Università
Ca'Foscari
Venezia

Corso di Laurea magistrale
in Economia - Economics

Tesi di Laurea

Ca' Foscari
Dorsoduro 3246
30123 Venezia

The determination of wages: from the Mincer equation to the scarring effects

Relatore
Prof.ssa Agar Brugiavini

Correlatori
Prof. Giacomo Pasini
Prof.ssa Anna Marenzi

Laureando
Marco Spinazzè
Matricola 806113

Anno Accademico
2011 / 2012

Acknowledgments

I am grateful to my supervisor Prof.ssa Agar Brugiavini whose guidance and suggestions allowed me to complete this thesis.
A very special thanks goes out to Dott. Yuri Pettinicchi for his support, advise and encouragement for the whole duration of the realization of this work.

The determination of wages: from the Mincer equation to the scarring effects

Abstract

The empirical estimation of the Mincer equation is a useful tool for the determination of wages in the labor market. As a matter of fact it allows to relate the level of education and work experience of an individual with the income that he perceives. This is part of the human capital theory which allows to evaluate what is the return on investment in education and training of an individual. Successively it is analyzed, for a given endowment of human capital, the extent to which an unemployment spell or a period of recession at school leaving age affects the ability of an individual to generate earnings. If the consequent ability of an individual to generate earnings provides negative (permanent or at least persistent) effects these are called scarring effects. The empirical application proposed to estimate the Mincer equation and the scarring effects is based on OLS models with cross section data obtained through the use of SHARE database for France, Germany, Italy and Spain.

Keywords: Human capital theory, Mincer equation, Scaring effects, Share

This thesis uses data from SHARELIFE release 1, as of November 24th 2010 and SHARE release 2.5.0, as of May 24th 2011. The SHARE data collection has been primarily funded by the European Commission through the 5th framework programme (project QLK6-CT-2001- 00360 in the thematic programme Quality of Life), through the 6th framework programme (projects SHARE-I3, RII-CT- 2006-062193, COMPARE, CIT5-CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th framework programme (SHARE-PREP, 211909 and SHARE-LEAP, 227822). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, Y1-AG-4553-01 and OGHA 04-064, IAG BSR06-11, R21 AG025169) as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions).

Contents

List of figures

List of tables

1. Introduction

The European Union indicates human capital as a strategic resource for the development of Europe and it establishes that especially in times of crisis the member States' education and training policies should be oriented towards enhancement of the personality of each individual throughout the whole course of his life. The reason why human capital has such an important response is due to the fact that this discipline assumes importance at both microeconomic and macroeconomic levels. From a macroeconomic perspective the development of human capital determines the level of the overall economic growth of a country by providing competitive advantages for international competition.

From a microeconomic point of view the human capital of an individual is considered one of the main determinants of success in the workplace and investment in education is relevant to the greater opportunities it offers individuals for access to and permanence in the labour market and for the higher earnings arising from it.

The approach used for the purpose of this thesis is microeconomic. The reason is that the major contributions of the human capital theory are related to the microeconomic approach. The most important approaches of the human capital theory developed in the sixties by Schultz (1961), Mincer (1958 and 1970) and Becker (1962 and 1964) predict that the level (or stock) and development of human capital determine a different level of earnings for individuals. Moreover the data used in this thesis are individual level micro data. This is fundamental because it allows to analyze the impact of the individual's choices.

The purpose of this thesis is to assess to what extent different endowments of individuals' human capital affect the determination of their remuneration. The aim of this thesis is also to verify to what extent a negative economic factor such as a recession or a period of unemployment has affected the career of a worker also in relation to his human capital endowment. The objective is to answer the following question: Does a recession or unemployment spell influence the capability of an individual to produce earnings? If these negative economic factors provide negative effects (permanent or at least persistent) on individual earnings, these are called scarring effects.

The first part of this thesis proposes a theoretical framework of the main implications of the human capital theory and scarring effects.
In particular, chapter 2 first focuses on an analysis of the costs and benefits an individual considers in order to decide whether to get a higher level of education or not, successively it refers to the importance of training provided at work, so that the individual can obtain the necessary tools to allow him to obtain the advantages from investments in human capital. Chapter 3 provides the theoretical implications relating to the vast existing literature on scarring effects.

The second part of the thesis contains the results of empirical estimations of human capital and scarring effects. In this way it has been possible to combine the various parts of the thesis through a theoretical discussion on the one hand and empirical evidence on the other.
The empirical analyses are conducted on SHARE data (Survey of Health, Ageing and Retirement in Europe), using in particular the data of the dataset of wave 2 and wave 3 (SHARELIFE). The data of the two datasets are used as a unique cross section where the dependent variables are earning expressed as annual net income from employment. Furthermore in appendix B there is a representation of the results of scarring effects with longitudinal data referring to wave 3 of SHARELIFE. The empirical estimation analyses have been done using Stata.

In particular the first part of Chapter 4 is dedicated to the connection between the level of earnings and investment in human capital, focusing on the education level and training in the workplace. The instrument used in this regard is the so-called Mincer equation, which in fact puts the level of earnings in relation to the years of education and the work experience of an individual. The standard formulation of this function dates back to Mincer (1974). Starting from this initial form, other factors have been added. In particular, in addition to level of education and years of work experience, it is considered that marital status and number of children may also have an important role in establishing the level of earnings. All the estimates also pay special attention to gender differences, since women in the labour market tend to be at a disadvantage compared to men. The estimation method used for empirical analysis of the Mincer

equation is that of the ordinary least squares (OLS) with cross section data (wave2). In this way it is possible to produce the estimates for a given period of time, considering a portion of the population.

In the second part of chapter 4 the focus shifts to the determination of scarring effects. In particular, scarring effects occur when there are negative effects on a worker's career due to unemployment spells or a period of recession at the time of entry into the labour market. Also in this case, the estimation method used is that of the ordinary least squares with cross section data. What is expected from these effects is the risk of a permanent reduction of income for individuals who have experienced a recession at school leaving age or an unemployment spell.

All these analyses are carried out for four countries: France, Germany, Italy and Spain. The choice of these countries is due to the following factors: the weight they hold within the European community, and their geographical position. As a matter of fact, in various studies it has been observed that the countries of southern Europe (Italy and Spain) have several differences compared to continental countries (France and Germany). Finally, these are the countries for which the SHARE database makes available the largest number of observations, making it possible to achieve more accurate estimates.

2. Human capital

To overcome the crisis currently affecting several European countries, the many policies undertaken should focus more strongly on research, innovation and human capital, i.e. the productive factor which represents a more valuable advantage for international competition of industrial countries.

As a matter of fact the process of accumulation of human capital is a fundamental growth factor in modern economies, especially in periods of great technological and organizational changes in production processes.

General knowledge and skills are important elements for productivity growth and the technological innovation of firms. Moreover, according to the human capital theory, they tend to reduce the risk of unemployment for workers. In this regard, the system of education and training has traditionally played a complementary role. Education provides general knowledge to the specific needs of the production system and it has an impact in the medium-long term on economic growth, being realized with overlapping generations with different levels of schooling in the labour market. On the other hand, training promotes skills and knowledge during the life cycle of individuals; generally it has the aim to develop functional skills to the specific needs of firms and workers.

In this perspective, investments in education and training have a significant impact on the dynamics of the microeconomic labour market, from the points of view of both demand and supply. For workers, the return on human capital can be realized in the form of a wage increase, greater employment stability or through a reduction in the duration of unemployment. The advantages for the firm consist of the possibility to have qualified personnel able to manage technology innovations or operate in a more international context.

Analysis of the relationship between economic development and human capital has great importance when it is considered according to the common trends in the most advanced economies over the last decades. New technologies, globalization and the process of institutional reforms have profoundly influenced the dynamics of the labour market and therefore the return on investment in human capital. Obviously each country has a specific production and institutional system that could have different

impact on macroeconomic changes and technological developments in the domestic labour market.

2.1. Existing literature

One of the most important discussions of economic research has explored the connections between income from employment and human capital. Economists have long debated the causes of inequality in personal income, but empirical analysis has recent origins.

The oldest theory of income inequality and human capital associated income distribution with the distribution of individual skills traditionally assumed by Gaussian or normal distribution. It was expected, therefore, that income distribution should also present normal distribution. On the contrary, in Western countries until 1900 and even now for poor countries, the shape of the income curve strictly decreased (with a high concentration of economic units to an income level close to zero, few classes of workers with medium-high income and few individuals with very high incomes).

Currently, the curve of income distribution for countries with developed economies and the developing countries have the usual unimodal skewed functional form, with relatively low frequency of very low incomes, a large part having medium-low income and a progressively decreasing rate for the higher incomes.

In this perspective, the pioneering work of Mincer (1958) was essential. He investigated the relationship between individual income and skills, introducing the concept of human capital as a factor of growth and therefore of inequality of earnings among workers throughout the whole of their lives.

Mincer's contribution forms part of the theories formulated by the School of Chicago (Schultz, Mincer and Becker.) In the second half of '900 it involved an improvement in the human capital concept. The authors analyzed in detail the economic concept of human capital (coinciding with years of schooling and work experience) providing important contributions related to the mechanisms of formation and accumulation.

The thesis of the Chicago School consists of the fact that years spent on education and professional experience acquired in the workplace (on-the-job training) are the main

factors explaining the evolution of labour income in a life cycle (gain functions or earnings profiles) for workers.

The authors of the Chicago School have provided fundamental contributions to the study of human capital, but at the same time their theory is incomplete in some aspects. The main limitation of their methodology consists of considering human capital with the binomial education-experience without including other aspects such as age, the number of children for each individual, the gender of the individual, the marital status of the individual, the geographical area, etc (Dagum, 1994). For this reason chapter 4 of this thesis provides these variables in addition to the classical ones (education and work experience) considered by Mincer. In any case, empirical evidence has confirmed the assumptions made by the earnings profile models of Becker and Mincer.

In this way, age, the number of children for each individual, the gender of the individual, the marital status of the individual and the geographical area, are identified as factors that can change the amount of human capital held in equal investment and these are context factors that can promote the accumulation of human capital (Jorgerson and Fraumeni, 1989; OECD 1998; Wößmann, 2003).

2.2. Education and training

According to the theory of human capital, of which Becker (1964) was one of the chief exponents, education is an investment that produces income in the future. The differences in individual productivity are influenced by investments in education or training provided by individuals throughout their lives. Understanding the differences in individual productivity of workers is crucial because it determines wage differentials.

To acquire the necessary skills the labour market will require, an individual must face the "costs of training" that will be a source of future income. These include the costs of study (university fees, accommodation and travel costs, purchase of books, etc.). It must also include opportunity cost, i.e. the potential loss of income for the time spent on study instead of allocating it as remunerated activity. However, at the same time investments in education can produce an accumulation of skills called "human capital" that will provide economic benefits in terms of higher returns.

Education and training are the most important investments in terms of human capital. In some countries (particularly those in northern Europe), high school and university education decisively increase the future income of an individual.

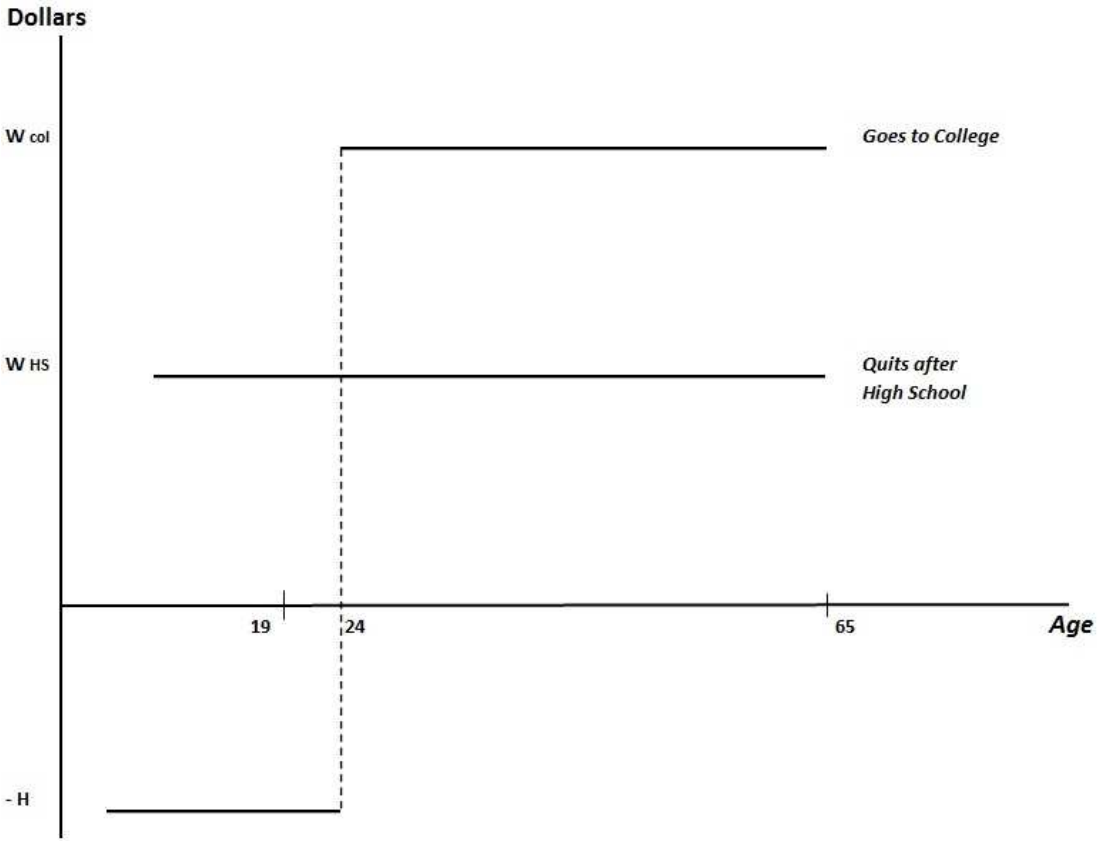
In these countries the earnings of people with higher education levels are almost always above the average. In Mediterranean countries this often does not happen and the years of education do not necessarily result in higher earnings. However, an analysis of human capital suggests that education increases the level of wages and the skills and ability to analyze problems.

Another theory, however, disputes the fact that education increases productivity but gives importance to "credentials" coming, for example, from achieving a degree. Grades and education in this case provide the right information on individual skills, determination and other personal characteristics (Spence, 1973). According to the most extreme versions of this type of analysis, the earnings of college graduates are higher than those of high school graduates, not because their education involves different levels of productivity, but simply because more productive students decide to go to college. There is no denying that this is so. However other evidence suggests that it cannot justify most of the positive correlation between earnings and education level. The main problem with this approach is that firms are not interested in the education people attain, but only in their capacity and performance in the workplace: the discipline imposed by firms, the need to satisfy customers and good relationships with its employees, etc.

High school and university education have been rapidly diffused in modern economies because the additional knowledge and information obtained during education are also fundamental in technologically advanced economies. However, learning and training also occur outside of schools and universities, particularly in the workplace. Even graduates are not ready for the labour market, into which they are inserted through training programmes. Although education reduces the probability of being unemployed and increases earnings, not all individuals obtain degrees or diplomas. It is therefore important to understand what factors motivate some individuals to stay at school, while others drop out.

The analysis of this important issue assumes that individuals acquire a skill level that maximizes the present value of lifetime earnings. Education and other forms of training are therefore considered important because it is assumed that they increase earnings.
Consider the situation of a nineteen year old who has finished high school and is deciding whether to enter the labour market or attend college, delaying entry onto the labour market by a further 5 years. It is assumed that there is no training on the job and that the skills learned at school do not devalue over time. These assumptions imply that the productivity of workers remains the same once they leave school, so that the real earnings (i.e. earnings after adjusting for inflation) are constant over the life cycle[1].

Figure 1: Earnings streams faced by a high school graduate



Source: George J. Borjas (1996)

[1] This is a simplification of the problem useful for a theoretical explanation because, as is demonstrated in Chapter 4, earnings tend to decrease after a certain age.

Figure 1 shows the economic trade-off that affects the worker's decision-making process. The figure indicates the earnings path during the life cycle attributed to each alternative.

The high school graduates that enter the labour market earn W_{HS} dollars annually until retirement age, which occurs when the worker is 65 years old. If a person decides to attend university he renounces W_{HS} dollars in labour income, and includes direct costs of H dollars to cover tuition, books, etc. Once he has achieved a degree, however, he earns W_{COL} dollars every year until retirement.

Firms wishing to attract a highly educated (and presumably more productive) person have to offer higher wages, so that $W_{COL} > W_{HS}$. The higher wages paid by firms for schooled workers, can be interpreted as compensation to workers for their training costs.

If graduates earned less than high school graduates, no one would attend university because it would mean that individuals would not receive economic benefits from further studies.

If an individual attends university the present value of earnings is the following:

$$PV_{COL} = -H - \frac{H}{(1+r)} - \frac{H}{(1+r)^2} - \frac{H}{(1+r)^3} + \frac{w_{COL}}{(1+r)^4} + \frac{w_{COL}}{(1+r)^5} + \dots + \frac{w_{COL}}{(1+r)^{46}} =$$

$$= -\sum_{t=0}^{3} \frac{H}{(1+r)^t} + \sum_{t=4}^{46} \frac{w_{COL}}{(1+r)^t} \qquad (1)$$

The parameter r is the worker's discount rate. The first four terms of this equation represent the present value of the direct costs of attending university, the remaining give the present value of earnings after graduation.

The present value of earnings if an individual attended only high school, is as follows:

$$PV_{HS} = w_{HS} + \frac{w_{HS}}{(1+r)} + \frac{w_{HS}}{(1+r)^2} + \dots + \frac{w_{HS}}{(1+r)^{46}} = \sum_{t=0}^{46} \frac{w_{HS}}{(1+r)^t} \qquad (2)$$

In this equation there are 47 terms, one for each year between the age of 19 (entry into the labour market) and 65 (retirement age).
Starting from these assumptions, an individual attains a degree if the present value of earnings arising from this level of education is greater than the present value of earnings deriving from obtaining a high school diploma. This means that $PV_{COL} > PV_{HS.}$
A school or university can be defined as an institution specialized in the production of training, which is something different from the production of training together with that of goods offered by a firm. The difference between what would be perceived and what is actually received should therefore be considered a direct cost of education for a student who decides to attend university.
In this way it is possible to define the net earnings as the difference between what would be received and the direct costs of education.

$$W = MP - k, \tag{3}$$

where W is the wage, MP is the effective marginal productivity (which is assumed equal to earnings) whereas k represents the direct costs. Indicating MP_0 as the marginal productivity that could be obtained, the equation (3) can be expressed as

$$W = MP_0 - (MP_0 - MP + k) = MP_0 - C, \tag{4}$$

where C is the sum of total costs (direct and indirect) and net earnings are therefore the difference between potential earning, represented by $MP_{0,}$ and total costs.

2.3. On-the-job Training

Training is more closely connected to the employment relationship than education. Training investments in adulthood often occur in the employment relationship. This is evident considering "learning by doing", i.e. the learning that occurs through work experience, as a result of the worker's participation in the workplace.

In general there are at least two parties involved in decisions relating to training, the employee and the employer, whereas in the case of education only the individual and perhaps his family decide.

In the case of training both parties are interested in investing because the benefits could be shared by the two parties. However, at the same time both parties sustain the costs on the basis that everyone shares the cost of the investment in proportion to the benefits that they could expect to obtain. Investment in human capital, therefore, consists of the participation and interaction of two parties (Becker, 1962; Hashimoto, 1981).

Often workers increase their productivity by acquiring new skills and improving those they already have on the job. It is assumed that future productivity can be increased only at a certain cost, otherwise there would be no limitations in the demand for training.

This cost includes the time and value attributed to the employment of those receiving training, the teaching provided and the materials used for the purpose. These are considered costs because the resources could have been employed in the production of current output if they had not been used to raise future output.

The cost and duration of the training period also depend on the type of training involved, since, for example, the cost for the specialization of an employee is greater than that of adapting a machine to a particular production process.

It is assumed for the moment that both the labour and product markets are perfectly competitive. It is also assumed that a firm is hiring employees in the same period. If there were no training on the job, the firm's wage levels would be independent from its actions.

A firm that maximizes profits will be in equilibrium if the marginal product equals wages, i.e. if marginal receipts equate marginal expenditures.

$$MP = W, \tag{5}$$

where W is the wage and MP is the marginal product. It is possible to assume therefore that employees all have the same marginal productivity (for given levels of other inputs) and receive the same wage in each period. These are respectively equal to the highest level of productivity in all cases.

A more complete set of equilibrium conditions is expressed as:

$$MPt = Wt, \tag{6}$$

Where t indicates the t-th period. The equilibrium level of each period will depend on the flows during the same period. These conditions change when the training acquired in the workplace and the connection it establishes between current and future earnings are considered.

Training could reduce wages and raise current expenses, therefore firms may have incentives to provide training if future earnings are high enough or if future expenditures are sufficiently reduced. The costs in each period will not necessarily be equal to the wages, the result obtained is that the expenditures and receipts of each period are now inevitably interrelated.

The equilibrium conditions derived from the equation (6) are replaced by equality between the present values of receipts and expenditures.

$$\sum_{t=0}^{n-1} \frac{R_t}{(1+i)^{t+1}} = \sum_{t=0}^{n-1} \frac{E_t}{(1+i)^{t+1}} \tag{7}$$

where Et is the expenditures and Rt is the receipts during the period t, while i is the market discount rate and n is the number of periods. Therefore, the equation (7) represents the equilibrium condition. The initial equilibrium equation is more general in the sense that if the marginal product is equal to the wage in each period the present value of marginal product flow should equal the flow of earnings. Obviously the opposite is not true.

If training were offered only during the initial period, expenditure during this period would be equal to the wages plus the cost related to training, whereas the expense of other periods would equal the level of marginal productivity.

The equation (7) therefore becomes:

$$MP_0 + \sum_{t=1}^{n-1} \frac{MP_t}{(1+i)^t} = W_0 + k + \sum_{t=1}^{n-1} \frac{W_t}{(1+i)^t} \quad (8)$$

where k indicates the outlay on training. Introducing the new term defined as:

$$G = \sum_{t=1}^{n-1} \frac{MP_t - W_t}{(1+i)^t} \quad (9)$$

the equation (8) can be written as:

$$MP_0 + G = W_0 + k, \quad (10)$$

the term k indicates only the actual expenditure on training, it does not consider all costs incurred. As such it does not take into consideration the time a person devotes to education, which could also have been employed in production under current conditions.

It is therefore necessary to consider the opportunity cost for training time that represents the difference between what could be produced, MP'_0 and what it is actually produced MP_0. If the sum of opportunity costs is represented by C, the result obtained becomes:

$$MP_0' + G = W_0 + C, \quad (11)$$

Value G is the excess of future receipts compared to costs. It measures the return to the firm resulting from the training offered. The difference between the return on and cost of training is given by the difference between G and C. The result that emerges from equation (11) is that the marginal product could be equal to the wages in the initial period only if the return equals costs, i.e. if G is equal to C.

According to Becker (1964) there are two types of on-the-job training, which he calls general training and specific training.

2.3.1. General training

General training is defined as the type of training that is useful (it increases productivity) in all other firms. The skills resulting from this type of training are diffused in the labour market. Specific training is the type of training that improves productivity only in the firm where the skills are acquired. The characteristic of this type of training is that its value decays once the employee leaves the firm. Instead perfect general training provides the same utility for different firms and the marginal productivity therefore grows by the same measure in every firm.

The result shows that the level of earnings would grow at the same level of marginal productivity and the firms that provide this training would not be able to obtain the corresponding return.

According to the theory of human capital, rational firms operating in competitive labour markets should provide general training if this does not imply any cost (Becker, 1964). On the other hand, the advantage for workers who receive general training derives from the fact that they are willing to bear these costs because training increases their future earnings.

The characteristics of general training can be expressed as follows:

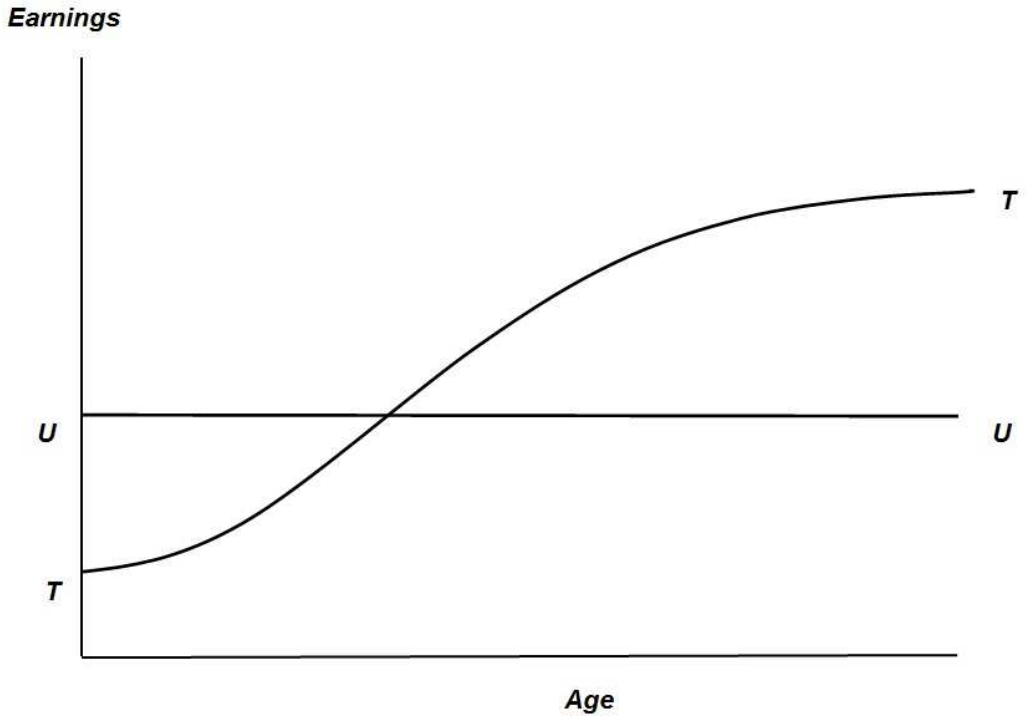
$$G = \sum_{t=1}^{n-1} \frac{MP_t - W_t}{(1+i)^t} = 0 \qquad (12)$$

Since wages and marginal productivity grow in the same measure, MPt will be equal to Wt for every t = 1, ... n-1. In this way the equation (11) becomes:

$$W_0 = MP'_0 - C, \qquad (13)$$

The earnings of those who acquire this training will not be equal to potential marginal productivity, but they will be lower because earnings include the cost of training. This means that workers would be willing to pay for a generic type of training by accepting wages lower than the level of potential productivity.
Training also considerably affects the relation between earnings and age as shown in Figure 2.

Figure 2: Relation of earnings to age



Source: Becker G. (1964)

It is assumed that a worker without training receives the same earnings at any age, as shown by the line UU in Figure 2.
However, people who obtained training would get a lower wage during the training period in which specialization is paid for and they will receive higher earnings in the future, when they receive the returns from training, as shown by line TT. The effect of training in line TT in Figure 2 is that the curve is steeper and also more concave. This means that the rate of growth of earnings is higher in youth than in old age.

Comparing education with training received in the workplace, generally only the direct costs of education are considered, although the opportunity costs are sometimes (especially in the case of university education) high. In the case of training on the job, the situation is reversed, where the costs are like forgone earnings.

Firms that maximize profits and operate in competitive labour markets would not bear the costs of general training and would pay the workers the going rate of remuneration. Otherwise if training costs were assumed by firms, many workers may leave and go to a firm where they could be productive in the same measure. In this way the costs of the firm that has provided specialization could be very high.

Finally if firms paid the costs of training but at the same time paid lower wages than those of the market, they would attract many workers looking for training but at the same time would have few trained persons (workers who have already received training).

General training has been considered until now in a perfectly competitive labour market. Now assume a different situation, in which there are some firms for which the training received by the worker (which, therefore, can be general) would be equally productive although competition between them is not full. At this point it is important to analyze the consequences of imperfect competition on investment in general training. General training that equally increases the productivity of a worker in any firm has a technological nature because it refers implicitly to the production function of firms, i.e. the ability of human capital to increase output.

However, in order for training to be general from an economic point of view, the labour market must also be perfectly competitive. In other terms, it is not enough for training to result in an equal increase of productivity everywhere, but it must also lead to an equal increase in workers' wages. In reality, however, there may be many situations where the labour market does not work in perfect competition. The difference between an increase in productivity and wages is an advantage for the firm. Consequently, also the firm, and not only the employee, is encouraged to support training costs. This contrasts with the result that was obtained by assuming a perfectly competitive labour market, according to which only employees have an interest in investing in their general training.

2.3.2. Specific training

Specific training increases productivity mainly in firms that supply it. Sometimes, after hiring a worker firms do not have the information they need about the new individual. So they try to increase the information available in various ways (tests, rotation between different offices, etc); it is a fact that greater knowledge allows better utilization of the workforce.

Most of the training provided at work is neither completely specific nor completely general, but since it promotes the increase of productivity in the firm that provides it, training tends to be mainly specific.

If all the training provided is specific, the wage an employee could earn in other firms would be independent of the level of training he had previously obtained. In this case, firms should bear the costs of training, otherwise no rational worker would pay for a type of training from which he could not acquire benefits.

The wage a worker might receive after specific training is therefore the same as that he would receive without training. In this way the firm that provides training can offer the same wage (or a slightly higher wage) to hold him.

The firm will therefore capture all the revenues because it can employ a qualified worker with a wage lower than productivity. In this way the firm is stimulated to support the cost of training. Summarizing, when there is specific training the firm captures all the revenues and therefore is also encouraged to support all the costs. However, the specific skills will be lost if the employee leaves the firm where they were acquired, since in another firm, his productivity is equal to that the worker had before training.

The possibility of taking advantage of specific training, therefore, depends on the continuation of the relationship between the worker and the firm.

In order to avoid a separation occurring, the firm in the second period finds it convenient to pay the worker a wage higher than the market, but in any case lower than the productivity it obtains from his employment. So the worker participates in the division of the gains from training and, consequently, will also be willing to share the costs in the first period. In fact, there are many unskilled workers who accept being

trained, in view of the higher wages they will earn after training. Inevitably competition between them pushes the wages of the first period below W_1. It is in this way that the worker participates in the investment: a reduction of the starting wage represents the portion of its costs.

In conclusion, to prevent labour mobility that would be inefficient, the firm will not achieve all investment but both parties will participate, sharing revenues and costs in proportions which may vary depending on the circumstances. In equilibrium, it is a profile wage increasing over time, i.e., $W_1 < W_2$, which has the effect of "binding" the worker to the firm, avoiding separations in the second period.

If training were not completely specific, productivity would increase in other firms and even the wages obtainable elsewhere.

Since firms do not bear any cost resulting from general training (in perfect competition) and they support only a portion of specific training, the fraction of the total actual costs paid by firms is inversely related to the importance of the general component or directly related to the specificity of the training provided. It can be expressed by starting from the previous equations:

$$MP' + G = W + C, \tag{14}$$

where G represents the present value of the return from training received by firms. Denoting the return received by employees with G', then the total return, G", is the sum of G and G'. Total return is equal to total costs in full equilibrium, i.e. G"= C. The equation (14) becomes:

$$W = MP' - (1-a)C, \tag{15}$$

Employees support a fraction of the costs (1-a) in the same measure as the returns. The result is that if training is completely general (a = 0), the above equation reduces to equation (13); instead, if firms obtain all the return resulting from training (a = 1), the equation (15) reduces to $MP'_0 = W_0$, finally if $0 < a < 1$, none of the above equation is satisfied.

