WAGE INEQUALITY TRENDS IN EUROPE AND THE USA

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

NAZMİ YÜKSELEN YAĞANOĞLU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
ECONOMICS

AUGUST 2007

Approval of the Graduate School of Social Sciences

Prof.Dr.Sencer Ayata
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy.

Prof.Dr. Haluk Erlat
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.

Assoc.Prof.Dr. Hakan Ercan
Supervisor

Examining Committee Members

Assoc.Prof.Dr. Hakan Ercan	(METU,ECON)	
Prof.Dr.Aysıt Tansel	(METU,ECON)	
Assoc.Prof.Dr. Burak Günalp	(HU,ECON)	
Assoc.Prof.Dr. Ramazan Sarı	(AIBU,ECON)	
Dr. Murat Kırdar	(METU,ECON)	

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name :

Signature :

ABSTRACT

WAGE INEQUALITY TRENDS IN EUROPE AND THE USA

Yağanoğlu, Nazmi Yükselen
Ph.D., Department of Economics
Supervisor : Assoc. Prof. Dr. Hakan Ercan

August 2007, 128 pages

There was a well documented surge of wage inequality in the US that started from mid-70s and continued in 80s, slowing down by mid-90s, caused by increased dispersion both between and within groups of people with similar personal characteristics and skills. We analyze the US wage inequality in the more recent years to see if this trend continues. We apply the decomposition technique of Juhn, Murphy and Pierce (1993) and quantile regression to March Current Population Survey data of the US Bureau of Labor Statistics data and Luxembourg Income Study data for a few selected European countries. We find that the increase in wage inequality continues during the 90s, especially in the second half. In addition, the focus of wage inequality shifts into the upper half of the wage distribution after mid-80s. The European countries do not show a common trend in the direction of wage inequality during the 90s. However, the focus of their wage inequality seems to be shifting towards the lower half of the wage distribution as opposed to that of US.

Keywords: Wage inequality, US, wage decomposition, quantile regression

ÖZ

AVRUPA VE ABD'DE ÜCRET EŞİTSİZLİĞİ TRENDLERİ

Yağanoğlu, Nazmi Yükselen
Doktora, İktisat Bölümü
Tez Yöneticisi : Doç. Dr. Hakan Ercan

Ağustos 2007, 128 sayfa

ABD'de 70'li yılların ortalarında başlayıp 80'lerde devam eden ve 90'ların ortalarında yavaşlayan, hem benzer hem de birbirinden farklı kişisel özellikler ve becerilere sahip insanlar arasındaki ücret farklılıklarının artmasıyla ilgili olarak ortaya çıkmış bir ücret eşitsizliği artışının varlığı, daha önceki çalışmalarda belgelenmişti. Bu çalışmada, yakın zamanlarda da bu trendin sürüp sürmediğini görmek için ABD'deki ücret eşitsizliğini inceliyoruz. Bu amaçla Juhn, Murphy and Pierce (1993) tarafından önerilen ayrıştırma tekniğini ve quantile regresyonu US Bureau of Labor Statistics'in Mart ayı Current Population Survey ve Luxembourg Income Study'nin birkaç Avrupa ülkesi için verilerine uyguluyoruz. ABD'de ücret eşitsizliğindeki artışın 90'larda, özellikle de ikinci yarısında devam ettiği görülmektedir. Ek olarak, ücret eşitsizliğinin odağı özellikle 80'lerin ortasından sonra ücret dağılımının üst yarısına kaymış bulunmaktadır. Avrupa ülkeleri 90'larda ücret eşitsizliğinin yönü ile ilgili ortak bir eğilim göstermemektedirler. Bununla birlikte, oradaki ücret eşitsizliğinin odağının ABD'dekinin tersine ücret dağılımının alt yarısına kaydığı gözlemlenmektedir.

Anahtar Kelimeler: Ücret eşitsizliği, ücret ayrıştırması, quantile regresyon

To my family, for their unconditional support and faith

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my supervisor Assoc.Prof.Dr. Hakan Ercan for his encouragement, guidance and support, which was at least as instrumental as the my own work in the completion of this thesis.

The constructive criticism and comments of Prof.Dr. Aysıt Tansel, Assoc.Prof.Dr. Burak Günalp, Assoc.Prof.Dr. Ramazan Sarı and Dr. Murat Kırdar have been very much appreciated.

Several people have made life much easier for me while dealing with the technical difficulties. I am grateful to Ms. Emilia Niskanen of Luxembourg Income Study for her quick solutions to problems related to the LISSY system, to Prof. Ben Jann of ETH Zurich for his help in overcoming some problems with the applications in software, and to Mr. Engin Akyürek of METU, Department of Economics, for his assistance with computer hardware.

This study was supported by BAP-2006-07-03-00-12 project.

TABLE OF CONTENTS

PLAGIARISM....iii

ABSTRACT....iv

ÖZ....v

DEDICATION....vi

ACKNOWLEDGMENTS....vii

TABLE OF CONTENTS....viii

LIST OF TABLES....x

LIST OF FIGURES....xiii

CHAPTERS

1 INTRODUCTION....1

2 US WAGE INEQUALITY: EMPIRICAL OBSERVATIONS....9

2.1 Data....9

2.2 An Overview of the Data....13

2.3 Evidence on Wage Inequality....19

2.4 Concluding Remarks....33

3 JUHN-MURPHY-PIERCE DECOMPOSITION....34

3.1 The Method....34

3.2 Application to US Data....38

3.3 Concluding Remarks....44

4 QUANTILE REGRESSION....46

4.1 The Method 46

4.2 Application to the US Data.. 50

4.3 Conditional Wage Distribution.. 59

4.4 JMP Decomposition and Quantile Regression 65

4.5 Concluding Remarks.. 67

5 A CROSS-COUNTRY ANALYSIS OF WAGE INEQUALITY USING ....
LUXEMBOURG INCOME STUDY 69

5.1 Data Set 70

5.2 Wage Inequality... 79

5.3 Education and Experience 82

5.4 JMP Decomposition... 89

5.5 Concluding Remarks.. 93

6 QUANTILE REGRESSION ANALYSIS OF THE EUROPEAN DATA........... 95

6.1 Experience 95

6.2 Education.. 101

6.3 Conditional and Counterfactual Distributions....................................... 107

6.4 Concluding Remarks.. 120

7 CONCLUSION 121

REFERENCES 124

CURRICULUM VITAE 129

LIST OF TABLES

Table 2.1 Descriptive Statistics 14

Table 2.2 Measures of Inequality 20

Table 2.3 Residual Inequality....................32

Table 3.1 JMP Decomposition for 5-year Intervals....................41

Table 3.2 Variance Decomposition of Wages For Industry Composition Effect 44

Table 4.1 Marginal Effects-Education....................51

Table 4.2 Impact Upon Dispersion-Education....................52

Table 4.3 College Premium....................54

Table 4.4 Marginal Effects-Experience 55

Table 4.5 Impact upon Dispersion-Experience 57

Table 4.6 Industries....................58

Table 4.7 Conditional Wage Distribution 59

Table 4.8 Comparison of Conditional and Empirical Wage Distributions................60

Table 4.9 Counterfactual Conditional Distributions-1968 Covariates 62

Table 4.10 Counterfactual Conditional Distributions-1968 Coefficients..................63

Table 4.11 Comparison of JMP and Quantile Regression....................66

Table 5.1. Waves and Sample Size....................74

Table 5.2. Education of the Wage-Earner Sample 77

Table 5.3 Effect of Education....................87

Table 5.4 Effect of Experience....................88

Table 5.5.a JMP Decomposition-Germany....................90

Table 5.5.b JMP Decomposition-Netherlands 91

Table 5.5.c JMP Decomposition-Sweden 92

Table 5.5.d JMP Decomposition-Spain 92

Table 5.5.e JMP Decomposition-Hungary..93

Table 6.1 Marginal Effect of Experience- Germany 96

Table 6.2 Impact upon Dispersion-Experience-Germany..96

Table 6.3 Marginal Effect of Experience- Netherlands 97

Table 6.4 Impact upon Dispersion-Experience-Netherlands 98

Table 6.5 Marginal Effect of Experience- Sweden 98

Table 6.6 Impact upon Dispersion-Experience-Sweden..99

Table 6.7 Marginal Effect of Experience- Spain...99

Table 6.8 Impact upon Dispersion-Experience-Spain 100

Table 6.9 Marginal Effect of Experience- Hungary 100

Table 6.10 Impact upon Dispersion-Experience-Hungary 101

Table 6.11 Marginal Effect of Education-Germany 102

Table 6.12 Impact upon Dispersion-Education-Germany 102

Table 6.13 Marginal Effect of Education-Netherlands..103

Table 6.14 Impact upon Dispersion-Education-Netherlands 104

Table 6.15 Marginal Effect of Education-Sweden 104

Table 6.16 Impact upon Dispersion-Education-Sweden 105

Table 6.17 Marginal Effect of Education-Spain..105

Table 6.18 Impact upon Dispersion-Education-Spain 106

Table 6.19 Marginal Effect of Education-Hungary 107

Table 6.20 Impact upon Dispersion-Education-Hungary 107

Table 6.21 Conditional Wage Distribution-Germany..108

Table 6.22 Counterfactual Distributions-Germany 110

Table 6.23 Conditional Wage Distribution-Netherlands 111

Table 6.24 Counterfactual Distributions-Netherlands..112

Table 6.25 Conditional Wage Distribution-Sweden...113

Table 6.26 Counterfactual Distributions-Sweden 115

Table 6.27. Conditional Wage Distribution-Spain...116

Table 6.28 Counterfactual Distributions-Spain..117

Table 6.29 Conditional Wage Distribution-Hungary 118

Table 6.30 Counterfactual Wage Distributions-Hungary 119

LIST OF FIGURES

Figure 2.1 Shares of Education Groups in Sample....15
Figure 2.2 Shares of Experience Groups in the Sample....17
Figure 2.3 Industries (Percentage of workers employed in the industry)....19
Figure 2.4 Gini Coefficient, Theil Index and Variance of Logs. 21
Figure 2.5 Indexed Log Weekly Wages by Percentile....23
Figure 2.6 Differences of Log Wage Percentiles....24
Figure 2.7 Changes in Percentiles of Wage Distribution over Time....26
Figure 2.8 Changes in Wage Percentiles, 1988-2005....27
Figure 2.9 Within Inequality....28
Figure 3.1 Decomposition of 90^{th}-10^{th} Log Wage Differential over the Years....39
Figure 4.1 Marginal Effect of Experience....56
Figure 4.2 Conditional Wage Distributions....59
Figure 4.3 Conditional vs. Counterfactual Wage Distributions....64
Figure 5.1 Wage Inequality....79
Figure 5.2. Change in Log Wages by Education....83
Figure 5.3. Change in Log Wages by Experience....86
Figure 6.1 Conditional Wage Distributions-Germany....109
Figure 6.2 Conditional Wage Distributions-Netherlands....112
Figure 6.3 Conditional Wage Distributions-Sweden....114
Figure 6.4 Conditional Wage Distributions-Spain....117
Figure 6.5 Conditional Wage Distributions-Hungary....119

