure reports the evolution of the females and males average wages. The average wage is defined for the two levels of career: high position (directors) and low position (individuals in a non supervisory position).

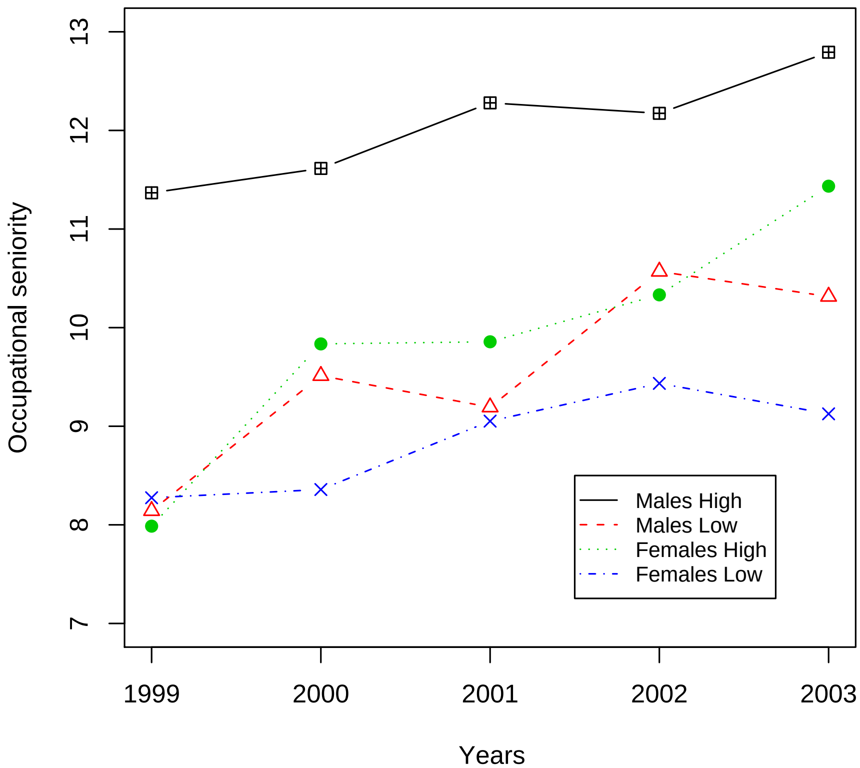

Figure 3.2: **Evolution of females and males average years of occupational seniority by career level**

The figure reports the evolution of females and males average years of occupational seniority. Note that the average wage is defined for two levels of career: high position (directors and individuals with supervisory position) and low position (individuals in a non supervisory position).

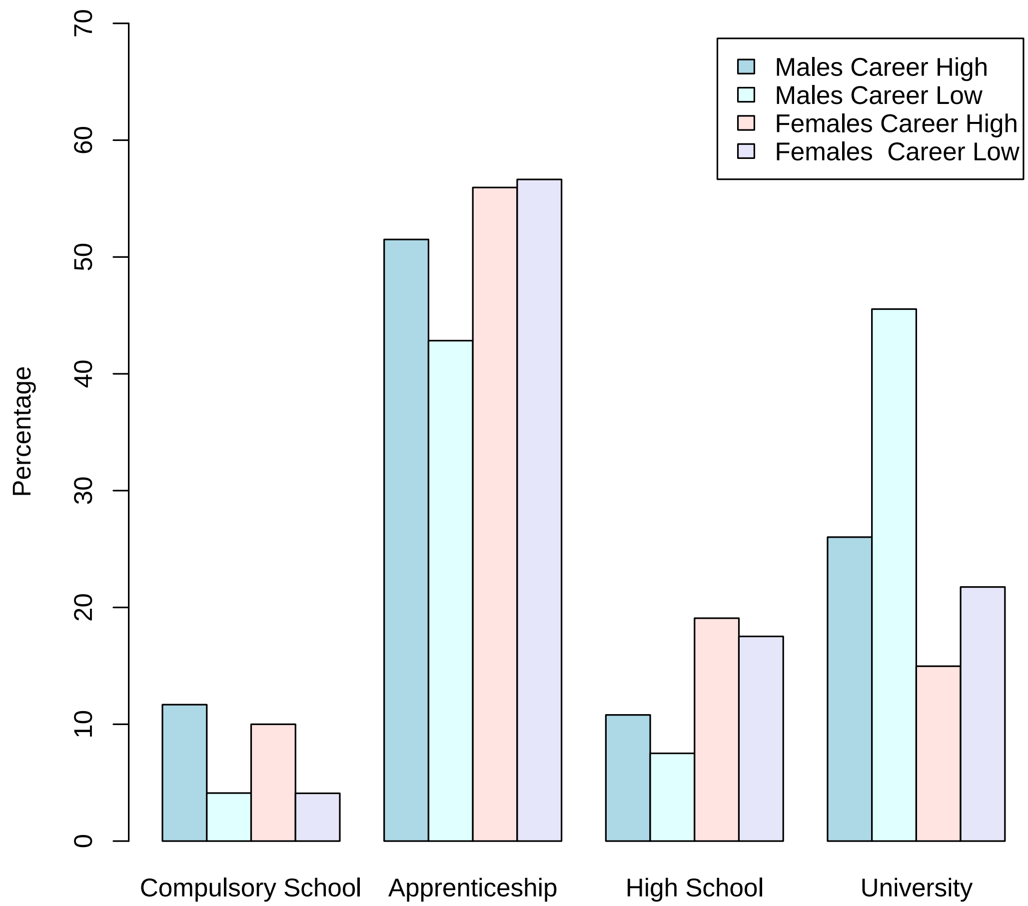

Figure 3.3: **Percentage of females and males in different formation level by career level**