2.4. Gender in the labour market

Among the objectives of the cohesion policies in the European Union, great importance is given to the reduction of economic and social disparities, in particular those concerning the access and permanence of women in the labour market and the improvement of their employment conditions.
The economic situation of women is not equivalent to that of men. The data on earnings and professional achievements show that women have a secondary economic position. This fact is verified in all countries for which data exist, but in different measures. The gap between women and men has been practically constant. Nevertheless, marital status, the family life cycle, etc. influence the size of wage differentials. It is important to understand why women seem to be relegated to a position of economic inferiority. In particular the lack of services aimed at reconciling work and family continues to be a determinant factor in the lower participation of women in the labour market in the early years of their motherhood. Moreover within the family household tasks and care weigh disproportionately on women, even among couples where both spouses work.

If these situations arise because of unequal opportunities due to unfair recruitment practices, then the economy fails to produce full and adequate productive personnel.
First it is useful to examine the possibility of discrimination in the labour market in assessing the causes of wage differences between men and women. In a labour market context, there is discrimination when men with the same productivity as women are paid more. If discrimination takes place in the firm, it is also called demand-side discrimination.
Women will receive lower pay than men with the same productivity, if there is prejudice against them by employers. However, in a competitive market, if men are paid more than women with similar productivity, some firms may prefer to hire women instead of men, in this way reducing their costs. The wages of men and women would be brought to parity.
Let us assume now that a firm begins to discriminate against women, for example it does not hire women. However if a firm hires only men, it reduces its possibilities and labour costs will be more expensive. Discriminators therefore pay for their

discrimination and the achievement of profit maximization will tend to drive out discrimination.

Another problem with discrimination is the motive, or why there should be discriminatory prejudices. The most likely reason is that some groups of men discriminate against women and do so because they "make money" from it in some way.

According to one school of thought, women earn less because they are in "women's jobs". This theory is called occupational segregation or crowding hypothesis. The thesis is that certain jobs are designed for women, whereas men can choose any occupation. What emerges is that women are crowded into a limited number of jobs, reducing wages in such occupations. One way to evaluate the crowding hypothesis is to compute the differences in occupational distribution and verify how they can explain the wage differentials.

However, it should be considered that given the same level of education, men's and women's wages are more or less balanced (Mincer and Polachek; 1974). This does not happen when women are married. Theories and empirical data show that married women with a family have less interest in investing in the labour market, preferring to devote their energy to other activities such as child care, especially at pre-school age. The allocation of time available to women between employment and family thus characterized the role of women in the labour market. An important consequence is an increasing reduction in the number of children. In many countries there is an increase in the number of women who prefer to pursue a career and receive good pay rather than choose a domestic role. There are generally few economic incentives provided to a large family. There are thus problems such as families without children with consequent ageing of the population, and women who have children preferring to invest initially in a working career.

2.4.1. Horizontal and vertical segregation

Women having the same investment in human capital as a man in theory should have the same right and opportunity to access the labour market. However, this does not occur, and women are often confined to specific areas and employed in positions with lower wages and fewer opportunities for employment growth. A group is subject to segregation when it is systematically over-represented in certain sectors and/or occupations. Segregation can be divided into:

- horizontal segregation: women are employed mainly in some specific sectors (crowding hypothesis);
- vertical segregation: women are over-represented in jobs in the lower part of the organization of a firm.

The causes of segregation can be identified in the supply side: by preference women choose specific professions and sectors, and in the demand side: employers prefer to hire men instead of women although they have the same characteristics.

Becker (1981) assumes that there are different preferences between men and women concerning the division of labour within the family. In this model the choice of women to specialize in domestic work and in certain types of occupations is the result of a rational decision based on an analysis of costs / benefits. As a matter of fact women have a comparative advantage in housework, for biological or discriminatory reasons. Division of labour within the family is also reflected in investment in human capital and the productive efficiency of men and women.
The basic assumption is that each individual has an incentive to invest in capital according to the sector where s/he will spend more time. Women therefore invest less in human capital or they choose to invest in specific sectors knowing that their participation in the labour market could be discontinuous as a result of caring for their family (Polacheck, 1976)[2]. These sectors correspond to those in which the amount of

[2] This is true with regard to training acquired in the workplace because women have about the same amount of investment in human capital in terms of education.

human capital required is lower, therefore, these are less qualified and less productive. The result is that these occupations are less remunerative.

2.4.2. A measure of wage discrimination

The measure of "wage discrimination" and the word 'discrimination' are two different concepts. Terms regarding discrimination and prejudice often refer to an aversion to join another person for reasons unrelated to the intrinsic aspects of productivity. Discrimination can also occur when a group (women) has average characteristics that are different from those of another group (men).

The former type of discrimination has been developed by Becker (1964), who first formalized it in a theory of discrimination; such discrimination is usually called "taste discrimination". The other type of discrimination is called "statistical discrimination" and it was developed by Kenneth Arrow (1973). For example in statistical discrimination the groups do not differ by features related to productivity. However, they differ in the talent of others to deduce the productivity of an award, such as education. The groups can differ in ability, education or skills, but small initial differences through feedback effects can lead to big differences over time.

The models of Becker's taste discrimination have been developed to hierarchically list forms of discrimination in which men have an aversion to working under the direction of a woman, but are willing to supervise women. In both cases highlighted, the presence of discrimination may cause differences in individuals' earnings and/or levels of employment. Competitive forces may help to attenuate the presence of discrimination and prejudice although these are widespread in society. If a substantial number of employers does not have prejudicial tastes, the wages obtained by the market forces cannot reveal the diversity of tastes.

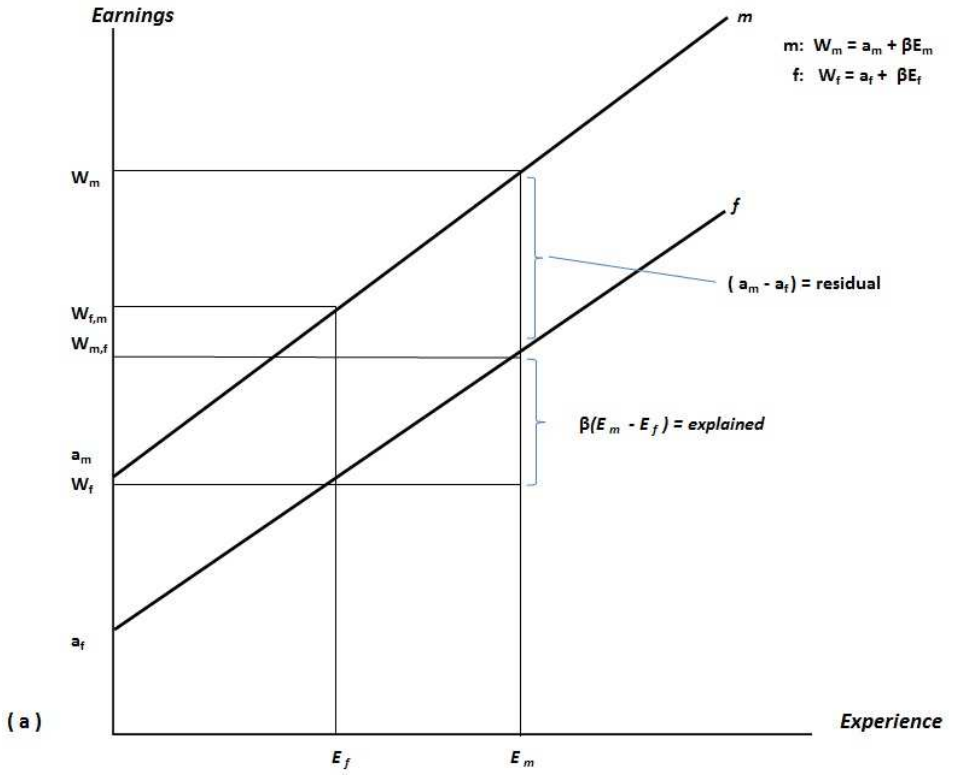
However, historical evolution shows the existence of both occupational segregation and wage discrimination, although they have had different effects over time. Where discrimination is due to tastes or statistical differences, the forces of competition have not eliminated the impact of discrimination. However there is a technique designed to assess the degree to which individuals with similar characteristics are paid differently.

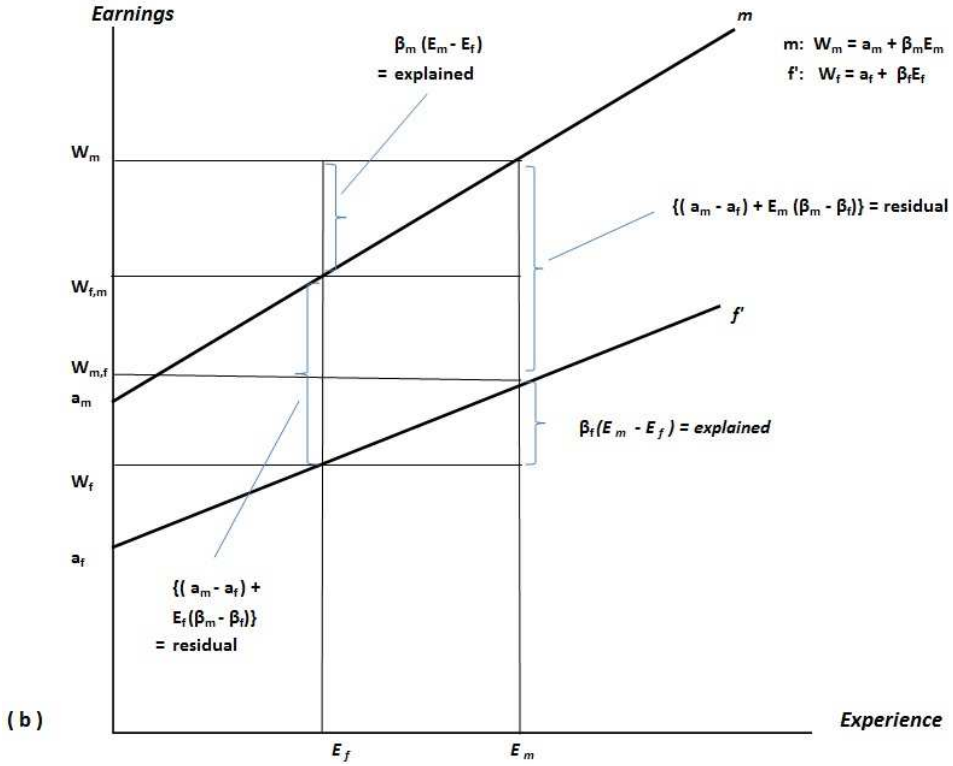
The technique is called "wage discrimination", because it can show the prejudices of one group against another. Specifically, this technique evaluates the dollar value of the various features of the labour market for each of the two groups (in this case males and females). At this point the amount a group (females) should receive if it is paid as the male group considered the reference group of the market is judged. The measure of wage discrimination therefore is given by the difference between the amount computed for women and that actually observed, divided by the difference in pay for males and females. In this way the technique evaluates the difference between male and female remuneration of that they had receive in the same measure and then it compares the result with the real difference.

Initially the case of a single feature rewarded in the same way is considered from the point of view of marginal increments for males and females, but whose earnings are different between an equally qualified man and woman. This case is represented in Figure 3a. If males and females are rewarded in the same measure for incremental amounts of total experience (E), the difference in wages between males and females (W_m - W_f) is due to a difference in their features, times the value of properties, [β (E_m - E_f)], plus an additional difference in the intercept (A_m - A_f) of the two lines, m and f. In the case in which the features between males and females are the same, the difference between the intercepts represents the "unexplained" difference in their wages.

Let us now tackle the case in which the reward to the feature is different for males and females. In Figure 3b, line f of figure 3a now becomes f' which has a different slope from line m. This means that in this case males and females are paid differently for increases in work experience. Hence, both the different slopes and the different intercepts cause differences in earnings between males and females. Considering point W_{fm} as the wage a female earns if she is paid the same wage as a male, the difference (W_m - W_{fm}) represents the difference in the level of experience (the males have more experience than the females), whereas the difference (W_{fm} - W_f) is the residual. The unexplained or residual value represents how females are disadvantaged because they start at a lower level of earnings and because they are paid less for the same level of experience compared to males. In the case represented in Figure 3a, returns for the attribute are equal by gender ($\beta_f = \beta m = \beta$), while in Figure 3b returns for the attribute differ by gender ($\beta_f < \beta_m = \beta$).

Figure 3: Earnings functions and "Wage discrimination"





Source: Goldin C. (1990)

2.5. The effect of the family context on earnings

A discussion of human capital must take into account the influence of families on the education, attitudes, skills and values of their children. The major differences among young children grow over time with age and education, because when they are better prepared they can learn more easily. Thus, during preparation provided by families to their children, even small differences among children imply major differences in adolescence. For this reason the labour market cannot do much for those who abandoned their studies or have never developed technical skills; it is therefore difficult to create actions to help these groups of people. Families play an important role in education and all the other characteristics of their children's lives.

Families classified as "underclass" are those in which low levels of education, welfare dependence and marital instability are transferred from parents to children. Despite this fact, although the incomes of parents and children are positively correlated, the relationship is not so strong. On the other hand, the income of parents and children is more closely related if the parents are poorer. This happens because the richer families can afford to spend more on educating their children, including the forgone income of the period of their education (opportunity costs).

Families share expenditure for their children on the basis of the number of children in the family. The amount spent for each child is generally negatively correlated with the number of children included in the family. This is because a large number of children increases the actual cost of the increased expenditure dedicated to each one. The negative relationship between the number of children and the expenditure available for each of them can be expanded at the aggregate level between investments in human capital and population growth. As a matter of fact, countries where on average there are small families generally spend large amounts on education and training for their children, while those countries where on average there are large families spend much less (Becker and Murphy, 1988).

The famous expectation of Malthus (1798), by which people marry earlier and birth rates grow as income rises was refuted by the industrial revolution. In fact, a short time later a second edition of his book on population was published; this led to a great text

being contradicted by the facts that occurred. In Western countries where per capita income increased fertility reduced rapidly instead of growing. Rapid progress in education and other types of training available to the population was accompanied by a drastic decline in fertility. Families began to spend more on their children as income rose (in agreement with the predictions of Malthus), but they spent much more for each child without increasing the number of them, as predicted by the theory of human capital.

2.6. Relationships between ability and the distribution of earnings

Economists have always been interested in income distribution in their studies, which go back to Smith, Mill, Ricardo and others who recognized that several issues of considerable economic importance are partly reflected in various aspects of income distribution.

The human capital theory helps to explain not only the differences in earnings over time and in different areas, but also among individuals within the same area[3]. Economists say that even if common measures of ability (intelligence tests, school grades and personality tests) may sometimes be indicative, they fail to reliably measure the characteristics needed to succeed in the economic field.

Therefore some authors argue that the only acceptable way to evaluate economic talent is by measuring economic performance (wages actually received). Individuals with higher earnings are therefore those more capable than others. The aim of linking capabilities to earnings is to distinguish its effects from differences in education, job training, health and all the factors of this kind. A solution to this problem would be to identify ability by earnings when the amount invested in these types of capital is kept constant. A rational approximation says that if two individuals have the same level of investment in human capital, the one who earns more is showing greater economic talent, although a full analysis should consider discrimination, luck and several other factors constant.

[3] This aspect of human capital theory is widely applied in Chapter 4.

Observed earnings are also influenced by variations in the amount and rate of return, because they are gross of the return from human capital. In effect, after a period of investment, earnings (Y) can simply be approximated by:

$$Y = X + rC, \tag{16}$$

where C is total investment costs, r represents the average rate of return, and X represents the earnings independent from investment in human capital. Y would depend only on r if C is constant and if X distribution is currently ignored, so the "ability" could be assessed by the average rate of return on human capital. The fact that the amount invested is not rigidly fixed to individuals and is different between them depends partially on the rate of return. In this way individuals with an higher marginal rate of return are encouraged to invest more than others. Since it is assumed that the average and marginal rates of return are positively correlated and since the ability is measured by the average rate, this implies that the abler individuals tend to invest more than others. From these implications it follows that there is a positive correlation between ability and the level of investment in human capital, and this fact has several important inferences. For example, the tendency of abler individuals to emigrate, to complete their education, etc. can be justified without the assumption of non-economic forces or market conditions. Another example is what distinguishes ability from education and context is difficult because high earnings may reflect both a higher level of ability and better environmental conditions.

The approach described has an implication that not only is consistent with economic analysis, but is based on one of its key aspects, i.e. that the invested amount is a function of the rate of expected return. Together with the effects of human capital on earnings, this approach is able to explain several characteristics of the distribution of earnings. In particular, according to the definition provided above, if all individuals invest the same amount in human capital, the distribution of earnings should follow the same distribution as ability; if abilities are symmetrically distributed, earnings should also be distributed in the same way.

According to equation (16) if all individuals were equally able, the distribution of earnings would be exactly the same as that of the investment.

On the other hand, if both the level of ability and that of investment change, then the earnings tend to be skewed even if ability and investment are not distorted. However, the skewness could be minimal if the invested amount is statistically independent of ability (Mincer, 1958).

However, individuals with higher ability would tend to invest in human capital more than others. In this way ability and investment would be strongly positively correlated. Changes in the variable X explain the difference among the categories of personal characteristics as regards the degree of skewness. If the total earnings depend to a lesser extent on the term rC with respect to X, i.e. investment in human capital compared to the case in which there is no investment in human capital, then the distribution of earnings will be dominated by the distribution of X. Categories characterized by a higher level of skills make higher average investment in human capital and therefore presumably a higher level of rC with respect to X.

The categories characterized by "unskilled ability" are dominated by X that establishes the distribution of earnings in this category; whereas the distribution of a product of ability and the amount invested, rC, tend to be dominant in the "skilled categories". Hence earnings are more symmetrically distributed among unskilled individuals than among skilled individuals when abilities are symmetrically distributed.

Equation (16) assumes value only in the case in which investment costs are small, but this usually occurs at later ages, presumably after thirty-five years of age.

Equation (17) gives the net earnings at earlier ages:

$$Y_j = X_j + \sum_{0}^{j-1} r_i C_i + (- C_j) \qquad (17)$$

where Ci indicates the investment cost of age i, and Cj refers to current costs; in particular j represents the current year and i refers to previous years, finally ri measures the rate of return on Cj. The distribution of –Cj is fundamental in determining the distribution of Yj, because investment is greater at this age. In this way it provides less

(positive) skewness at a younger age because X would result in youth greater than Σ riCi.

This analysis shows that investment in human capital can explain the fact that distribution of earnings is more skewed than distribution of abilities, and it also shows that earnings are indeed more skewed among older skilled individuals than among younger less skilled ones.

2.7. The theoretical model of Mincer

One of the most important authors to have theorized about the relationship between education and success in the labour market is Jacob Mincer (1958). Together with Gary Becker, winner of the Nobel Prize for Economics, Mincer has made a particular contribution, revolutionary in some aspects of the human capital theory.

The concepts previously explained in paragraphs 2.2 and 2.3 are used to explain the theory expounded by Mincer.

Mincer's theoretical model tries to explain the connection between human capital and the effect it produces on the earnings received by individuals. The gain then depends on the choice of work. There is therefore a significant connection between an analysis of education, age of individuals and work experience. Mincer's model starts from the assumption that in nature each individual has the same ability and each one has an equal chance of finding any work activity. Occupations differ according to the amount of training required to perform them. This training represents the time spent in years of education and the preparation for entry into the labour market (Mincer called this type of training "formal training"). Instead the period of time dedicated to work experience Mincer calls "informal training". Each year dedicated by the individual to formal training corresponds to a year in loss of earnings (opportunity cost). Only when the individual chooses his occupation, will he be rewarded for costs incurred during the formal training period. According to this hypothesis the years devoted to formal training are a form of investment that will be reflected in future employment. There will therefore be a possibility for advancement in the labour market by improving the economic benefits received. Mincer assumes that individuals with a higher level of

education earn more than individuals with less education. What has been said was transformed by Mincer into a "Mincer equation" (Mincer, 1974). In particular, the Mincer equation makes it possible to measure the effects of education on earnings taking into account observable individual characteristics. In its classical form the equation can be expressed as:

$$ln\ Y_i = a + bS_i + cEX_i + dEX_i^2 + e_i\ , \qquad (18)$$

where Y_i represents the natural logarithm of the wages of individual i, S_i are the years of schooling, EX_i are the years of work experience and e_i is the statistical error.
The log-linear functional form allows us to evaluate coefficient b as the percentage variation in wages associated with an additional year of education. In addition, Mincer showed that b can be interpreted as the private return on a year of investment in education, defined as the discount rate that equals the expected value of the net economic benefits of its costs, if the sole cost of attending school or university for another year delays entry into the labour market (opportunity cost) and if the increase in wages due to years of education is constant throughout working life.

According to the theory of human capital, since investment can be made in different types of education, in the presence of perfect markets there will be a tendency to somewhat level the returns on education in all sectors and the current values of future earnings will be equivalent, whereas the return on investment in human resources cannot fall below the interest rate of the market, which measures the result of alternative investments by private individuals.
The econometric tests should therefore show parameters b not too different for different types of education within the same country. Differences in returns would be interpreted as a consequence of market imperfections.
Chapter 4 contains the empirical analysis of the function, with regard to rates of return to education.
Nevertheless, it is appropriate here to point out some general difficulties that such studies need to address. First, many methodological pitfalls should be avoided in the use of econometric tools, especially in terms of specification errors and in the presence of

factors that distort the estimates. Furthermore the use of indirect variables (proxy) should be avoided, as they are unsuitable in the absence of information on the crucial variables of the model. When setting the relationship between earnings and the independent variables which influence it, some relevant factors can be omitted, or important links between the variables themselves can be neglected.

A good example is the case of workers' abilities, a variable of difficult observation that some authors neglect and which could in turn influence schooling, since the ablest individuals may study more to earn more.

3. The scarring effects

This chapter deals with the implications of scarring effects that can be connected with the main aspects of the human capital theory explained in the previous chapter. In particular "scarring effects" are the negative effects on the future career of a worker arising from a period of recession or long unemployment. These effects occur mainly in younger generations and can be identified especially in the following cases:

- Effects resulting from graduating in a recession period;
- Effects resulting from youth unemployment.

Scarring effects may occur primarily in wasting new generations excluded from the labour market, especially in the countries of Southern Europe.
The economic crisis has to a greater extent caused the weakness of young people in the European labour markets. They are the most affected by unemployment and so are unable to enter the labour market or are employed in marginal, temporary and precarious positions. Moreover, the employment situation of young individuals is influenced by the norms that regulate labour markets and education and training systems in each country, all factors that affect the quality of the transition from school to work. The unemployment experience is particularly worrying because it can generate scarring effects that have permanent negative effects which reduce the prospects for the future level of employment and income. Due to this mechanism, the effects of a prolonged crisis on youth employment risk becoming structural problems that could be difficult to overcome, and the generation most affected by the crisis risks becoming a "lost generation" which will not be able to compete with future generations in the labour market. These young individuals may be trapped in unemployment with a high probability of social exclusion.
In addition, the need to reduce public expenditure and transition to a system of social protection increasingly of the insurance type, penalizes further the current young generations, who risk being unprotected with a high risk of poverty in the future. At the collective level, this involves a loss of potentially productive human resources, which

inevitably leads to a weakening of economic growth and to an enlargement of the poverty range with inequalities of income among generations. Moreover it generates high costs for the State budget due to lower income tax and increases in expenditure for welfare policies.
For these reasons, one of the main objectives of the European Commission is the reduction of youth unemployment, and the improvement of the level of qualification of young people (education and training).

3.1. Existing literature

Obviously a period of job loss reduces peoples' current income, but the damage may be greater if it involves a scar. Existing literature on the effects of scarring shows a dual impact: the future employment prospects of workers worsen and/or decrease their future income and these effects can potentially persist for the remaining working life of an individual. However, there are economic theories that can predict scarring. Following the intuition of Becker (1975) explained in subsection 2.2, if firms provide specific skills, these skills are not transferable and therefore increase the marginal productivity of workers only in the firm that provides the investment, differently from a firm with general skills that increase the marginal productivity of a worker in all the different firms. In this way the consequence of unemployment is depreciation of general skills and the loss of specific ones. Therefore, when a worker returns to the labour market he will obtain a lower wage than he received before the unemployment spell. At the same time, returning to the labour market will provide further accumulation of human capital, hence the scarring effects will be only temporary.

According to Search Theory, unemployment periods are used for job seeking and to improve the probability of finding a better job in future periods.
In recent years, empirical economic studies have attempted to evaluate the scarring effects of unemployment by looking at wages in periods immediately before and after the unemployment spell. Significant long-term effects on wages due to unemployment were found by Rhum (1991). In a comparison of displaced workers in the three previous years of displacement and the four following years, Rhum realized that in the long run

the disadvantage of employment decreases, while the disadvantage of wages remains large and persistent.
At the time of displacement there will be a strong decrease, but in the five years following, after a slight recovery, displaced workers will perceive 25% lower earnings than non-displaced workers (Jacobson et al, 1993).

According to Stevens (1997), it is important to assess the permanent scarring effects because if individuals can avoid being unemployed more than once, they will do so if there is a good chance of recovery. In this way multiple job loss is a fundamental element behind the permanent scarring effects of unemployment.

Gregory and Jukes (2001), analyzing the variations among age groups, suggest that the impact of scarring effects is greater for older workers, but the duration of the effects is more substantial for the young.

The general opinion of these authors is that a period of unemployment constantly imposes a wage scar on income individuals. In particular, Stevens (1997) says that a substantial part of the reason for the persistent effects is that repeated incidence inhibits the recovery of wages and Gregg (2001) says that a spell of unemployment increases the probability of repeated job loss.

3.2. Scarring effects resulting from graduating in a recession period

In times of economic crisis and subsequent recession, young individuals in particular are more exposed to the effects of the crisis, because they are often employed with atypical forms of contract. The evaluation of the labour market consequences of a recession can be made by evaluating the determinants of the probability of being employed or unemployed in periods pre- and post-crisis. The purpose is to try to identify which features of an individual are significant in assessing transitions among states of employment, and especially to assess how these features change from the pre-crisis to post-crisis period.

Analyzing the long-term effects on future earnings of graduating in a recession and the future level of employment of young individuals according to different studies, it can be observed that young graduates in a recession suffer significant initial losses in the labour market, which tend to disappear on average after 8-10 years (Oreopoulos and Von Wachter, 2008). However, it should be considered that more highly skilled graduates suffer less in a recession, because they can choose to go to better firms more rapidly than less skilled graduates who might be permanently affected by low earnings.
Even short-term shocks on the labour market may have substantial and different long term effects on the future careers of workers (Kahn, 2006; Kondo, 2008).
Four main implications can result from temporary negative shocks on the labour market (Oreopoulos and Von Wachter, 2008):

1. **Time dependent search**: when age or a permanent job increase, the negative scarring effects deriving from a labour market shock decrease. Workers reduce their search effort and consequently their benefits deriving from searching, as they accumulate specific human capital in a firm.
2. **Differences by skill group**: the more skilled workers recover more quickly from adverse initial conditions on the labour market through a higher rate of mobility toward firms that offer a higher wage. The intensity of searching increases not only the wages of workers, but also their abilities.
3. **Catch-up on-the-job:** when a worker finds a job in a high-wage firm, the earnings converge to that of similar workers already present in the same structure as the worker accumulates specific skills in that firm. This fact is due to the concave tenure-wage profile.
4. **Zero Search:** an increasing fraction of workers tends to reduce job seeking over time. The reduction of seeking with the passage of time is a common factor, but it seems to be more widespread for low-skilled workers. This means that the process of catch-up tends to be incomplete for low-skilled workers.

The intuition from these four implications consists of the fact that the effects of initial unemployment rates provide permanent earnings differences only if combined with search frictions which increase with age or the quality of jobs. In addition, these

processes can be different according to graduates' different ability levels. Low-skilled graduates risk being most affected by time increasing mobility costs, and being cataloged by firms as low paid. The result that emerges is that low-skilled graduates are more susceptible to adverse effects resulting from the scarring effects of the initial conditions of the labour market.

3.3. The youth labour market problem and its scarring effects

To understand the causes of young peoples' success in the labour market, we must consider the difficult choice all young people face between continuing their studies on the one hand, or entering the labour market on the other. Participation in the labour market obviously brings with it the hope of finding employment, but also the risk of being in a state of unemployment. There is also the risk that an increase in the duration of unemployment and its episodes increase the probability of being unemployed for a long time in the rest of the worker's working life (scarring effects). In this case there is dependence of unemployment on its duration (Mroz and Savage, 2006; Doiron and Gørgens, 2008). Youth unemployment also generates other negative consequences. Indeed, it reduces the average level of community welfare and has a negative effect on the birth rate, because it affects the age when female fertility is higher. It leads in the long-run to dramatic stagnation of population growth in advanced countries, a phenomenon that involves not only academic, but also economic and political debate. For this reason the major international organizations give great importance to youth problems.

The level of unemployment is different in different countries, but the common aspect is that unemployment is higher for young people than adults (in Europe youth unemployment is generally higher in Mediterranean countries). In general, the higher average unemployment is, the higher youth unemployment is. For example, youth unemployment has a strong cyclical component, which also influences unemployment at the aggregate level (O'Higgins, 2005).

This does not imply that if adult unemployment is higher, then that of young individuals must also be higher. Conversely, it could happen that the young may be at a relative

disadvantage even in the presence of a situation of economic growth. In this case, the main causes of youth unemployment must be sought in the institutional factors that influence school-work transitions (Ryan, 2001; Christopoulou, 2008).

One of the characteristics of the youth labour market on which economists have concentrated their attention is the low rate of activity of young people compared to adults (Ryan, 2001). It should be noted, however, that the lack of youth participation in the labour market should not necessarily be seen as a negative phenomenon. As a matter of fact it may be caused by two different factors: participation in the education system or discouragement. A situation of discouragement refers to those individuals who are without a job and who are not actively looking for one.
The causes of youth unemployment compared to that of adults have to be found primarily in the factors at the aggregate level (low level of aggregate demand). In contrast to adults, whose employment is somewhat stable over the economic cycle, youth employment is anti-cyclical (Jimeno and Rodriguez-Palenzuela, 2002). During mass dismissals, however, young people and adults are equally likely to lose their jobs. Indeed, as often occurs in old sectors, the youth workforce is a minority, due to low recruitment of labour, and the probability of losing jobs is higher for adults than for young individuals.

The employment situation of young people is very different among countries, but it is possible to observe two distinct types, toward which European youth tends. As explained in section 2.4, there are those who possess high qualifications. They usually have a large family background, both in terms of education and employment and their transition from school to work should generally be made without significant problems. Other young individuals, mainly from difficult family circumstances, tend to remain low-skilled, like their parents, and usually have continuous interruptions in their school-work, with obvious consequences also on their future earnings.
To understand the reason why in every country young people have an unemployment rate greater than that of adults, we must consider the difference in work experience. This difference makes the human capital of young people less than that of adults despite the presence of increasing levels of education. It is precisely the need to overcome this difference that drives young people to "experiment" in the labour market, frequently

passing from one job to another. Analyzing the flows in the youth labour market, young individuals are constantly engaged in finding the best job to which they can aspire given their ability (Clark and Summers, 1982). One consequence of this approach to research and experimentation is that their flows in the labour market between the states of employment and unemployment are much higher than the average of the population. When seeking occupation, they often fall in the condition of unemployment. The reasons for the high mobility of youth employment, sometimes even called job shopping, are the following (Topel and Ward, 1992):

- young people look for the best job solution, but they do not know in detail either the nature of the best workplace for them or their own abilities: the only way to obtain this information is to try different work experiences;
- at the same time, they want to accumulate human capital in the form of more work experience, so they look for easier access to information on jobs available on the market. The purpose of this research is to put young people on a par with adults;
- employers also look for workers more productive and suitable to their needs, so they need flexible selection tools: on-the-job training, trial periods, etc;
- young individuals, especially those with a low level of qualification, in some cases go back into the education system or training on the job to overcome the training gaps accumulated.