CHAPTER 1

INTRODUCTION

Following a long period of relative stability in the post-war era, the male wage inequality[1] began to rise in the US in the second half of the 1970s. This trend continued during the 80s, especially in the first half. Then, it slowed down by mid-90s, beginning to increase again in the second half of the decade. Most of this increase has been attributed to an increase in demand for people with skills to use of high-technology tools and personal computers that have become common in many industries since the beginning of 80s. The European countries did not respond to this increase during the 80s, with the exception of United Kingdom. Even though some showed wage inequality increases, they were of a smaller magnitude than that of the US. This difference was mostly attributed to the institutional difference between US and most European countries that let the US labor market adapt to changes in supply and demand via changes in wages, while the European labor market could not do this because of their highly regulated labor markets and typically strong unions. The aim of this work is to analyze and determine the components and dynamics of the changes in wage inequality in the USA and a selection of European countries that includes Germany, Netherlands, Sweden, Spain and Hungary to explore how wage inequality changed in the recent years.

Wage inequality can be described as the differences in wages due to some measureable and unmeasurable (or observable and unobservable) characteristics of earners. We could talk about two types of wage inequality depending on the source of it. The "between" inequality is simply the wage inequality casued by differences

---

1 From this point on, whenever we mention wage inequality, it means male wage inequality.

in observable skills and characteristics. It can be seen easily by comparing the wage distributions of people with different characteristics, for example high school graduates vs. college graduates or workers with higher experience vs. workers with lower experience. One could also look for gender, race, location or industry for such comparisons.

On the other hand, "within" inequality represents the wage inequality within narrowly defined groups of similar characteristics. Even though these groups of people have seemingly similar observable skills (such as education and experience), usually they still display wage inequality that cannot be explained by their observed skills. If we consider two workers with similar years of experience and education, if one of them can do certain things that the other cannot (unmeasured skills), a change in demand for the things the former worker can do but the latter cannot will result in higher wages for the former. Obviously it is not an easy task to measure the effect of something that is unobservable or unmeasurable. For one thing, we cannot argue that the data sets available to us gives all the details about a person. Thus, even "narrowly defined groups" might not be as similar as we expect. Levy and Murnane (1992) give supply and demand shifts for worker characteristics that are not observable from the standard data sets available to us as well as industry and plant specific characteristics as examples of unobservables that lead to within inequality. A generally accepted measure of within inequality is the inequality observed among OLS residuals obtained from a Mincerian wage regression (Mincer (1974)), due to their representation of unobserved or unmeasurable skills and characteristics. This measure is also known as "residual" inequality.

The literature on overall male wage inequality in the US generally agrees that it recorded an increase starting from mid-1970s and gained pace during the 1980s, slowing down toward the mid-90s. These changes were accompanied by three important facts (See Juhn, Murphy and Pierce (1993), Murphy and Welch (1992), Katz and Murphy (1992), Katz and Autor (1999)):

1) The returns to education expressed as years of schooling showed a changing pattern from 1960s to the end of 1980s. It increased during the 60s, then fell during the 70s, and rapidly increased during the 80s. There was also an increase in the gap between the returns to high school and college degrees.

2) The returns to experience also increased, especially during the 80s, with more experienced workers gaining on the younger workers especially on the low education levels.

3) There was also an increase in the wage inequality among the workers with similar characteristics such as experience and education.

The first two of the above are related to "between" inequality, while the last one is an indicator of "within" or "residual" inequality. Within inequality carries considerable weight in the debates over reasons for these changes because of the difficulty of measuring it. It was found by Juhn, Murphy And Pierce (JMP here after) (1993) that the change in residual inequality explained a much higher portion of the changes in wage inequality than that of education, experience or other observable demographic characteristics.

A number of studies pointed towards the changes in workplace during the 80s that resulted in increased use of skill-intensive technology. Katz and Murphy(1992) mention the rapid growth in demand for more educated and "more skilled" workers among the reasons for changing wage inequality. Berman, Bound and Griliches(1994) point towards increased employment of skilled workers and relate this to investment in computers and research and development facilities. Autor, Katz and Krueger(1998) also confirm the effect of computers and reveal that during 1940-1996 period, the relative demand for college graduates grew strongly and persistently. It was generally argued that such changes favored high-skilled workers over the low-skilled ones. During the 80s, it was observed that the increase in wage inequality was mostly in the shape of higher demand for the more skilled workers and somewhat lower demand for the less skilled ones, the former having a more dominant effect (JMP(1993)). The increase in the wage inequality favoring college graduates over high school graduates and new college graduates over older ones confirm this view. The increase in within inequality, as explained above, showed that the workers with newly needed skills were also favored over the ones that do not have those skills. All in all, the argument that the increased demand for high-skilled workers caused an increase in wage inequality is generally known as "Skill-biased Technological Change Hypothesis" (SBTC). Acemoglu(2002) concludes that most of the technological change in the 20th century is skill-biased, this kind of technological change gained pace during the 80s with the increased use of computers and demand for people with higher education and required skills increased, and the increase in wage inequality is closely related to an acceleration in this trend.

Another attempt at explaining the increases in wage inequality, yet did not find as many followers as the first one, was the effect of increased trade with the developing countries which shifted the focus of production towards exports and high-skill oriented products, while increased imports of cheaper low-skill products hit the demand for less skilled workers Wood, (1995), Leamer(1992) and Berman et al, (1994) all focus on the importance of increased international trade and its such effects.

Some institutional factors were also mentioned among the causes of this increase in wage inequality. Dinardo, Fortin and Lemieux (1994) found that ,as well as some supply and demand shocks and de-unionization, declining real value of the minimum wage appeared to explain a serious amount of the rising inequality, especially at the lower end. Lee(1999), Card(2001) and Teulings(2003) are among the researchers who found the importance of labor market institutions such as unions and minimum wage to be important. Demographic factors like immigration were also seemed to have some explanatory power (Morris and Western, 1997)

While labor economists were focused on explaining the reasons for this change in wage inequality, studies on 1990s data showed a relative slowdown, even a stabilization in the wage inequality around mid-90s (Card and Dinardo (2002)-CPS data through 2001, Acemoglu (2003a)-LIS data through 1997, among many others)) in the US. This new trend, which appeared despite continuing improvement in technology, brought out some doubts over the validity of SBTC thesis (Card and Dinardo (2002). The main criticism was that despite the skill-biased technological change continued, the wage inequality did not keep pace with it.

Revisiting the historical change in wage inequality, Autor, Katz and Kearney (2005) conclude that none of the explanations that have been listed above can be sufficient to explain the changes in wage inequality in the last few decades, since different measures and components of it have changed in different times themselves, making one dominant explanation unrealistic.

The European experience during the same period proved different in more ways than one. In major European countries, there was little increase in wage inequality during the 1980s, with the exception of Italy and Germany (Gottschalk and Smeeding (1997)). Also the wage inequality, calculated as the log difference of 90th-10th percentile shares of the wage distribution, was clearly less than the US in countries like France, Germany, Sweden etc. This measure of wage inequality actually fell even further in the mid-1990s for a number of countries in Europe (Acemoglu (2003)). Here, The United Kingdom stands apart from the rest of Europe, in that it followed a path more similar to the US than to the rest of Europe, with considerable increase in inequality, rates of return for skill and education.

Another point of comparison between the US and Europe is unemployment. The unemployment levels in the European part of the OECD surpassed those in the US in the mid-80s and the gap kept opening throught the 90s. Siebert(1997) and Nickel and Bell(1996) attributed this trend to the inflexibility of labor markets in most European countries, which in turn prevented them from acting to counter the effects of changes in demand for skills. A number of other studies challenged this theory by arguing that the unemployment was not only a matter of the lower part of the distribution, but a more widespread phenomenon (see, Gregg and Manning (1997) among others). Acemoglu (2003) attributed the differences in inequality to a combination of unemployment, differences in the rate of increase in supply of skills, and more importantly, differences in demand for skills. Recently there have been calls to change the things in the European Union. During his discussion of the concept of "flexicurity", Wilthagen (2004) points that there is a strong need to make European Labor markets more flexible while providing security to vulnerable employee groups. Ebbinghaus & Kittel(2005) challenge the general assumption that European labor markets are too rigid by showing that countries have different experiences under different structures.

Some of the East Asian economies saw an improvement in wage inequality during 60s and 70s. The trend seems to have continued in Taiwan during the 80s and 90s. (Tsou (2002), Lin and Orazem (2004)). Fan and Cheung (2004) found an increasing wage inequality in Hong-Kong in 1982-1994 and claimed that this can be largely attributed to increasing trade with China. Kijima(2005) found that the wage inequality increase in India between 1983 and 1999 was mainly due to increases in returns to skills which was caused by higher demand for skilled labor. Skoufias and Suryahadi(2002) reveal that the wage inequality increased in Indonesia between 1986 and 1998.

When we look at the Latin American countries, we see widening wage inequality. Attanasio, Goldberg and Pavcnik(2004) found increasing within-group inequality in Colombia over 1984-1998. Beyer, Rojas and Vergara(1999) conclude that opening to international trade increased wage inequality in Chile. Green, Dickerson and Arbache (2001) did not find a strong relationship between trade and wage inequality for Brazil. They concluded that the reason for increasing wage inequality in Brazil is the relatively small share of college graduates in the work force. On the other hand, Feliciano (2001) found that the lowering of import license coverage increased the wage inequality in Mexico. Wage inequality in Argentina increased sharply during the 90s (Galiani and Sanguinetti (2003)).