The figure reports the percentage of females and males in different formation level by career level. The first two columns refer to males (high and low career level, respectively) and the second two to females.

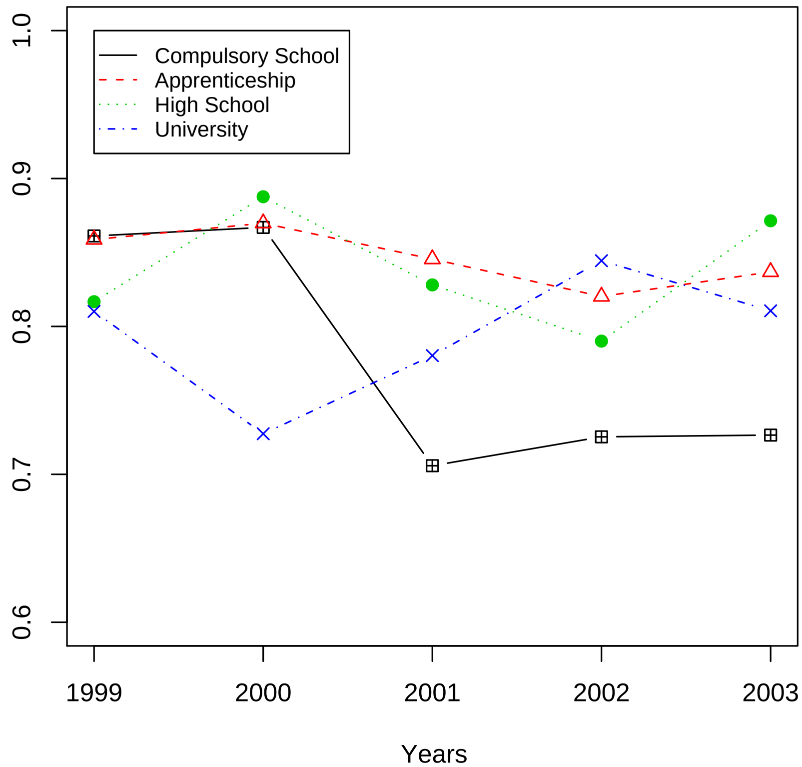

Figure 3.4: **Ratio of female wages to male wages, by educational level**

The figure reports the ratio of female wages to male wages by educational level in Switzerland for the years 1999-2003.

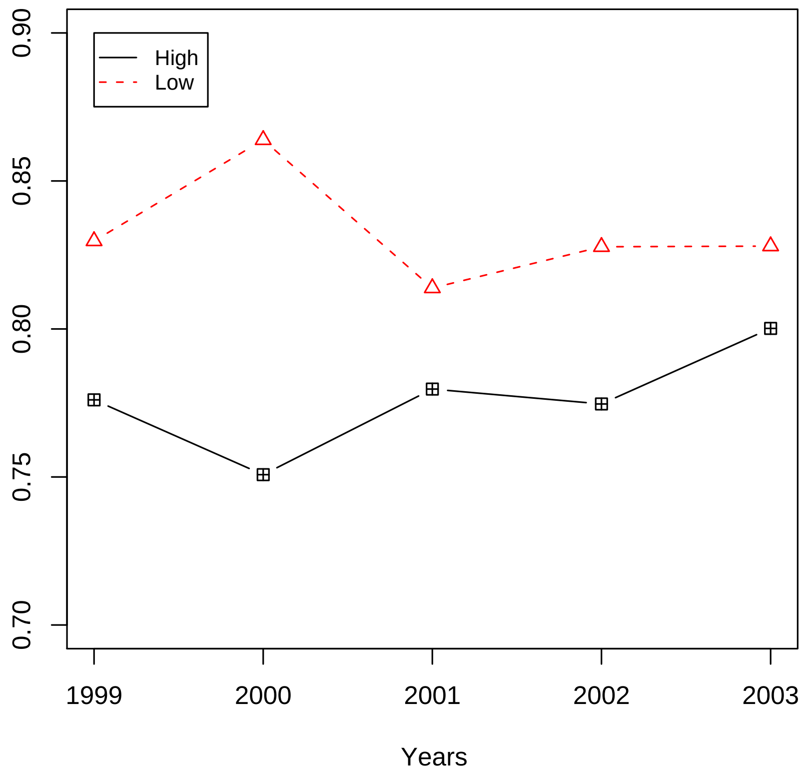

Figure 3.5: **Ratio of female wages to male wages, by career level**

The figure reports the ratio of female wages to male wages by career level in Switzerland for the years 1999-2003.