Considering the lack of work experience of young individuals and their quest for the best solution, it is easier to understand their high unemployment rates.
Some economists believe that unemployment would be a logical consequence for young individuals who will overcome it with the passage of time, becoming adults. The main objective of economic policy in favour of the young unemployed should consist, therefore, of making the labour market more flexible. In effect, greater flexibility can allow young workers to increase work experience more quickly and give them easier transitions from school to work. According to this school of thought, the need for mobility among different occupations is considered the main solution to decrease the

dependence of unemployment on its duration. As noted by some empirical studies (Lancaster, 1979; Nickell, 1979), the probability of the unemployed finding a job is reduced rather than increased by the increasing duration of the episode of unemployment. The assumption is that the longer a worker is unemployed, the more difficult is for him to find work, since, on the supply side, he will suffer a loss of his abilities to work due to unemployment, and on the demand side employers tend to prefer short-term unemployed because they consider the duration of unemployment a sign of their lack of motivation and discipline at work. According to Heckman (1984) there are four different types of dependency on unemployment (scarring effects):

1. the probability that an unemployed person will remain unemployed is greater than that of an employed person becoming unemployed. There are several factors that explain this scarring effect, but high transaction costs are one of the most important;
2. the probability of finding a job is lower for an unemployed person who has experienced a high number of episodes of unemployment. Employers consider large numbers of episodes of unemployment low motivation to work;
3. the duration of the last episode of unemployment leads to an effect of human capital loss.
4. the total duration of all episodes of unemployment involves a loss of work experience.

In this way, in addition to an immediate effect on earnings and on the probability of finding a job, unemployment may also cause a permanent scarring effect, i.e. a long-term negative effect on income and the type of employment. This problem would be particularly severe and worrying in the case of young people, given their longer life perspective in the labour market (Ellwood, 1982; Ruhm, 1991). Firms tend to introduce the lowest wage for those who have recently entered the labour market. In a market where the employer is free to pay each worker according to his level of productivity, the wages of entry should be lower in the case of young individuals, to take account of their work experience gap and, therefore, more generally of human capital with respect to adults. If the human capital of the younger is lower than that of the older, they are also

less productive and to convince employers to hire them they should be allowed to pay them lower wages. As shown previously, for a firm to hire a young individual is an opportunity for exchange: the young worker provides the firm with his work activity and the firm provides the young worker with the training (general or specific) that he needs to increase his human capital (Becker, 1962). In other words, the firm provides a young worker with non-monetary remuneration. However the most obvious obstacle to realization of an entry wage is the presence of the legal minimum wage, which is a constraint to the free determination of wages by firms. The minimum wage legislation requires firms to pay a wage higher than what the firm would pay given the production level of the worker.

The search to increase work experience through temporary work shows that if, on the one hand, temporary work increases the probability of finding permanent work, on the other hand, when there are strong rigidities in the labour market, there is a real risk that young people will fall into a trap that relegates them, usually for a very long period, in working circuits characterized by temporary and/or part-time work, or in long-term unemployment with serious negative social effects.

The policy implications are clear: to decrease long-term unemployment and precarious work, it is not sufficient to increase flows in the labour market, but it is necessary to affect the low motivation and professional abilities of the weaker individuals with targeted measures.

If it is true that the longer the duration of unemployment, the higher the probability of remaining unemployed, then it should be an objective of economic policy to attempt to interrupt long unemployment episodes. The flexibility of the labour market may reduce the duration of unemployment in order to equal motivation and ability. However, active policy should be added for employment, from the point of view of training in favour of long term unemployed and precarious workers. Moreover, as argued in the European Employment Strategy, training must be conducted on a large scale to be effective and produce additional jobs rather than replacing existing ones.

The causes and consequences of youth unemployment are not yet well defined by economists because youth unemployment is sometimes in conflict with economic

analysis.[4] Before explaining the scarring effects, it is useful to start by considering the implications of youth unemployment, according to some of the most common labour theories. Human capital is probably the main theory that explains the long term effects of youth unemployment. Human investments for job careers are useful to explain the concave pattern of aggregate age-earning profiles. This fact implicitly imposes heavy costs on the unlucky young individual who misses an initial investment opportunity. If no investment is made by the individual during the period without a job, the whole profile is shifted back. The net present value of earnings flows must also be discounted in wasted time, even if in this case retirement is delayed.[5] Moreover, the period of discouragement can generate weak attachment to the labour force. This involves a vicious cycle of unemployment. Young individuals face only a limited number of initial jobs that lead to better jobs. Workers risk being permanently situated on an inferior scale if they lack good initial jobs.

Another issue concerns the serious information problems of young individuals on the labour market. Workers and employers are involved in a process of matching skills and jobs, but there is very little information available on both sides. Employers rely on the workers' past experiences to make hiring decisions because they need to categorize people according to their abilities. These problems may be greater in times of recession because if employers do not adapt their expectations to those of the young labour force, those who enter such a labour market will suffer. As a matter of fact, early unemployment has considerable long-term effects (Becker and Hills, 1978; Stevenson, 1978).[6]

Summarizing, there are two major economic policy options in favour of young people:

- to increase labour market flexibility;
- to reform the education and training system;

[4] For example according to the theory of human capital expounded in Chapter 2, an investment in education should provide higher earnings and a higher level of employment in the future, but this does not always occur.

[5] In Section 2.1 it has been shown how the net present value is computed in the case of investment in human capital (direct costs and opportunity costs) resulting in higher earnings in the future. In this case, however, unemployment does not provide more earnings in the future, because there is no investment in human capital in the period of unemployment.

[6] All the empirical evidence of scarring effects is demonstrated in Chapter 4.

Both options generally have the objective to increase the competitiveness of the economy. However, the choice of flexibility tends to increase price competitiveness trying to reduce labour costs for firms, whereas the choice to reform education and training aims to increase competitiveness through growth in labour productivity and the quality of production and employment. A workforce more and better educated, able to have faster school-work transitions, can accumulate human capital and contribute to a greater extent to the overall growth of the country. In this way, a flexible education system must be created that combines on-the-job training with an efficient system of general education, together with a flexible labour market (well-regulated), thus constituting an ideal policy mix. To be truly efficient in promoting the accumulation of human capital of young individuals, labour market flexibility must provide rules to discourage the precariousness of employment.
The European Commission identifies the rules on flexibility as a fundamental objective.

It must be considered, however, that in the last 15-20 years various measures of labour market flexibility have been adopted in several European countries, until now without achieving appreciable results in terms of reduction of their respective unemployment rates. Therefore, to produce good results policies that aim at a flexible labour market have to be well regulated. Only in the United Kingdom have these policies provided appreciable significant results. This is presumably due to the fact that this country has adopted a process of labour market reform more radical than other European countries: for example there has been the introduction of legislation that restricts the right to strike, decentralizes wage bargaining and liberalizes the possibility to dismiss. Moreover, this country has abolished the organizations responsible for setting minimum wages, and it has reduced the period for concession of unemployment benefits, making the rules on granting social benefits to the community more stringent. At the same time, however, some institutions have been established to facilitate job seeking and, more generally, to facilitate matching labour demand with supply. This result was achieved without substantially affecting the main benefits of the welfare state beyond the above mentioned measures, and no tax reform was introduced that somehow compensates for the effects of income redistribution resulting from the operations of the flexible labour market.

The British experience constitutes a very important test for the rest of Europe. It agrees with those who argue that marginal labour market reforms do not achieve any appreciable effect on employment (Lindbeck, 1996). The labour market reforms introduced in European countries, except for Great Britain, were only marginal and they did not substantially affect the high goodness of welfare systems that in some countries have distorted their original social purpose and have also become unsustainable in financial terms. Consequently, a portion of unemployment in these countries is simply masked by very generous social assistance schemes, which makes it preferable in some cases for some workers to be unemployed, instead of actively looking for a job (Coe and Snower, 1997).

3.3.1. Education and unemployment

As explained in the previous subsection, one of the economic policy options for young people is to create a workforce more and better educated in order to reduce unemployment. Education has a positive influence on the stability of the state of employment, meaning that workers with higher levels of human capital have less risk of falling into unemployment. In addition to the higher earnings and greater upward mobility, both in terms of qualifications and wages, the benefits of higher education also provide a higher probability to maintain a job longer and make it easier to find another job once the first is abandoned.

The characteristics and possible motivations of the links between education and unemployment are highlighted in a study by Mincer (1991). In general, the likelihood or incidence of unemployment depends on the probability of being separated from a previous job and the probability of becoming unemployed after the separation. In particular this fact can be described as follows:

$$q\,(U) = q\,(s)\; q\,(U \mid S), \tag{19}$$

where q (U) is the incidence of unemployment, and the two factors q (S) and q (U | S), respectively indicate the probability of being separated from a previous job and the one

of being unemployed after the separation, i.e. the conditional probability of unemployment given a separation. Moreover, it also seems that for workers with a better education the frequency of resignation is slightly higher than that of dismissals.

In turn, these phenomena depend largely on the presence in the firm of specific human capital created by on-the-job training in the workplace. The investment made in forms of on-the-job training pushes the employer to keep the trained employee in the firm as long as possible, together with other factors such as the presence of high fixed selection costs, recruitment, etc. Taken together, these factors could provide a possible explanation for the lower risk of unemployment which affects workers with higher levels of education.

With specific regard to the female workforce, these results should be modified in the following terms. First we should consider a structural difference compared to the male labour force as shown in section 2.3. Moreover, the timing of women's work in view of their life cycle is more complicated due to family needs and child care. Despite this, even for the female labour force there is a negative correlation between education and unemployment, in the sense that women with higher levels of education bear fewer risks of unemployment. However, in the case of men the connection between education and unemployment goes mainly through training in the workplace, whereas for women the most important factors are participation rates. In other words, men with higher education levels are less affected by unemployment than the rest of the male labour force thanks to the specific human capital acquired in the workplace, while women with more education have a similar condition to that of other women thanks to their higher participation in the labour market. To recover investment in human capital requires long working hours, with the result that women with higher education express higher participation rates than women with lower education. Moreover, for the female workforce training activities in the workplace are fewer than those provided for men.

4. Empirical analysis

This section contains an empirical analysis of the concepts described in the first two chapters. In particular, first there are graphic representations of the various concepts described in the human capital theory. Successively, the different aspects of the Mincer equation are estimated through different regressions by the method of ordinary least squares (OLS). Finally, there are estimates for the evaluation of scarring effects with the same method used for the Mincer equation. All these results are implemented using Stata (Statistics / Data Analysis) software for the following four countries: France, Germany, Italy and Spain.

4.1. Data description

The datasets used are provided by "SHARE - Survey of Health, Ageing and Retirement in Europe". In particular SHARE is a survey of Health, Ageing and Retirement in Europe that consists of a multidisciplinary and multi database of individual data on health, socio-economic status and family relationships of people over fifty.

The data obtained by SHARE have been collected from three interviews carried out in three different years: 2004, 2006 and 2008. These interviews were implemented through a number of questionnaires structured into several parts for each of the three surveys, according to the different areas of interest. The collected data include health variables, psychological variables, economic variables and social interaction variables.

Eleven European countries participated in the first survey of 2004 (wave 1): Austria, Belgium, Denmark, France, Germany, Greece, the Netherlands, Italy, Spain, Sweden, and Switzerland. Other countries were included in 2006 (wave 2): Czech Republic, Ireland, Israel and Poland, which participated in the second survey in 2007-2008. Finally, the third survey (wave 3), called SHARELIFE, took place in 2008-09 and collected retrospective information on the entire life cycle in sixteen countries with the addition of Slovenia. SHARE is coordinated by the Munich Center for the Economics of Aging (MEA), Max-Planck-Institute for Social Law and Social Policy. The SHARE

data collection was financed mainly by the European Commission, while other funds came from various institutions in each participating country.
For the purposes of this thesis the datasets obtained by SHARE are structured as follows: of the sixteen countries included in SHARE France, Germany Italy and Spain have been selected. The regressions of the Mincer equation are carried out with the use of the dataset from wave2. The reason is that this dataset contains cross section data. In this way the characteristics of a number of individuals at a given moment can be observed.

The analysis of scarring effects has also been implemented using the dataset of wave2 with a further specification with the dataset of SHARELIFE in appendix B.
The reason is that SHARELIFE makes it possible to check the entire life cycle of an individual and this is especially useful in determining the scarring effects, because it allows us to verify whether there were periods of unemployment or recession in the past of an individual. In this way the dataset of SHARELIFE provides the use of longitudinal data. Longitudinal data refer to a structure in which there is generally a large number of individuals and there are repeated observations relating to the same individual for a certain number of periods. However the data of the two dataset are used mainly as a unique cross section because the data of the entire wage profile of the individuals are not available.
In particular, Table 1 shows how the dataset of wave 2 is composed:

Table 1: Composition of the wave 2 dataset

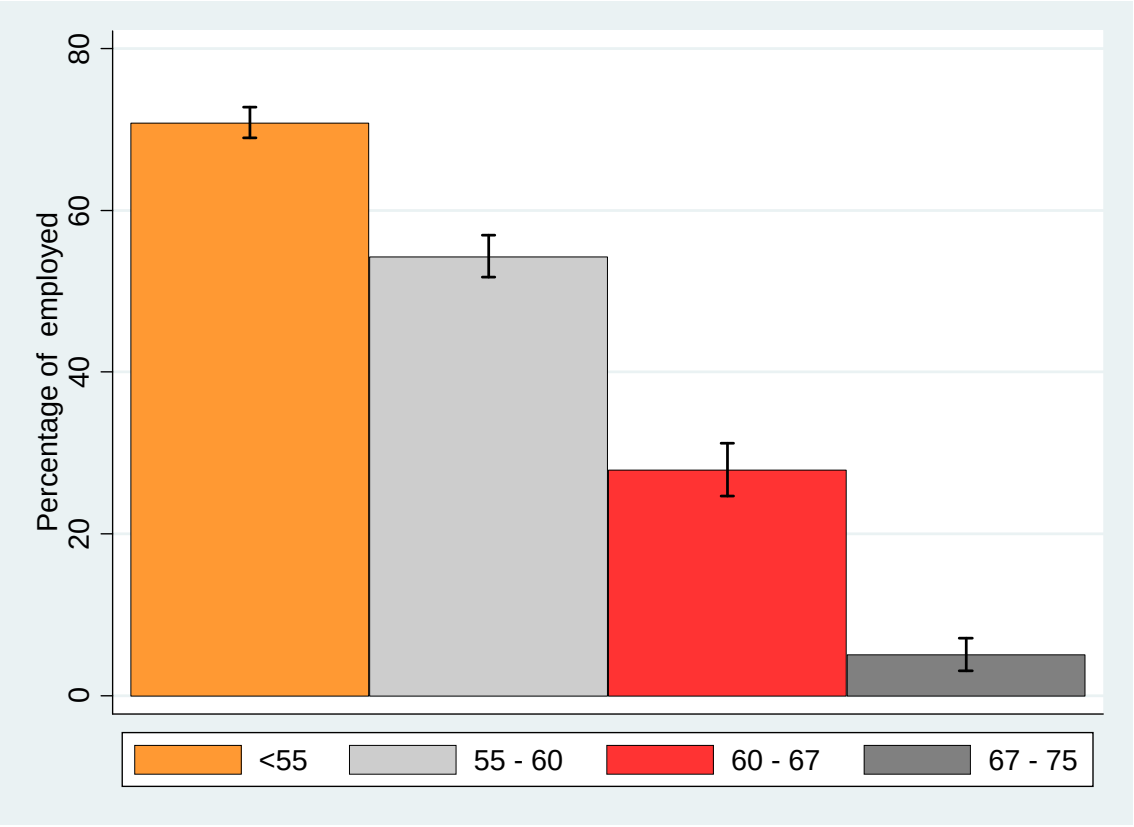
	Total	Male	Female	Under50	50-60	65-74	75+
France	2968	1273	1695	122	1200	1036	610
Germany	2568	1184	1384	52	948	1150	418
Italy	2983	1345	1638	78	1050	1331	524
Spain	2228	1003	1225	57	745	903	523
Total	*10747*	*4805*	*5942*	*309*	*3943*	*4420*	*2075*

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

4.1.1. Empirical evidence of the implications of the human capital theory

This subsection provides a descriptive analysis of the main implications of the human capital theory. In particular, the following descriptive analysis highlights the levels of employment and earnings according to the statements of the human capital theory described in chapter 2, providing graphical representations for the four countries considered[7].

Figure 4. Percentage of total employment



Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

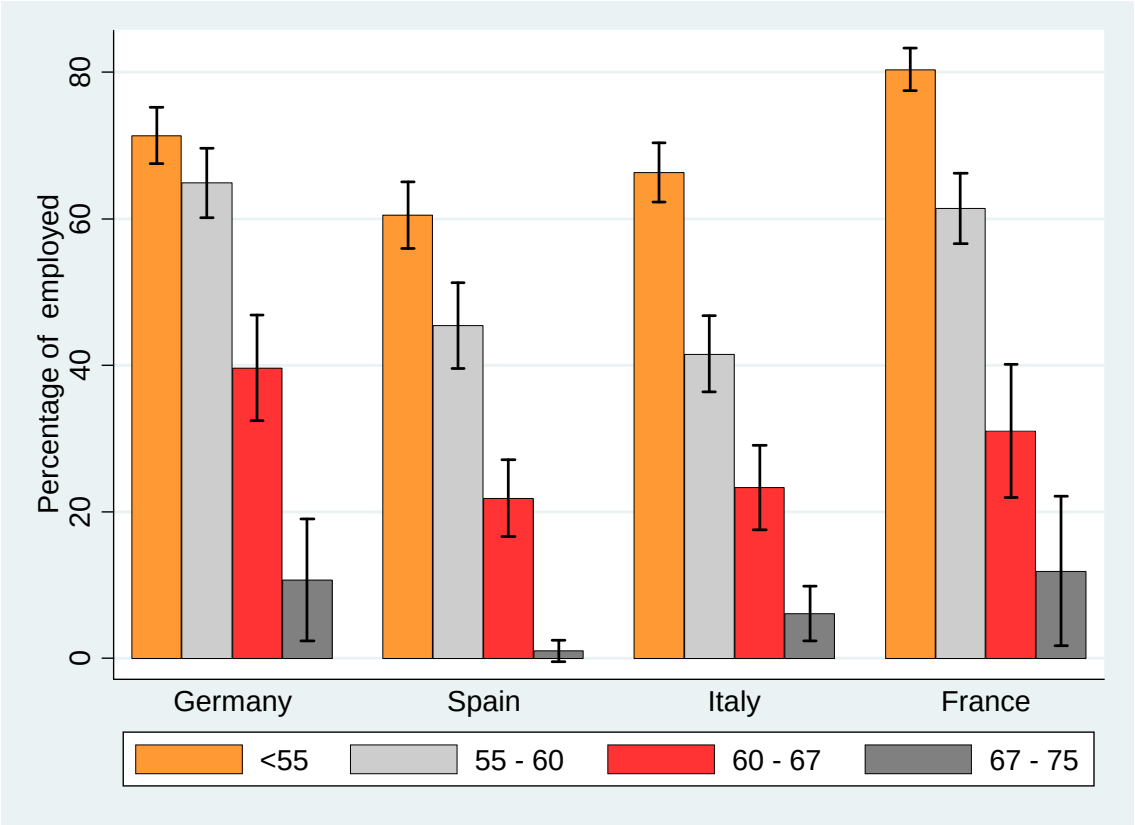
Figure 4 has been built starting from the variable *ep005* obtained from the questionnaires of wave2. This variable describes the individual's current job situation.

[7] All graphs show error bars and confidence intervals highlighting where the estimates are more precise.

However, to exclude retired individuals from this variable, another variable has been created called *workforce*, which does not include the retired. In this way it was possible to obtain the percentage of employed in the labour force instead of the entire population of the sample. The age classes are particularly high; the lowest age group includes individuals who are under 55, which is due to the fact that the dataset includes individuals who are principally over fifty, as can be observed in Table 1.

The result obtained, however, reflects the fact that the level of employment tends to decrease when the individual is closer to retirement age. Moreover, it has been observed in previous studies conducted on SHARE data that people who enjoy their work retire later and this explains why the graph shows some workers who work after 67, which is the retirement age in most of the countries considered.

Figure 5. Percentage of total employment by country

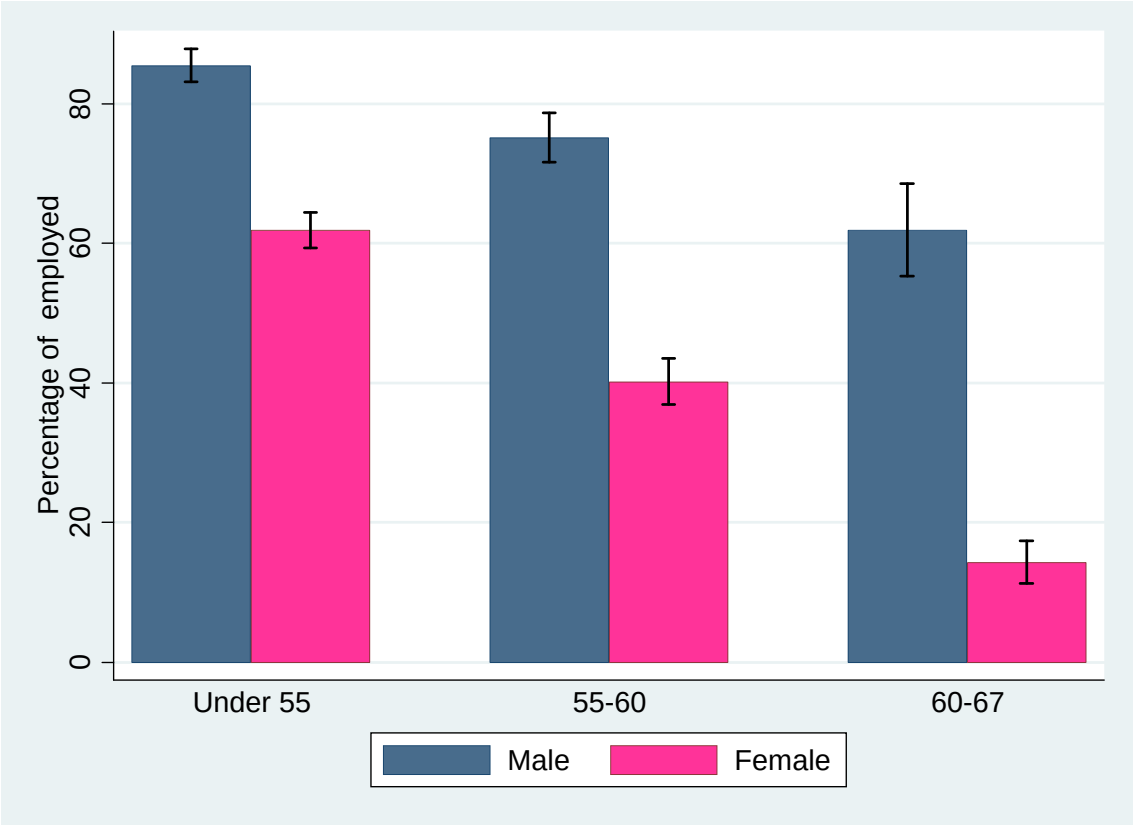


Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

The graph in Figure 5 confirms the result obtained in Figure 4 for the four countries, i.e. that the level of employment decreases when the individual is closer to retirement age.
In particular, the graph in Figure 4 shows that on average for the four countries the level of employment for individuals under 55 years old is 70% and for those who are between 55 and 60 years old it is 54%. This means that looking at the graph in Figure 5, Germany and France are above the average with respectively 71% and 80% of employed for individuals under 55 years old; and 65% and 61% for those who are between 55-60 years old.
Conversely Italy and Spain have a percentage of employed below the average, as shown in Figure 5. In particular, in Italy the level of employment for individuals under 55 years old is 66%, whereas for those between 55-60 years old it is 42%. In Italy, therefore, the difference with the average level of employment is -4% for individuals who are less than 55 years old, and -12% for those who are between 55-65 years old. In Spain the situation is even more worrying and it is possible to observe that in this country there is the lowest percentage of employment in the labour force, so the sample reflects the high unemployment rate existing in Spain. In this country in particular, the employment rate for individuals under 55 years old is 60% whereas for those between 55-60 years old it is 45%. This means that in Spain the difference in the average level of employment is -10% for individuals younger than 55 and -9% for those between 55-60 years old.

Making the analysis even more detailed, it is possible to distinguish between gender differences to verify the concepts included in subsection 2.3, as shown in Figure 6. The graph shows the level of total employment for the four countries distinguishing the level of employment by gender, excluding the retired as in previous graphs. In particular, Figure 6 has been constructed with the purpose of obtaining the percentage of males employed on the total of the males in the sample, and the total of females employed on the total of the females in the sample by age group.

Figure 6. Percentage of total employment by gender

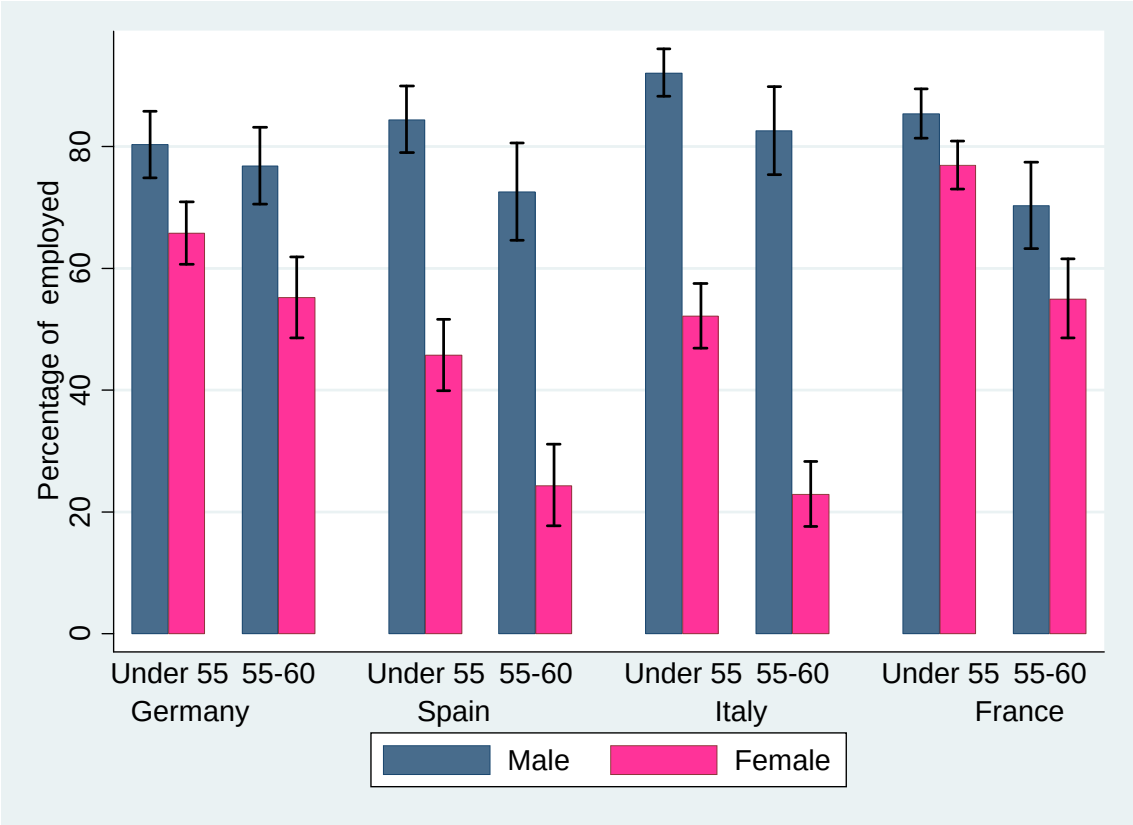


Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

The result in Figure 6 shows that the trend in the average level of employment among the four countries is the same for men and women, i.e. that employment tends to decrease when individuals are closer to retirement age. What can be seen is that there are fewer women employed in each age group than men. The reasons may be associated with the factors described in subsection 2.3.

The graph in Figure 7 allows us to obtain a distinction between men and women by country in order to place greater emphasis on the differences in each of them.

Figure 7. Percentage of total employment by gender and country



Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

What emerges from Figure7 is that in all countries more men are employed than women for each age group, but to a different extent. For example, looking at this graph, the level of employment in Italy is above the average of the four countries with respect to men, but it is below the average for women. The dataset used provides a level of male employment in Italy of 95%, but for women it is 51% for individuals under 55 years old. This means that in Italy the gender gap for the employment rate is 44% for this age group.

Also in Spain there is a similar trend to that of Italy. In Spain the male employment rate is 92%, above the average of the four countries considered, but the percentage of female employed is 49%. This means that in this country the gender gap for the employment rate is 43% for individuals under 55 years old.

This explains why in the graph in Figure 5 the percentage of employment is greater for Germany and France although the percentage of men employed is greater in Italy.

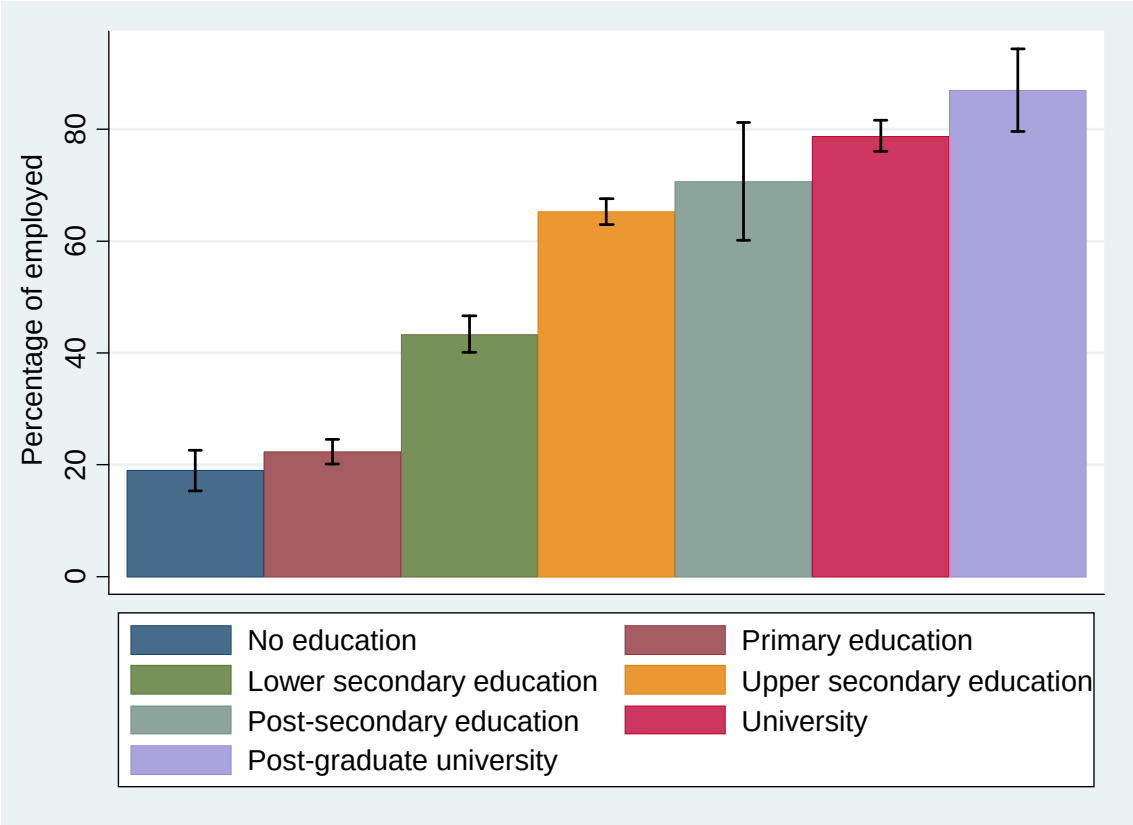
Since the percentage of women employed is greater in these two countries, the result that emerges is that the total percentage of people employed in Germany and France is higher than in Italy and Spain, as shown previously. As a matter of fact, although there is a gender gap in Germany and France, this is lower than the other two countries, in particular 16% in Germany and 9% in France for individuals under 55 years old. The differences are even greater considering the age group between 55-60 years old. Even in this case, the countries with the largest gender gap are Italy and Spain.
Observing Figure 7, therefore, Italy is the country with the largest gender difference due to the high percentage of men employed.