Due to the lack of available data, wage inequality studies have been limited in Turkey. One notable study by Tansel and Bircan (2006) compares the effect of education on the wage inequality picture for the years 1994 and 2002, revealing the existence of between and within inequality.

The decomposition method that was suggested and used by JMP (1993) has been used extensively since its introduction. In recent years, quantile regression methods were started to be used to analyze the trends in wage inequality. This study takes stock by applying both methods to the data of the US and selected European countries, namely Germany, Netherlands, Sweden, Spain and Hungary.

After their two initial papers (1991 and 1993), JMP's decomposition method and its versions (see Yun (2006) for a recent version) have been used in several studies to analyze wage inequality in the US and different countries. The method has been applied to analyze the wage inequality in different areas and countries. Blau and Kahn (1994, 1996 and some later studies) used it mainly to analyze the dynamics of gender wage gap in the United States, as well as some international comparisons on overall wage inequality. Margo (1995) followed the technique to look into racial wage inequality in the US during 40s. Labor economists used the method in analyzing the wage inequality picture in a number of other countries (Orazem & Vodopivec (2000) for Estonia and Slovenia, and Kijima (2005) for India are examples). JMP method's advantage is that it decomposes the change in wage inequality into observable prices, observable quantities and unobservable prices and quantities. This contribution is important in the sense that it gives an account of the contribution of "within" inequality (represented by unobservables) to overall wage inequality separately from the contribution of "between" inequality (represented by observables).

Some later studies used quantile regression techniques (see, for example, Buchinsky, 1994 and 1998; Machado and Mata, 2001; Martins and Pereira, 2004). This technique is useful since it shows the effects of covariates on the distribution of dependent variable at different quantiles. Running a number of quantile regressions, one can obtain a very useful tool of observing the relative importance of covariates at different parts of the wage distribution. Although the JMP method also compares the conditional quantiles of wage distribution, their technique is based upon ordinary least squares estimation, and thus the conditional mean. Naturally, they assign the same coefficients for the covariates for any point on the distribution of wages and sometimes have problems with identifying the changes in the tails of the distribution. In this study, we will also make a comparison of the results obtained by JMP and quantile regression methods

Using March Current Population Survey (CPS) data of the US Bureau of Labor Statistics (BLS), we find that the wage inequality in the US, which increased remarkably during the 80s, has been still increasing at a slower rate. Although most of this increase came from the lower half of the wage distribution until the mid-80s, the change in the later years came almost exclusively from the dispersion in the upper half of the wage distribution. The role of within inequality in these changes are reflected in our JMP decompositions and it looks still strong. We also generate the conditional quantile comparisons of JMP method by using quantile regression and compare the results. We find that they might report similar values if the conditional wage distribution does not change much at different quantiles.

Our analysis concerning the European countries shows that the wage inequality experience was not uniform in Europe. We go into details about the wage inequality dynamics in each country.

The rest of this study will be as follows: In Chapter 2 we give technical details about the US data and describe the historical progress of some key variables. In Chapter 3 we give wage inequality measures and use the decomposition method proposed by JMP (1993). In Chapter 4 we go over the wage inequality picture using a quantile regression approach and make a comparison of JMP and quantile regression results. In Chapter 5, we describe the European data that comes from Luxembourg Income Study (LIS from this point onwards) and apply the JMP method to a selection of European countries. In Chapter 6 we apply the quantile regression techniques to go into more details of European wage inequality. In Chapter 7 we conclude.

CHAPTER 2

US WAGE INEQUALITY: EMPIRICAL OBSERVATIONS

The detail, availability and reliability of wage and demographic data for the US have been one of the main reasons for abundance of studies on wage inequality. It is possible to do such analysis for periods as early as 1940s (Goldin & Margo (1991)) or even nineteenth century (Atack, Bateman & Margo (2004) using different sources of data.

Before going into a detailed analysis of US wage inequality, it is a good idea to look closely into the data and understand the main characteristics of our wage-earner sample. As we will see in the later chapters, a sizable part of wage inequality is related to personal characteristics. Then we introduce some observations and measures of wage inequality to establish the basis for our later analysis.

2.1 Data

In this study we have used the Current Population Survey (CPS) March Annual Social and Economics Supplement (ASEC) for years 1967 to 2005. Although some data is available from earlier years for the US from the same survey, our decision to start with 1967 is directly related to the change in formation of data in 1967, which especially makes comparisons with earlier years a bit tricky.

The CPS data is collected monthly by interviewing a large number of households (it was around 57,000 households for 2005). Each household is interviewed once a month for four months in a row every year, then interviewed again next year during the same four months.

The main purpose of the survey is to collect employment information within the United States of America. However, it also contains data on demographic characteristics of the population such as age, race, number of children, area of living etc. The data is organized to give three perspectives: household, family and person. CPS is conducted among the civilian, non-institutional population in every state of the United States of America and the District of Columbia. ASEC is released for the month of March every year, including everything the other months' data have as well as additional information on work experience, income and migration.

The related raw data files concerning 1967 to 2005 can be downloaded from National Bureau of Economic Research (NBER) web site, http://www.nber.org/data/cps.html. The data in the site is named in a somewhat misleading way. Because of the way the survey is conducted, a given year's data is published during the next year, and labeled as the year of publication. That is, the data labeled as 2006 in the above link actually belongs to year 2005. Throughout our analysis, whenever we mention years, we mean the actual year that the data belongs.

The structure of CPS yearly files is hierarchical. One can think of this structure as a huge matrix. Each row consists of a household, a family, or a person record. There are indicators showing which type of record the row holds. Each column represents one digit. So, a person's record is reported in his row by the columns that make up numbers of different length depending on the detail of the corresponding statistic. For example, a person's age in 2005 data is reported by the number that starts at the 15th column and ends at the 16th one. Thus, age is given in 2-digits. Thus, the length of a row is, in a way, an indication of the amount of detail the data includes about persons, families or households (of course there are "fillers", or blanks in some places). To give an idea about the size of the data, we can say that a total of around 210,000 separate persons (rows) were covered in the data for the year 2005. When we add the information for families and households that these people belonged to, this number increases to 396,000. This number was around 200,000 for 1967. For 1967, when our data begins, the length of a row in the giant matrix was 360 columns. This number increased to 396 in 1988. The latest data that is available, the 2005 data, has 976 columns. This increase has been achieved not only by increasing the number of variables, but in many cases the detail level of a variable has been increased as well. Industry data is a good example of this. During the period we are interested in, the data structure was changed significantly in 1975 and 1987, with changes of smaller scale in 1981 and 2002, as well as some minor changes almost every year.

The data is organized as follows: the household record is followed by a family record, which in turn is followed by the record of each person in that family (even if it is a non-family household, the persons are still categorized under the first family record as a family). Then the records of related subfamilies, unrelated subfamilies and non-relatives follow as family records, each followed by person records categorized under them. Once every family under the same household is covered, a new household record comes up. In our analysis we only concentrate on person records, using household and family data attributed to them whenever necessary. For example, information on the type of living quarters and region come from the household record which includes the person under consideration.

We have limited our analysis to the males who are between the ages of 16-64, work full-time and full year (defined as working at least 40 hours a week and 35 weeks a year), earn at least $67 in 1982 dollars per week (half of the real minimum wage based on 40-hour week in 1982 dollars), have at least 1 year of potential labor market experience (defined as: *age - years of education - 6*), not living in group quarters, not self-employed or working without pay.

We concentrate on log values of weekly wage and salary income throughout our studies. The annual wage and salary income entries in ASEC are given in top coded form. We impute the top coded values as 1.33 times the reported maximum value of the variable for that year. Then we deflate the annual wage and salary income to 1982 values, using the personal consumption expenditure deflator from National Income and Product Accounts (NIPA) which can be downloaded from the website of the Bureau of Economic Analysis (BEA, www.bea.gov/bea/dn1.htm). The natural logarithm of this value divided by the number of weeks worked during the reference year gives us average log weekly wage and salary income. In none of the years top coded (and imputed) values are numerous enough to affect the percentiles we used to analyze wage inequality seriously.

Wage and salary income was reported as a single entry in CPS March data before 1987. However, primary and secondary job earnings started to be reported separately after this year. We imputed the top coded values with 1.33 of the top coding value separately before adding them up to find the total wage and salary income. One obvious problem for these years is that top coded values show big fluctuations between years, especially for the secondary job earnings.

Wage and salary income is available in six digits for the entire range of years. However, we have some trouble with the weekly wages. Number of weeks worked during the reference year began to be reported as the exact number of weeks only from the 1975 data. It was a recoded variable before that, giving a value 1 to 7 representing a number of weeks in each group. We used an average of the number of weeks falling into each of these groups for 1975, 1976 and 1977, weighted by March Supplement Weights, to recode the previous years' data into the actual number of weeks worked.

The educational attainment variable sees a few changes in its coding within the range of years we use. However, there is one major point of change: 1990. Before this year, the data was gathered from two separate variables, one that gives the highest grade of the school attended and another that states whether it is completed or not. If a person did not complete his/her last year of education, we simply assumed the years of education without that year. However, starting with 1991, the reporting of the years of education changed significantly. The variable became a recoded one that focuses not on the years of education, but on the degree received. Some values of the new variable represent several years of education grouped together. Some others report only a degree received (for example high school graduate). It became quite hard to follow the exact years of education. Following the general approach by other researchers in this area, we decided not to use the years of education in our regression due to this problem. Instead, we divided our sample into four groups of educational attainment: Less than high school (lowedu-less than 12 years of education), high school graduate (HS-12 years of education), some college (scedu-between 12 and 16 years of education) and college graduate (colledu-16 years or higher). We used four dummies in our regression indicating whether a person belonged to a group or not.

Of course there is still the question of calculating the potential labor market experience variable for each person. We made this calculation by using the average years of schooling for each group from 1988, 1980 and 1990. Park (1994) indicates this method as one of the plausible alternatives of dealing with the change in structure of this variable.