	Male		Female	
Variables	Mean	Std. Dev.	Mean	Std. Dev.
Wage Monthly	5775.892	2290.366	3408.642	2057.719
Married	0.6472532	-	0.464108	-
Formation1	0.0764147	-	0.081716	-
Formation2	0.4688145	-	0.561625	-
Formation3	0.0904585	-	0.186005	-
Formation4	0.3643123	-	0.170655	-
Foreign	0.1606774	-	0.085327	-
Experience	21.63982	10.10032	23.18104	10.88289
Experience squared	570.2561	497.0925	655.7445	540.2948
Occupational seniority	10.70441	9.561288	8.926122	8.5197
Occupational seniority squared	205.9649	334.1595	152.2282	275.7192
Part-time	0.0574143	-	0.564786	-
Children $<$ 5 years	0.2011565	-	0.117833	-
Number of observations	2420		2215	

Table 3.2: **Descriptive Statistics by gender**

Educational level	Men $n = 484$ (%)	Women $n = 443$ (%)	Both $n = 927$ (%)	Share of women (%)
Compulsory School	7.6	8.2	7.9	49.5
Apprenticeship	46.9	56.2	51.3	52.3
High School	9.0	18.6	13.6	65.3
University	36.4	17.1	27.2	30.0
Total	100.0	100.0	100.0	47.8

Table 3.3: **The share of employees in the sample, by educational level**

Career level	Men $n = 484$ (%)	Women $n = 443$ (%)	Both $n = 927$ (%)	Share of women (%)
High	53.3	30.9	42.6	34.7
Low	46.7	69.1	57.4	57.5
Total	100.0	100.0	100.0	47.8

Table 3.4: **The share of employees in the sample, by career level**

	Males		Females	
Variables	Est.	Std. Err.	Est.	Std. Err.
Constant term	7.786164**	0.0581161	7.631576**	0.106037
Married	0.0485276**	0.0197155	-0.220826**	0.036910
Formation2	0.1409613**	0.0332014	0.210335**	0.062473
Formation3	0.1581395**	0.0389132	0.279298**	0.068824
Formation4	0.3313344**	0.0356583	0.419415**	0.072042
Foreign	-0.1176634**	0.0333989	0.091177	0.073418
Experience	0.0486955**	0.0042652	0.037417**	0.007056
Experience squared	-0.0007616**	0.0000887	-0.000741**	0.000136
Occupational seniority	-0.0017758	0.0022332	0.015058**	0.004911
Occupational seniority squared	0.0000621	0.0000703	-0.000396**	0.000162
Part-time	-0.3455685**	0.0279431	-0.480195**	0.031508
Children < 5 years	0.0236556	0.0171646	-0.306642**	0.043174
σ_u	0.26775685		0.420079	
σ_w	0.15678049		0.330879	
ρ	0.74468539		0.617129	
N	2206		2029	

* Values significant at 5%, ** values significant at 1%

Table 3.5: **Panel model with random effects**

The dependent variable is the natural log of monthly earnings. The estimation is made for the two samples of males and females separately.

Variables	Males Est.	Males Std. Err.	Females Est.	Females Std. Err.
Constant term	7.788972**	0.057671	7.620852**	0.105273
Career	0.029269**	0.012263	0.058271**	0.024352
Married	0.046012**	0.019675	-0.222661**	0.036712
Formation2	0.137321**	0.033182	0.203025**	0.062274
Formation3	0.154661**	0.038922	0.272788**	0.068614
Formation4	0.326943**	0.035688	0.410123**	0.071903
Foreign	-0.116262**	0.032931	0.092863	0.072608
Experience	0.047773**	0.004244	0.037728**	0.007006
Experience squared	-0.000751**	0.000081	-0.00075**	0.000135
Occupational seniority	-0.001962	0.002234	0.014443**	0.004905
Occupational seniority squared	0.000065	0.000071	-0.000371**	0.000161
Part-time	-0.345162**	0.027981	-0.477832**	0.031548
Children < 5 years	0.026131	0.017212	-0.30463**	0.043145
σ_u	0.260422		0.411194	
σ_w	0.156825		0.330976	
ρ	0.733871		0.606837	
N	2206			2029

* Values significant at 5%, ** values significant at 1%

Table 3.6: **Panel model with random effects and career variable**

The dependent variable is the natural log of monthly earnings. We control for vertical segregation by introducing the variable *career* among the regressors.