To make a comparison of employment levels at European level, the indicators proposed by the European Commission have been used, especially taking into consideration the employment rates of people with different qualifications (Isced levels 0-2, 3-4 and 5-6). In particular, the education levels for the variable *isced* are the following:

- primary education: no education, primary and lower secondary education are catalogued in levels 0-2 (Isced 1997);
- secondary and post-secondary education: upper secondary and post-secondary education are catalogued in levels 3-4 (Isced 1997);
- tertiary education: university is catalogued in levels 5-6 (Isced 1997).

In this way it was possible to create the graph in Figure 8, which provides the average rates of employment for the four countries according to education levels.

Figure 8: Percentage of total employment by education levels

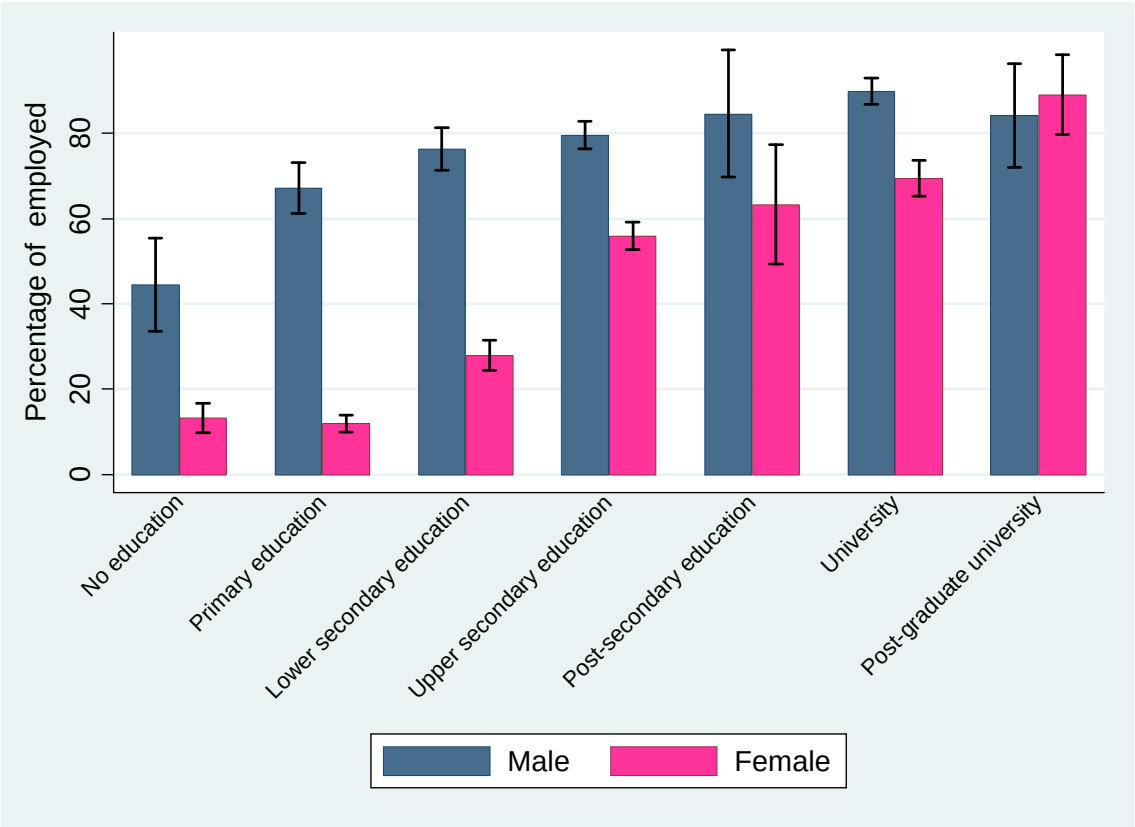


Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

It can be seen that Figure 8 confirms the theory of human capital as demonstrated in subsection 2.1, i.e. that there are increases in employment levels with increasing education levels. As a matter of fact the employment rates of individuals with university degrees are much higher than the employment rates of people with lower qualifications, confirming the importance of investment in education in terms of opportunities in the labour market.

Figure 9 represents an extension of this analysis to the gender differences for all the countries, highlighting the percentage of men employed on the total of men by education levels and the percentage of women employed on the total of women by education levels.

Figure 9: Percentage of total employment by education levels and gender

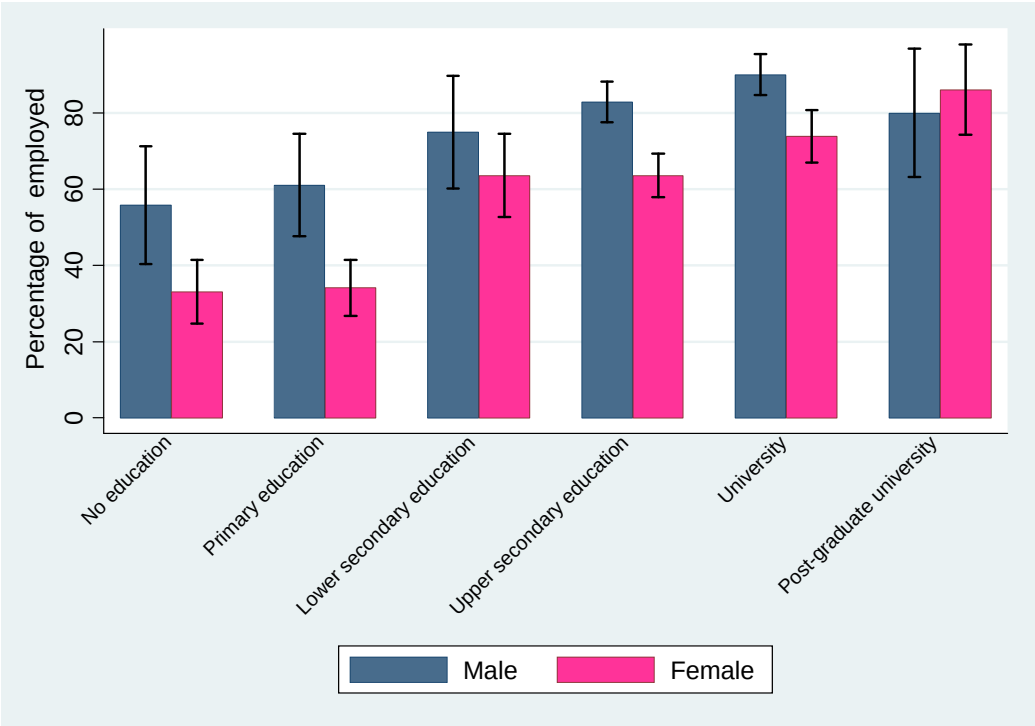


Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

As illustrated in Figure 9, also for women the employment rate increases with an increase in education levels, but to a lesser extent than for men. However, in the presence of high levels of education there is a significant reduction in the gender gap. For isced 5 (university), the gender gap is only 22% compared to the differences resulting from lower educational qualifications and it is null for isced 6 (post-graduate university).

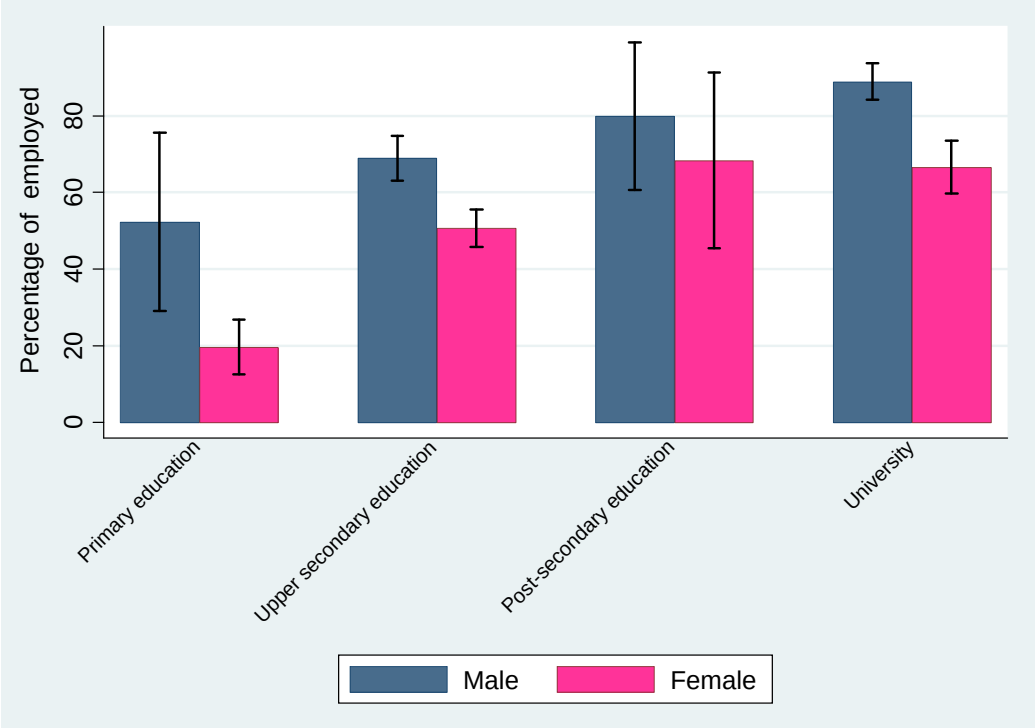
The following graphs represent the same analysis for each country, so that it is possible to verify in each of them if people who possess higher qualifications have higher employment rates according to the human capital theory.

Figure 10: Percentage of total employment by education levels and gender in France



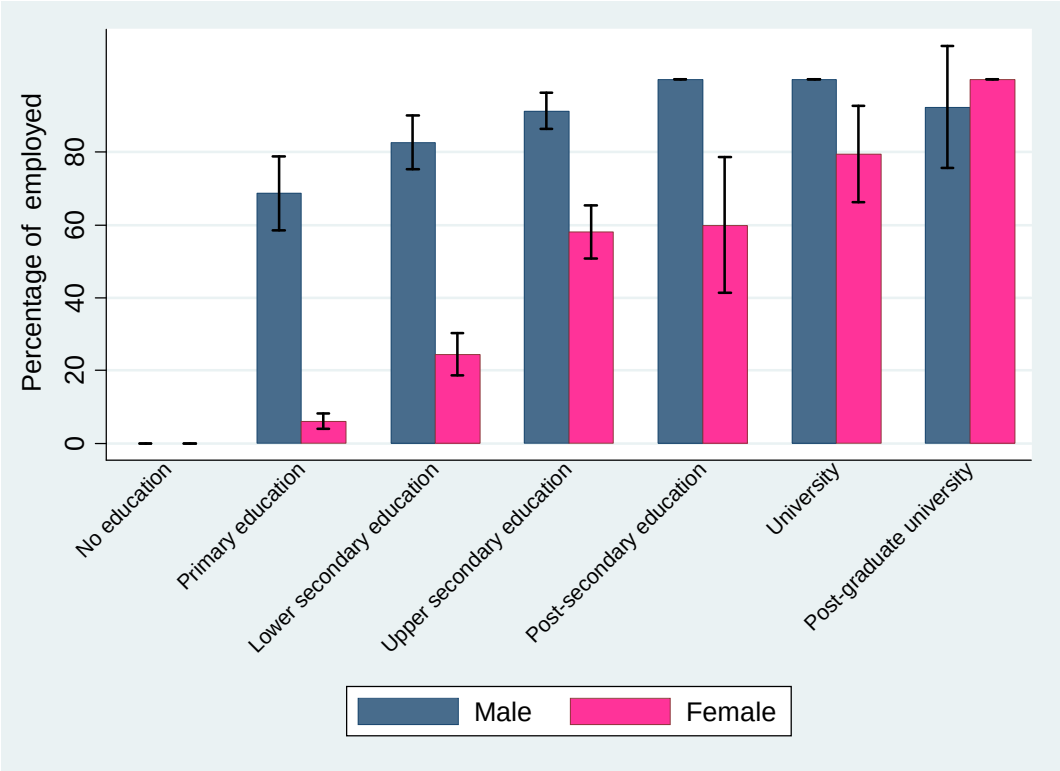
Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

Figure 11: Percentage of total employment by education levels and gender in Germany



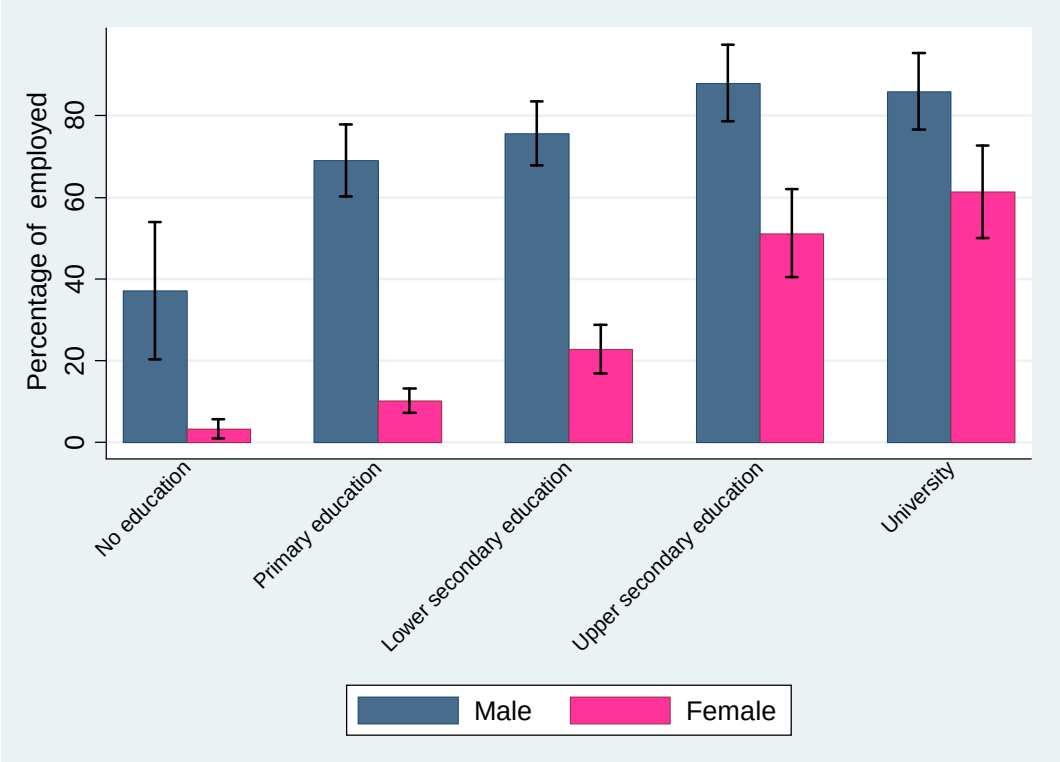
Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

Figure 12: Percentage of total employment by education levels and gender in Italy



Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author

Figure 13: Percentage of total employment by education levels and gender in Spain



Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author

The school system in different countries has differences, as can be seen in the graphs, but the isced indicators proposed by the European Commission make it possible to equate levels of education among different countries. For example, Italy is the country with the largest number of classifications. In fact isced has seven classifications (0-6), but other countries such as Spain have only five classifications. However it is possible to assign the same type of education in each country to each level of isced, as shown in Figures 10 and following.

Moreover it is possible to note in Figure 12 that Italy does not have employed people at isced 0 corresponding to "no education", but this is due to the fact that the dataset available for this category has 35 observations in all between men and women and of these 35 observations the majority is retired or not employed. The same happens in Figure 11 for Germany at isced 0 "no education". In this category there are no employed, but the reasons are attributable to those explained for Italy. The main reason is due to the fact that the category "no education" is composed mainly of older people who are already retired and therefore outside the workforce. However, young people are hardly included in that category because compared to the past the years of compulsory schooling have increased significantly in all the countries considered.

Looking at gender differences in the graphs of the four countries it can be noted that all four observe the average trend shown in Figure 9. As a matter of fact the percentage of men and women employed increases with increases of education levels and the differences between them decrease with higher levels of education. In France and Italy the gender gap is null for university degrees and in Germany and Spain it is reduced compared to the differences resulting from lower educational qualifications. Only in Germany the gender gap increases slightly for university graduates, but simply because the number of men employed for this level of education increases significantly.

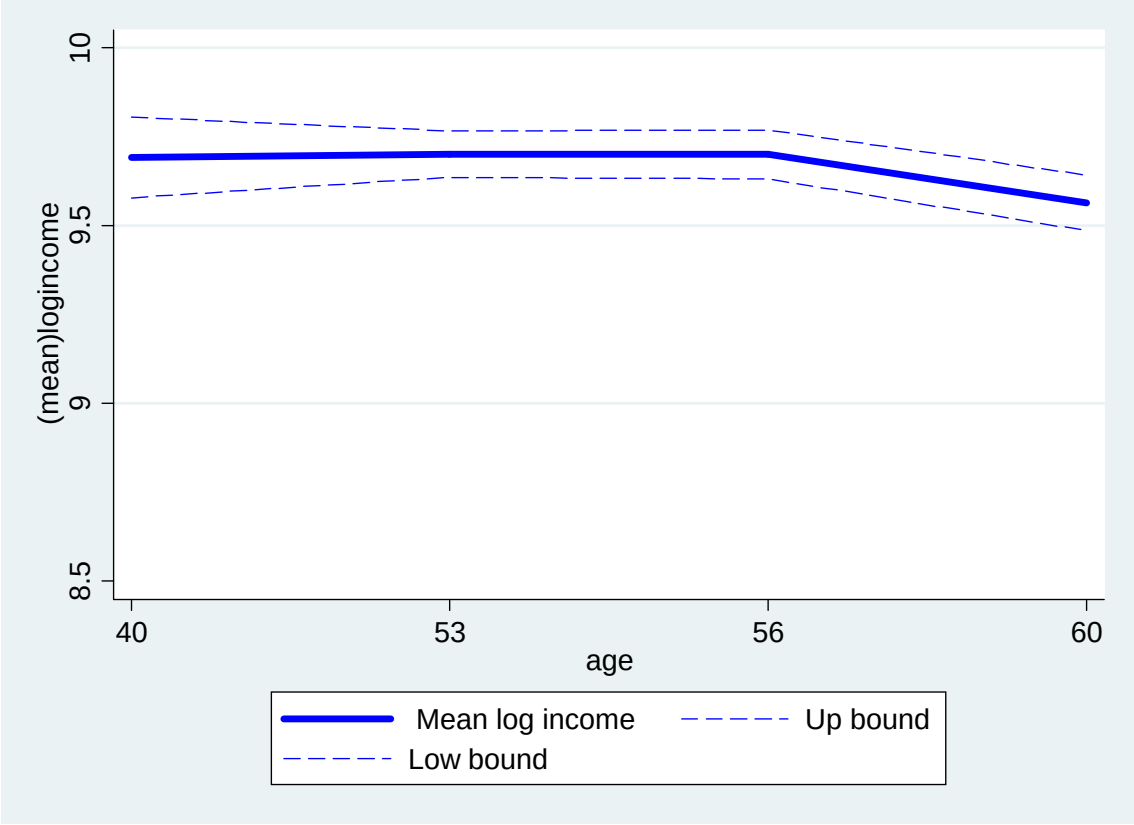
Spain is the country with the highest gender gap for the highest isced classifications. The gender gap for this level of education is 25% with 86% of men and 61% of women employed, as illustrated in Figure 13. However the gender gap in Spain as explained above and illustrated in Figure 7 is 38% for individuals younger than 55. This means that in this country although there is still a gender gap for high levels of education, this is 13% less than the overall difference independently of qualifications. The most significant aspect, however, seems to be represented by Italy. Figure 7, as previously

illustrated and explained, shows that the gender gap for individuals under 55 years old is 42% and this makes Italy the country with the largest gender gap. However, in Figure 12 the gender gap for isced 5 "university" is 21%, but for isced 6 (post-graduate university) it appears that women are more employed than men. This means that the gender gap in Figure 7 comes from individuals who have low levels of education.

The previous figures provide a graphical representation of employment levels, first by age groups and then by education level for all the countries considered, focusing on the gender gap in each of them. The next graphs provide a descriptive analysis that shows income levels for the individuals of each country with gender differences; finally graphs will be provided showing the level of individuals' incomes by years of work experience.

In particular Figure 14 represents the mean income considered with increases in age for all the countries.

Figure 14: Level of income



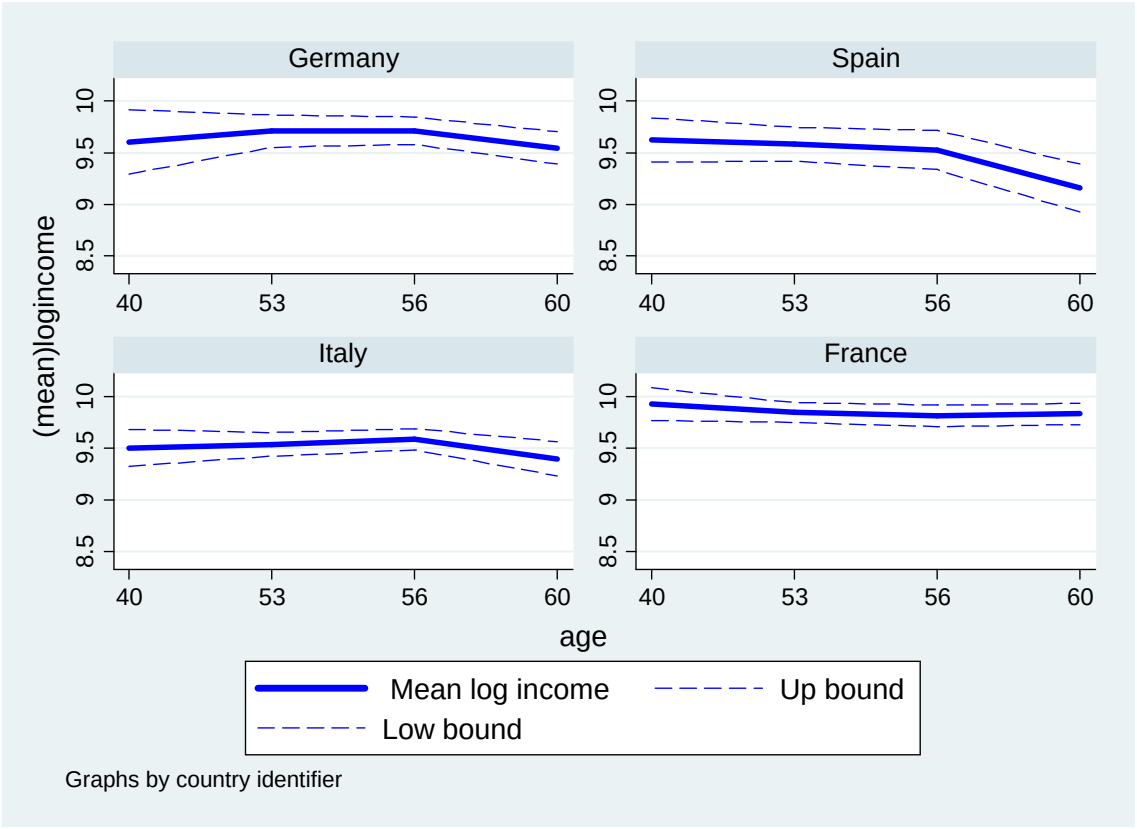
Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author

The graph is created by including in the age group individuals between 40 and 60 years old. This is due to the fact that in the dataset there are few observations for those under 40. The income variable used is *ydipv* representing the annual net income from employment. In this way retirement income is not included because the variable *ydipv* considers only income from employment. In particular, the logarithm of income was used to make the data more normalized, so the variable used in the graph is *logincome*.

As can be seen in Figure 14 the mean income of individuals tends to increase with increasing age and then it decreases slightly after 56 when the individual is closer to retirement age. This is due to the fact that, as explained in section 2.2, income is not only positively correlated with the level of education, but a fundamental role is also provided by work experience which increases with age and can be achieved through training (general and/or specific) provided by the firm.

The following figures show first the differences among countries and then among countries and gender. In particular Figure 15 shows the mean income in each country.

Figure 15: Level of income by country

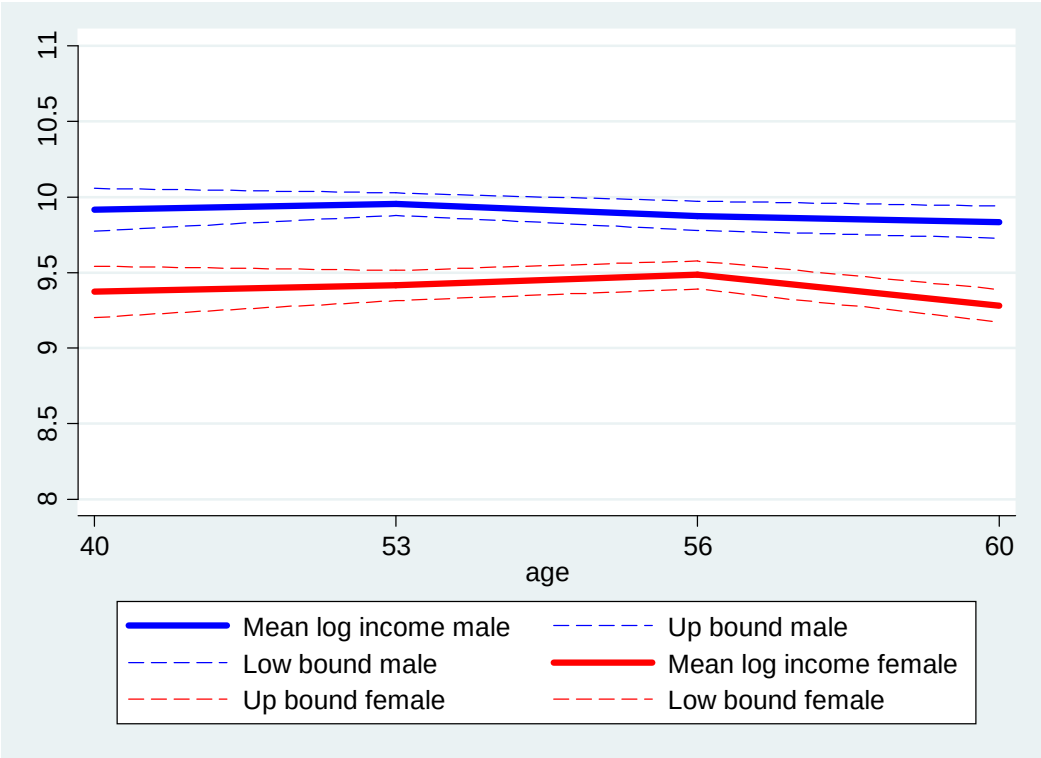


Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author

The procedure used to obtain the graph is the same as that for Figure 14 but in this case the mean income for each country has been considered in order to show the differences in each of them. The trend in income in the four countries seems to be more or less similar, but there are some differences that cannot be neglected. The main similarities seem to be in Italy and Germany, where income increases with age and decreases slightly when the individual is close to retirement age. In Germany however, incomes appear to be on average greater than in Italy, although the trend of these two countries is similar. Conversely, in France and Spain the trend seems to be different because income does not increase with increasing age. However, France appears to be the country with a greater mean income compared to the other countries. Germany is the country with the second highest level of income, while Italy and Spain have more or less the same level of income, but with a different trend. In Spain income seems already high for the younger age groups, but the mean income seems comparable to that of Italy during working life.

Figure 16 shows the overall mean income for the four countries reporting a gender gap.

Figure 16: level of income by gender

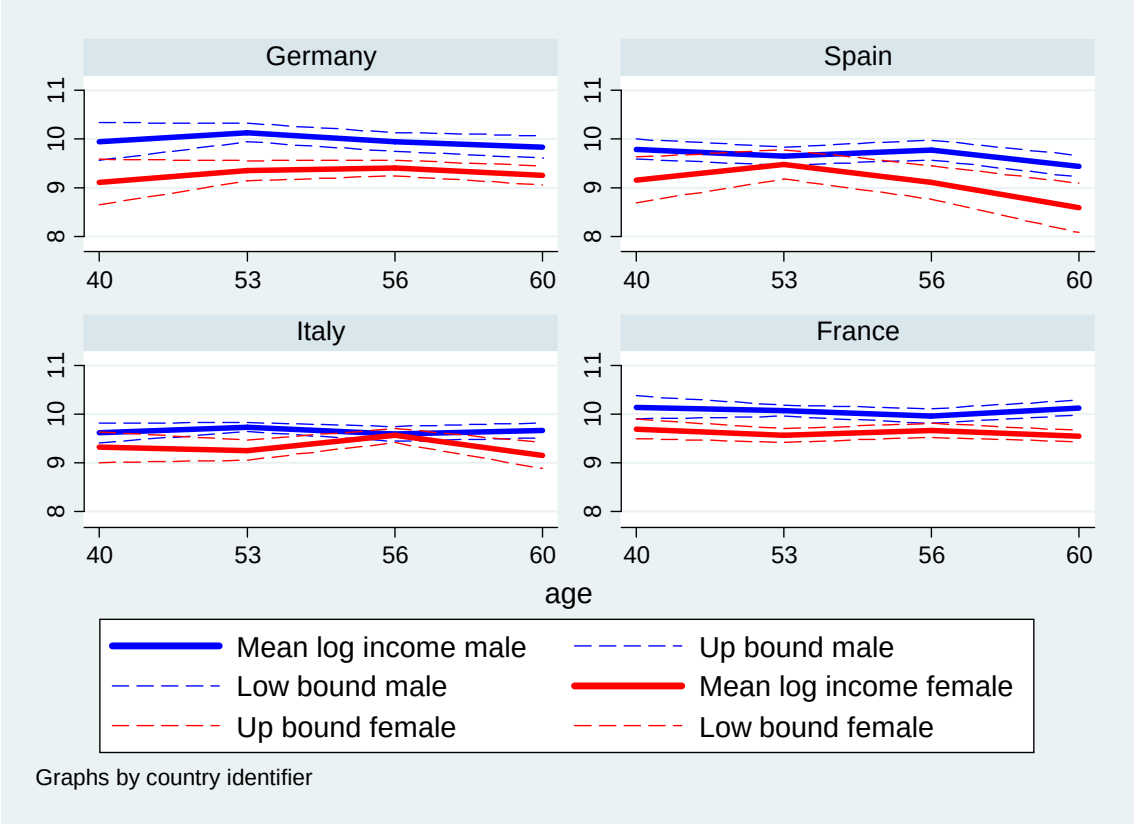


Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author

As discussed in subsection 2.3.1, the gender gap does not exist only for employment levels[8], but also for wage discrimination. In Figure 16 the mean income of men (blue line) is greater than the overall mean income shown in Figure 14, whereas the mean income of women is below this. However, the trend in Figure 16 over the twenty years of analysis is similar to that of Figure 14, and in fact the level of income for men and women increases with age and tends to decrease when the individual is closer to retirement age, although it occurs in different measures for men and women.[9]

Figure 17 represents the gender differences in each country. The figure shows that in each country the mean income of men is greater than that of women.