2.2 An Overview of the Data

Table 2.1 gives us some information concerning the demographic details of the sample we are working with, as well as the education and industry information, for the years 1967, 1990 and 2004.

We see first a decrease, then an increase in the mean age of our sample. This is not a surprise to anyone who is familiar with the baby-boom, a term which marks the sudden increase in the population of the US following the Second World War which saw its peak around 1960. One can easily guess that babies born during this era will have some effect on the average and median age of the working population of the country starting from late 60s. The tide turns around in 80s and 90s, as this group grows towards middle age. The implications of this trend for the future, although not part of our analysis, are a line of work that is gaining interest (Baker, 2001).

We also see a decrease in the share of white people in our sample. This is no doubt due to migration from South and Central American countries, as well as the increased number of people who left their ex-communist homes to live in the US within the last decade and half.

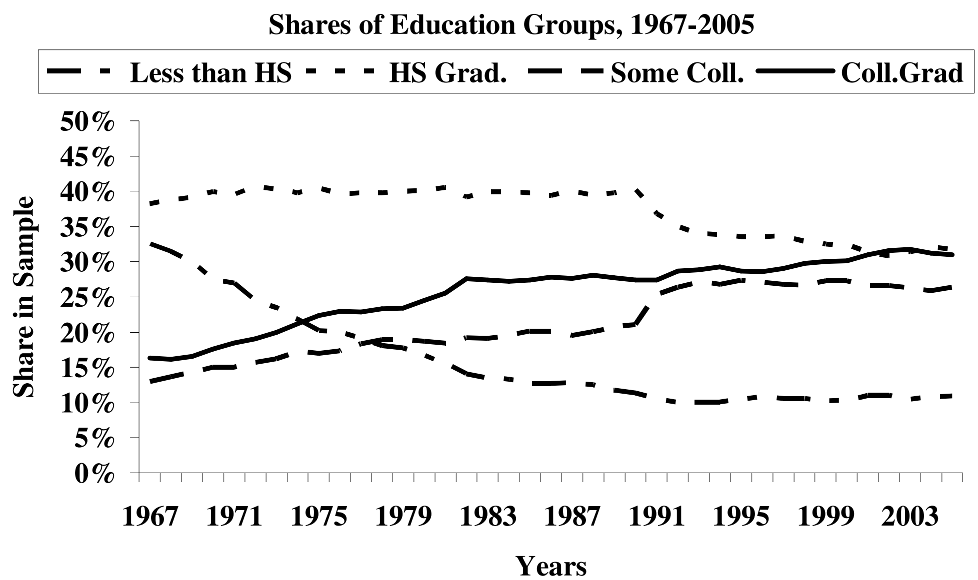
The share of married males also decreases from about 90% in 1967 to 66% in 2005, marking the results of a great social transformation in the country in this time period, especially in the 70s and 80s.

Table 2.1 Descriptive Statistics

	1968 n = 67,852		1990 n = 77,797		2005 n = 106,302	
	Mean	Std.	Mean	Std.	Mean	Std.
Weekly Pay	408.55	234.08	456.47	298.75	552.06	546.39
Demographics						
Age	37.67	10.40	36.86	9.86	39.18	10.43
White	0.91	0.29	0.87	0.34	0.83	0.37
Married	0.95	0.22	0.66	0.47	0.62	0.49
SMSA St.	0.70	0.46	0.65	0.48	0.71	0.46
Education and Experience						
Less than HS	0.31	0.46	0.11	0.32	0.11	0.31
High School	0.39	0.49	0.39	0.49	0.32	0.47
Some Coll.	0.14	0.34	0.22	0.42	0.26	0.44
College Grad.	0.16	0.37	0.28	0.45	0.31	0.46
E<=10	0.26	0.44	0.28	0.45	0.23	0.42
10<E<=20	0.27	0.44	0.36	0.48	0.29	0.45
20<E<=30	0.27	0.44	0.24	0.42	0.30	0.46
30<E<=40	0.20	0.40	0.13	0.34	0.19	0.39
Industry						
Agriculture	0.02	0.12	0.02	0.13	0.01	0.11
Mining	0.01	0.11	0.01	0.11	0.01	0.09
Constr.	0.08	0.28	0.09	0.28	0.12	0.32
Manufacturing	0.38	0.49	0.27	0.44	0.18	0.39
Tr.Com&P.U.	0.10	0.30	0.11	0.31	0.11	0.31
Trade	0.15	0.36	0.17	0.38	0.15	0.36
Finance&Serv.	0.18	0.38	0.25	0.43	0.35	0.48
Government	0.08	0.27	0.09	0.28	0.06	0.24

3-year averages are given, centered on the specified year

The share of high school and lower degree holders seems to have dropped between 1967 and 2005, although with different timing. The former falls before 1990 and remains level after that while the latter is the same level as it was in 1990 but decreases after that. The drop in the share of people who have lesser than high school degree is much sharper, from 31% to 11%. The share of high school graduates, on the other hand, saw a relatively milder decrease, from 39% to 32%. The share of people going to college (either 4 year or lower) increased from 1968 to 1990 and from 1990 to 2005. It nearly doubled both for people with college degree or for people with some college education but not a 4-year college degree.



Shares of four education groups in the wage-earner sample.

Figure 2.1 Shares of Education Groups in Sample

Since we will be talking about the educational attainment quite a bit, it is worth spending some more time on this subject. Figure 2.1 gives the shares of all 4 categories of education we use for each year in percentages, enabling us to see a much more detailed picture of educational attainment in the US over this period.

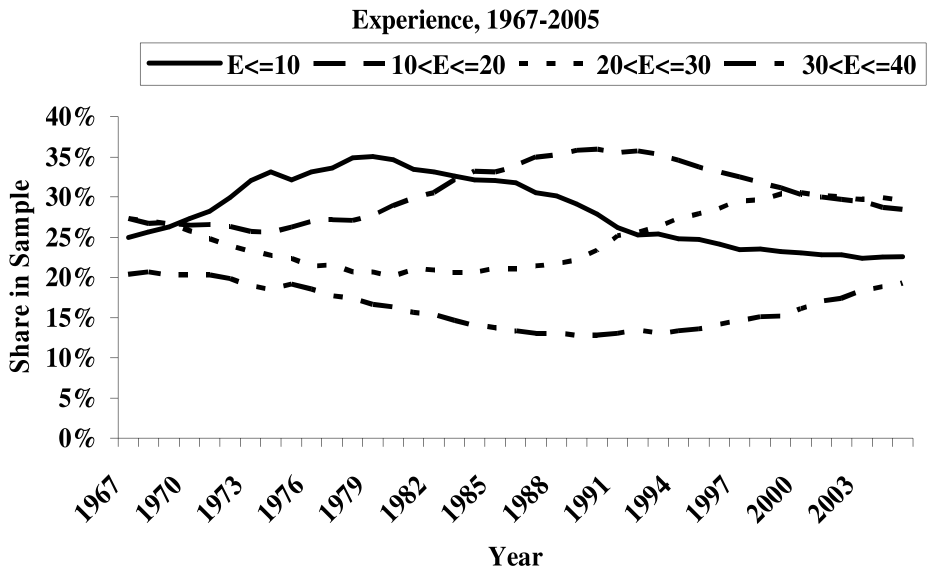
It is clear from the figure that the share of college graduates has been increasing since the beginning of our data. However one has to identify 3 different periods: 1967-1982, 1982-1996, and 1996-2005. The first period is when the share of college graduates increased the fastest. In just 14 years of time, their share increased from 16% to 28% of the sample. Then during the next 15 years it increased just by 1%. In the last period we see that the share of college graduates in the whole sample is increasing again. The rate is slower (4% in about 10 years) but still it is noteworthy.

Share of the people with some college education but not a 4-year degree keeps increasing until 1990 (21%), and then it suddenly jumps in 1991 (25%). This jump is very likely to be a result of the change in educational attainment coding, since the group under consideration includes every kind of education between high school diploma and a 4-year college diploma. After 1991, the share of this group remains more or less the same.

When we look at the high school graduates line, we notice that it remains level until 1990(40%) and then it drops in a matching move with the some college group in 1991 (37%). After that point, it remains level again with one or two minor fluctuations. From the way high school and some college lines match, one is led to consider the possibility that some people that formerly had been in the high school category were transferred to some college category in 1991. Unfortunately the new data structure does not give us a way to prove this. However it is a point that should be kept in mind while analyzing the relationship between wage inequality and education.

The low education group (less than high school graduate) decreases from 1967 (35%) until 1992 (10%) and stays there. They lose about 2/3 of their share and become the smallest of all categories by far.

All of these changes tell us that the employees of 2005 are much better educated than those of 1967. The fast increase in the supply of college graduates during the 70s attracted attention as an indicator of an increase in supply of skills. The stabilizing of college graduate supply at the beginning of 80s is seen as one of the reasons for the increase in wage inequality during the 80s, combined with the increase in demand.



Shares of three experience groups in the wage-earner sample.