Variables	Males Est.	Males Std. Err.	Females Est.	Females Std. Err.
Constant term	-3.12058**	0.526108	-1.30477**	0.428991
Married	0.606469**	0.185818	0.125583	0.147113
Formation2	1.110444**	0.32075	0.75939**	0.271404
Formation3	1.373995**	0.38618	0.867423**	0.30141
Formation4	1.972684**	0.3441	1.181876**	0.307426
Foreign	0.034704	0.260838	-0.34047	0.283771
Experience	0.122121**	0.037594	-0.02025	0.029147
Experience squared	-0.00244**	0.000778	0.000263	0.000585
Occupational seniority	0.056039**	0.021073	0.059265**	0.021282
Occupational seniority squared	-0.00069	0.000659	-0.00195**	0.0007
Part-time	-0.9258**	0.292536	-0.55057**	0.133155
Children $<$ 5 years	-0.37655*	0.172065	-0.2545	0.19813
σ_u	1.878373	0.103947	1.426627	0.09343
ρ	0.779166	0.019044	0.670539	0.028936
N	2421		2215	

* Values significant at 5%, ** values significant at 1%

Table 3.7: **Probit Model with Random Effects**

The dependent variable in the probit model takes the value of 1 if the individual i in year t is in the high level of job hierarchy, 0 otherwise.

Probability	1999	2000	2001	2002	2003
Males	0.5222	0.5315	0.5449	0.5497	0.5606
Females	0.3052	0.3049	0.3053	0.3057	0.3302

Table 3.8: **Probability of an individual to be in a high level of the job hierarchy during the period 1999-2003**

Variables	Males dy/dx	Males Std. Err	Females dy/dx	Females Std. Err
Married	0.2362223	0.0711000	0.0338108	0.0398100
Formation2	0.3610290	0.0208315	0.0513365	0.0101136
Formation3	0.4653261	0.0216109	0.0809212	0.0112750
Formation4	0.6703305	0.0221590	0.1818415	0.0079938
Foreign	0.0134670	0.1009300	-0.0797412	0.0573800
Experience	0.0475192	0.0146500	-0.0054296	0.0078300
Experience squared	-0.0009478	0.0003000	0.0000706	0.0001600
Occupational seniority	0.0218056	0.0082100	0.0158945	0.0057700
Occupational seniority squared	-0.0002689	0.0002600	-0.0005224	0.0001900
Part-time	-0.3508453	0.0956900	-0.1518085	0.0386200
Children $<$ 5 years	-0.1484352	0.0678700	-0.0622877	0.0440200

Table 3.9: **Marginal effects of the probit model for males and females**

Variables	Males with mean of females dy/dx	Males with mean of females Std. Err	Females with mean of males dy/dx	Females with mean of males Std. Err
Married	0.2069610	0.0129427	0.0430493	0.0446069
Formation2	0.2152727	0.0609769	0.2178028	0.0878911
Formation3	0.3063246	0.0650712	0.2537385	0.0921502
Formation4	0.5404756	0.0687397	0.3695426	0.082297
Foreign	0.0118430	0.0038131	-0.1167116	0.0956717
Experience	0.0416745	0.0021781	-0.0069400	0.0102073
Experience squared	-0.0008312	0.0000447	0.0000902	0.0001918
Occupational seniority	0.0191236	0.0012025	0.0203160	0.0083427
Occupational seniority squared	-0.0002358	0.0000175	-0.0006678	0.0002745
Part-time	-0.3159349	0.0209401	-0.1887331	0.0606227
Children $<$ 5 years	-0.1284995	0.0083486	-0.0872408	0.0676057

Table 3.10: **Marginal effects of the probit model for males and females with different means**

Variables	Males Est.	Males Std. Err.	Females Est.	Females Std. Err.
Constant for high livel	7.972446**	0.044585	6.090929**	0.535943
Constant for low level	7.85675**	0.044135	8.176573**	0.195675
Married	0.0468832**	0.016101	-0.12856**	0.032813
Formation2	0.1691681**	0.02734	0.4954**	0.101536
Formation3	0.3198541**	0.03329	0.698473**	0.116215
Formation4	0.4827371**	0.028212	1.092858**	0.161025
Foreign	-0.0917396**	0.019586	-0.1183	0.066375
Experience	0.0285255**	0.00316	0.036932**	0.005727
Experience squared	-0.0004348**	0.000064	-0.00097**	0.000106
Occupational seniority	0.0039634**	0.002125	0.045075**	0.008943
Occupational seniority squared	-0.0000971	0.000063	-0.00109**	0.000286
Part-time	-0.5409551	0.028605**	-0.9345**	0.083445
Children $<$ 5 years	0.0283529	0.018799	-0.47257**	0.054617
Correction	-0.0073226**	0.003624	1.358621**	0.430972
N	2206		2029	

* Values significant at 5%, ** values significant at 1%

Table 3.11: **Panel switching regression for vertical segregation**

The dependent variable is the natural log of monthly earnings.We have two different constants, the first for the high level of job hierarchy, the second for the low level.

GAP	1999	2000	2001	2002	2003
High	0.2177	0.2378	0.2423	0.2218	0.2574
Low	0.3247	0.3159	0.3371	0.3546	0.3590

Table 3.12: **Wage gap for the two career levels**

Chapter 4

Conclusions

During the last years, occupational segregation has re-gained popularity in the analysis of gender relations at work. This interest has developed two directions. First, interest in measuring the level of occupational segregation by gender over time, using various indices of segregation. And second, the focus on the causes and the consequences of the occupational segregation by gender. From this starting point, this thesis has addressed several issues related to the development of models to detect and measure occupational segregation and gender wage gap.