Figure 17: Level of income by gender and country



Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

[8] The gender gap for the employment levels is shown from Figure 9 to Figure 13.
[9] As described in Figure 3, the difference in the income of men and women may be due to men's actual greater experience or to an unexplained difference (residual), depending on the circumstances.

It can be seen that Italy is not the country with the highest gender gap with regard to income levels, whereas it is the country with the largest gender gap with regard to the level of employment, as shown in Figure 7. This means that the gender gap in each country changes depending on whether there is a difference in the level of employment or in the level of income for individuals. Moreover, it emerges that the mean incomes of men and women in Italy are equivalent at the age of 56 and then they diverge again until retirement age.

Comparing Figure 17 with Figure 15, in which income was reported for each country without gender gap, it can be seen that German men are those who earn more. In this country, therefore, the younger age groups seem to have the largest gender gap. In Figure 15 the level of income in Germany is lower than that of France, but this due to the fact that German women earn less than French women. If it is true that the gender gap is greater in Germany for the younger age classes, it is also true that this difference tends to decrease with increasing age. This does not occur in Spain, where according to the graph, the mean income of men and women tends to be equal at about 53, whereas the differences increase when the individual is closer to retirement age.

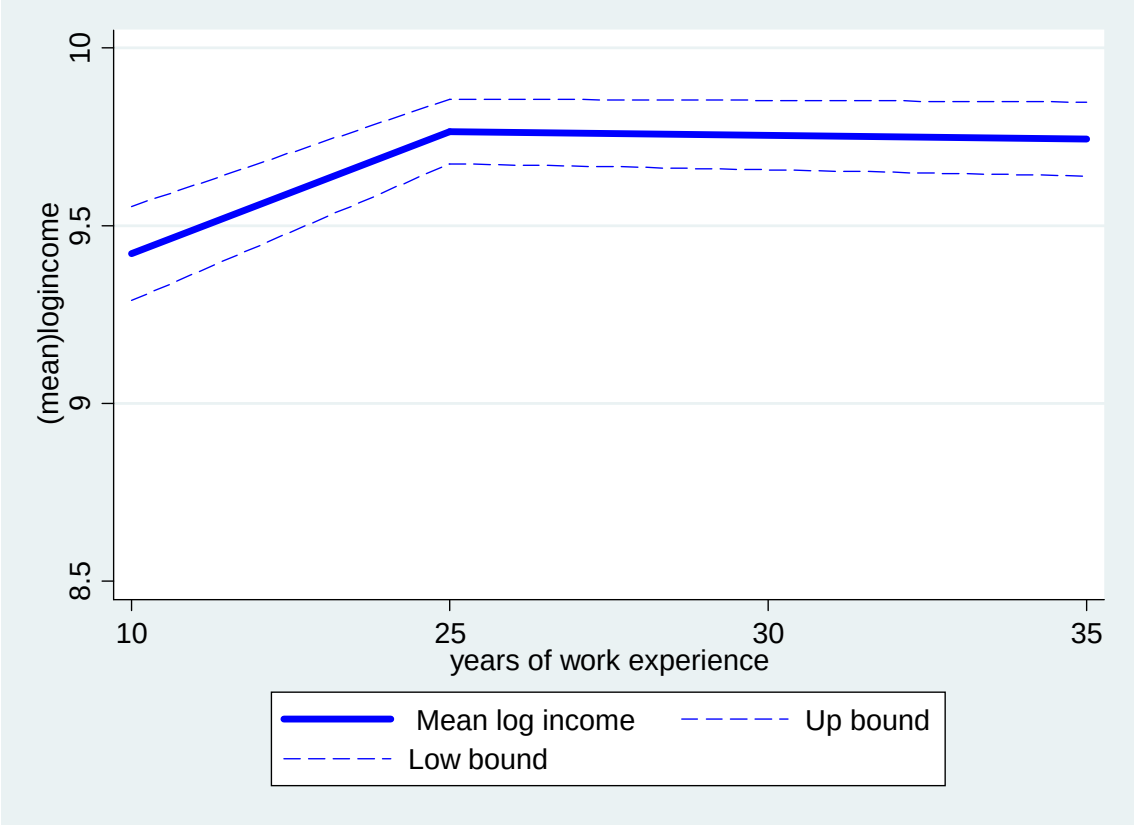
To carry out a more detailed analysis, the following graphs show the levels of income for years of work experience. As in the previous graphs, the analysis is performed first globally and then by country and gender differences. Analyzing the income trend by years of work experience, it is possible to check the influence of unemployment spells, seeing how these curves differ from those in which age was used. For example, an individual who had several unemployment spells may have less work experience than a younger individual.

The years of work experience are computed starting from variable *y_work_w2*. This variable represents potential experience. Potential experience (*y_work_w2*) is obtained by excluding unemployment spells from the worker's working life, considering actual jobs (startjob_ `i '), years of education and the age at which he started school.

$$y\_work\_w2 = startjob\_`i' - yedu - age\_schooling - yrbirth \quad (20)$$

Figure 18 shows the total income for the four countries with increases in the years of work experience. The years of work considered are between 10 and 35 (the years of work assumed close to retirement age).

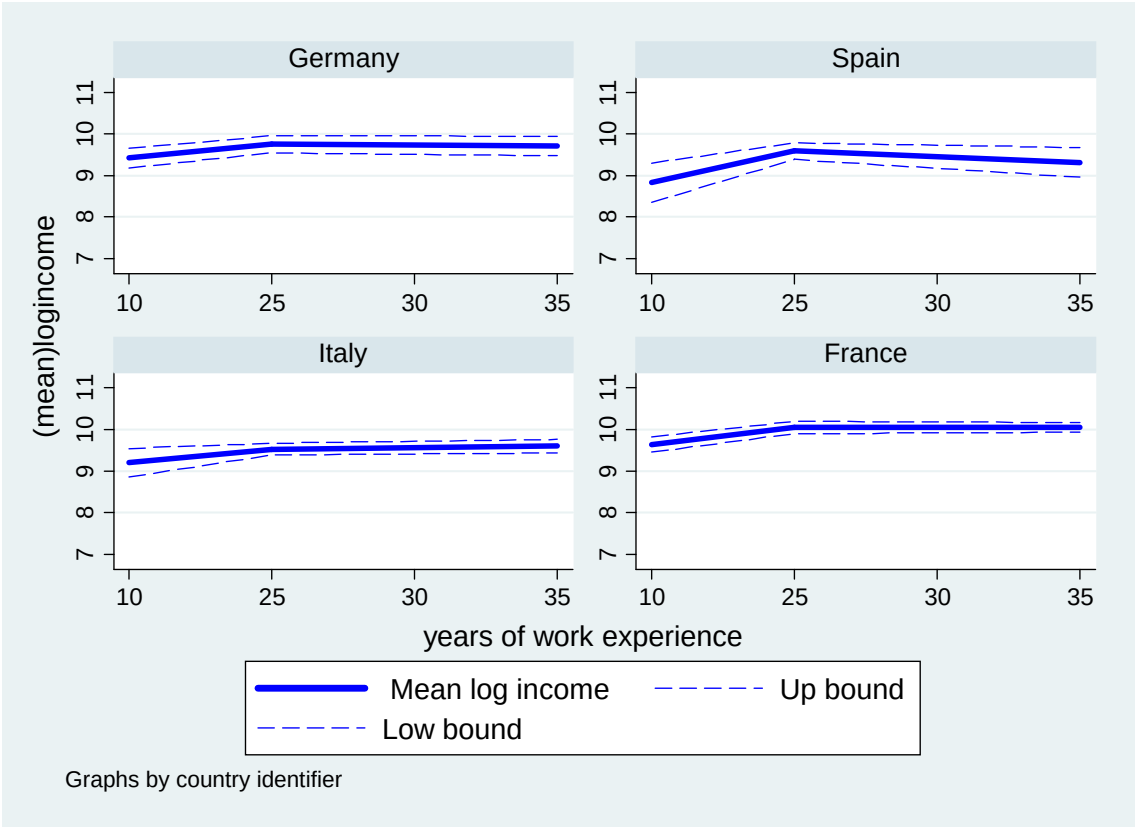
Figure 18: Level of income by years of work experience



Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

It can be seen from the graph that income gradually increases with additional years of work experience as a result of training received from a firm. Figure 19 shows the level of income for work experience in each country without gender differences.

Figure 19: Level of income by years of work experience and country

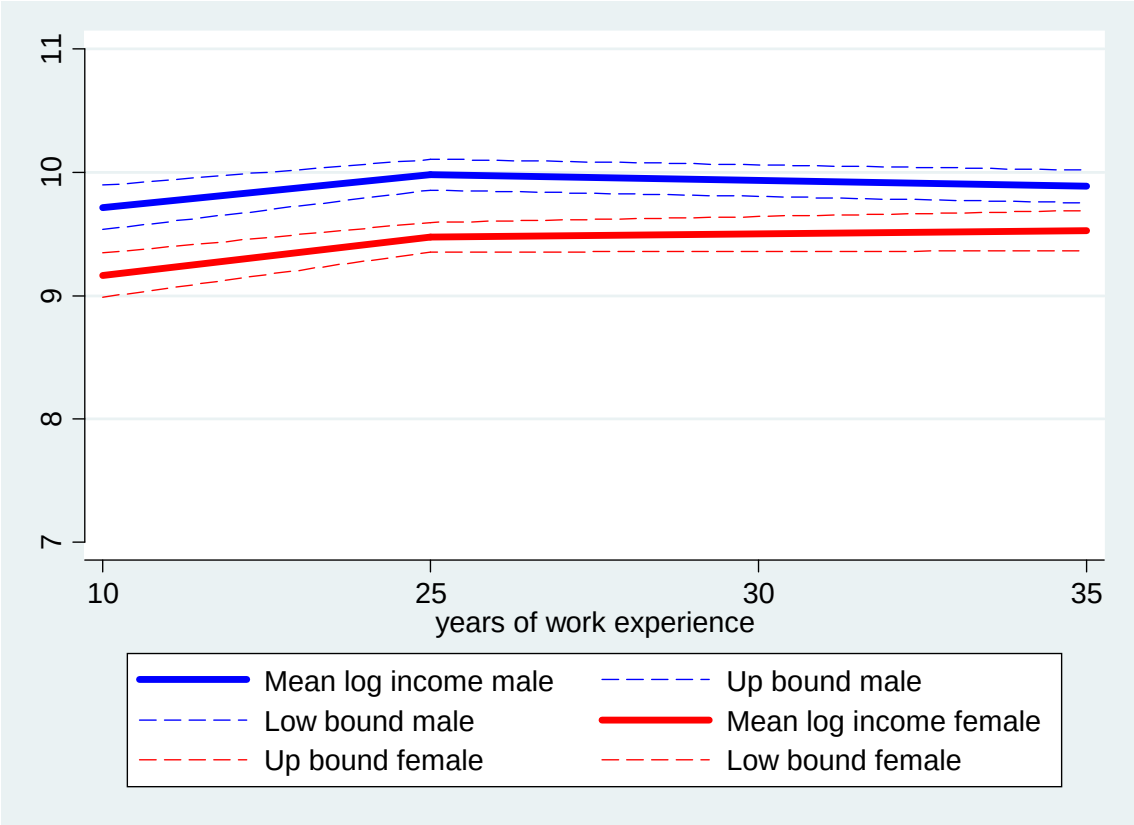


Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

The graph in Figure 19 shows that all countries follow the average trend of Figure 18, except for Spain, where after 25 years of work experience income tends to decrease, albeit slightly. The cause can be compared with the graphs of level of income by age shown above, where income decreases as the individual approaches retirement age.

The next two graphs represent gender differences first globally and then for each country.

Figure 20: Level of income by years of work experience and gender

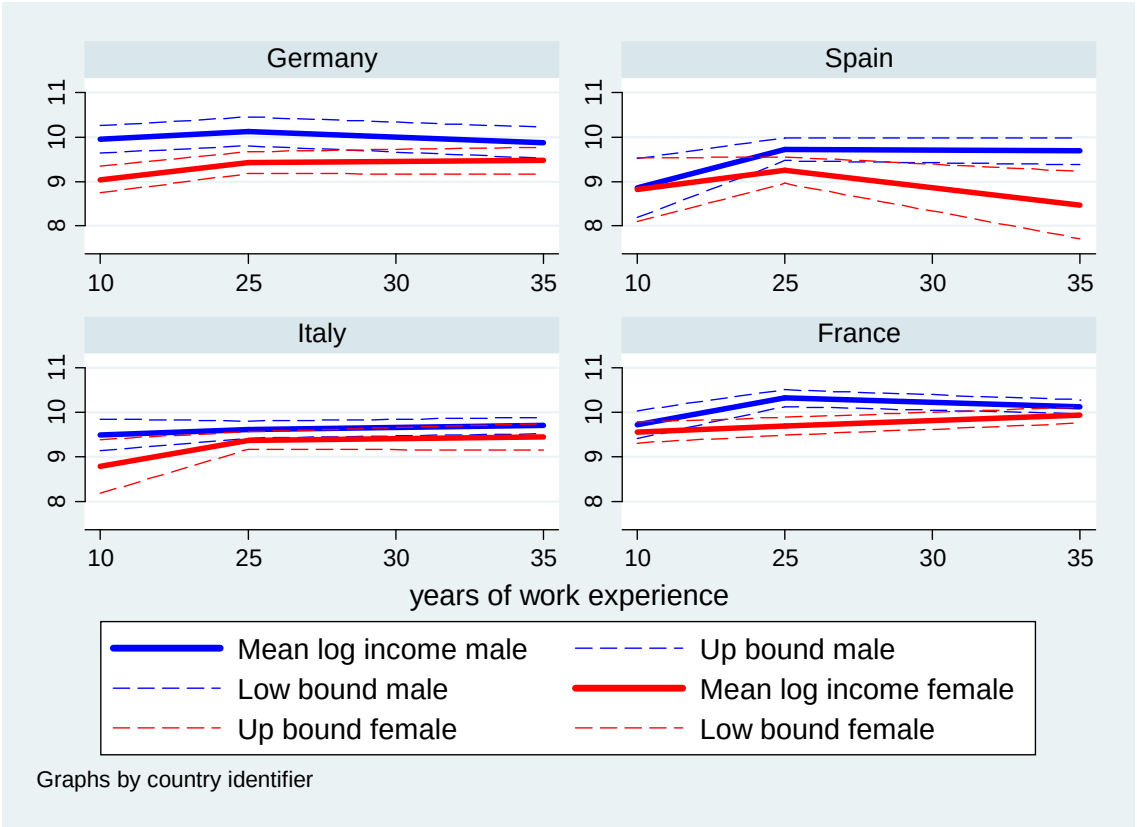


Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

In Figure 20 the mean income of men and women follows a very similar trend. The level of income increases up to 25 years of work experience and remains steady in subsequent years. However, although the trend is similar, the mean income of men is greater than that of women for their entire working life.

Figure 21 shows the mean income of men and women in each country. The graph shows that the gender gap is greater in the first years of work experience and then tends to decrease in subsequent years. This trend is particularly visible in Italy and Germany.
Spain seems to be the only country that follows a different trend. In this country men and women on average seem to have the same initial level of income in the early years of work, whereas the gender gap increases in subsequent years. This may be due to the fact that men's wage increases for career advancement are greater than the wage increases for women

Figure 21: level of income by years of work experience, gender and country



Source: This figure uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

Summarizing, on average income for the different countries tends to increase till 53-56 years of age and it increases till 25 years of work experience. Successively earnings tend to decrease in the last age class. This could probably be due to a process of selective retirement before the age of 67, because the individuals whose earnings tend to decline more rapidly may choose to retire earlier (Mincer, 1958). The alternative is to remain in the labour market, adapting to what the employer decides to offer. Usually it is a lower wage because the employer has the advantage of investing in training to increase a younger person's human capital that may include the use of new technologies more difficult to learn by an older worker.

Moreover earnings are reduced during the investment period because there are investment costs, and successively earnings increase when the investment terminates. For this reason the age-earnings profile is concave. The consequence is that the total

amount invested is positively correlated with education, and better educated individuals have a more concave age-earnings profile.
Therefore the human capital theory can explain the differences concerning both the concavity and the steepness of the age-earnings profile of individuals.

4.2. The Mincer equation: results of Empirical Estimations

In the previous subsection a descriptive analysis of the wave 2 dataset has been provided regarding the employment levels and income levels of individuals.
This subsection involves the empirical estimations of the Mincer equation on the basis of what is described in subsection 2.6.
According to the human capital theory widely described in Chapter 2, on average, individuals with higher education receive higher earnings.
Starting from the general Mincer equation described in subsection 2.6, a more detailed analysis was made including other variables that describe the characteristics of individuals. In addition to education and work experience, we have taken into account: the gender of the individual, the status of the individual (whether married or single), the number of children, and the geographical area. There are two different types of estimates. The first model allows us to estimate the level of income, including among the several variables the dummy of the four countries (i); the second estimate is made in each country to determine how each variable in each country affects the level of income (ii).

Starting from equation (18) with the inclusion of the additional variables, (i) the variables of the dataset for the first model are therefore: *logincome* (dependent variable), *yedu, y_work_w2, female, single, nchild, Germany, Spain, Italy and France*. In particular *logincome* is the logarithm of *ydipv* (annual net income from employment), which represents the same variable used in the descriptive analysis for the graphs in the previous subsection. The reason for using the logarithm of income is that through this operator it is possible to reduce the problem of heteroscedasticity. The variable *y_edu* is the level of education, expressed in years of study for each individual. The variable

y_work_w2 is potential experience, computed as described in detail in the previous subsection, excluding unemployment spells and the time spent on education for each individual. The variable *female* is a dummy variable that describes the gender; it assumes value one if the individual is a woman and zero if he is a man. The variable *single* describes the marital status of the individual, in particular it is a dummy variable that takes value one if the individual is single and zero if s/he is married. The variable *nchild* indicates the number of children for each individual. Finally the variables *Germany*, *Spain*, *Italy* and *France* are dummy variables that take value one if the individual is part of the given country and zero otherwise.

The sample of the dataset for this loglinear estimation has been reduced. The new sample includes individuals aged between 21 and 67 years. In this way it is possible to consider only individuals who are that age within the working life cycle. The variable *ydipv* does not consider retirement income because it refers only to income from employment, but proceeding in this way has excluded from the estimate individuals who continue to work after retirement age. The consequent result is a more easily comparable estimate of the countries, because not in all countries are older workers over 67 equally employed, as shown in Figure 5. For this purpose, the sub-sample of regression used is composed of 1589 individuals for the four countries, with 787 men and 802 women. The dataset used refers to wave 2 because for the regressions related to the Mincer equation there are cross section data.

Table 2 shows the results obtained from this regression.

Table 2: Results of OLS estimation of the Mincer equation for the first specification

	logincome
yedu	0.0758***
	(0.00547)
y_work_w2	0.0131***
	(0.00184)
female	-0.387***
	(0.0442)
single	-0.0587
	(0.0599)
nchild	-0.0121
	(0.0196)
Germany	8.340***
	(0.118)
Spain	8.389***
	(0.110)
Italy	8.452***
	(0.112)
France	8.771***
	(0.114)
N	1589
r2	0.992

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

The regression makes it possible to estimate coefficients b, c and d of equation (18). For example, in Table 2 *yedu* corresponds to b and *w_work_w2* corresponds to c.

One of the advantages of transforming income into logarithmic form, is to directly interpret coefficient estimates in terms of percentage changes, since, for example, b = (δ lny/ δ S); this is the relative increase in income following an increase in S (the rate of return to an additional year of schooling).

Looking at the coefficients, it can be noted that having an additional year of education implies on average an increase in income of 7.58% and the coefficient is statistically significant.

This result confirms what has been widely described in subsection 2.1. on the human capital theory. Workers with higher education levels are more productive than those with lower levels of education and therefore the better educated workers have higher levels of income.

Another important result is given by potential experience (*y_work_w2*). An additional year of work experience leads to an increase in logincome of 1.31%. Also in this case it is confirmed that work experience has a positive effect on income, as expressed above in the graphs of the descriptive analysis and in subsection 2.2. The two variables examined represent the explanatory variables of the classical Mincer equation, the subsequent variables provided represent the additional variables needed for the purpose of estimating this model.

In particular, the *female* coefficient measures the relative difference between the expected value of income for men and women. In particular, the difference in the expected value of men's and women's income logarithm is equal to -0.387. It is a common practice for the linear models of the logarithms to interpret coefficient estimates in terms of percentage changes; in this way a coefficient of -0.387 for women is interpreted as a difference in expected income of about -38%. This explains that 38% is the difference between the expected income of a man and a woman having the same level of education, years of work experience, number of children, marital status and geographic area.

The variable *nchild* has a coefficient equal to -0.0121. This means that an individual who has children earns on average 1.21% less than an individual without children. However, it must be considered that this coefficient is not significant[10].

The variable *single* has a negative coefficient, but the inclusion of this variable does not affect the determination of the model because it is not significant. In this way even if it was expected to obtain a positive coefficient by the *single* variable (because being married and having a family is assumed to reduce earnings, especially for women), this is not a negative result since this variable is not determinant in this case to explain the model.

[10] Appendix A includes estimates with the variable *nchild_fem* (a woman who has children) so that it is possible to see the gender differences arising from having children. What emerges is that the gender gap in earnings is due to having children.

Finally, the four dummy variables of the countries indicate the expected mean income for each of them. It may be noted that on average France is the country where mean income is greater, but these four variables compared with each other are not significantly different. To see in more detail the effect on income in each country further regressions are provided, one for each of them. Moreover in this regression the constant term was not included in order to give more evidence on the average income expected in each country.

Although the *age* variable is interesting in the determination of income, it is not included in this regression to avoid multicollinearity problems. In general, there is nothing wrong with including variables correlated to each other in the model. In an equation of this type, for example, both age and experience may be included, even if it is reasonably expected that older individuals also have, on average, more work experience. If the correlation between the two variables is too high, however, some problem could arise. This may lead to unreliable estimates with high standard errors, or coefficients with a negative sign or an unexpected value. In a case like this it may be possible to solve the problem and receive reasonable answers using a large number of observations that are characterized by sufficient variability in age and work experience. However, this solution has not been applied, because the problems could be aggravated by the inclusion in this model of education measured in years of schooling. For these reasons, it has been established to analyze the effect of age in a different regression, to avoid the problem of multicollinearity arising from the existence of an approximate linear relationship between the explanatory variables that could generate unreliable estimates.

The use of too many dummy variables (which take only value 0 or 1) is a typical cause of multicollinearity. In this case the risk concerns the inclusion of the dummy for the four countries. The problem could be solved by not including all the variables, or including all the dummies, but without the constant term as it is carried out in this regression. This latter approach, however, may create problems in the determination of R^2, because without the constant term, the R^2 and F statistics tend to be computed in a different way. For this reason the R^2 of the estimated model is equal to 0.992, which implies that 99% of the variability in individual incomes may depend (in a linear way)

on gender differences, work experience and education levels. However it must be emphasized that its value is overestimated by the absence of the constant term.

The aim of the next regression is to analyze the effect of age on income with the introduction of the *age* variable, but in this case excluding the years of work experience in order to avoid the problem of multicollinearity as explained previously, whereas the other variables are the same as the previous regression. The results of this new regression are shown in Table 3.

Table 3: Results of OLS estimation of the Mincer equation for the second specification

	(1) loginc ome	(2) loginc ome
yedu	0.0758***	0.0711***
	(0.00547)	(0.00553)
y_work_w2	0.0131***	
	(0.00184)	
female	-0.387***	-0.494***
	(0.0442)	(0.0441)
single	-0.0587	-0.0223
	(0.0599)	(0.0612)
nchild	-0.0121	-0.0241
	(0.0196)	(0.0197)
age		-0.0195***
		(0.00486)
N	1589	1589
r2	0.992	0.992

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

In Table 3 the results of the new regression with the introduction of the *age* variable (column "loginocme 2") are compared with the previous regression having the variable *y_work_w2*. For greater simplicity of comparison Table 3 does not include the dummy variables of the four countries. The *age* variable represents the individual's age at the

time of the interview for wave2. The coefficient of the *age* variable is negative, in contrast to what was expected, but this may be due to the fact that, as the descriptive analysis shows, income tends to decline when the individual is closer to retirement age. Subsequently a further analysis will be provided including the variable of income with increasing age. It can be seen that the variables *y_work_w2* and *age* affect income in different ways. This could mean that older age does not always correspond with many years of work experience, which could be caused by unemployment spells. Regarding the other variables, the trend is similar, although to a different extent. For example, in logincome2 the coefficient of the *female* variable gives a greater difference between the expected income of a woman compared to a man having the same education level, age, number of children and marital status. This means that the introduction of the *age* variable implies that women have an expected income -49% compared to men. The possible interpretation is that for a given level of education, the years of work experience affect the income to a greater extent than age even in terms of gender gap, thus, given the same levels of education, increased age does not necessarily consistently affect income if it is not combined with an effective increase in the years of work experience.

Even in the second regression at global level for the four countries, the variable *nchild* and *single* are not significant. Finally, the R^2 is unchanged at a high level, but this is due to lack of the constant term.

A further consideration might suggest that the effect of age and years of work experience on the income of an individual is non-linear. This means that for advanced age of individuals and years of work experience, the effect on income of an additional year of age or an additional year of work experience may become smaller, as previously demonstrated with the graphs of the descriptive analysis. To describe this phenomenon the square of age and the square of years of work experience can be included in the model, expecting negative coefficients. This estimation is made separately for age and years of work experience in order to avoid problems of multicollinearity and is shown in Table 4[11].

[11] With the aim of obtaining a more effective comparison, the dummy variables of the four countries in these two regressions are not considered.

Table 4: Results of OLS estimation of the Mincer equation with the non-linear effect of age and work experience on income

	(1) logincome	(2) logincome
yedu	0.0774***	0.0749***
	(0.00539)	(0.00525)
y_work_w2	0.0177*	
	(0.00721)	
y_work_w2_sq	-0.000140	
	(0.000157)	
female	-0.382***	-0.469***
	(0.0451)	(0.0442)
single	-0.00134	0.0146
	(0.0607)	(0.0608)
nchild	0.00960	-0.000204
	(0.0198)	(0.0195)
age		0.411***
		(0.0756)
age2		-0.00392***
		(0.000681)
_cons	8.437***	-1.806
	(0.114)	(2.101)
N	1589	1589
r2	0.183	0.192

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

Table 4 introduces new variables: in the first regression *y_work_w2_sq* represents the square of work experience, whereas the second regression *age2* indicates the square of the age of individuals.

Looking at column 1 (logincome 1) of Table 4, the estimate of the coefficient of the variable *y_work_w2_sq* is negative, as was expected. The t-test suggests that the square of work experience is not significant, which means that this variable is not decisive in defining the model. Even *y_work_w2* is not significant for a pvalue of 0.001 and 0.01, but this variable is still significant for pvalue 0.05, thus *y_work_w2* is considered

significant (given the high number of observations, a test dimension of 5% can be considered more appropriate).

Since work experience in the regression appears to be linear with *y_work_w2* and non-linear with *y_work_w2_sq*, the two coefficients cannot be interpreted separately. One way to describe the effect of work experience is to consider that the difference in expected income due to a marginal increase in work experience is given by:

$$0.0177 - (y\_work\_w2\_sq \cdot 2 \cdot years\ of\ work\ experience) \quad (21)$$

This equation is obtained by making the derivative with respect to *y_work_w2* and it shows how the effect of experience on income varies with changes of this variable. Initially the effect is 1.77% for an additional year of work experience, but this is reduced to 0.93% for an individual who has 30 years of work experience.

The other variables of column 1 appear to be similar to those of Table 2 where the estimate was made with a single variable for years of work experience.

The same analysis can be performed in the case provided by column 2 of Table 4, where instead of considering the years of work experience, age and squared age is considered. In particular the square of age is represented by the variable *age2*. The interesting aspect is that the *age* variable in this case assumes a value of 0.411, while in Table 3 its value was -0.0195. This means, as described previously, that the value of the *age* variable in Table 3 is affected by a decrease in the income of older individuals. The method applied for equation 20 can be extended also in this case. Thus, for istance, if initially one additional year may result in a higher income by 41% (although in this case, this variable has a relatively high standard error of 7% and this could mean that the estimate is not precise, perhaps due to a sampling error) for an individual of 50 the increase in income is only 1.9%.

Also in this case, the other variables have similar values to those of Table 3. Finally R^2 decreases in both regressions, but this is due to the fact that the constant term has been introduced and it does not mean that the model fails to provide good results. In Table 5 there is an overview of the previous regressions. However, with the aim of making the comparison more effective it also includes the constant term in the regressions of Table

3, but as can be seen the coefficients are not significantly different from the previous cases. In particular the first two columns refer respectively to regressions with the inclusion of years of work experience and the effect of work experience with increasing years of work (*y_work_w2_sq*), whereas the last two columns refer to the effects of age.

Table 5: Final results of OLS estimation of the Mincer equation at global level

	(1) logincome	(2) logincome	(3) logincome	(4) logincome
yedu	0.0783***	0.0774***	0.0742***	0.0749***
	(0.00529)	(0.00539)	(0.00531)	(0.00525)
y_work_w2	0.0115***	0.0177*		
	(0.00186)	(0.00721)		
female	-0.378***	-0.382***	-0.483***	-0.469***
	(0.0449)	(0.0451)	(0.0445)	(0.0442)
single	-0.00383	-0.00134	0.0296	0.0146
	(0.0606)	(0.0607)	(0.0614)	(0.0608)
nchild	0.00866	0.00960	-0.00354	-0.000204
	(0.0198)	(0.0198)	(0.0197)	(0.0195)
y_work_w2_sq		-0.000140		
		(0.000157)		
age			-0.0233***	0.411***
			(0.00487)	(0.0756)
age2				-0.00392***
				(0.000681)
_cons	8.473***	8.437***	10.18***	-1.806
	(0.107)	(0.114)	(0.285)	(2.101)
N	1589	1589	1589	1589
r2	0.183	0.183	0.175	0.192

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

Successively the same regressions are made for each country. (ii) The advantage obtained is that it is possible to verify the results of the estimates in each country, but at the same time regressions are performed with a smaller number of observations,

therefore the risk could arise that the estimates may be less precise than the previous cases. The procedure followed is the same as shown above: a first estimate containing education level, years of work experience, gender, number of children and marital status; the second estimate contains age instead of years of work experience, to avoid multicollinearity. Finally there are two regressions to verify the influence on income of high age and many years of work experience. The results are given in Table 6.