Figure 2.2 Shares of Experience Groups in the Sample

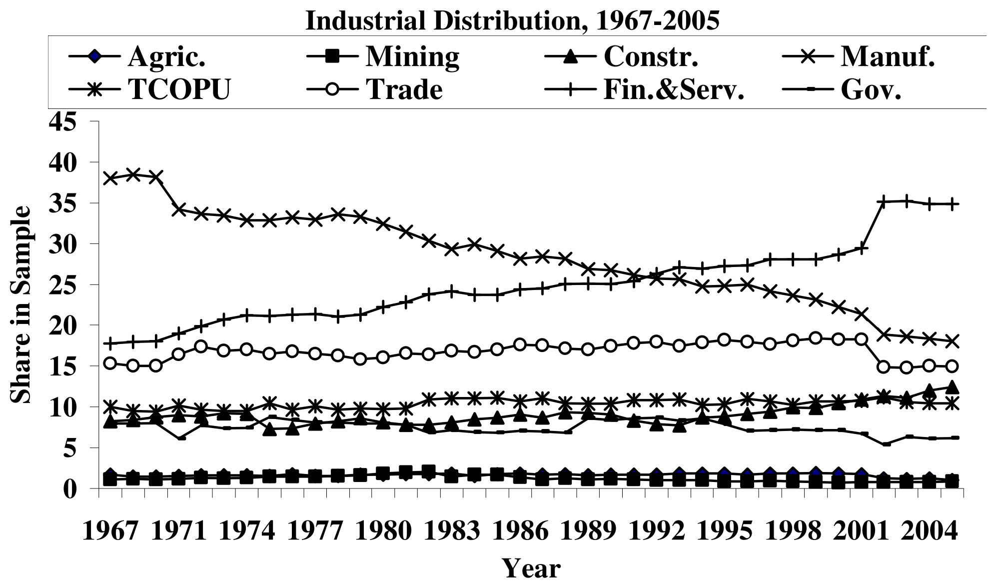
The composition of potential labor market experience was also reported in Table 2.1. We see a more detailed picture in Figure 2.2. The figure shows the percentages of 4 groups of workers in the wage-earner sample: less than 10, between 10 and 20(not including 10), between 20 and 30(not including 30), and between 30 and 40 (not including 40). Of course the shares of these groups are related in a much closer way than those of the education groups. Since all persons in our group are full time and full year employees, and their experience increases as years pass, obviously a new entrant in 1967 will pass from the lowest experience group to the second lowest one in 1977. Even though some people might drop out of the labor force before they reach the highest experience group, there is still a strong sense of continuity here.

Although this looks quite trivial, there is still something to be learned from here. The group with the lowest experience, which is the persons with 10 years or less potential labor market experience, grow compared to other groups until 1980(35%). Then their share starts falling and never recovers again. The share of the 10-20 group begins to increase mid-70s (25%) and this trend goes on until 1990(36%). Then it enters its own free fall which still continues. The third experience group (20-30 years of experience) recovers from a fall in 1985 (21.1%) and rises until 2002 (30%). Then it begins to fall too. And the last group, people with 30-40 years experience, lose their share in the work force until 1990 (13%), then their share begins t rise in a move that still continues (19.3% in 2005).

These transitions are obviously affected by the demographic development of baby-boomer population. Someone who was born in 1955 would be in the work force by 1975, adding to the numbers of the lowest experience group. These people would have ten years of experience during around mid-80s and 20 years of experience around mid-90s. This obviously fits with our peaks. However, our main focus is not the baby boomer population. We are more interested in the dates of these transitions. For our later discussion in wage inequality, it might be a good idea to remember that the lowest experience group had its highest share at the beginning of 80s then started decreasing. And also the third group, that is the people with 20 to 30 years o experience, started increasing their share after 1985. We will refer to these two groups as “low experience” and “high experience” groups in our discussions.

Figure 2.3 gives us details about the distribution of our sample over major industry groups. The numbers are given in percentage of people employed in each industry in a given year.

Looking at Figure 2.3, we see the transformation of US economy in the last three decades. Manufacturing industry had 38% of all employees that met our criteria of selection in 1967. In 2005, only 18% of our sample was working in this industry. Of course there was also productivity increase, however it only partly explains the change here. This big drop is no surprise to anyone who remembers how some manufacturing sectors like automobile industry have moved their production plants outside the US. With all of the other industries keeping more or less the same share as they had in 1967, it appears that the loss of jobs in manufacturing sector was balanced by more job openings in finance and services sector which doubles its share after three decades of steady increase. The only other noteworthy change is the increasing trend in construction share during the nineties and into the new century.



Shares of major industry groups in the wage-earner sample. Agric: Agriculture. Constr: Construction. Manuf: Manufacturing. TCOPU: Transportation, Communication and Public Utilities. Fin. &Serv: Finance and Services. Gov: Government.

Figure 2.3 Industries (Percentage of workers employed in the industry)

2.3 Evidence on Wage Inequality

Now it is time to see how these characteristics of the wage-earner sample are related to wage inequality. Three of the fairly standard measures of inequality are the Gini coefficient, the Theil Index and variance of logs. They do not give a detailed analysis of the nature of wage inequality beyond stating whether it improved or worsened.

However we feel it will be a good starting point to see just what they say about the wage distribution of US between 1967 and 2005.

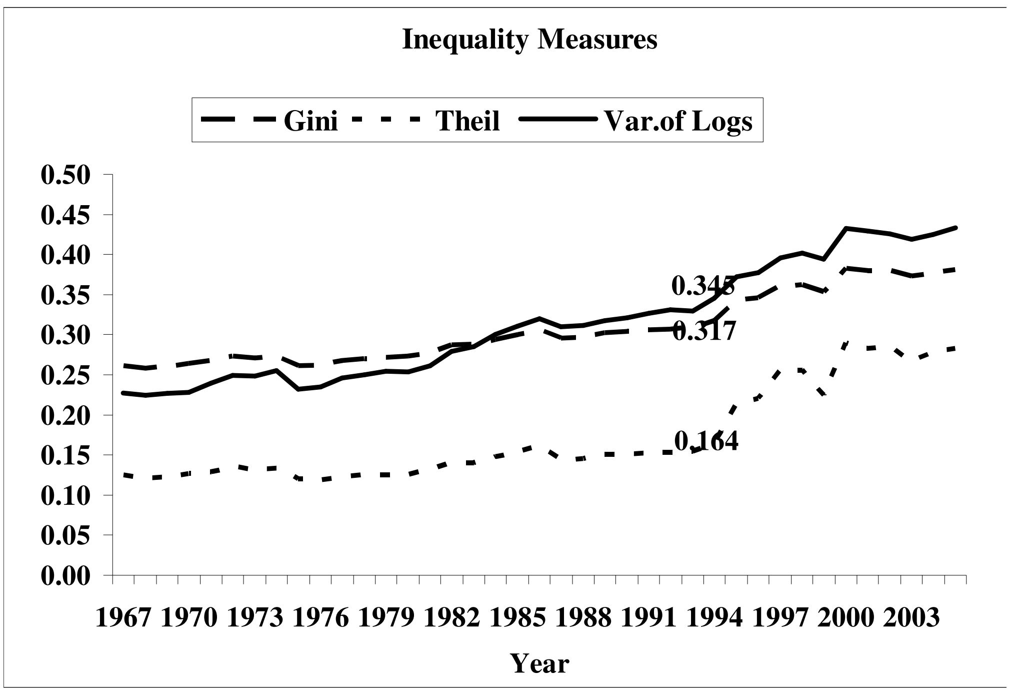
Reported in table 2.2 are the values of these measures corresponding to the real weekly wages of our sample for each year in 1982 dollars. We will not go into a deep analysis of these values since we present them only as an introduction to wage inequality picture. In general, one can say that each of these measures represent "no inequality" when they take 0 value. Then, there is higher inequality as their values increase. Gini coefficient is more sensitive to the changes around the middle of the distribution (the mode). The Theil Index is sensitive to the size of the sample, so it should be used with caution since, as one can see from Table 1.1, our samples are of quite different sizes each year.

Table 2.2 Measures of Inequality

Years	Measures of Inequality		
	Gini Coeff.	**Theil Index**	**Var.of Logs**
1967	0.261 *(0.002)*	0.125 *(0.003)*	0.228 *(0.003)*
1975	0.261 *(0.002)*	0.120 *(0.002)*	0.232 *(0.003)*
1980	0.273 *(0.001)*	0.126 *(0.001)*	0.254 *(0.002)*
1985	0.300 *(0.002)*	0.155 *(0.002)*	0.311 *(0.003)*
1990	0.304 *(0.001)*	0.151 *(0.001)*	0.321 *(0.003)*
1995	0.343 *(0.002)*	0.215 *(0.003)*	0.372 *(0.004)*
2000	0.383 *(0.003)*	0.289 *(0.005)*	0.432 *(0.005)*
2005	0.381 *(0.002)*	0.283 *(0.004)*	0.433 *(0.004)*

Bootstrapped standard errors are given in parentheses.

All three measures show that wage inequality is much more serious in the US in 2005 than it was in 1967. The Theil Index and variance of logs more than double their measures; there is a serious increase in the Gini coefficient, if not as big as the others. The first big jump in inequality seems to be during the first half of 80s. In the second half of the same decade we see a slowdown. Then, the 90s, quite surprisingly if we remember what we read about the literature, pass with remarkable increases. After the turn of the century, the tide is reversed and there is a drop in inequality for the Gini coefficient and the Theil index, while the variance of logs shows only minimal increase.



Values of three inequality measures. The values given on the figure are for 1994, and should be compared to the values for 1995 in Table 2.2.

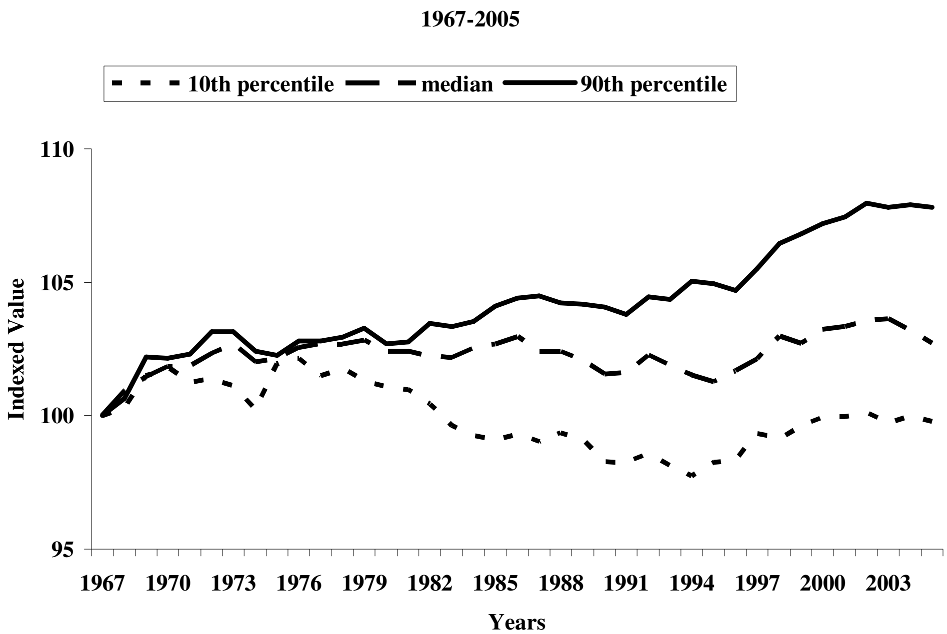
Figure 2.4 Gini Coefficient, Theil Index and Variance of Logs.