In the main part of the thesis a major focus was given on the occupational segregation. More precisely, in Chapter 2have considered the horizontal occupational segregation, which occurs when there is a concentration of women or men in a given field or occupation. By following the approach introduced by Hirsch (1995) and Hansen and Wahlberg (2000), a new explanatory variable FEM^1 was included in the wage equation, which allow to control for occupational segregation. However, this new variable can cause a problem of endogeneity. In fact, it is difficult to ensure that there is no correlation between the density of females in an occupation and the error term of the

[1]This variable measures the percentage of females in an occupational specification.

wage equation. To overcome this problem, an ordered probit model was considered, in which the dependent variable admits three values (0, for male dominated occupations, 1 for intermediate occupations and 2 for female dominated occupations). Separate wage regressions are specified for each value of the discrete variable. I have estimated this model with a version of Heckman two step approach. Two main results can be drawn from this analysis. First, in accordance with previous studies, I have found that a high concentration of female workers in an occupation in Switzerland has a negative effect on male and female estimated wages. This confirms the hypothesis of wage penalty for working in female dominated occupations. Second, I have found that the smallest gender wage gap occurs in male dominated occupations.

In chapter 3, I have considered the vertical occupational segregation, that occurs when there is a concentration of women and men in a given level of responsibility. To study this problematic, we have proposed a general linear model where we have introduced a dummy variable for being in the high level of the job hierarchy, or in the low level. The estimation methodology was a version of Heckman two step procedure. The main results from the analysis of this model are the following. First, the level of formation has a relevant impact on the vertical segregation, especially for women, where a higher formation increases the probability to be in the high level of the job hierarchy. Second, instead, for women being married or having children decreases the probability to be in the high level of the job hierarchy. For men, the opposite holds. Finally, the seniority plays also an important role (more than the experience), but this only in the female case.

The limitations of these models concern primarily the problem of sample selection and of endogeneity. Regarding sample selection, in both chapters 2 and 3, we don't consider the individual choice to enter or not in the job market, or the choice between being a self-employee or an employee. In Chapter 2, we have a limitation on the data. The survey gives information only on individuals who are working. We don't have any information on individuals who decide not to enter in the job market. At the contrary, in principle Chapter 3 allows for a more sophisticated model. The survey, in fact, contains information on individuals who are working and not, and who are self-employee and not.

The decision to disregard these other layers of individual choice is motivated by the desire to focus on the simplest specification to highlight the main issues.

Concerning the endogeneity, besides the problem considered for the two variables FEM and $career$, we have neglected the endogeneity of some others variable such as experience. In fact, in our thesis, this variable is not the real experience of an individual (namely the real number of years spent in the job market), but it has been constructed from other variables. Especially for women, this is a limitation on the results. In fact, frequently for females this variable was not significant. To overcome this limitation, an instrumental variable approach could be used.

The extension of the methods presented in this thesis to carefully account for the endogeneity and sample selection issues is left for future research. Another interesting direction of research is to apply the model of Chapter 3 in a dynamic setting, to study the different chances of career promotion of women and men.

Appendix A

Computation of the Standard Error

In our model we consider 11 exogenous variables, where 8 are continuous and 3 are discrete variables (namely $formation2$, $formation3$ and $formation4$).

The vector of the exogenous variables is defined as follow (the individual and the period indices are suppressed here and in the following to simplify the exposition):

$$\begin{bmatrix} X' & z_1 & z_2 & z_3 \end{bmatrix} = \begin{bmatrix} X' & Z' \end{bmatrix} \tag{A.1}$$

where z_1, z_2 and z_3 are the three discrete variables and $X' = \begin{pmatrix} x_1 & \dots & x_8 \end{pmatrix}$ the vector of continuous variables. The corresponding regression coefficients are denoted

$$\theta' = \begin{bmatrix} \beta' & \delta_1 & \delta_2 & \delta_3 \end{bmatrix} = \begin{bmatrix} \beta' & \delta' \end{bmatrix} \tag{A.2}$$

where δ_1, δ_2 and δ_3 are the parameters for the discrete variables and β is the 8×1 vector of coefficients for the continuous variables.

The marginal effects for the continuous variables x_j are given by

$$\begin{aligned}\gamma_j(\theta) &= \frac{\partial[P(d=1 \mid X, Z, u=0)]}{\partial x_j} \\ &= \frac{\partial\Phi(X'\beta+\delta_1 z_1+\delta_2 z_2+\delta_3 z_3)}{\partial x_j} = \phi(\zeta)\beta_j, \quad j=1,\ldots,8 \end{aligned} \tag{A.3}$$

where u is the individual specific effect and $\zeta\!:= X'\beta+\delta_1 z_1+\delta_2 z_2+\delta_3 z_3$. Considering that the marginal effects are evaluated at the means of the variables, we can rewrite the expression (3.12) as,

$$\gamma_j(\theta) = \phi(\bar{X}'\beta+\delta\bar{z}_1+\delta\bar{z}_2+\delta\bar{z}_3)\beta_j = \phi(\bar{\zeta})\beta_j. \tag{A.4}$$