Table 6: Results of OLS estimation of the Mincer equation by country for the first specification

	France logincome	Germany logincome	Italy logincome	Spain logincome
yedu	0.0637***	0.118***	0.0841***	0.0550***
	(0.00900)	(0.0149)	(0.00973)	(0.0118)
y_work_w2	0.0130***	0.0184***	0.0138***	0.00669
	(0.00272)	(0.00406)	(0.00389)	(0.00460)
female	-0.356***	-0.390***	-0.283***	-0.439***
	(0.0677)	(0.0970)	(0.0816)	(0.127)
single	-0.0754	-0.0962	-0.159	0.154
	(0.0828)	(0.134)	(0.131)	(0.165)
nchild	-0.0187	-0.0292	-0.0381	0.0628
	(0.0283)	(0.0426)	(0.0443)	(0.0508)
_cons	8.930***	7.654***	8.342***	8.600***
	(0.169)	(0.290)	(0.205)	(0.235)
N	536	452	339	262
r2	0.181	0.237	0.247	0.129

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

At this point it is possible to verify the effects of each variable in each country. In particular, the *yedu* variable implies on average a return to education approximately equal to 6.37% in France, 11.8% in Germany, 8.41% in Italy and 5.5% in Spain, considering the other variables constant. Compared to the average value of 7.83% for the four countries shown in Table 5, it can be seen that Germany has a return to school a

lot higher than the average value, whereas Italy is located just above the average and France just below. Spain is the country that seems to deviate more from the average value but this may be due to the fact that the sample[12] is composed of only 262 observations for this country.

Looking at years of work experience, an additional year leads to an increase in expected income equal to 1.3% in France, 1.84% in Germany and 1.38% in Italy, whereas in Spain this variable is not significant even for a significance level of 5%. This means that in this country the effect of work experience on income is not as relevant as it is in other countries. In Table 5 the average return on work experience was 1.15%, so France, Germany and Italy have an expected return that is higher than the average value. Moreover the results indicate that in Spain there is a larger gender gap in terms of income levels. Comparing a woman with a man having the same level of education and experience, in this country the expected income of a woman is -43.9%, whereas the average value for the four countries in Table 5 is -37.8%. On the contrary, in Italy there is a lower gender gap and a woman compared with a man has an expected income equal to -28.3%. This smaller gender gap in Italy is a good result for the gender gap in income levels, but in Figure 7 Italy is the country with the largest gender gap with regard to employment levels. France and Germany have gender gap values of -35.6% and -39% respectively, approximated to the average value of -37.8%. All these results are also confirmed in the descriptive analysis in Figure 21.

The *nchild* variable has a coefficient equal to -1.87% in France, -2.92% in Germany, -3.81% in Italy and 6.28% in Spain. It therefore seems that Italy is the country where an individual who has children is more penalized, but these results are not statistically significant[13]. Moreover in Spain being single seems to imply an increase in income of 15.4%, but also in this case the coefficient is not significant.

Finally, observing the values of the constant terms, it can be seen that again France is the country with the highest expected mean logincome of 8.93.

[12] As previously the sample for these regressions is created by excluding individuals who are over 67 years old.

[13] Also in this case more detailed results are provided for each country regarding the number of children in order to verify the gender gap.

The next regressions are carried out performing the same analysis with the use of the *age* variable instead of *y_work_w2.*

Table 7 shows that the coefficients relating to education levels are similar to those of the previous regressions. The results that have changed most are those related to the *female* variable, but the reasons may be associated with those explained in Table 3. Also in this case the coefficients of the *age* variable are negative for all the countries, in contrast to what was expected, but this may be due to the fact that, as shown in the descriptive analysis, income tends to decline when the individual is closer to retirement age. In France, however, this variable is not significant and it is therefore not considered essential in the determination of the model, whereas in Italy and Spain, the variable is not significant for a significance level of 0.001 and 0.01, but it is significant for a significance level of 0.05.

Table 7: Results of OLS estimation of the Mincer equation by country for the second specification

	France logincome	Germany logincome	Italy logincome	Spain logincome
age	-0.00531	-0.0353***	-0.0177*	-0.0286*
	(0.00809)	(0.0104)	(0.00896)	(0.0121)
yedu	0.0600***	0.106***	0.0806***	0.0505***
	(0.00916)	(0.0149)	(0.00982)	(0.0119)
female	-0.419***	-0.588***	-0.384***	-0.567***
	(0.0681)	(0.0958)	(0.0836)	(0.122)
single	-0.0647	-0.0825	-0.0587	0.198
	(0.0858)	(0.136)	(0.135)	(0.165)
nchild	-0.0355	-0.0663	-0.0326	0.0768
	(0.0287)	(0.0420)	(0.0451)	(0.0510)
_cons	9.638***	10.45***	9.759***	10.41***
	(0.468)	(0.647)	(0.516)	(0.706)
N	536	452	339	262
r2	0.147	0.222	0.227	0.140

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

The next regressions introduce the variables *age2* and *y_work_w2_sq* as previously to verify the non-linear effect of the years of work experience and age of the individual.
In particular, in Table 8 it can be seen that the introduction of the variable *y_work_w2_sq* (the square of the years of work experience) is not significant for any country. This means that the introduction of this variable is not significant for a better explanation of the model. The other variables appear to be similar to those obtained in Table 6 without the specification of the variable *y_work_w2_sq*. In this way it seems that many years of work experience do not affect specification of income in each country, but this may be due to the fact that there are few observations available in each of them. This effect is more visible in Table 9 for the *age2* variable. The *age* variable increases income in all countries, while the *age2* variable decreases income when age increases (or the increase is gradually decreasing). Spain is the only country in which the *age2* variable is not significant.

Table 8: Results of OLS estimation of the Mincer equation by country with the non-linear effect of years of work experience on income

	France logincome	Germany logincome	Italy logincome	Spain logincome
yedu	0.0657***	0.114***	0.0869***	0.0531***
	(0.00922)	(0.0151)	(0.00997)	(0.0119)
y_work_w2	0.00271	0.0423**	-0.00482	0.0316
	(0.0109)	(0.0162)	(0.0157)	(0.0174)
y_work_w2_sq	0.000252	-0.000522	0.000394	-0.000539
	(0.000259)	(0.000344)	(0.000321)	(0.000363)
female	-0.352***	-0.417***	-0.287***	-0.465***
	(0.0679)	(0.0984)	(0.0816)	(0.127)
single	-0.0784	-0.0866	-0.168	0.163
	(0.0829)	(0.134)	(0.131)	(0.165)
nchild	-0.0182	-0.0300	-0.0445	0.0747
	(0.0283)	(0.0426)	(0.0446)	(0.0513)
_cons	8.967***	7.535***	8.497***	8.418***
	(0.173)	(0.300)	(0.241)	(0.265)
N	536	452	339	262
r2	0.183	0.241	0.250	0.136

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

Table 9: Results of OLS estimation of the Mincer equation by country with the non-linear effect of age on income

	France logincome	Germany logincome	Italy logincome	Spain logincome
age	0.433***	0.578**	0.394**	0.176
	(0.127)	(0.194)	(0.139)	(0.163)
age2	-0.00403***	-0.00542**	-0.00370**	-0.00185
	(0.00117)	(0.00171)	(0.00125)	(0.00148)
yedu	0.0630***	0.110***	0.0831***	0.0499***
	(0.00911)	(0.0148)	(0.00974)	(0.0119)
female	-0.383***	-0.566***	-0.379***	-0.556***
	(0.0682)	(0.0951)	(0.0826)	(0.122)
single	-0.0671	-0.0891	-0.0823	0.188
	(0.0849)	(0.135)	(0.134)	(0.165)
nchild	-0.0364	-0.0558	-0.0317	0.0822
	(0.0284)	(0.0417)	(0.0446)	(0.0511)
N	536	452	339	262
r2	0.165	0.239	0.247	0.146

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

Thus, for instance, applying equation 20 in France a 50-year-old individual has an expected increase in income equal to 3%, in Germany an individual of the same age has an expected increase equal to 3.6% and in Italy it is equal to 2.4%. This means that Germany and France are countries in which individuals suffer less from a remunerative point of view from being closer to retirement age. This may explain why in Figure 5 France and Germany are the countries where there is the largest number of individuals employed after retirement age.

Equation (18) is used to describe with reasonable success the relationship between years of education and income. According to a study of Psacharopoulos (1985), even in countries with different economies and education systems, the estimated coefficients b are relatively homogeneous and they are included between 0.05 and 0.15, with an average of about 0.10 (Psacharopoulos, 1985; Psacharopoulos, 1994; Psacharopoulos and Patrinos, 2004).

In this empirical analysis the results show that: France has a coefficient of 0.0637 (6.37%), Germany has a coefficient of 0.118 (11.8%), Italy has a coefficient of 0.0841 (8.41%) and Spain has a coefficient of 0.0550 (5.5%)[14]. According to this methodology, therefore, a year of education has on average an additional economic output of about 0.0783 (7.83%) as shown in Table 5. Since these estimates are the result of the Mincer equation (18), they refer to the case in which regression includes work experience and not age. Age has been included separately to make the analysis more complete but as final results for the purposes of this analysis those from the solution of the Mincer equation are taken into account.

4.2.1. Extension of the model

The objective of this subsection is to highlight the impact of different levels of education on income given the other conditions. Three dummy variables have been added to the model, corresponding to levels of education. In particular, the three dummy variables are: *educ_3* (low education level), *educ_4* (medium education level), *educ_5* (high education level). These variables are created grouping isced levels 0-2 in *educ_3*, *educ_4* includes isced levels 3-4 and *educ_5* incudes isced levels 5-6. In this way *educ_3* represents primary education, *educ_4* is secondary education and *educ_5* is university education (tertiary education). In the case of the three classifications, the inclusion of two dummy variables is sufficient to capture all the effects (Verbeek, 2004). The inclusion of three dummy variables generates the risk of perfect multicollinearity. The choice of the variable to be excluded is irrelevant, the only consequence concerns the economic interpretation of the coefficients of the other dummy variables. The omitted category is a reference group, and all the effects are measured with respect to this group. In this case the reference category is represented by the variable *educ_3*. The results are estimated considering years of work experience as an expression of the Mincer equation. The results obtained with this specification are shown in Table 10 at global level, and in Table 11 for each country.

[14] The results refer to Table 6 because the addition of the variable *y_work_w2_sq* in Table 8 is not significant.

Table 10: Results of OLS estimation of the Mincer equation with classification by education level

	(1) logincome
educ_4	0.361***
	(0.0529)
educ_5	0.791***
	(0.0575)
y_work_w2	0.00970***
	(0.00188)
female	-0.397***
	(0.0454)
single	-0.0113
	(0.0616)
nchild	-0.00240
	(0.0199)
_cons	9.143***
	(0.0868)
N	1579
r2	0.169

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

The results of Table 10 show that high school education (secondary education level, variable *educ_4*) leads to an average expected increase in income equal to 36% compared to individuals who have a low level of education (variable *educ_3*). The increase is about 79% if the individual has a university education (tertiary education, variable *educ_5*) compared to those who have a low level of education (primary education). The results of Table 10 also show that given the same educational qualifications, a woman earns on average 39.7% less than a man. These coefficients are considered average values for the four countries. In Table 11 there are the returns to education for isced classifications in each country. Table 11 shows that returns to university in Germany are above the average value of 79% shown in Table 10. In this

country an individual with a tertiary education earns on average 100% more than those who have a low level of education (primary education). The difference with those who have secondary education is 57.4%.

The other countries also show good results that confirm the assumptions of Mincer explained in chapter 2. Tertiary education provides an increase in income of 65.6% with respect to primary education in France, 74.8% in Italy and 69.5% in Spain.

Table 11: Results of OLS estimation of the Mincer equation with classification by education level and country

	France logincome	Germany logincome	Italy logincome	Spain logincome
educ_4	0.304***	0.427*	0.471***	0.419**
	(0.0824)	(0.187)	(0.0926)	(0.150)
educ_5	0.656***	1.001***	0.748***	0.695***
	(0.0852)	(0.193)	(0.125)	(0.155)
y_work_w2	0.0106***	0.0162***	0.0118**	0.00642
	(0.00269)	(0.00411)	(0.00404)	(0.00466)
female	-0.367***	-0.480***	-0.327***	-0.421**
	(0.0674)	(0.0966)	(0.0858)	(0.128)
single	-0.0589	-0.116	-0.162	0.121
	(0.0831)	(0.137)	(0.136)	(0.168)
nchild	-0.0236	-0.0224	-0.0715	0.0458
	(0.0281)	(0.0432)	(0.0458)	(0.0511)
_cons	9.483***	8.762***	9.106***	8.985***
	(0.125)	(0.257)	(0.170)	(0.208)
N	530	451	339	259
r2	0.196	0.214	0.186	0.130

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

4.2.2. The issues of interpretation and the ability bias

The results arising from the estimates of the Mincer equation (equation 18) provide a positive correlation between education and the performance of the labour market, as has been shown in previous subsections. However, it is more problematic to interpret this correlation.

First, if private returns to education are actually so high, it is important to understand why so few individuals decide to continue their studies. One reason shared by the publication of a famous essay by Milton Friedman (1955) is that of imperfections in capital markets: the costs of studying (not only in terms of opportunity costs) in terms of difficulty for people to borrow the funds necessary to attend university and support themselves during their studies, even though it is highly likely that, given returns to education, the loans would be repaid. One of the reasons for the market imperfections lies in the uncertainty that characterizes future returns to education.

Secondly, the advantages of higher education may be heterogeneous among individuals with different characteristics. In particular, for those who graduate returns to education are not necessarily always higher than for those who have lower levels of education.

Thirdly, it is possible that the estimated returns to university education do not depend on the greater productivity of graduates, but they are the result of credentials that indicate skills hardly observable in the labour market and these skills were not acquired at university (Spence, 1973). In this case, returns to university reflect the inefficiencies of the labour market and not the human capital acquired by graduates.

Finally, there is the problem of how to control the unobservable individual characteristics, which are relevant in the determination of income: if these are correlated with years of education, the estimation equation (18) using the OLS technique of ordinary least squares would provide the ability bias (Griliches, 1977). From a theoretical point of view it is often assumed that this bias leads to over-estimation of the actual effect of education, since some hardly observable characteristics such as abilities and personal motivations have a positive effect on individual income and these are probably positively correlated with years of education. The fundamental issue is that educational choices depend on individual characteristics that are hardly observable. The

ideal conditions for measuring the impact of education are therefore those in which the choices of further studies are taken randomly with respect to the individual characteristics, as has been done in the estimates of the Mincer equation of this thesis.
In conclusion, the results obtained suggest that, despite the possibility that the equation (18) may have some interpretation problems, parameter b of the equation reliably measure the effects of education on individual incomes.

4.3. The scarring effects: results of Empirical Estimations

This subsection involves the results of the estimates to assess the different size of scarring effects in the four countries. Also in this case the dataset of wave2 is used as was done in the previous subsection for the estimates of the Mincer equation.
However, to estimate scarring effects, in addition to the level of education and demographic characteristics (gender, marital status, number of children and geographic area), variables relating to the recession period and unemployment spells have been introduced. This is necessary in the determination of scarring effects which imply verifying whether there was a period of recession at school leaving age for every worker or whether there were unemployment spells in the past in order to test their effects on the individual's working career.
The following equation is used for estimating scarring effects:

$$Y_i = \beta X_i + \gamma Z_i + f(U_{i0},\ldots,U_{it}) + \lambda_i + \mu_i \qquad (21)$$

where Y_i is income from employment, X_i is a vector of personal features, Z_i is a vector of the labour market characteristics, Ui_0 ... U_t are characteristics of periods of unemployment observed in t previous time intervals, λ_i is a measure of heterogeneity for each individual and μ_i is the disturbance term.
The regressions are structured first at global level for the four countries and successively there is one regression for each country showing for each of them the internal geographical differences between North and South.

In particular for global regression, the variables within the dataset used, which refer to equation (21) are: *log_income, yedu, y_work_w2, female, rec_edu, rec_fem, ever_unemp, nchildren, single, Germany, Spain, Italy, France*. Where *log_income, yedu, female, nchildren, single, Germany, Spain, Italy* and *France.* These are the same variables used for estimating the regressions carried out for the Mincer equation. The new variables for estimating scarring effects are *rec_edu, rec_fem* and *ever_unemp*. In particular *rec_edu* is a dummy variable indicating whether there was a period of recession when the individual entered the labour market and it takes on value one if there was recession and zero otherwise. The variable *rec_fem* is a variable of interaction. This means that the variable *rec_fem* is the product of two variables. This makes it possible to obtain the effect of one variable on *log_income*, depending on the other variable. In this case, the interaction terms are *rec_edu* and *female*. The result is that *rec_fem* indicates the extent to which a recession period has affected the income of women.

Finally the variable *ever_unemp* indicates whether the individual has been unemployed at least once during his working life.

Table 12 provides a representation of the empirical results obtained at global level for the four countries. In particular in column 1 there are the effects of recession and unemployment to verify whether there are scarring effects, while in column 2 there is the effect of recession on women with the inclusion of the interaction variable *rec_fem*. This sample is greater than that of the Mincer equation in Table 2, in fact now there are 1865 observations, 908 males and 957 females. The reason is that the sample is not restricted to a retirement age of 67, but all individuals who receive income from work are considered. In this way it is possible to evaluate the scarring effects also for older individuals and at the same time obtain a larger sample.

As can be seen in both columns of the Table, the effect of a recession at school leaving age is not significant for both men and women at global level[15]. On the other hand, significant results are obtained for the variable *ever_unemp*. An individual who has experienced a period of unemployment earns -24.5% annually compared to another individual having the same characteristics (level of education, years of work experience,

[15] As will be demonstrated below, this result does not mean that the effect of the recession is not significant in each country.

gender, etc.) but without an unemployment spell. This result confirms the negative effects that may arise from a period of unemployment. Finally looking at column 2 the *female* variable indicates that a woman who did not experience a period of recession at school leaving age earns on average 32.1% less than a man, which is in line with the results obtained for the Mincer equation in Table 6. However, a woman who has experienced a recession earns on average: -0.281 – 0.321 +0.161 = -44.1%, but the data of *rec_fem* is not significant.

Table 12: Results of OLS estimation of scarring effects at global level

	(1) logincome	(2) logincome
yedu	0.0773***	0.0775***
	(0.00505)	(0.00505)
y_work_w2	0.0171***	0.0170***
	(0.00222)	(0.00221)
female	-0.335***	-0.321***
	(0.0421)	(0.0430)
rec_edu	0.0134	0.161
	(0.0920)	(0.133)
ever_unemp	-0.245***	-0.244***
	(0.0559)	(0.0559)
nchild	-0.00829	-0.00927
	(0.0178)	(0.0178)
single	-0.0996	-0.101
	(0.0553)	(0.0553)
Germany	8.124***	8.115***
	(0.128)	(0.128)
Spain	8.156***	8.148***
	(0.118)	(0.118)
Italy	8.284***	8.279***
	(0.118)	(0.118)
France	8.562***	8.556***
	(0.123)	(0.123)
rec_fem		-0.281
		(0.183)
N	1865	1865
r2	0.992	0.992

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

Table 13 shows the estimates of scarring effects in each country. The variables used are the same as in the previous regression, but in addition for each country there are dummy variables indicating the geographical area. These variables are *North*, *South* and *Centre*. They indicate in each country the amount of the mean expected income of individuals. These variables are created starting from the *Nuts1* variable of the dataset. In order to avoid problems of multicollinearity the *North* variable is excluded and the results of the other two geographical areas consider the *North* variable a benchmark.

Table 13: Results of OLS estimation of scarring effects by country

	France logincome	Germany logincome	Italy logincome	Spain logincome
yedu	0.0720***	0.116***	0.0858***	0.0523***
	(0.00759)	(0.0144)	(0.00959)	(0.0115)
y_work_w2	0.0247***	0.0190***	0.0160***	0.00841
	(0.00334)	(0.00526)	(0.00420)	(0.00560)
female	-0.304***	-0.310**	-0.278***	-0.428**
	(0.0591)	(0.0944)	(0.0812)	(0.132)
rec_edu	-0.129	0.345	0.217	-0.701*
	(0.133)	(0.199)	(0.160)	(0.303)
ever_unemp	-0.179*	-0.487***	-0.150	-0.0195
	(0.0761)	(0.111)	(0.131)	(0.171)
nchild	0.0176	-0.0776	-0.0423	0.0721
	(0.0243)	(0.0396)	(0.0414)	(0.0500)
single	-0.129	-0.141	-0.110	0.0379
	(0.0726)	(0.126)	(0.121)	(0.163)
Center_fra	-0.117			
	(0.0658)			
South_fr	-0.173*			
	(0.0746)			
Center_ger		0.194		
		(0.0996)		
South_ger		0.0774		
		(0.112)		
Center_ita			-0.105	
			(0.0967)	
South_ita			-0.0406	
			(0.0959)	
Center_esp				0.0860
				(0.136)
South_esp				-0.156
				(0.146)
_cons	8.425***	7.586***	8.273***	8.527***
	(0.178)	(0.337)	(0.218)	(0.271)
N	675	511	369	310
r2	0.236	0.243	0.254	0.146

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

As can be seen from Table 13 the variable *ever_unemp* appears to be significant in France and Germany. In particular, in France the effect of unemployment implies a reduction in income of 17.9% and in Germany the reduction is equal to 48.7%. In Italy and Spain the variable is not significant in any case providing negative coefficients (-15% in Italy and -1.95% in Spain). The effect that emerges is that the unemployment dummy is significant in the continental countries (France and Germany). In particular in Germany the scarring effect of unemployment seems to be very large, but this may partly be due to the fact that estimating the effects by countries generates a reduced size of the sample.

With regard to the variable *rec_edu*, it seems that it is not significant in all countries except for Spain where a recession at school leaving age implies a reduction in income equal to 70%. Even in this case the result could be less clear because in Spain there are only 310 observations, and the estimate is significant only for p-value 0.05. Moreover the variable has a large standard error (0.303). This could mean that the estimate is over-estimated, although in this country there is a scarring effect of recession. In France it seems that there is a reduction in income of 12.9% for individuals who have experienced a recession at school leaving age although it is not significant.

Finally, the estimates of the dummy variables indicating geographical areas do not provide significant results. However, we can note that, with the exception of Germany there are negative coefficients for the variables referring to the Centre and South of the countries, confirming the existence of geographical differences between these areas and the North.

An interesting result emerges from Italy where an unemployment spell and the recession are not significant. However this does not mean that scarring effects do not exist in this country. As Chapter 3 explains, unemployment or recession could provide negative effects in terms of higher future unemployment or low-skilled jobs even for those individuals who have high levels of education. Furthermore, as mentioned previously, although the variable *ever_unemp* is not significant in this country, it has a negative coefficient equivalent to -15%. Finally looking at the *female* variable, we can see that a woman having the same characteristics as a man earns less than a male by an amount equal to 30.4% in France, 31% in Germany, 27.8% in Italy and 42.8% in Spain.

In conclusion, the estimates in Tables 12 and 13 show that the effect of unemployment is large and significant for the Continental countries (France and Germany), whereas Spain seems to be more affected by recession periods. In general what has emerged is that an unemployment spell for an individual negatively affects income to a greater extent than a recession at school leaving age. However, in this thesis it is not possible to make a complete analysis of scarring effects because the entire wage profile of individuals is not available [16].

4.4. Differences among countries and possible measures

The empirical results obtained in the previous subsections show that individuals of different countries receive different wages according to the different variables used. The reason is mainly that each country has differences in the labour market, social insurance and the education system. However, in all the countries analyzed Mincer's assumptions and the consequences arising from scarring effects are satisfied.

In particular within the European Union the resources allocated for possible action are very different and they vary depending on the importance given in each country. In this way, the different welfare models are useful to better focus the framework at European level.

In Germany and France, although there has been some recession in the past, the youth unemployment rate has increased to a lesser extent than in Italy and Spain. This performance is certainly to be attributed to the particularly elastic transitions between school and work and the apprenticeship system of mass.

Educational programmes constitute a guarantee of inclusion in the labour market at the end of the education phase, also thanks to specific instruments such as, for example, an institutional network with firms. The choices of each country are contextualized within the different welfare systems in each country. Since the eighties, the difficulties faced by welfare systems (in terms of financial sustainability, ageing population, different

[16] In order to provide a further specification of the scarring effects in Appendix B estimates with longitudinal data are included. What emerges from this additional specification is that in Italy there is a negative effect of recession on income equal to 33.3%.

allocation of public expenditure determined by both economic difficulties and structural changes) have led to a reduction in allocated resources, employment policies in the form of unemployment benefits and specific measures aimed at the youth population (training contracts, support for entrepreneurship, etc.). Reduction of these costs has affected mostly the young individuals residing in the countries of Southern Europe (Italy and Spain). The regulation seems to offer more protection to insiders (workers employed with full-time and permanent contracts) compared to young outsiders (Lindbeck and Snower, 1988)[17].

On the other hand, social spending in France and Germany is mainly financed by contributions from workers and employers and acts as a protection against unemployment, but despite this, as demonstrated in chapter 4, even in these countries there are no satisfactory results in terms of the female employment rate.

For example, Germany finances several measures to support employment in the labour market and the long-term unemployed. Social spending in Italy is lower than continental countries and in particular it is differently distributed across sectors (Italy is certainly characterized by the importance given to pensions). Spain, given the high unemployment rate, is investing a lot in policies to support the labour market.

With regard to the education system, to ensure that education and training can play a decisive role in achieving the strategic goal set at the Lisbon European Council (to make the European Union the most competitive and dynamic 'knowledge economy" in the world), member States have to invest sufficient resources and ensure that these are addressed and managed in the most efficient way. In a particularly competitive and dynamic international context the investment policy in education and training should take into account the changing needs of society. In order to make the financial allocations for education more efficient, countries should address the areas of inefficiency of expense, such as high rates of school failure, early school leaving, graduate unemployment, etc.

[17] Insiders are employed in a stable way and are those who benefit from a more favourable and protected position than outsiders (the unemployed, inactive and looking for employment).

5. Conclusion

Through the use of "earnings functions" this thesis analyzes the incidence of several variables in the determination of wages for individuals in the labour market at global level and then splits the sample by countries. The countries examined in this thesis are France, Germany, Italy and Spain. The tool used is that of the Mincer equation (Mincer, 1974). Given the characteristics of this equation, the human capital theory requires special attention because the Mincer equation refers to investment in education and the work experience of individuals. Mincer identifies these elements of human capital as formal training (individual's education level) and informal training (human capital acquired on the job).
Successively the effects on earnings of an individual who has experienced a recession at the time of entry into the labour market or has experienced an unemployment spell have been analyzed in relation to human capital endowments. These adverse episodes provide scarring effects if there are negative effects on individual earnings.

The empirical estimation analyses are made using Stata through OLS estimations with cross section data. In addition to the classic variables of the Mincer equation relating to level of education and years of work experience, the following individual demographic variables have been introduced: gender, marital status, number of children and the geographic area of the individual. These same variables are used to determine the scarring effects in addition to those specific to their determination as a dummy variable indicating whether the individual has experienced a recession at school leaving age and a dummy variable indicating whether the individual has experienced an unemployment spell. Moreover, estimates of regressions pay great attention to gender differences, confirming that women are in a position of disadvantage compared to men.
The empirical analyses are conducted on SHARE data (Survey of Health, Ageing and Retirement in Europe), using in particular the data of the dataset in wave 2.

The results show differences between the different countries, according to existing differences in the labour market, social insurance and the education system.

However, in each country the expected results have been obtained according to statements from the human capital theory and the characteristics of scarring effects.
In effect, according to the human capital theory, investments in education and training have significant impact on the dynamics of the microeconomic labour market. This implies for workers that a return on human capital can be realized in the form of wage increases and greater employment stability.
In this regard the results obtained show that on average for the four countries the return to education expressed as an increase in income (log-income) resulting from an additional year of education is 7.83%. In particular it is 6.37% in France, 11.8% in Germany, 8.41% in Italy and 5.5% in Spain. On the other hand, an additional year of work experience implies on average an increase in income of 1.31%. In particular it is 1.3% in France, 1.84% in Germany, 1.38% in Italy and 0.67% in Spain. The results of the demographic variables are also in line with the relevant literature; on average women have lower earnings than men, the number of children has a negative effect on earnings for women, etc. The greater gender differences in earnings are in Spain where a woman having the same characteristics as a man earns on average 43.9% less than he. This difference is -35.6% in France, -39% in Germany and -28.3% in Italy, even if, as has been demonstrated, this difference is due mainly to the existence of children.
However, Italy is the country where there is the greatest gender gap in terms of employment levels and where the percentage of men employed is much higher than that of women employed. The countries where there is a lower gender difference in terms of employment levels are France and Germany. Moreover these are the countries where there are the highest percentages of total employed. Although the gender gap is consistent, it has been observed that the gap between women's and men's employment rates decreases as the education level of the individual grows.
Referring again to the Mincer equation, a return of educational qualifications in each country has been proposed. What has emerged is that the tertiary level of education (university) implies the highest return in each country.

Regarding scarring effects, the results obtained show greater evidence of adverse effects on income arising from having experienced an unemployment spell rather than a recession at school leaving age. The effect of unemployment is large and significant for

the Continental countries (France and Germany). In France an individual who has experienced an unemployment spell earns on average 17.9% less than an individual without unemployment spells and in Germany the reduction is 48.7%. Even in Italy there is a reduction of 15% and in Spain the earnings reduction is 1.95%, but these do not seem to be significant.
Having experienced a recession at school leaving age seems to be worrying in Spain, where the reduction of earnings is 70.1% and partly in France, where the reduction is 12.9%, but in this country the value is not significant.

This thesis provides an important contribution to existing literature. Most of the existing studies, especially those related to scarring effects, refer to the United States and there are few empirical estimates of this type for European countries. Moreover the use of the data of SHARE and SHARELIFE has provided a lot of specific information.
However, in this thesis it was not possible to make a complete analysis of scarring effects because the entire wage profile of individuals is not available.
For this reason, the thesis examines the scarring effects on earnings much later in an individual's life. In order to provide a further specification, Appendix B contains estimates of scarring effects made with longitudinal data with reference to SHARELIFE (wave3) data.
The reason why the scarring effects resulting from recession at school leaving age are not verified in all countries may reflect the literature. According to Oreopoulos (2008) these effects should decrease after 8-10 years.
The main causes of scarring effects can be explained by the concepts of the human capital theory: the duration of the last episode of unemployment leads to an effect of human capital loss and the total duration of all episodes of unemployment involves a loss of work experience, moreover employers could consider a large number of episodes of unemployment low motivation to work. In this way, in addition to an immediate effect on earnings and the probability of finding a job, unemployment may also cause a scarring effect, i.e. a long-term negative effect on income and the type of employment. This problem would be particularly severe and worrying in the case of young people, given their longer life perspective in the labour market. However the intuition coming from the statements by Oreopoulos lies in the fact that the effects of initial negative

shocks in the labour market provide permanent earnings differences only if combined with search frictions which increase with age or the quality of jobs.