Looking at 5-year intervals is no doubt helpful for following how things develop over a period of time. However, as we see in Figure 2.4, it might be a bit misleading as to the timing of changes. Although the jump that was seen in between 1980 and 1985 was a gradual process that was distributed over the years within the interval, we see a different situation in the first half of the nineties. A huge chunk of the big increase that we saw between 1990 and 1995 happened actually between 1994 and 1995. Year-by-year values, however, confirm that there was another tide of increase in wage inequality during the second half of nineties. This is a bit of a surprise, since the wage inequality in the US was supposed to have been stabilized by mid-90s. Wage inequality stabilizes at a higher level by the turn of the new century.

Cumulative measures of inequality have their uses, however one has to look at different points of a distribution to see what makes it change. There are often movements in the tails that are not easily captured by measures like Gini coefficient, if not completely missed. One can learn a great deal by looking at the percentiles of a wage distribution. Based on the log weekly wages, Figure 2.5 shows the change in 90th, 50th (median) and 10th percentiles over the years. To make a clearer comparison possible, the values are indexed to 1967=100.

We see in the figure that the 10th percentile starts losing ground in real terms at the beginning of 80s and actually goes below 1967 level after 1982, staying below that level for a long time. It rebounds back around mid-nineties, no doubt thanks to the live economy of the nineties which saw an increase in demand for all sorts of workers. It only catches up with the 1967 level in 2002. The recovery of the lower tail of the distribution in the second half of the nineties is noteworthy and points to an important change in the wage inequality pattern of the US. We will return to this point later.

Log weekly wage of the ninetieth percentile tells a different story overall. After an intial jump from 1967 to 1975, it remains steady until the 80s. Then it grows until 1987. After staying at this level for about a decade with minor fluctuations, it jumps up during the mid- and late-nineties, ending up at 8 indexed points gain compared to its 1967 level. The interesting story here is that while the median keeps a steady course especially after early 70s, the 90th percentile seems to increase while 10th percentile decreases in real terms. One also has to note that the indexed values of 90th percentile and the median go close to each other until the mid-80s, then they start spreading from each other. Especially during the second half of 90s there is an obvious divergence, while both keep their distance with the 10th percentile throughout the 90s.

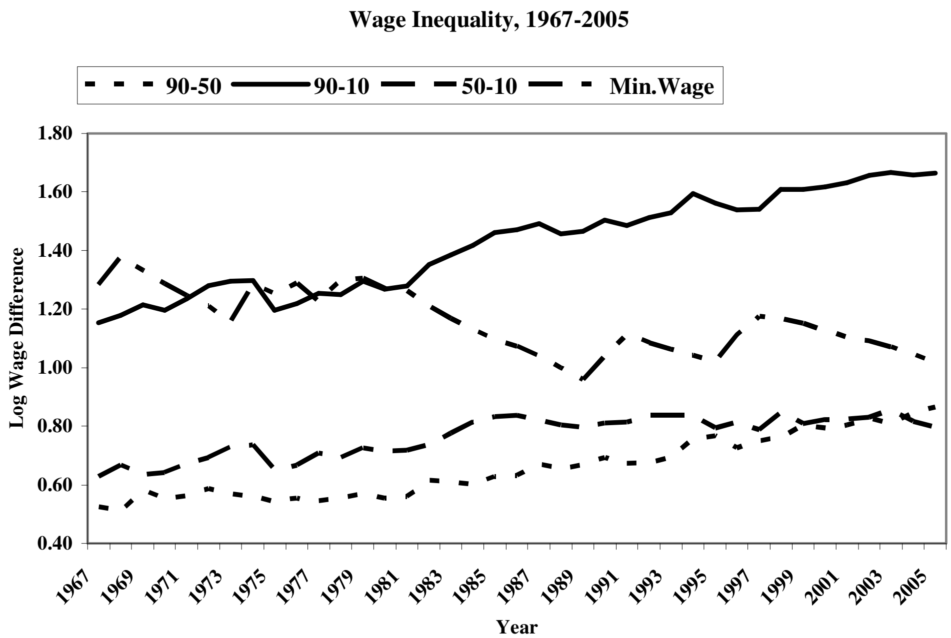


Three percentiles of log weekly wages indexed to 1967=100.

Figure 2.5 Indexed Log Weekly Wages by Percentile

Overall, we notice two trends here. Firstly, from mid-seventies until the 90s the gap between 50th and 10th percentiles increases, then stabilizes. Secondly, the gap between 90th and 50th percentile, not big until the mid-80s, starts expanding afterwards, especially during the nineties. We notice that all three measures remain a level at 2002 from which they do not increase any more, probably due to the slowdown in the US economy and uncertainty related to the unfortunate events in 2001.

In Figure 2.5, we saw the individual paths of these three measures, making comparisons between their relative positions when there is a clear divergence and convergence. We get a better picture of their relative growth in Figure 2.6, which shows the difference between 90th-10th, 50th-10th and 90th-50th percentiles of log weekly wages.



Differences of 90th, 50th and 10th percentiles of log wages for each year. The minimum wage is reported in log 1982 values.

Figure 2.6 Differences of Log Wage Percentiles

90th-10th percentile difference can be thought of as a measure of overall wage inequality, which can be compared to the three general inequality measures that we observed before. It has been increasing since the mid-70s with a few fluctuations.

The period of the most dramatic increase seems to be the 80s, more or less the whole decade. Then it stabilizes for a few years, and it starts increasing around mid-90s at a slower rate than 80s. Then it looks to have stabilized since 2003.

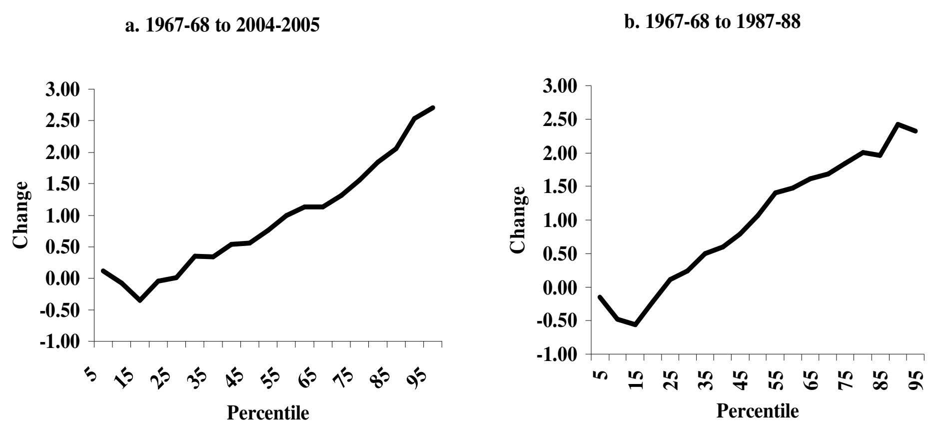
The 50-10 difference, in other words the inequality at the lower half of the distribution, grows parallel to the 90-10 difference until mid-80s, and then reaches at a level that has been kept more or less the same until today. On the other hand, the 90-50 difference follows a different path. It stays steady until 80s, then starts an increasing trend which still continues in 2005. Actually, the 90-50 difference become bigger than the 50-10 difference in 2005 for the only time in our range of years. Also it grows the fastest during the second half of the nineties. Similar to the 50-10 difference, one can think of 90-50 difference as a measure of wage inequality in the upper half of the wage distribution.

We also see from the graph that the log real minimum wage values remain relatively stable over the same period in which the 50-10 difference remains steady, while the jump during the mid-80s coincides with a long drop in the real value of minimum wage. This indicates that one of the reasons for the stability of dispersion in the lower half of the wage distribution in the recent years could be the increases in minimum wage during the nineties which kept the real value relatively stable.

One can think of 90-50 and 50-10 measures as two pieces that add up to create a new one, which is the 90-10 difference. We see from the figure that the increase in 90-10 difference between mid-70s and 80s was fueled by the 50-10 difference. The remarkable increase in 90-10 during the 80s was a result of increased dispersion in both halves of the distribution. For the rest of our time range, 90-10 and 90-50 follow very parallel moves of increase, with 50-10 remaining at the same level. Thus, although we say that 90-10 difference has kept growing for the better part of our range of years, this growth is a result of 2 different trends: Until the mid-80s it is mostly a product of lower-half inequality increase, while in the nineties and to the day, it is closely related to the upper-half inequality.

We observed from Figure 2.5 that the 90th percentile of log real weekly wages increased faster than the 10th percentile over the period 1967-2005, pointing towards an increase in overall wage inequality. This is a general trend, pointed out by JMP(1993) as well. They showed that there is a clearly positive relationship between the percentiles and the change in real wages for the period 1964-88, with the change in the real wage increasing as one goes higher in percentiles. Obviously this implies increased wage inequality over the whole distribution of wages.