The marginal effects for the discrete variables z_1, z_2, z_3 are

$$\alpha(\theta) = \begin{bmatrix} \Phi(\bar{X}'\beta+\delta_1)-\Phi(\bar{X}'\beta) \\ \Phi(\bar{X}'\beta+\delta_2)-\Phi(\bar{X}'\beta) \\ \Phi(\bar{X}'\beta+\delta_3)-\Phi(\bar{X}'\beta) \end{bmatrix} = \begin{bmatrix} \alpha_1(\theta) \\ \alpha_2(\theta) \\ \alpha_3(\theta) \end{bmatrix}. \tag{A.5}$$

Stacking the marginal effects of the continuous and discrete variable we get

$$g\left(\theta\right) = \begin{bmatrix} \gamma\left(\theta\right) \\ \alpha\left(\theta\right) \end{bmatrix}. \tag{A.6}$$

The associated covariance matrix of the estimated marginal effects can be computed using the Delta-Method. The basic idea is that the estimated marginal effects $g\left(\hat{\theta}\right)$ are a non-linear transformations of the vector $\hat{\theta}$ of estimated parameters. Since the (asymptotic) variance-covariance matrix of $\hat{\theta}$ is known, we can obtain the (asymptotic) variance-covariance matrix of $g\left(\hat{\theta}\right)$ by a first-order Taylor series approximation of function $g\left(\theta\right)$.

We get

$$AsVar(\hat{g}) = GVG', \tag{A.7}$$

where $V = AsVar(\hat{\theta})$ and $G = \frac{\partial g}{\partial \theta'}$. The matrix of derivatives is defined as follow,

$$G = \begin{bmatrix} \frac{\partial \gamma}{\partial \beta'} & \frac{\partial \gamma}{\partial \delta'} \\ \frac{\partial \alpha}{\partial \beta'} & \frac{\partial \alpha}{\partial \delta'} \end{bmatrix},$$

where,

$$\frac{\partial \gamma}{\partial \beta'} = \phi(\overline{\zeta})I - \phi(\overline{\zeta})\overline{\zeta}\widehat{\beta \overline{X}}' \tag{A.8}$$

$$\frac{\partial \alpha}{\partial \delta'} = \begin{bmatrix} \frac{\partial \alpha_1}{\partial \delta_1} & 0 & 0 \\ 0 & \frac{\partial \alpha_2}{\partial \delta_2} & 0 \\ 0 & 0 & \frac{\partial \alpha_3}{\partial \delta_3} \end{bmatrix}, \quad \frac{\partial \alpha}{\partial \delta_i} = \phi(\overline{X}'\beta + \delta_i), \quad i = 1, 2, 3 \tag{A.9}$$

$$\frac{\partial \alpha}{\partial \beta'} = \begin{bmatrix} \frac{\partial \alpha_1}{\partial \beta'} \\ \frac{\partial \alpha_2}{\partial \beta'} \\ \frac{\partial \alpha_2}{\partial \beta'} \end{bmatrix}, \quad \frac{\partial \alpha_i}{\partial \beta'} = \left[\phi\left(\overline{X}'\beta + \delta_i\right) - \phi\left(\overline{X}'\beta\right)\right]\overline{X}', \quad i = 1, 2, 3 \tag{A.10}$$

$$\frac{\partial \gamma}{\partial \delta'} = \begin{bmatrix} \frac{\partial \gamma}{\partial \delta_1} & \frac{\partial \gamma}{\partial \delta_2} & \frac{\partial \gamma}{\partial \delta_3} \end{bmatrix}, \quad \frac{\partial \gamma}{\partial \delta_i} = -\phi(\overline{\zeta})\overline{\zeta}\overline{z}_i\beta, \quad i = 1, 2, 3. \tag{A.11}$$

Bibliography

[1] Acker,J., 1990. *Hierarchies, job and bodies: a theory of gendered organizations.* Gender and Society, Vol. 4, No. 2, 139-158.

[2] Anker, R., 1997. *Theories of occupational segregation by sex: an overview.* International Labour Review, Vol. 136.

[3] Balestra, P. and Nerlove, M., 1966. *Pooling cross section and time series data in the estimation of a dynamic model: The demand for natural gas.* Econometrica, Vol. 34, No. 3, 585-612.

[4] Baker, M. and Fortin, N., 2000. *Occupational and gender composition and wages in Canada: 1987-1988.* CIRANO, 2000s-48.

[5] Bayard, L., Hellerstein, J., Neumark, D. and Troske, K., 1999. *New evidence on sex segregation and sex differences in wages from matched employee-employer data.* Journal of Labor Economics, Vol. 21, 887-922.

[6] Banfi S., Farsi M., Filippini M., 2006. *An empirical analysis of child care demand in Switzerland.* Quaderno 06-03, Facolt di scienze economiche, Universit della Svizzera Italiana.