The empirical results obtained in Chapter 4 therefore respect the theoretical assumptions underlying the human capital theory and scarring effects. However, interpreting the differences that exist between countries is not easy. A possible explanation could be given by labour market characteristics and the types of education systems adopted in each country. In a particularly competitive, dynamic international context, investment policies for education and training should take into account the changing needs of society. In order to make financial allocations for education more efficient, countries should address the areas of expense inefficiency, such as high rates of school failure, early school leaving, graduate unemployment, etc. The choice of policies to be adopted are contextualized within the different welfare systems in each country. For example Germany finances several measures to support employment in the labour market and for the long-term unemployed. Social spending in Italy is lower than in continental countries and it is particularly differently distributed across sectors (Italy is certainly characterized by the importance given to pensions). Given its high unemployment rate, Spain is investing a lot in policies to support the labour market.

The results obtained in this thesis suggest that understanding the origins of wage differences arising from a return to schooling and the consequences on wages of the adverse conditions in the labour market for different countries should be a particular focus for further research and policy development.

6. Appendices

The appendices provide some deeper aspects about the different concepts discussed in the course of the thesis. In particular Appendix A introduces estimates of the Mincer equation, highlighting to what extent having children affects women. In Appendix B there are estimates of scarring effects with the use of longitudinal data in reference to the SHARELIFE (wave3) dataset.

6.1. Appendix A

This appendix develops in detail the regressions in Chapter 4, including in the most important parts the interaction variable between the number of children (*nchild*) and the *female* variable. In this way it is possible to test the effect of having children for women.

For example, Table 14 refers to the regressions in Table 5 with the Mincer equation at global level, where column 1 shows years of work experience (*y_work_w2*), column 2 the non-linear variable *y_work_w2_sq*, column 3 the *age* variable instead of years of work experience and column 4 the non-linear variable age2. This table introduces the interaction variable *nchild_fem* which indicates the effect on earnings of women who have at least one child. In this way, the coefficient of the interaction terms measure the difference of the effects on earnings between men and women. The female coefficient does not currently reflect the effect determined by gender, given that another variable is also a function of this feature. For example in column 3 the estimate of the difference in income between a man and a woman who have two children can be computed as: – 0.187 – (0.153*2) + (0.0666*2) = -0.359. This means that a woman with the same characteristics as a man and two children earns on average 35.98% less than he. In the case in which a woman has only one child, she gains 27.34% less than a man with the same characteristics, whereas in the case of no children the gain is -18.7% (*female* variable). This means that the value equivalent to -48.3% obtained in column 3 of Table 5 for the *female* variable, includes the effect of having children. Another interesting aspect that emerges from Table 14 is that considering years of work experience in

columns 1 and 2, including the variable *nchild_fem* means that the *female* variable becomes insignificant. The fact that the *female* variable is not significant means that being a female is important only in relation to the existence (number) of children and not as a female.

However, even if the *female* variable is not significant in column 1, results similar to those described previously are obtained: if the woman has two children, she earns -25.82% and -19.81% in the case of one child, and -13.8% if she has no children, even if in this last case there is no strong evidence because the *female* variable is not significant.

.

Table 14: Results of OLS estimation of the Mincer equation for the first specification with the interaction variable

	(1) logincome	(2) logincome	(3) logincome	(4) logincome
yedu	0.0782***	0.0771***	0.0741***	0.0749***
	(0.00528)	(0.00537)	(0.00528)	(0.00523)
y_work_w2	0.0111***	0.0188**		
	(0.00186)	(0.00719)		
female	-0.138	-0.138	-0.187*	-0.179*
	(0.0861)	(0.0861)	(0.0862)	(0.0854)
single	0.00256	0.00577	0.0396	0.0245
	(0.0604)	(0.0605)	(0.0612)	(0.0606)
nchild	0.0649*	0.0673*	0.0666*	0.0684**
	(0.0261)	(0.0262)	(0.0263)	(0.0260)
nchild_fem	-0.125**	-0.127***	-0.153***	-0.150***
	(0.0381)	(0.0382)	(0.0382)	(0.0378)
y_work_w2_sq		-0.000173		
		(0.000157)		
age			-0.0245***	0.405***
			(0.00485)	(0.0753)
age2				-0.00388***
				(0.000678)
_cons	8.375***	8.328***	10.10***	-1.758
	(0.111)	(0.119)	(0.284)	(2.091)
N	1589	1589	1589	1589
r2	0.188	0.189	0.183	0.200

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

Finally, as can be seen from the results in Table 5, the other variables have not changed significantly.

Table 15 shows the results obtained by including the interaction variable by country. The variable *nchild_fem* has a higher incidence in the Continental countries (France and Germany), whereas in Italy and Spain the effect is not significant. The results show that in France women earn on average -27.34% than men if they have two children and -19.52% if they have one child. In Germany with two children, the reduction in earnings for women is -28.59% and -14.11% in the case of one child.

On the other hand, in Italy the reduction in earnings for women who have two children is -34.92% and in the case of one child -30.26%, whereas in Spain the reduction is -23.6% in the case of two children and -23.5% in the case of one child.

Table 15: Results of OLS estimation of the Mincer equation for the first specification with the interaction variable by country

	France logincome	Germany logincome	Italy logincome	Spain logincome
yedu	0.0630***	0.117***	0.0842***	0.0552***
	(0.00898)	(0.0148)	(0.00974)	(0.0118)
y_work_w2	0.0125***	0.0183***	0.0137***	0.00605
	(0.00272)	(0.00404)	(0.00393)	(0.00463)
female	-0.117	0.00363	-0.256	-0.234
	(0.135)	(0.178)	(0.168)	(0.219)
single	-0.0720	-0.0784	-0.157	0.147
	(0.0826)	(0.134)	(0.131)	(0.165)
nchild	0.0348	0.0692	-0.0310	0.106
	(0.0385)	(0.0564)	(0.0588)	(0.0630)
nchild_fem	-0.113*	-0.214**	-0.0156	-0.107
	(0.0550)	(0.0811)	(0.0850)	(0.0933)
_cons	8.834***	7.493***	8.331***	8.529***
	(0.175)	(0.294)	(0.213)	(0.243)
N	536	452	339	262
r2	0.188	0.249	0.247	0.133

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARE release 2.5.0 (wave 2, 2006/07); computations by the author.

The result obtained in Spain is interesting because it seems that there is no difference related to the number of children, which means that on average the reduction in earnings is due to having children, independently of the number. However, the results obtained in these two countries are not significant.

6.2. Appendix B

This appendix provides estimates of scarring effects with longitudinal data using the dataset of Sharelife (wave3). In this way it is possible to give a further interpretation thanks to the available information about the entire past of an individual. In this case, with longitudinal data, the estimation method is different. The longitudinal data or panel data combine information relating to the characteristics of N individuals in the same time instant with those observed for the same individuals in different periods of time (Verbeek, 2004).

$$\underset{(N\,x\,T)}{Y} = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{i1} & \cdots & y_{n1} \\ y_{12} & y_{22} & \cdots & y_{i2} & \cdots & y_{N2} \\ \cdots & \cdots & & \cdots & & \cdots \\ y_{1t} & y_{2t} & \cdots & y_{it} & \cdots & y_{Nt} \\ \cdots & \cdots & & \cdots & & \cdots \\ y_{1T} & y_{2T} & \cdots & y_{iT} & \cdots & y_{NT} \end{pmatrix} \quad (22)$$

As this type of regression is characterized by longitudinal data, it requires a different estimation method from that used for the regressions carried out with cross section data. The estimation method used is a random effects model due to the fact that the random effects model treats individual effects as part of the error term, therefore the individual effects are treated as stochastic components uncorrelated with the regressors. In this way it is possible to include variables that change among individuals, despite the fact that they remain constant within the T observations related to the individual.

The estimation is first carried out at global level, then the estimates are split by countries. The variables are the same as those used in the regressions with cross section data, but in this case the number of observations is much greater because for each variable there is an observation for every year of the individual's past.
In particular Table 16 provides empirical results at global level.

Table 16: Results of the estimation of scarring effects with longitudinal data

	(1) log_wage	(2) log_wage
yedu	0.0612***	0.0612***
	(0.00472)	(0.00472)
female	-0.554***	-0.546***
	(0.0395)	(0.0416)
rec_edu	-0.0717	-0.0721
	(0.0654)	(0.0654)
ever_unemp	0.147*	0.189*
	(0.0650)	(0.0922)
nchildren	0.0159	0.0160
	(0.0151)	(0.0151)
single	-0.0693	-0.0698
	(0.0482)	(0.0482)
Spain	-0.782***	-0.782***
	(0.0652)	(0.0652)
Italy	-0.863***	-0.863***
	(0.0591)	(0.0591)
France	0.433***	0.433***
	(0.0561)	(0.0561)
unemp_fem		-0.0836
		(0.130)
_cons	8.699***	8.695***
	(0.0841)	(0.0844)
N	162898	162898

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARELIFE release 1 (wave 3, 2008/09); computations by the author.

In column 1 of Table 16 the variable *rec_edu* is not significant. However, it indicates a reduction in income resulting from a recession at school leaving age equal to 7.17%. The variable *ever_unemp* instead is significant for p-value 0.05. The interesting fact is that this variable has a positive coefficient, whereas a negative coefficient was expected. The reason could arise from the fact that, as explained in Chapter 3, sometimes a period of unemployment can lead to positive results, because it makes it possible to find a better occupation, especially for the most qualified individuals.
However, to verify whether the positive coefficient exists only for men, column 2 includes the interaction variable *unemp_fem*. This variable has been generated by interaction between the variable *ever_unemp* and the *female* variable, in order to verify the effect of unemployment on women only. In this a case negative coefficient is actually obtained, but it is not significant enough to be included in the model. It would suggest that the negative effect of unemployment has a negative impact on women.

Table 17 shows the empirical results of scarring effects with longitudinal data by country. The interesting data emerging from this table seems to concern France and Italy for which the regressions with cross-section data of Table 13 do not provide significant results. In particular, in Italy an individual who has experienced a recession at school leaving age earns on average 33.3% less than an individual having the same characteristics, but who has not experienced a period of recession. In France the variable *rec_edu* has a positive coefficient, as for the variable *ever_unemp* in Italy and Spain. Although not significantly, Germany shows scarring effects of recession equal to -12% and negative effects of unemployment equal to -1.74%.
Finally Table 17 shows that the geographic variables are not significant, except for the South of Italy where an individual earns on average 37.2% less than an individual in Northern Italy having the same characteristics, confirming the strong geographical disparities that exist in this country.

Table 17: Results of the estimation of scarring effects with longitudinal data by country

	France log_wage	Germany log_wage	Italy log_wage	Spain log_wage
yedu	0.0424***	0.0790***	0.0630***	0.0617***
	(0.00893)	(0.00639)	(0.0102)	(0.0100)
female	-0.655***	-0.669***	-0.216*	-0.833***
	(0.0703)	(0.0405)	(0.0899)	(0.0979)
rec_edu	0.294**	-0.120	-0.333*	-0.237
	(0.113)	(0.0907)	(0.148)	(0.134)
ever_unemp	-0.0362	-0.0174	0.473**	0.443*
	(0.106)	(0.0604)	(0.165)	(0.181)
nchildren	0.0189	0.0282*	0.00929	0.0290
	(0.0246)	(0.0142)	(0.0440)	(0.0346)
single	0.131	-0.0712	-0.323**	-0.115
	(0.0765)	(0.0487)	(0.120)	(0.121)
Center_fra	-0.0960			
	(0.0814)			
South_fra	0.0969			
	(0.0851)			
Center_ger		0.0550		
		(0.0457)		
South_ger		-0.0486		
		(0.0528)		
Center_ita			-0.0163	
			(0.109)	
South_ita			-0.372***	
			(0.106)	
Center_esp				0.118
				(0.122)
South_esp				-0.113
				(0.114)
_cons	9.310***	8.518***	7.856***	8.021***
	(0.141)	(0.0972)	(0.153)	(0.147)
N	46299	35647	45353	28884

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: This table uses data from SHARELIFE release 1 (wave 3, 2008/09); computations by the author.

7. References

Arrow K.J., (1973), "The Theory of Discrimination", in O. Ashenfelter and A. Rees (Eds.), *Discrimination in Labor Markets*, Princeton, NJ: Princeton University Press, 1-24.

Becker E. and S. Hills., (1979), "Teenage unemployment: Some evidence of the long run effects", *Journal of Human Resources 15(3),* 354-372.

Becker G., (1964), *Human Capital,* New York, Columbia University Press.

Becker G. and K. Murphy., (1988), *Economic Growth, Human Capital and Population Growth*, New York, Buffalo.

Becker G., (1993), *Human capital: a theoretical and empirical analysis, with special reference to education,* 3nd ed. Chicago ; London: University of Chicago press, 390 pages.

Ben-Porath Y., (1967), "The Production of Human Capital and the Life Cycle of Earnings", *Journal of Political Economy 75,* 352-365.

Berkovich E., (1985), "Reputation Effect in Equilibrium Search and Bargaining. A Stigma Theory of Unemployment Duration", in *Northwestern University, Center for Mathematical Studies in Economics and Management Science, Discussion paper n. 668,* 43 pages.

Blanchard O.J. and P. Diamond., (1994), "Ranking, Unemployment, Duration, and Wages", *Review of Economic Studies 61,* 417-434.

Borjas G.J., (1996), *Labor economics,* New York, McGraw-Hill, 220-248.

Brunello G. and S. Comi., (2000), "Education and Earnings Growth: Evidence from 11 European Countries", *Institute for the Study of Labor, Discussion paper n. 140,* 2-23.

Budria S. and P.T. Pereira., (2005), "Educational Qualifications and wage inequality: evidence for Europe", *Institute for the Study of Labor, Discussion paper n. 1763*.

Clark K.B. and L.H. Summers., (1982), "The Dynamics of Youth Unemployment", in R. Freeman and D. Wise (Eds.), *The Youth Labor Market Problem: Its Nature, Causes, and Consequences,* Chicago; London: The university of Chicago press, 199-234.

Cockx B. and M. Picchio., (2011), "Scarring Effects of Remaining Unemployed for Long-Term Unemployed School-Leavers", *Institute for the Study of Labor, Discussion paper n. 5937,* 4-33.

Coe D.T. and D.J. Snower., (1997), "Policy Complementarities: the Case for Fundamental Labor Market Reform", *IMF Staff Papers 44(1)*, 1-35.

Christopoulou R., (2008), "The Youth Labor Market Problem in Cross-Country Perspective" in G. DeFreittas (Ed.), *Youths in the Global Economy*, New York: Edward Elgar.

Dagum C., (1994), "Human Capital, Income and Wealth Distribution Models and their Applications to the U.S.A.", *Business and Economics Section 154th meeting*, Toronto, 253-258.

Doiron D. and T. Gørgens., (2008), "State Dependence in Youth Labour Market Experiences, and the Evaluation of Policy Interventions", *Journal of Econometrics 145(1/2)*, 81-97.

Ellwood D., (1982), "Teenage Unemployment: Permanent Scars or Temporary Blemishes", in R. Freeman and D. Wise (Eds.), *The Youth Labor Market Problem: Its Nature, Causes, and Consequences*, Chicago; London: The university of Chicago press, 349-390.

Freeman R., (1986), "Demand for education", in O.Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics vol. I*, North-Holland, New York, 357-386.

Friedman M., (1955), "The role of government in education", in R. Solo (Ed.), *Economics and the Public Interest*, New Brunswick, Rutgers University Press, 123-144.

Garrouste C., (2010), "100 Years of Educational Reforms in Europe: a contextual database", *JRC Scientific and Technical Reports*, Luxembourg, Publications Office of the European Union, 349 pages.

Goldin C., (1990), *Understanding the gender gap: an economic history of American women*, New York ; Oxford: Oxford university press, 83-117.

Gornick J.C., (1999), "Gender Equality in the Labour Market: Women's Employment and Earnings", *Luxembourg income study, Working paper n. 206*, 1-48.

Gregg P. and E. Tominey, (2004), "The Wage Scar from Youth Unemployment", *Working Paper CMPO n. 04(097)*, 2-24.

Gregory M. and R. Jukes., (2001), "Unemployment and subsequent earnings: estimating scarring among British men 1984-94", *Economic Journal 111(475)*, 607-25.

Hashimoto M., (1981), "Firm-specific human capital as a shared investment", *American economic Review 71*, 475-482.

Heckman J.J. and B. Singer., (1984), "A Method of Minimizing the Impact of Distributional Assumptions for Duration Data", *Econometrica 52*, 271–320.

Jimeno J.F. and D. Rodriguez-Palenzuela., (2002), "Youth Unemployment in the OECD: Demographic Shifts, Labour Market Institutions and Macroeconomic Shocks", *European Central Bank, Working paper n. 155*, 17-30.

Jorgenson D.W. and B.M. Fraumeni., (1989), "The accumulation of human and nonhuman capital", in Lipsey R.E. and T.H. Stone (Eds.), *The Measurement of Saving, Investment, and Wealth*, Chicago, University of Chicago Press, 227-282.

Kahn L., (2006), "The Long-Term Labor Market Consequences of Graduating College in a Bad Economy", *Labour Economics 17(2)*, 3-47.

Kondo A., (2008), *Differential Effects of Graduating during Recessions across Race and Gender*, Mimeo, Columbia University.

Lancaster T., (1979), "Econometrics Methods for the Duration of Unemployment", *Econometrica 47(4)*, 939-956.

Lindbeck A. and D. Snower., (1988), *The insider-outsider theory of employment and unemployment*, Mitpress, Cambridge.

Lindbeck A., (1996), "The West European Employment Problem", *Weltwirtschaftliches Archiv vol. 132*, 609-37.

Miller H., (1960), "Annual and Lifetime Income in Relation to Education", *American Economic Review 50*, 962-986.

Mincer J., (1958), "Investment in Human Capital and Personal Income Distribution", *Journal of Political Economy 66*, 281-302.

Mincer J., (1974), *Schooling experience and earnings*, Columbia University Press, New York, 41-96.

Mincer J., (1991), "Education and Unemployment of Women", *Working Paper NBER, n. 3837*, 3-23.

Mincer J., (1991), "Education and Unemployment", *Working Paper NBER n. 3838*, 2-32.

Mroz T.A. and T.H. Savage., (2006), "The Long-Term Effects of Youth Unemployment", *Journal of Human Resources 41(2)*, 259-293.

Nickell S., (1979), "Estimating the Probability of Leaving Unemployment", *Econometrica 47(5)*, 1249-1266.

Nilsen Ø.A. and K.H. Reiso,. (2011), "Scarring Effects of Unemployment", *Institute for the Study of Labor, Discussion paper n. 6198*, 1-33.

OECD documents., (1996), *Education and training : learning and working in a society in flux*, Paris, Organisation for economic co-operation and development, 40 pages.

OECD documents., (1998), *Human Capital Investment, An International Comparison*, Paris, Organisation for economic co-operation and development, 7-92.

Oreopoulos P., T. Von Wachter and A. Heisz., (2008), "The Short- and Long-Term Career Effects of Graduating in a Recession: Hysteresis and Heterogeneity in the Market for College Graduates", *Institute for the Study of Labor, Discussion paper n. 3578*, 2-38.

O'Higgins N., (2005) "Trends in the Youth Labour Market in Developing and Transition Countries," *EconWPA 0507002*, 1-45.

Ouyang M., (2005), *The scarring effect of recessions*, University of Maryland at College Park, 2-46.

Pissarides C., (1994), "Search unemployment with on-the-job search", *Review of Economic Studies vol. 61*, 457-75.

Polacheck S.W., (1974), "Differences in expected post-school investment as determinant of market wage differentials", in *International Economic review, vol. 16*, 451-470.

Polachek S.W. and W.S. Siebert., (1993), *The economics of earnings*, Cambridge: Cambridge university press, 367 pages.

Psacharopoulos G., (1985), "Returns to education: a further international update and implications", *Journal of Human Resources 20*, 583-604.

Psacharopoulos G., (1994), "Returns to investment in education: a global update", *World Development 22(9)*, 1325-43.

Psacharopoulos G. and H.A. Patrinos., (2004), "Returns to investments in education. A further update", *Education Economics 12(2)*, 111-34.

Rubery J., C. Fagan and M. Smith., (1995), "Changing Patterns of Work and Working-time in the European Union and the Impact on Gender Divisions", *report for the Equal Opportunities Unit, DGV*, Brussels, European Commission.

Ruhm C., (1991), "Are Workers Permanently Scarred by Job Displacement?", *American Economic Review 81(1)*, 319-324.

Ryan P., (2001), "The school-to-work transition. A Cross-National Perspective", *Journal of Economic Literature vol. 39*, 3-67.

Schmid G, J. O'Reilly and K. Schomann., (1997), *International handbook of labour market policy and evaluation*, Cheltenham, UK ; Northampton, Mass: Edward Elgar, 954 pages.

Schultz T., (1963), *The Economic Value of Education*, Columbia University Press, New York, 92 pages.

Spence M., (1973), "Job Market Signaling", *Quarterly Journal of Economics 87(3)*, 355-74.

Stevens A.H., (1997), "Persistent effects of job displacement: the importance of multiple job losses", *Journal of Labor Economics vol. 15(1)*, 165-88.

Stevenson W., (1978), "The relationship between early work experience and future employability", in Adams A. and G. Mangum (Eds.), *The lingering crisis of youth unemployment*, Kalamazoo: Upjohn Institute for Empirical Research.

Topel R.H. and M.P. Ward., (1992), "Job Mobility and the Careers of Young Men", *Quarterly Journal of Economics 107(2)*, 439-479.

Verbeek M., (2004), *A Guide to Modern Econometrics*, 2nd Edition**,** John Wiley & Sons.

Wasmer E., P. Fredriksson, A. Lamo, J. Messina and G. Peri., (2007), "The Macroeconomics of Education", in G. Brunello, P. Garibaldi and E. Wasmer (Eds.), *Education and training in Europe*, Oxford; New York: Oxford university press, 36-118.

Wößmann L., (2003), "Specifying Human Capital", *Journal of Economic Surveys 17(3)*, 239-70.

8. The Stata codes

This section of the thesis contains the main parts of the code implemented with Stata to achieve empirical results.