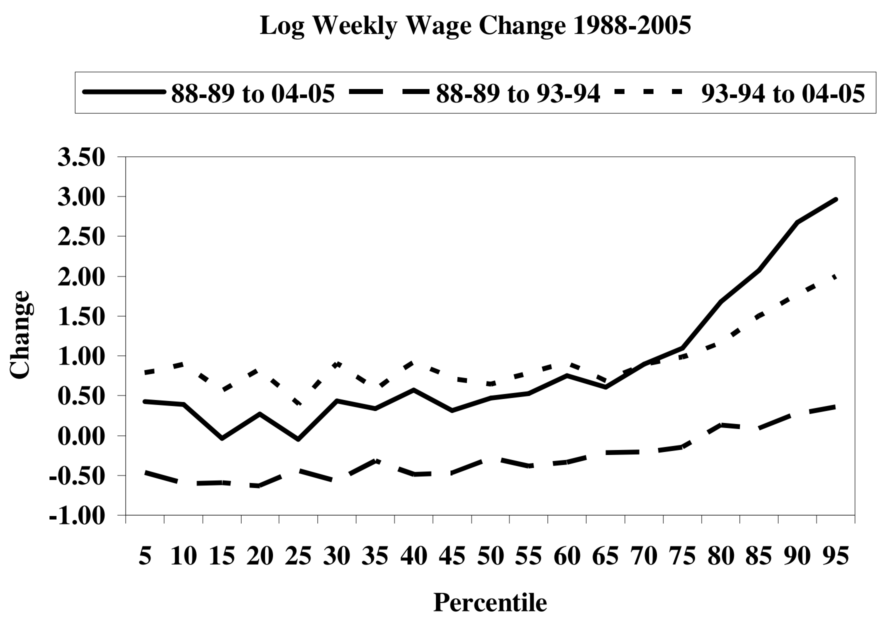
JMP's figure of percentiles and wage changes was similar to Figure 2.7.b except that their data started from 1964. In Figure 2.7.a we have the same picture for the whole time period of analysis. In both figures, earners over the 20th percentile recorded increases in their wages, while the ones below this level earned less than what they earned in 1967-68, of course in real terms. Obviously there is a striking similarity between the two figures that could trick someone into thinking that nothing much changed between 1988 and 2005. From our previous figures and tables we have a feeling that this might not be the case. In fact the two figures represent quite different stories, as we will see in our next figure.



Change in percentiles of wage distribution between two years. The change is calculated as the difference of corresponding percentile between two points in time. Two-year pooled data is used to avoid measurement error.

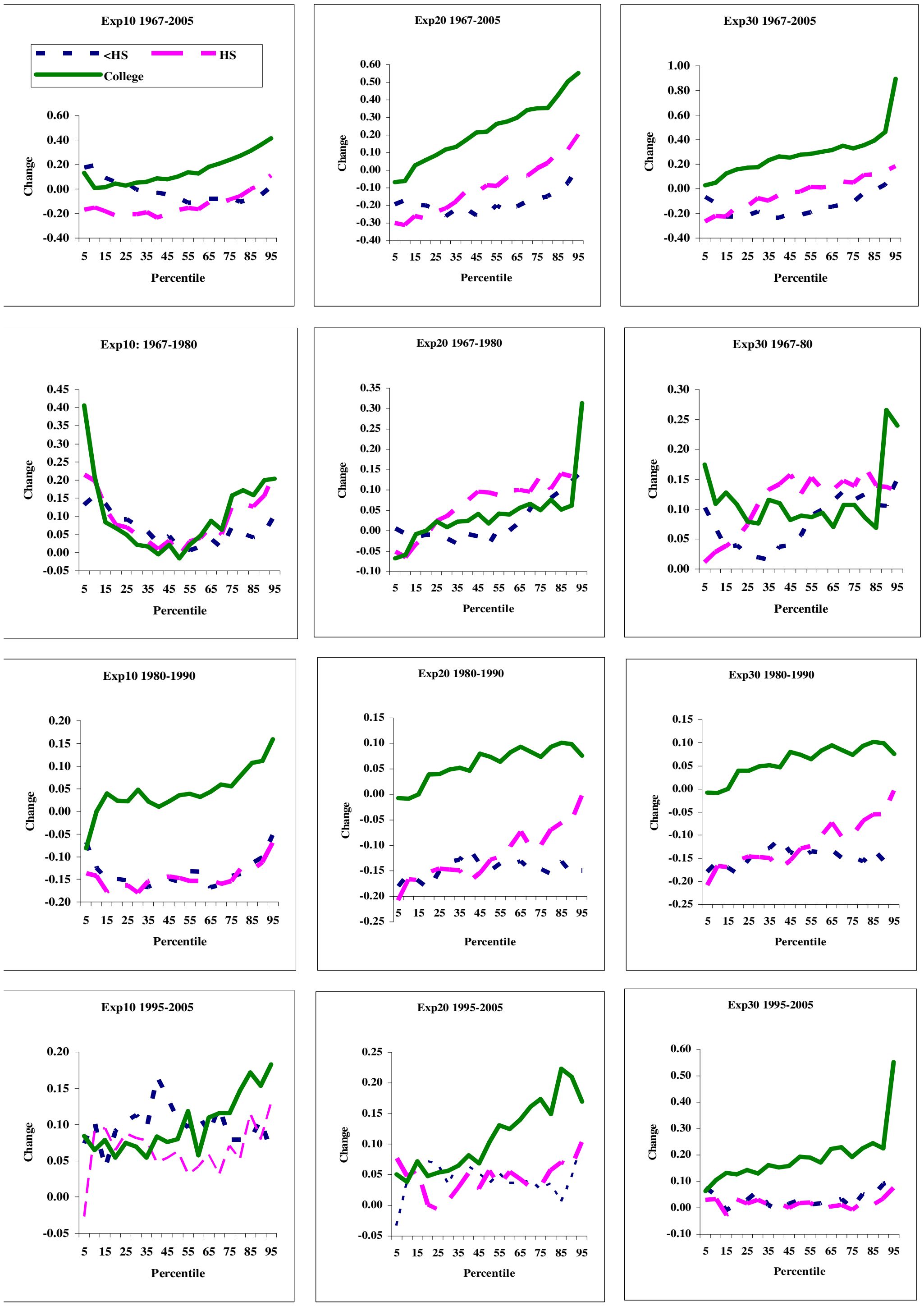
Figure 2.7 Changes in Percentiles of Wage Distribution over Time

To show that 1988-2005 period was different from the 1967-1988 period, we break it into smaller sub-periods in Figure 2.8: 1988-89 to 1993-94 and 1993-94 to 2004-2005. In Graph.2.7, we saw that the lowest 20% of the distribution actually lost in real terms for 1967-1988. For 1988-2005 period, there is a gain at every percentile, although very small in 15th and 25th percentiles. Also, the relationship betwen percentiles and change in wages is not as clear as it was in the previous figure. Although there is some positive relationship between the percentile and the change here as well, it is mostly in the upper half of the distribution, and stronger after the 60th percentile. For the lower levels, there is no such clear trend. However we can say that the highest 25% of the distribution gained more than the rest. This should mean more inequality, both in the upper half and overall.



Change in percentiles of wage distribution between two years. The change is calculated as the difference of corresponding percentile between two points in time. Two-year pooled data is used to avoid measurement error.

Figure 2.8 Changes in Wage Percentiles, 1988-2005



Change in percentiles of wage distribution between two years. The change is calculated as the difference of corresponding percentile between two points in time. Two-year pooled data is used to avoid measurement error. Due to space limitations, the legend is given only for the first figure.

Figure 2.9 Within Inequality

We had observed a considerable jump in overall inequality after 1994. In this respect, comparing the sub-periods in Figure 2.8 should be interesting. The first thing that draws attention is that the differences between 93-94 and 2004-2005 are all positive, meaning that all earners recorded real gains during the this period. The increased demand for labor in 90s obviously meant gains for all percentiles as we can see from the the line corresponding to 93-94 to 2004-2005, benefiting 75th and higher percentiles more than the others. On the other hand, 1988-89 to 1993-94 period sees a loss in most percentiles, except for the highest few. It is no longer possible to observe clearly positive relationship overall betwen the percentile and the change in either case. However, we also observe that in both sub-periods, and in the whole range as well, there is increased inequality in the upper half of the wage distribution. This is more severe in the later period. These all confirm our findings earlier that especially from the mid-nineties it is mostly a matter of upper-half inequality growth.

Another aspect of wage inequality growth in the United States that has been drawing attention is the so-called within inequality. Our analysis so far has focused on wage inequality in general sense, that is, the size of the gap between high earners and low earners. There is more to this story than just that. The wage inequality among the people with similar observable qualities such as education and work experience has also shown some interesting tendencies over the years. Figure 2.9 gives us an interesting picture of this phenomenon. We have created 3 groups of workers based on their years of potential labor market experience: Exp10 represents those with 10 years or less experience, Exp20 represents those with between 10 and 20 years of experience (not including 20) and Exp30 represents those with between 20 and 30 years of experience (not including 30). We have also created 3 subgroups of workers according to their educational attainment: <HS stands for education less than high school degree, HS stands for HS graduate and no more education, and College stands for 4-year college graduates and higher levels than that. Then we organized each experience group according to their education levels. We calculated the change in wages by percentile (same as in figures 2.4 and 2.5) within these groups for the periods 1967-2005, 1967-1980, 1980-1990 and 1995-2005. Each row of figures

shows the same time period for 3 different experience levels. When calculating the percentiles, we took averages of the last 2 years on each boundary of the time period. For example, 1967-2005 change is the difference of the average 2004 and 2005 from that of 1967 and 1968.

The figures for 1967-2005 period show that people with college degree not only earn better wages in 1967 than in 2005, their gain is also better than high school graduates on all percentiles (except for the Exp10 group where they gain more than high school graduates after the 20th percentile). So we can say that there is increased wage inequality due to educational differences. We also observe that there is an increase in wage inequality among the college graduates and among the high school graduates themselves, higher percentiles gaining more than lower percentiles. This is an evidence of within inequality among the earners with similar education and experience profiles. The same cannot be said for the third education group, those with less than high school degree. For one thing its seems like they earn less than 1967 except for the highest 5th or 10th percentiles. It is also noteworthy that the ones on the edges of wage distribution for this group are better off than the ones nearer to the median.

1967-80 pictures do not give a clear picture about between inequality, but we notice that there is an increase in lower-half inequality for the high school graduates in Exp20 and Exp30 groups.

Naturally, we are interested in the 1980-1990 period which showed increased wage inequality in our earlier analysis as well as in the literature. We notice the same trend here as we saw in 1967-2005 picture. There is an increase in inequality generated by college graduates gaining more than both of the other two groups. Also high school graduates gain more than the lower education group especially in upper percentile. There is also evidence of within inequality here, especially for Exp20 and Exp30 groups, which are the more experienced earners.

The 1995-2005 figures give a different picture. For one thing, the college graduates' earnings increased more than those of the high school graduates in each experience group. This is especially so for the workers with higher experience. We also see what we pointed out before, upper-half wage inequality growing within both college and high school categories especially for Exp10 and Exp20 groups. The low education group with 10 years or less experience makes small gains for this time period throughout its wage distribution, but there is no clear sign of any sort of within inequality there. So, although the between inequality picture is not as clear for 1995-2005 period as it was for 1908-1990 period, we can see that it is still there, noticeably in the upper percentiles. There is also a growing within inequality which shows itself especially in the upper percentiles.