[7] Baron, J. and Newman A., 1990. *For what it's worth: organizations, occupations and the value of work done by women and nonwithes.* American Sociological Review, Vol. 55, 155-175.

[8] Becker, G., 1971.*The economics of discrimination. II edition.* Chicago, University of Chicago Press.

[9] Berrebi, Z. and Silber, J., 1987. *Dispersion, asymmetry and the Gini index of inequality.* International Economic Review, Vol. 28, No. 2, 331-338.

[10] Blau, F., 1998. *Trends in the well-being of american women, 1970-1995.* Journal of Economic Literature, Vol. 36, No. 1, 112-165.

[11] Blinder, A., 1973. *Wage discrimination: reduced form and structural estimates.* The Journal of Human Resources, Vol. 8, No. 4, 436-455.

[12] Bonjour, D., 1997. *Lohndiskrimierung in der Schweiz: Eine ökonometrische Untersuchung.* Berner Beiträge zur Nationalökonomie, Bd. 83, Bern - Stuttgart - Wien: Paul Haupt.

[13] Brüderl, J., Dieckmann, A. and Engelhardt, H. 1993. *Einkommensunterchieds zwischen Frauen and Männer in der Schweiz.* Schweiterische Zietschrift fü Soziologie, Vol. 19, 573-588.

[14] Cattaneo, A. and Winkelmann R., 2005. *Earnings differentials between german and french speakers in Switzerland.* Schweizerische Zeitschrift fur Volkswirtschaft und Statistik.

[15] Charles, M., 1992. *Cross-national variation in occupational sex segregation.* American Sociological Review, Vol. 57, 483-502

[16] Charles, M. and Buchmann M., 1994. *Assessing micro-level explanations of occupational sex segregation: human-capital development and labor market opportunities in Switzerland.* Schwiez Z. Soziol. /Rev. Suisse sociol., Vol. 20, 595-620.

[17] Conway, D. A., 1994. *Analysis of employment discrimination through homogeneous job groups.* Journal of Econometrics, Vol. 61, No. 1, 101-131.

[18] Cotton, J., 1988. *On the decomposition of wage differentials.* The Review of Economics and Statistics, Vol. 70, No. 2, 236-43.

[19] Deutsh, J., Flückiger, Y. and Silber, J., 1994. *Measuring occupational segregation: summary statistics and the impact of classification errors and aggregation.* Journal of Econometrics, Vol. 61, 133-146.

[20] Dolton, P. and Makepeace, G., 1986. *Sample selection and male-female earnings differentials in the graduate labour market.* Oxford Economic Papers, Vol. 38, No. 2, 317-341.

[21] Duncan, O. and Duncan, B., 1955. *A methodological analysis of segregation indexes.* American Sociological Review, Vol. 20, No. 2, 210-217.

[22] England, P. and others, 1986. *Explaining occupational sex segregation and wages: finding from a model with fixed effects.* American Sociological Review, Vol. 53, 544-558.

[23] Flückiger, Y. and Ramirez, J. 2000. *Analyse comparative des salaires entre les hommes et les femmes sur la base de la LE 1994 et 1996.* Observatoire Universitaire de l'Emploi, University of Geneva.

[24] Green,William H., 2000. *Econometric Analysis, IV edition.* Prentice Hall.

[25] Greenberger, E. and Steinberg, L., 1983. *Sex differences in early labour force experience: harbinger of things to come.* The University of North California Press.

[26] Gong, X. and van Soest, A., 2002. *Wage differentials and mobility in the urban labour market: a panel data analysis for Mexico.* Labour Economics, Vol. 9, 513-529.

[27] Hansen, J. and Wahlberg, R., 2000. *Occupational gender composition and wages in Sweden.* IZA and CERP, Discussion Paper No.217.

[28] Heckman, J., (1976). *The common structure of statistical models of truncation, sample selection and limited dependent variables and a simpler estimator for such models.* Annals of Economic and Social Measurement, Vol. 5, 475-492.

[29] Hirsch, B. and Macpherson, D., 1994. *Wages, racial composition, and quality sorting in labor markets.* Institute for Research on Poverty. Discussion Paper No. 1038-94.

[30] Hirsch, B. and Macpherson, D., 1995. *Wages and gender composition: Why do women's job pay less?* Journal of Labor Economics, Vol. 13, No. 3, 426-471.

[31] Hirsch, B. and Schumacher, E., 1992. *Labor earnings, discrimination, and the racial composition of jobs.*Journal of Human Resources, Vol. 27, 602-28.

[32] Hirschberg, J. and Slottje, D., 1994. *An empirical Bayes approach to analyzing earning functions for various occupations and industries.* Journal of Econometrics, Vol. 61, No. 1, 65-79.

[33] Hirschberg, J., Slottje, D., and Hayes, K., 1994. *A new method for detecting individual and group labor market discrimination.* Journal of Econometrics, Vol. 61, No. 1, 43-64.

[34] Hutchens, R., 1991. *Segregation curves, Lorenz curves, and inequality in the distribution of people across occupations.* Mathematical Social Sciences, Vol. 21, No. 1, 31-51.