In the first part there are the necessary commands to produce a descriptive analysis and the results of the Mincer equation, and in the second part there are the commands related to the determination of scarring effects. Both parts are combined with the necessary comments.

```
clear all
set mem 800m
set more off
capture log close

cd "C:\Users\utente\Desktop\Esami\TesiLaurea\workspace\Nuovacartella"

log using mincer_equation.log,replace
use wave2_ms.dta, clear
*-------------------------------------------------------------------------------

tab country
tab country, nolabel

gen     iso = "GER" if country == 12
replace iso = "ESP" if country == 15
replace iso = "ITA" if country == 16
replace iso = "FRA" if country == 17

gen age = 2006-yrbirth if yrbirth !=.

*generation of the age groups
gen     clage = 1 if age<=55
replace clage = 2 if age>55 & age<=60
replace clage = 3 if age>60 & age<=67
replace clage = 4 if age>67 & age<=75

gen workforce =(ep005 >1 & ep005 !=.)

gen     worker = 1 if ep005 == 2
replace worker = 0 if worker==.

gen logincome = log(ydipv)

tabulate gender, generate(sex)

rename sex1 male
rename sex female

gen logincomeM = log(ydipv) if gender == 1
gen logincomeF = log(ydipv) if gender == 2

*isced classifications
gen     isced = 0 if isced_r == 0
replace isced = 1 if isced_r == 1
```

```
46 replace isced = 2 if isced_r == 2
47 replace isced = 3 if isced_r == 3
48 replace isced = 4 if isced_r == 4
49 replace isced = 5 if isced_r == 5
50 replace isced = 6 if isced_r == 6
51
52 *years of work experience
53 gen     age_exp = 1 if y_work>10 & y_work<=20
54 replace age_exp = 2 if y_work>20 & y_work<=25
55 replace age_exp = 2 if y_work>25 & y_work<=30
56 replace age_exp = 4 if y_work>30 & y_work<=35
57 *-------------------------------------------------------------------------
58
59 ****************************************************
60 *The following section shows the descriptive analysis*
61 ****************************************************
62
63 *1°: descriptive analysis by age groups and levels of employment
64
65 preserve
66 *in this way it is considered the percentage of employed on the labor force
   and the retired are excluded
67 keep if workforce == 1
68
69 *percentage of total employment
70 collapse (mean) meanworker= worker (sd) sdworker=worker (count) n=worker,
   by(clage)
71
72 generate hiworker = meanworker + invttail(n-1,0.025)*(sdworker / sqrt(n))
73 generate loworker = meanworker - invttail(n-1,0.025)*(sdworker / sqrt(n))
74
75
76 twoway (bar meanworker clage if clage==1, fcolor(orange)lcolor(black)) ///
77 (bar meanworker clage if clage==2, fcolor(gs12)lcolor(black)) ///
78 (bar meanworker clage if clage==3, fcolor(red)lcolor(black)) ///
79 (bar meanworker clage if clage==4, fcolor(gs6)lcolor(black)) ///
80 (rcap hiworker loworker clage), ///
81 ylabel(0 .2 "20" .4 "40" .6 "60" .8 "80") ytitle(Percentage of employed)
   xtitle("") ///
82 xlabel(, nolabels noticks) ///
83 legend( rows(1)order(1 "<55" 2 "55 - 60" 3 "60 - 67" 4 "67 - 75"))
84 graph save percentage_total_employment.gph, replace
85
86 restore
87
88 preserve
89 keep if clage !=.
90 keep if workforce == 1
91
92 *percentage of total employment by country
93 collapse (mean) meanworker= worker (sd) sdworker=worker (count) n=worker,
   by(clage country)
94
95 generate hiworker = meanworker + invttail(n-1,0.025)*(sdworker / sqrt(n))
96 generate loworker = meanworker - invttail(n-1,0.025)*(sdworker / sqrt(n))
97
98 generate agecountry = clage    if country == 12
99 replace  agecountry = clage+5  if country == 15
100 replace agecountry = clage+10 if country == 16
101 replace agecountry = clage+15 if country == 17
102 sort agecountry
103
104
105 twoway (bar meanworker agecountry if clage==1,
    fcolor(orange)lcolor(black)) ///
```

```
106 (bar meanworker agecountry if clage==2, fcolor(gs12)lcolor(black)) ///
107 (bar meanworker agecountry if clage==3, fcolor(red)lcolor(black)) ///
108 (bar meanworker agecountry if clage==4, fcolor(gs6)lcolor(black)) ///
109 (rcap hiworker loworker agecountry), ///
110 ylabel(0 .2 "20" .4 "40" .6 "60" .8 "80") ytitle(Percentage of employed)
    xtitle("") ///
111 legend(row(1) order(1 "<55" 2 "55 - 60" 3 "60 - 67" 4 "67 - 75") ) ///
112 xlabel( 2.5 "Germany" 7.5 "Spain" 12.5 "Italy" 17.5 "France", noticks)
113 graph save percentage_employment_country.gph, replace
114
115 restore
116
117 preserve
118 keep if clage !=.
119 drop if clage == 4
120 keep if workforce == 1
121
122 *percentage of total employment by gender
123 collapse (mean) meanworker= worker (sd) sdworker=worker (count) n=worker,
    by(gender clage)
124
125 generate hiworker = meanworker + invttail(n-1,0.025)*(sdworker / sqrt(n))
126 generate loworker = meanworker - invttail(n-1,0.025)*(sdworker / sqrt(n))
127
128 generate genderage = gender  if clage == 1
129 replace genderage = gender+3 if clage == 2
130 replace genderage = gender+6 if clage == 3
131 sort    genderage
132
133
134 twoway (bar meanworker genderage if gender==1) ///
135 (bar meanworker genderage if gender==2, fcolor(pink)) ///
136 (rcap hiworker loworker genderage), ///
137 ylabel(0 .2 "20" .4 "40" .6 "60" .8 "80") ytitle(Percentage of employed)
    xtitle("") ///
138 legend(row(1) order(1 "Male" 2 "Female") ) ///
139 xlabel( 1.5 "Under 55" 4.5 "55-60" 7.5 "60-67", noticks)
140 graph save percentage_employment_gender.gph, replace
141
142 restore
143
144 preserve
145 keep if clage !=.
146 drop if clage == 3
147 drop if clage == 4
148
149 *percentage of total employment by gender and country
150 keep if workforce == 1
151
152 collapse (mean) meanworker= worker (sd) sdworker=worker (count) n=worker,
    by(gender clage country)
153
154 generate hiworker = meanworker + invttail(n-1,0.025)*(sdworker / sqrt(n))
155 generate loworker = meanworker - invttail(n-1,0.025)*(sdworker / sqrt(n))
156
157 generate gendercountry = 1  if gender == 1 & clage == 1 & country == 12
158 replace  gendercountry = 2  if gender == 2 & clage == 1 & country == 12
159 replace  gendercountry = 4  if gender == 1 & clage == 2 & country == 12
160 replace  gendercountry = 5  if gender == 2 & clage == 2 & country == 12
161 replace  gendercountry = 8  if gender == 1 & clage == 1 & country == 15
162 replace  gendercountry = 9  if gender == 2 & clage == 1 & country == 15
163 replace  gendercountry = 11 if gender == 1 & clage == 2 & country == 15
164 replace  gendercountry = 12 if gender == 2 & clage == 2 & country == 15
165 replace  gendercountry = 15 if gender == 1 & clage == 1 & country == 16
166 replace  gendercountry = 16 if gender == 2 & clage == 1 & country == 16
```

```
167 replace  gendercountry = 18 if gender == 1 & clage == 2 & country == 16
168 replace  gendercountry = 19 if gender == 2 & clage == 2 & country == 16
169 replace  gendercountry = 22 if gender == 1 & clage == 1 & country == 17
170 replace  gendercountry = 23 if gender == 2 & clage == 1 & country == 17
171 replace  gendercountry = 25 if gender == 1 & clage == 2 & country == 17
172 replace  gendercountry = 26 if gender == 2 & clage == 2 & country == 17
173
174
175 twoway (bar meanworker gendercountry if gender==1) ///
176 (bar meanworker gendercountry if gender==2, fcolor(pink)) ///
177 (rcap hiworker loworker gendercountry), ///
178 ylabel(0 .2 "20" .4 "40" .6 "60" .8 "80") ytitle(Percentage of employed)
    xtitle("") ///
179 legend(row(1) order(1 "Male" 2 "Female") ) ///
180 xlabel( 1.5 "Under 55" 4.5 "55-60" 8.5 "Under 55" 11.5 "55-60" 15.5 "Under
   55" 18.5 "55-60" 22.5 "Under 55" 25.5 "55-60" , noticks)
181 graph save percentage_employment_country&gender.gph, replace
182
183 restore
184
185 *2°: descriptive analysis by education and levels of employment
186
187 preserve
188 keep if workforce == 1
189
190 *percentage of total employment by education levels and gender
191 collapse (mean) meanworker= worker (sd) sdworker=worker (count) n=worker,
    by(gender isced)
192
193 generate hiworker = meanworker + invttail(n-1,0.025)*(sdworker / sqrt(n))
194 generate loworker = meanworker - invttail(n-1,0.025)*(sdworker / sqrt(n))
195
196
197 generate genderisced = gender    if isced == 0
198 replace  genderisced = gender+3  if isced == 1
199 replace  genderisced = gender+6  if isced == 2
200 replace  genderisced = gender+9  if isced == 3
201 replace  genderisced = gender+12 if isced == 4
202 replace  genderisced = gender+15 if isced == 5
203 replace  genderisced = gender+18 if isced == 6
204 sort     genderisced
205
206
207 twoway (bar meanworker genderisced if gender==1) ///
208 (bar meanworker genderisced if gender==2, fcolor(pink)) ///
209 (rcap hiworker loworker genderisced), ///
210 ylabel(0 .2 "20" .4 "40" .6 "60" .8 "80") ytitle(Percentage of employed)
    xtitle("") ///
211 legend(row(1) order(1 "Male" 2 "Female") ) ///
212 xlabel( 1.5 "No education" 4.5 "Primary education" 7.5 "Lower secondary
    education" 10.5 "Upper secondary education" 13.5 "Post-secondary
    education" 16.5 "University" 19.5 "Post-graduate university", noticks)
213 graph save percentage_employment_education&gender.gph, replace
214
215 restore
216
217 preserve
218 keep if country == 17
219 keep if workforce == 1
220
221 *Percentage of total employment by education levels and gender in France
222 gen     isced_fra = 0 if isced == 0
223 replace isced_fra = 1 if isced == 1
224 replace isced_fra = 2 if isced == 2
225 replace isced_fra = 3 if isced == 3
```

```
replace isced_fra = 4 if isced == 5
replace isced_fra = 5 if isced == 6

keep if isced_fra !=.

collapse (mean) meanworker= worker (sd) sdworker=worker (count) n=worker, by(gender isced_fra)

generate hiworker = meanworker + invttail(n-1,0.025)*(sdworker / sqrt(n))
generate loworker = meanworker - invttail(n-1,0.025)*(sdworker / sqrt(n))

generate genderisced = gender    if isced_fra == 0
replace  genderisced = gender+3  if isced_fra == 1
replace  genderisced = gender+6  if isced_fra == 2
replace  genderisced = gender+9  if isced_fra == 3
replace  genderisced = gender+12 if isced_fra == 4
replace  genderisced = gender+15 if isced_fra == 5
sort     genderisced


twoway (bar meanworker genderisced if gender==1) ///
(bar meanworker genderisced if gender==2, fcolor(pink)) ///
(rcap hiworker loworker genderisced), ///
ylabel(0 .2 "20" .4 "40" .6 "60" .8 "80") ytitle(Percentage of employed) xtitle("") ///
legend(row(1) order(1 "Male" 2 "Female") ) ///
xlabel( 1.5 "No education" 4.5 "Primary education" 7.5 "Lower secondary education" 10.5 "Upper secondary education" 13.5 "University" 16.5 "Post-graduate university", noticks)
graph save percentage_employment_education_france.gph, replace

restore

preserve
keep if country == 12
keep if workforce == 1

*percentage of total employment by education levels and gender in Germany
gen isced_ger = 0     if isced == 2
replace isced_ger = 1 if isced == 3
replace isced_ger = 2 if isced == 4
replace isced_ger = 3 if isced == 5

keep if isced_ger !=.
keep if gender !=.
keep if workforce !=.

collapse (mean) meanworker= worker (sd) sdworker=worker (count) n=worker, by(gender isced_ger)

generate hiworker = meanworker + invttail(n-1,0.025)*(sdworker / sqrt(n))
generate loworker = meanworker - invttail(n-1,0.025)*(sdworker / sqrt(n))

generate genderisced = gender   if isced_ger == 0
replace  genderisced = gender+3 if isced_ger == 1
replace  genderisced = gender+6 if isced_ger == 2
replace  genderisced = gender+9 if isced_ger == 3
sort     genderisced



twoway (bar meanworker genderisced if gender==1) ///
(bar meanworker genderisced if gender==2, fcolor(pink)) ///
(rcap hiworker loworker genderisced), ///
```

```
285 ylabel(0 .2 "20" .4 "40" .6 "60" .8 "80") ytitle(Percentage of employed)
    xtitle("") ///
286 legend(row(1) order(1 "Male" 2 "Female") ) ///
287 xlabel( 1.5 "Primary education" 4.5 "Upper secondary education" 7.5 "Post-
    secondary education" 10.5 "University", noticks)
288 graph save percentage_employment_education_germany.gph, replace
289
290 restore
291
292 preserve
293
294 keep if isced !=.
295 keep if country == 16
296 keep if workforce == 1
297
298 *percentage of total employment by education levels and gender in Italy
299 collapse (mean) meanworker= worker (sd) sdworker=worker (count) n=worker,
    by(gender isced)
300
301 generate hiworker = meanworker + invttail(n-1,0.025)*(sdworker / sqrt(n))
302 generate loworker = meanworker - invttail(n-1,0.025)*(sdworker / sqrt(n))
303
304 generate genderisced = gender    if isced == 0
305 replace  genderisced = gender+3  if isced == 1
306 replace  genderisced = gender+6  if isced == 2
307 replace  genderisced = gender+9  if isced == 3
308 replace  genderisced = gender+12 if isced == 4
309 replace  genderisced = gender+15 if isced == 5
310 replace  genderisced = gender+18 if isced == 6
311 sort     genderisced
312
313
314 twoway (bar meanworker genderisced if gender==1) ///
315 (bar meanworker genderisced if gender==2, fcolor(pink)) ///
316 (rcap hiworker loworker genderisced), ///
317 ylabel(0 .2 "20" .4 "40" .6 "60" .8 "80") ytitle(Percentage of employed)
    xtitle("") ///
318 legend(row(1) order(1 "Male" 2 "Female") ) ///
319 xlabel( 1.5 "No education" 4.5 "Primary education" 7.5 "Lower secondary
    education" 10.5 "Upper secondary education" 13.5 "Post-secondary
    education" 16.5 "University" 19.5 "Post-graduate university", noticks)
320 graph save percentage_employment_education_germany.gph, replace
321
322 restore
323
324 preserve
325 keep if country == 15
326 keep if workforce == 1
327
328 *percentage of total employment by education levels and gender in Spain
329 gen     isced_esp = 0 if isced == 0
330 replace isced_esp = 1 if isced == 1
331 replace isced_esp = 2 if isced == 2
332 replace isced_esp = 3 if isced == 3
333 replace isced_esp = 4 if isced == 5
334
335 keep if isced_esp !=.
336
337 collapse (mean) meanworker= worker (sd) sdworker=worker (count) n=worker,
    by(gender isced_esp)
338
339 generate hiworker = meanworker + invttail(n-1,0.025)*(sdworker / sqrt(n))
340 generate loworker = meanworker - invttail(n-1,0.025)*(sdworker / sqrt(n))
341
342 generate genderisced = gender    if isced_esp == 0
```

```
343 replace  genderisced = gender+3  if isced_esp == 1
344 replace  genderisced = gender+6  if isced_esp == 2
345 replace  genderisced = gender+9  if isced_esp == 3
346 replace  genderisced = gender+12 if isced_esp == 4
347 sort     genderisced
348
349
350 twoway (bar meanworker genderisced if gender==1) ///
351 (bar meanworker genderisced if gender==2, fcolor(pink)) ///
352 (rcap hiworker loworker genderisced), ///
353 ylabel(0 .2 "20" .4 "40" .6 "60" .8 "80") ytitle(Percentage of employed)
    xtitle("") ///
354 legend(row(1) order(1 "Male" 2 "Female") ) ///
355 xlabel( 1.5 "No education" 4.5 "Primary education" 7.5 "Lower secondary
    education" 10.5 "Upper secondary education" 13.5 "University", noticks)
356 graph save percentage_employment_education_spain.gph, replace
357
358 restore
359
360 *3°: descriptive analysis by age/years of work experience and level of
    income
361
362 cap drop clage
363 gen     clage = 1 if age>40 & age<=50
364 replace clage = 2 if age>50 & age<=53
365 replace clage = 3 if age>53 & age<=56
366 replace clage = 4 if age>56 & age<=60
367
368 label define clagel 1 "40-50"
369 label define clagel 2 "50-53", add
370 label define clagel 3 "53-56", add
371 label define clagel 4 "56-60", add
372
373 label values clage clagel
374
375 *level of income
376 preserve
377
378 collapse (mean) meanlogincome = logincome (sd) sdlogincome=logincome
   (count) n=logincome [aweight=w2aci], by(clage)
379
380 generate hilogincome = meanlogincome + invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
381 generate lologincome = meanlogincome - invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
382
383 lab var meanlogincome " Mean log income"
384 lab var hilogincome "Up bound"
385 lab var lologincome "Low bound"
386
387
388 twoway (line meanlogincome clage, lcolor(blue) lwidth(thick)
    lpattern(solid)) (line hilogincome clage, lcolor(blue) lwidth(thin)
    lpattern(dash)) ///
389 (line lologincome clage, lcolor(blue) lwidth(thin) lpattern(dash)),
    xlabel(, labels valuelabel) ///
390 ytitle((mean)logincome) ///
391 xtitle(age)
392 graph save level_income.gph, replace
393
394 restore
395
396 *level of income by country
397 preserve
398
```

```
399 collapse (mean) meanlogincome = logincome (sd) sdlogincome=logincome
   (count) n=logincome [aweight=w2aci], by(clage country)
400
401 generate hilogincome = meanlogincome + invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
402 generate lologincome = meanlogincome - invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
403
404 lab var meanlogincome " Mean log income"
405 lab var hilogincome "Up bound"
406 lab var lologincome "Low bound"
407
408
409 twoway (line meanlogincome clage, lcolor(blue) lwidth(thick)
    lpattern(solid)) (line hilogincome clage, lcolor(blue) lwidth(thin)
    lpattern(dash)) ///
410 (line lologincome clage, lcolor(blue) lwidth(thin) lpattern(dash)),
    by(country, iyaxes ixaxes iylabel ixlabel) ///
411 ytitle((mean)logincome) ///
412 xlabel(, labels valuelabel) xtitle(age)
413 graph save income_country.gph, replace
414
415 restore
416
417 *level of income by gender
418 preserve
419
420 collapse (mean) meanlogincome = logincome (sd) sdlogincome=logincome
   (count) n=logincome [aweight=w2aci], by(clage gender)
421
422 generate hilogincome = meanlogincome + invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
423 generate lologincome = meanlogincome - invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
424
425 twoway (line meanlogincome clage if gender == 1, lcolor(blue)
    lwidth(thick) lpattern(solid)) ///
426 (line hilogincome clage if gender == 1, lcolor(blue) lwidth(thin)
    lpattern(dash)) ///
427 (line lologincome clage if gender == 1, lcolor(blue) lwidth(thin)
    lpattern(dash)) ///
428 (line meanlogincome clage if gender == 2, lcolor(red) lwidth(thick)
    lpattern(solid)) ///
429 (line hilogincome clage if gender == 2, lcolor(red) lwidth(thin)
    lpattern(dash)) ///
430 (line lologincome clage if gender == 2, lcolor(red) lwidth(thin)
    lpattern(dash)), yscale(range(8 11)) yscale(noextend nofextend) ///
431 xlabel(, labels valuelabel)
432 graph save income_gender.gph, replace
433
434 restore
435
436 *level of income by gender and country
437 preserve
438
439 collapse (mean) meanlogincome = logincome (sd) sdlogincome=logincome
   (count) n=logincome [aweight=w2aci], by(clage gender country)
440
441 generate hilogincome = meanlogincome + invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
442 generate lologincome = meanlogincome - invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
443
444 twoway (line meanlogincome clage if gender == 1, lcolor(blue)
    lwidth(thick) lpattern(solid)) ///
```

```
445 (line hilogincome clage if gender == 1, lcolor(blue) lwidth(thin)
    lpattern(dash)) ///
446 (line logincome clage if gender == 1, lcolor(blue) lwidth(thin)
    lpattern(dash)) ///
447 (line meanlogincome clage if gender == 2, lcolor(red) lwidth(thick)
    lpattern(solid)) ///
448 (line hilogincome clage if gender == 2, lcolor(red) lwidth(thin)
    lpattern(dash)) ///
449 (line logincome clage if gender == 2, lcolor(red) lwidth(thin)
    lpattern(dash)), by(country, iyaxes ixaxes iylabel ixlabel) ///
450 yscale(range(8 11)) yscale(noextend nofextend)
451 graph save gender_income_country.gph, replace
452
453 restore
454
455 *level of income by years of work experience
456 label define age_expl 1 "10" 2 "25" 3 "30" 4 "35"
457 label values age_exp age_expl
458
459 preserve
460
461 collapse (mean) meanlogincome = logincome (sd) sdlogincome=logincome
   (count) n=logincome [aweight=w2aci], by(age_exp)
462
463 generate hilogincome = meanlogincome + invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
464 generate logincome = meanlogincome - invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
465
466 lab var meanlogincome " Mean log income"
467 lab var hilogincome "Up bound"
468 lab var logincome "Low bound"
469
470
471 twoway (line meanlogincome age_exp, lcolor(blue) lwidth(thick)
    lpattern(solid)) (line hilogincome age_exp, lcolor(blue) lwidth(thin)
    lpattern(dash)) ///
472 (line logincome age_exp, lcolor(blue) lwidth(thin) lpattern(dash)),
    xlabel(, labels valuelabel) ///
473 ytitle((mean)logincome) ///
474 xtitle(years of work experience)
475 graph save experience_income.gph, replace
476
477 restore
478
479 *level of income by years of work experience and country
480 preserve
481
482 collapse (mean) meanlogincome = logincome (sd) sdlogincome=logincome
   (count) n=logincome [aweight=w2aci], by(age_exp country)
483
484 generate hilogincome = meanlogincome + invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
485 generate logincome = meanlogincome - invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
486
487 lab var meanlogincome " Mean log income"
488 lab var hilogincome "Up bound"
489 lab var logincome "Low bound"
490
491
492 twoway (line meanlogincome age_exp, lcolor(blue) lwidth(thick)
    lpattern(solid)) (line hilogincome age_exp, lcolor(blue) lwidth(thin)
    lpattern(dash)) ///
```

```
493 (line logincome age_exp, lcolor(blue) lwidth(thin) lpattern(dash)),
    by(country, iyaxes ixaxes iylabel ixlabel) ///
494 ytitle((mean)logincome) ///
495 xlabel(, labels valuelabel) xtitle(years of work experience)
496 graph save experience_country.gph, replace
497
498 restore
499
500 *level of income by years of work experience and gender
501 preserve
502
503 collapse (mean) meanlogincome = logincome (sd) sdlogincome=logincome
   (count) n=logincome [aweight=w2aci], by(age_exp gender)
504
505 generate hilogincome = meanlogincome + invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
506 generate logincome = meanlogincome - invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
507
508
509 twoway (line meanlogincome age_exp if gender == 1, lcolor(blue)
    lwidth(thick) lpattern(solid)) ///
510 (line hilogincome age_exp if gender == 1, lcolor(blue) lwidth(thin)
    lpattern(dash)) ///
511 (line logincome age_exp if gender == 1, lcolor(blue) lwidth(thin)
    lpattern(dash)) ///
512 (line meanlogincome age_exp if gender == 2, lcolor(red) lwidth(thick)
    lpattern(solid)) ///
513 (line hilogincome age_exp if gender == 2, lcolor(red) lwidth(thin)
    lpattern(dash)) ///
514 (line logincome age_exp if gender == 2, lcolor(red) lwidth(thin)
    lpattern(dash)), yscale(range(8 11)) yscale(noextend nofextend) ///
515 xlabel(, labels valuelabel) xtitle(years of work experience)
516 graph save experience_gender.gph, replace
517
518 restore
519
520 *level of income by years of work experience, gender and country
521 preserve
522
523 collapse (mean) meanlogincome = logincome (sd) sdlogincome=logincome
   (count) n=logincome [aweight=w2aci], by(age_exp gender country)
524
525 generate hilogincome = meanlogincome + invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
526 generate logincome = meanlogincome - invttail(n-1,0.025)*(sdlogincome /
    sqrt(n))
527
528 twoway (line meanlogincome age_exp if gender == 1, lcolor(blue)
    lwidth(thick) lpattern(solid)) ///
529 (line hilogincome age_exp if gender == 1, lcolor(blue) lwidth(thin)
    lpattern(dash)) ///
530 (line logincome age_exp if gender == 1, lcolor(blue) lwidth(thin)
    lpattern(dash)) ///
531 (line meanlogincome age_exp if gender == 2, lcolor(red) lwidth(thick)
    lpattern(solid)) ///
532 (line hilogincome age_exp if gender == 2, lcolor(red) lwidth(thin)
    lpattern(dash)) ///
533 (line logincome age_exp if gender == 2, lcolor(red) lwidth(thin)
    lpattern(dash)),
    by(country, iyaxes ixaxes iylabel ixlabel) ///
534 yscale(range(8 11)) yscale(noextend nofextend)
535 graph save experience_gender_country.gph, replace
536
537 restore
```

```
538
539 ********************************************************************************
540 *The following section shows the results of the empirical estimations of*
    the Mincer equation
541 ********************************************************************************
542
543 tab country, gen(paese)
544
545 rename paese1 Germany
546 rename paese2 Spain
547 rename paese3 Italy
548 rename paese4 France
549
550 *restriction of the sample with the exclusion of the older workers
551 gen sampleMS = (yedu != . & y_work _w2 != . & age >= 21 & age <= 67)
552
553 *generation of the dummy variable single
554 tabulate mstat, generate(single)
555 drop single1 single2
556 rename single3 single
557
558 *generations of the non-linear variables
559 gen age2 = age^2
560 gen y_work_w2_sq = y_work_w2^2
561
562 *interaction variable
563 gen nchild_fem = nchild*female
564
565
566 *globale results of OLS estimation of the Mincer equation
567 regress logincome yedu y_work_w2 female single nchild Germany Spain Italy
    France if sampleMS == 1, noc
568
569 estimates store REG1
570
571 regress logincome age yedu female single nchild Germany Spain Italy France
    if sampleMS == 1, noc
572
573 estimates store REG2
574
575 esttab REG1 REG2, keep(age yedu y_work_w2 female single nchild) se
    scalars(N r2)
576
577
578 *results of OLS estimation of the Mincer equation with the non-linear
    effect of age and work experience on income
579 regress logincome yedu y_work_w2 y_work_w2_sq female single nchild if
    sampleMS == 1
580
581 estimates store REG3
582
583 regress logincome age age2 yedu female single nchild if sampleMS == 1
584
585 estimates store REG4
586
587 esttab REG3 REG4, se scalars(N r2)
588
589
590 regress logincome yedu y_work_w2 female single nchild if sampleMS == 1
591
592 estimates store REG5
593
594 regress logincome age yedu female single nchild if sampleMS == 1
595
596 estimates store REG6
```

```
597
598 esttab REG5 REG3 REG6 REG4, se scalars (N r2)
599
600
601 *inclusion of the interaction variable
602 regress logincome yedu y_work_w2 y_work_w2_sq female single nchild
    nchild_fem if sampleMS == 1
603
604 estimates store REG3
605
606 regress logincome age age2 yedu female single nchild nchild_fem if
    sampleMS == 1
607
608 estimates store REG4
609
610 esttab REG3 REG4, se scalars(N r2)
611
612
613 regress logincome yedu y_work_w2 female single nchild nchild_fem if
    sampleMS == 1
614
615 estimates store REG5
616
617 regress logincome age yedu female single nchild nchild_fem if sampleMS ==
    1
618
619 estimates store REG6
620
621 esttab REG5 REG3 REG6 REG4, se scalars (N r2)
622
623
624 *results of OLS estimation of the Mincer equation by country
625 regress logincome yedu y_work_w2 female single nchild if sampleMS == 1 &
    iso == "FRA"
626
627 estimates store REG7
628
629 regress logincome yedu y_work_w2 female single nchild if sampleMS == 1 &
    iso == "GER"
630
631 estimates store REG8
632
633 regress logincome yedu y_work_w2 female single nchild if sampleMS == 1 &
    iso == "ITA"
634
635 estimates store REG9
636
637 regress logincome yedu y_work_w2 female single nchild if sampleMS == 1 &
    iso == "ESP"
638
639 estimates store REG10
640
641 esttab REG7 REG8 REG9 REG10, se scalars (N r2)
642
643
644 *inclusion of the interaction variable by country
645 regress logincome yedu y_work_w2 female single nchild nchild_fem if
    sampleMS == 1 & iso == "FRA"
646
647 estimates store REG7
648
649 regress logincome yedu y_work_w2 female single nchild nchild_fem if
    sampleMS == 1 & iso == "GER"
650
651 estimates store REG8
```

```
652
653 regress logincome yedu y_work_w2 female single nchild nchild_fem if
    sampleMS == 1 & iso == "ITA"
654
655 estimates store REG9
656
657 regress logincome yedu y_work_w2 female single nchild nchild_fem if
    sampleMS == 1 & iso == "ESP"
658
659 estimates store REG10
660
661 esttab REG7 REG8 REG9 REG10, se scalars (N r2)
662 *-------------------------------------------------------------------------
663
664 regress logincome age yedu female single nchild if sampleMS == 1 & iso ==
    "FRA"
665
666 estimates store REG11
667
668 regress logincome age yedu female single nchild if sampleMS == 1 & iso ==
    "GER"
669
670 estimates store REG12
671
672 regress logincome age yedu female single nchild if sampleMS == 1 & iso ==
    "ITA"
673
674 estimates store REG13
675
676 regress logincome age yedu female single nchild if sampleMS == 1 & iso ==
    "ESP"
677
678 estimates store REG14
679
680 esttab REG11 REG12 REG13 REG14, se scalars (N r2)
681
682
683 *results of OLS estimation of the Mincer equation by country with the non-
    linear effect of the work experience and age
684 regress logincome yedu y_work_w2 y_work_w2_sq female single nchild if
    sampleMS == 1 & iso == "FRA"
685
686 estimates store REG15
687
688 regress logincome yedu y_work_w2 y_work_w2_sq female single nchild if
    sampleMS == 1 & iso == "GER"
689
690 estimates store REG16
691
692 regress logincome yedu y_work_w2 y_work_w2_sq female single nchild if
    sampleMS == 1 & iso == "ITA"
693
694 estimates store REG17
695
696 regress logincome yedu y_work_w2 y_work_w2_sq female single nchild if
    sampleMS == 1 & iso == "ESP"
697
698 estimates store REG18
699
700 esttab REG15 REG16 REG17 REG18, se scalars (N r2)
701 *-------------------------------------------------------------------------
702
703 regress logincome age age2 yedu female single nchild if sampleMS == 1 &
    iso == "FRA"
704
```

```
705 estimates store REG19
706
707 regress logincome age age2 yedu female single nchild if sampleMS == 1 &
    iso == "GER"
708
709 estimates store REG20
710
711 regress logincome age age2 female single nchild if sampleMS == 1 & iso ==
    "ITA"
712
713 estimates store REG21
714
715 regress logincome age age2 yedu female single nchild if sampleMS == 1 &
    iso == "ESP"
716
717 estimates store REG22
718
719 esttab REG19 REG20 REG21 REG22, keep(age age2 yedu female single nchild)
    se scalars (N r2)
720
721
722 *generation of the variables of the education levels from the variable
    isced
723 tabulate isced, generate(educ)
724
725 gen low_educ = educ1 + educ2 + educ3
726 gen educsup = educ4 + educ5
727 gen uni = educ6 + educ7
728
729 rename low_educ educ_3
730 rename educsup educ_4
731 rename uni educ_5
732
733 *results of OLS estimation of the Mincer equation with the classification
    by education level
734 regress logincome educ_4 educ_5 y_work_w2 female single nchild if sampleMS
    == 1
735
736 esttab, se scalars (N r2)
737
738 *results of OLS estimation of the Mincer equation with the classification
    by education level and country
739 regress logincome educ_4 educ_5 y_work_w2 female single nchild if sampleMS
    == 1 & iso == "FRA"
740
741 estimates store REG23
742
743 regress logincome educ_4 educ_5 y_work_w2 female single nchild if sampleMS
    == 1 & iso == "GER"
744
745 estimates store REG24
746
747 regress logincome educ_4 educ_5 y_work_w2 female single nchild if sampleMS
    == 1 & iso == "ITA"
748
749 estimates store REG25
750
751 regress logincome educ_4 educ_5 y_work_w2 female single nchild if sampleMS
    == 1 & iso == "ESP"
752
753 estimates store REG26
754
755 esttab REG23 REG24 REG25 REG26, se scalars (N r2)
756 *-------------------------------------------------------------------------
757
```

```
758 ********************************************************************************
759 The following section shows the results of the empirical estimations of *
    the scarring effects
760********************************************************************************
761
762 clear all
763 set mem 800m
764 set more off
765 capture log close
766
767 cd "D:\Tesilaurea\workspace\Nuovacartella"
768
769 log using scarring_effects2.log,replace
770 use wave2_ms_ultimo.dta, clear
771 *-------------------------------------------------------------------------------
772
773 tab country
774 tab country, nolabel
775 gen     iso = "GER" if country == 12
776 replace iso = "ESP" if country == 15
777 replace iso = "ITA" if country == 16
778 replace iso = "FRA" if country == 17
779
780 gen logincome = log(ydipv)
781
782 tabulate gender, generate(sex)
783 rename sex1 male
784 rename sex female
785
786 tabulate mstat, generate(single)
787 drop single1 single2
788 rename single3 single
789
790 tab country, gen(paese)
791
792 rename paese1 Germany
793 rename paese2 Spain
794 rename paese3 Italy
795 rename paese4 France
796
797 *generation of the geographical distinctions from the variable Nuts1
798 tabulate Nuts1, generate(geographic_area)
799
800 *Italy
801 gen North_ita = geographic_area30 + geographic_area31
802 gen Center_ita = geographic_area32
803 gen South_ita = geographic_area33 + geographic_area34
804
805 *France
806 gen North_fra = geographic_area24 + geographic_area25
807 gen Center_fra = geographic_area26 + geographic_area28
808 gen South_fra = geographic_area29 + geographic_area27
809
810 *Spain
811 gen North_esp = geographic_area17 + geographic_area18
812 gen Center_esp = geographic_area19 + geographic_area20
813 gen South_esp = geographic_area22 + geographic_area23
814
815 *Germany
816 gen North_ger = geographic_area3 + geographic_area4 + geographic_area5 +
    geographic_area6 + geographic_area8 + geographic_area9 + geographic_area14
    + geographic_area15
817 gen Center_ger = geographic_area7 + geographic_area10 + geographic_area11
    + geographic_area13 + geographic_area16
818 gen South_ger = geographic_area1 + geographic_area2 + geographic_area12
```

```
819 *-------------------------------------------------------------------------
820
821 *generation of the interaction variable
822 gen rec_fem = rec_edu*female
823
824 *results of OLS estimation of the scarring effects at global level
825 regress logincome yedu y_work_w2 female rec_edu ever_unemp nchild single
    Germany Spain Italy France, noc
826
827 estimates store REG1
828
829 regress logincome yedu y_work_w2 female rec_edu rec_fem ever_unemp nchild
    single Germany Spain Italy France, noc
830
831 estimates store REG2
832
833 esttab REG1 REG2, se scalars(N r2)
834 *-------------------------------------------------------------------------
835
836 *results of OLS estimation of the scarring effects by country
837 regress logincome yedu y_work_w2 female rec_edu ever_unemp nchild single
    Center_fra South_fra if iso == "FRA"
838
839 estimates store REG3
840
841 regress logincome yedu y_work_w2 female rec_edu ever_unemp nchild single
    Center_ger South_ger if iso == "GER"
842
843 estimates store REG4
844
845 regress logincome yedu y_work_w2 female rec_edu ever_unemp nchild single
    Center_ita South_ita if iso == "ITA"
846
847 estimates store REG5
848
849 regress logincome yedu y_work_w2 female rec_edu ever_unemp nchild single
    Center_esp South_esp if iso == "ESP"
850
851 estimates store REG6
852
853 esttab REG3 REG4 REG5 REG6, se scalars(N r2)
854 *-------------------------------------------------------------------------
855
856 ***************************************************************************
857 The following section shows the results of the empirical estimations of *
    the scarring effects with longitudinal data
858***************************************************************************
859
860 clear all
861 set mem 800m
862 set more off
863 capture log close
864
865 cd "D:\Tesilaurea\workspace\Nuovacartella"
866
867 log using scarring_effects3.log,replace
868 use scarring_effects.dta, clear
869 *-------------------------------------------------------------------------
870
871 tab country
872 tab country, nolabel
873 gen     iso = "GER" if country == 12
874 replace iso = "ESP" if country == 15
875 replace iso = "ITA" if country == 16
876 replace iso = "FRA" if country == 17
```

```
877
878 tab country, gen(paese)
879
880 rename paese1 Germany
881 rename paese2 Spain
882 rename paese3 Italy
883 rename paese4 France
884
885 gen rec_fem = rec_edu*female
886 gen unemp_fem = ever_unemp*female
887
888 *generation of the geographical distinctions from the variable Nuts1
889 tabulate Nuts1, generate(geographic_area)
890
891 *Italy
892 gen North_ita = geographic_area30 + geographic_area31
893 gen Center_ita = geographic_area32
894 gen South_ita = geographic_area33 + geographic_area34
895
896 *France
897 gen North_fra = geographic_area24 + geographic_area25
898 gen Center_fra = geographic_area26 + geographic_area28
899 gen South_fra = geographic_area29 + geographic_area27
900
901 *Spain
902 gen North_esp = geographic_area17 + geographic_area18
903 gen Center_esp = geographic_area19 + geographic_area20
904 gen South_esp = geographic_area22 + geographic_area23
905
906 *Germany
907 gen North_ger = geographic_area3 + geographic_area4 + geographic_area5 +
    geographic_area6 + geographic_area8 + geographic_area9 + geographic_area14
    + geographic_area15
908 gen Center_ger = geographic_area7 + geographic_area10 + geographic_area11
    + geographic_area13 + geographic_area16
909 gen South_ger = geographic_area1 + geographic_area2 + geographic_area12
910 *-------------------------------------------------------------------------
911
912 *generation of the panelvar
913 egen individual = group(mergeid)
914
915 xtset individual yr_at, yearly
916
917 *results of the estimation of the Scarring effects with longitudinal data
918 xtreg log_wage yedu female rec_edu ever_unemp nchildren single Spain Italy
    France
919
920 estimates store REG1
921
922 xtreg log_wage yedu female rec_edu ever_unemp unemp_fem nchildren single
    Spain Italy France
923
924 estimates store REG2
925
926 esttab REG1 REG2, se scalars(N)
927
928
929 *results of the estimation of the Scarring effects with longitudinal data
    by country
930 xtreg log_wage yedu female rec_edu ever_unemp nchildren single Center_fra
    South_fra if iso == "FRA", re
931
932 estimates store REG1
933
```

```
934 xtreg log_wage yedu female rec_edu ever_unemp nchildren single Center_ger
    South_ger if iso == "GER", re
935
936 estimates store REG2
937
938 xtreg log_wage yedu female rec_edu ever_unemp nchildren single Center_ita
    South_ita if iso == "ITA", re
939
940 estimates store REG3
941
942 xtreg log_wage yedu female rec_edu ever_unemp nchildren single Center_esp
    South_esp if iso == "ESP", re
943
944 estimates store REG4
945
946 esttab REG1 REG2 REG3 REG4, se scalars(N)
947 *-------------------------------------------------------------------------
948
949 xtreg log_wage yedu female rec_edu rec_fem ever_unemp nchildren single
    Center_fra South_fra if iso == "FRA", re
950
951 estimates store REG5
952
953 xtreg log_wage yedu female rec_edu ever_unemp unemp_fem nchildren single
    Center_ita South_ita if iso == "ITA", re
954
955 estimates store REG6
956
957 xtreg log_wage yedu female rec_edu ever_unemp unemp_fem nchildren single
    Center_esp South_esp if iso == "ESP", re
958
959 estimates store REG7
960
961 esttab REG5 REG6 REG7, se scalars(N)
```