[35] Jenkins, S., 1994. *Earnings discrimination measurement. A distributional approach.* Journal of Econometrics, Vol. 61, No. 1, 81-102.

[36] Kanter, R., 1977. *Men and women of the corporation.* Basic Books, New York.

[37] Killingsworth, M., 1990. *The economics of comparable worth.* Institute for Employment Research. Working paper.

[38] Kugler, P., 1988. *Lhondiskriminierung in der Schweiz. Evidenz von Mikrodaten.* Schwiezerische Zeitschrift für Volkswirtschaft und Statistik, Vol. 124, 23-47.

[39] Lazear, E. and Rosen, S., 1990. *Male-female wage differentials in job ladders.* Journal of Labor Economics, Vol. 8, No. 1, 106-123.

[40] Le Grand, C., 1991. *Explaining the male-female wage gap: Job segregation and solidarity wage bargaining in Sweden.* Acta Sociologica, Vol. 34, No. 4.

[41] Maddala, G., 1999. *Limited-dependent and qualitative variables in econometrics.* Econometric Society Monographs.

[42] Matyas, L. and Sevestre, P., 1996. *The econometrics of panel data. A handbook of the theory with applications. II edition.* Kluwer accademic publishers.

[43] Mincer, J. and Polachek, S., 1978. *Women's earnings reexamined.* The Journal of Human Resources, Vol. 13, No. 1, 118-134.

[44] Miller, P., 1987. *The wage effect of the occupational segregation of women in Britain.* The Economic Journal, Vol. 97, No. 388, 885-896.

[45] Moulton, B., 1990. *An illustration of a pitfall in estimating the effects of aggregate variables on micro units.* The Journal of Human Resources, Vol. 71, No. 2, 334-338.

[46] Neumark, D., 1988. *Employers' discriminatory behaviour and the estimation of wage discrimination.* Review of Economics and Statistics, Vol. 23, No. 3, 279-295.

[47] Oaxaca, R., 1973. *Male-female wage differentials in urban labor markets.* International Economic Review, Vol. 14, No. 3, 693-709.

[48] Oaxaca, R. and Ransom, M., 1994. *On discrimination and the decomposition of wage differentials.* Journal of Econometrics, Vol. 61, 5-21.

[49] Polachek, S. and Kim, M., 1994. *Panel estimates of the gender earnings gap.* Journal of Econometrics, Vol. 61, No. 1, 23-42.

[50] Polachek, S. and Kamalich, R.F., 1985. *On estimating discrimination in the economy.* Southern Economic Journal, Vol. 51 , No. 4 , 1227-1229.

[51] Polachek, S. and Kamalich, R.F., 1982. *Discrimination: Fact or fiction? An examination using an alternative approach.* Southern Economic Journal, Vol. 49, No. 2, 450-461.

[52] Polachek, S., 1975. *Differences in expected post-school investment as a determinant of market wage differentials.* International Economic Review, Vol. 16, No. 2, 451-470.

[53] Reskin, B., 1988. *Bringing the men back in: Sex differentiation and the devaluation of women's work.* Gender and Society, Vol. 2, 58-81.

[54] Reskin, B., 1984. *Sex segregation in the workplace. Trends, explanation, remidies.* National Academy Press, Washington, DC, National Accademy Press.

[55] Ridder, G., 1990. *Attrition in multi-wave panel data.* In "Panel data and labor market studies", ed. J. Hartog, G. Ridder, and J. Theeuwes. Elsevier: North-Holland.

[56] Robinson, D., 1998. *Differences in occupational earnings by sex.* International Labour Review, 1998

[57] Rosholm, M. and Smith, N.,1996. *The danish gender wage gap in the 1980s: a panel study.*Oxford Economic Papers, Vol. 48, No. 2, 254-279.

[58] Silber, J., 1989. *Factor components, population subgroups and the computation of the Gini index of inequality.* The Review of Economics Statistics, Vol. 7, No. 1, 107-115.

[59] Sorensen, E., 1989. *Measuring the pay disparity between typically female ocupations and other jobs: A bivariate selectivity approach.* Industrial and Labor Relations Review, 57-79.

[60] Sorensen, E., 1990. *The crowding hypothesis and comparable worth issue: A survey and new results.* Journal of Human Resources, Vol. 25, 55-89.

[61] Sousa-Poza, A. 2003. *The gender wage gap and occupational segregation in Switzerland, 1991-2001.* Discussion Paper, No. 89.

[62] Vella, F. and Verbeek, M., 1999. *Two-step estimation of panel data models with censored endogemous variables and selection bias.* Journal of Econometrics, Vol. 90, 239-263.

[63] Vella, F., 1998. *Estimating models with sample selection bias: A surevey.* The Journal of Human Resources, Vol. 33, No. 1, 127-169.

[64] Williams, G., 1979. *The changing U.S. labor force and occupational differentiation by sex.* Demography, Vol. 16, No. 1, 73-87.

[65] Zellner, H., 1972. *Discrimination against women, occupational segregation, and the relative wage.* The American Economic Review, Vol. 62, No. 1/2, 157-160.