It is also informative to compare this figure with Graphs 2.1 and 2.2. We see that the 1980-1990 period falls into the time period when the share of college graduates in the wage-earner sample was more or less stable. In Figure 2.9 we see that this period is when the college graduates obtained the biggest gains over the other groups, probably due to increased demand and stable supply.

Another, and quite popular, way of getting an idea about the evolution of within inequality is by observing regression residuals, which partly represent the effects of unobservable characteristics on wages. Following this logic, the inequality in the distribution of regression residuals (residual inequality) for a given year can be seen as an indicator of within inequality. On Table 2.3 we see the 90-10, 90-50 and 50-10 percentile differences for 5-year intervals. Panel A gives these values for log weekly wages that is the empirical data. Panel B reports the same values from the residual distribution from a wage regression. In panel C we see how much of the values in Panel A are explained by the values in Panel B, in percentage terms. In other words, Panel C shows how the "residual inequality" compares to the inequality measures from empirical data.

One can see the existence of strong within inequality from the table. The residual distribution seems to explain about ¾ of the inequality in the empirical data. The overall inequality (90-10 difference) and upper half inequality (90-50 difference) values, residual distribution has highest match in 1980. For the lower half inequality, the best match is 2005. It is also interesting to note that the portion of 90-10 difference explained by the residuals remained mostly the same since 1980. However, this is a result of two opposite trends: The portion of 90-50 inequality explained by residuals has decreased, while that of 50-10 increased. There seems to be evidence of an increase in within inequality in the lower half, while it decreases in the upper half. Of course one needs to approach these values with a little bit of care, as the residuals represent a number of other things than unobserved variables.

Table 2.3 Residual Inequality

A.LOG WEEKLY WAGES

Percentile	1968	1975	1980	1985	1990	1995	2000	2005
90-10	1.18	1.24	1.28	1.45	1.49	1.57	1.62	1.66
90-50	0.54	0.55	0.56	0.62	0.68	0.75	0.80	0.85
50-10	0.64	0.68	0.72	0.83	0.81	0.82	0.82	0.81
std	0.50	0.51	0.52	0.57	0.59	0.62	0.65	0.66

B.WEEKLY WAGE RESIDUALS

Percentile	*1968*	1975	1980	1985	1990	1995	2000	2005
90-10	0.93	0.95	1.00	1.08	1.11	1.18	1.23	1.26
90-50	0.44	0.44	0.47	0.51	0.52	0.57	0.61	0.61
50-10	0.49	0.50	0.53	0.57	0.58	0.61	0.62	0.64
std	0.40	0.40	0.41	0.44	0.46	0.50	0.53	0.53

C.% OF DIFFERENCE EXPLAINED BY RESIDUALS

Percentile	*1968*	1975	1980	1985	1990	1995	2000	2005
90-10	78%	76%	78%	74%	75%	75%	76%	76%
90-50	81%	80%	84%	82%	77%	75%	76%	72%
50-10	76%	73%	73%	69%	72%	75%	76%	79%
std	80%	79%	79%	78%	78%	80%	81%	80%

The residuals are obtained from a regression of log weekly wages on a quadratic of experience, education dummies for less than high school, high school graduate, some college and college graduate, industry dummies and demographic dummies like married, white, metropolitan area and living in the south. All regressions are 3-year pooled regressions centered on the indicated year except 2005, which is a 2-year pooled regression of 2004 and 2005.

2.4 Concluding Remarks

There have been considerable changes in the structure of the wage-earner sample in nearly four decades. Notably, the percentage of people who attended college increased, while the percentage of people with lower degrees, especially of those with less than high education, decreased in dramatic rates. A part of employment seems to have shifted from manufacturing to finance and services which usually demand people with computing skills.

The increase in wage inequality from mid-70s can easily be observed from the data. Although the increase slowed down a bit after mid-80s, it continued until 2003. After mid-80s, the composition of wage inequality seems to have changed. The dispersion in the lower half stabilized while the dispersion in the upper half has been increasing ever since, fueling the increase in wage inequality. There is also evidence of strong within inequality across the education and experience groups.

CHAPTER 3

JUHN-MURPHY-PİERCE DECOMPOSİTİON

We have seen so far that the wage inequality increased for the better part of the last three decades. Some of this increase can be explained by the changes in some observable characteristics such as education, experience and industry of employment that reflect ownership of certain skills. There is also some inequality that cannot be explained by observable skill differences, that is the inequality among people with similar characteristics. Although the figures and tables in the previous chapter gave us an indication about how these factors work toward inequality, they told us little in terms of the relative magnitudes of their contribution.

Developed by JMP (1993), this method decomposes the changes in wages into the effects of changes in observable individual characteristics (education, experience etc.), the effects of changing prices of these observable skills, and changes in the distribution of residuals using a wage equation and comparing for different percentiles.

3.1 The Method

It starts with a wage equation such as:

$$Y_{it} = X_{it}\beta_t + u_{it} \tag{1}$$

where Y_{it} is the log weekly wage for individual i in year t, X_{it} is a vector of individual characteristics and u_{it} is the part of wages accounted for by the unobservable characteristics, defined by an individual's percentile in the residual distribution, θ_{it}, and the distribution function of the wage equation residuals $F_t(\ .)$. These residuals can be expressed as:

$$u_{it} = F_t^{-1}(\theta_{it})$$

The wage equation can be manipulated so as to capture three sources of inequality: changes in the distribution of X's, changes in β's and changes in $F_t(\ .)$. Suppose that $\overline{\beta}$ symbolizes the average coefficients vector and $\overline{F}^{-1}(\theta_{it})$ symbolizes average inverse distribution function of the wage equation residuals over the two periods in question. Then, by adding and subtracting $X_{it}\overline{\beta}$ and $\overline{F}^{-1}(\theta_{it})$ on the right hand side, we have the following:

$$Y_{it} = X_{it}\,\overline{\beta} + X_{it}(\beta_t - \overline{\beta}) + \overline{F}^{-1}(\theta_{it}) + \left[F_t^{-1}(\theta_{it}) - \overline{F}^{-1}(\theta_{it})\right] \tag{2}$$

The first term on the right-hand side captures the effect of changes in observable characteristics, the second one captures the effect of changing skill prices of these observable characteristics, and the last one captures the effect of changes in the distribution of wage residuals. Then, one can restrict (2) to find different wage distributions.

For example, with average observable skill prices and average residual distribution, we have

$$Y_{it}^1 = X_{it}\,\overline{\beta} + \overline{F}^{-1}(\theta_{it}). \tag{3}$$

Equation (3) attributes the change in wage distribution from one year to the other only to the changes in X's, observable characteristics. On the other hand, if we allow the quantities and prices of the observable characteristics change over time, the wage distribution can be generated by

$Y_{it}^2 = X_{it}\,\beta_t + \overline{F}^{-1}(\theta_{it})$

(4)

Finally, allowing observable prices, quantities and the distribution of residuals to change in time, we have

$$Y_{it}^3 = X_{it}\,\beta_t + F_t^{-1}(\theta_{it}) = X_{it}\,\beta_t + u_{it} \qquad (5)$$

Obviously (5) is the same as (1). In practice, one generates these hypothetical distributions as follows:

a) A pooled OLS regression of wages on individual characteristics over the whole sample is run. The estimated coefficients from this regression are used as $\overline{\beta}$, the average coefficients vector. The residuals from the same regression are ranked by percentiles to construct $\overline{F}^{-1}(\theta_{it})$

b) Separate OLS regressions are run for each year. Coefficients are recorded as β_t, coefficient vector for each year. Residual distributions $F_t^{-1}(\theta_{it})$ are constructed for each year as in (a).

c) Hypothetical distribution (3) is constructed to calculate Y_{it}^1 for each year. The average coefficients vector is used as coefficients for individual characteristics vector $\overline{\beta}$. Each individual is assigned the residual from $\overline{F}^{-1}(\theta_{it})$ that is of the same percentile as her rank in the residual distribution from the regression run for year t.

d) Hypothetical distribution (4) is constructed to calculate Y_{it}^2 for each year. This time, β_t are used as the coefficient vector for the individual characteristics vector, X_{it}. Residuals are assigned in the same way as (c).

e) Hypothetical distribution (5) is simply constructed from the estimated coefficients and residuals from the separate regressions run for each year.

f) Then, the inequality measure that needs to be decomposed (90th-10th percentile difference, for example) is calculated for each of these six distributions. After this point, one proceeds as follows to decompose the change in wage distributions:

The change in inequality measures of Y_{it}^{1} from one year to the other is attributed to the change in observable characteristics.

The change in inequality measures of Y_{it}^{2} over time is calculated and compared to the change in Y_{it}^{1}. Any difference is attributed to the change in the coefficients of observable characteristics (skill prices).

The change in inequality measures of Y_{it}^{3} from year to year is calculated and compared to the change in Y_{it}^{2}. Any further increase or decrease is attributed to changes in unobservable characteristics.

The difference of this change from the change in (4) is attributed to the change in the coefficients of observable characteristics (skill prices); and the difference between the change in (5) and the change in (4) is attributed to changes in unobservable characteristics, or unmeasured skills and their prices.

The main advantage of the method is that it decomposes the change in wage inequality into observable prices, observable quantities and unobservable prices and quantities. This contribution is important in the sense that it actually gives an account of the share of "residual" inequality in the overall wage inequality, enabling us to compare it to the effects of changing observable skill composition and prices of these skills. As a result, we get a picture of the change in "within" inequality that has made many people busy since the first signs of increase in wage inequality in the US during 80s.

JMP's decomposition technique was criticized for the sensitivity of its results to the order of decomposition (Autor and Katz, (1999)). Another criticism brought against the application of the method was about the way the changes in the distribution of residuals were modeled, which sometimes makes the total of decomposed components to be slightly different from the observed total change (Lemieux, (2002)).

3.2 Application to US Data

At this point, it will be a good idea to have a look at Figure 3.1 to see what kind of information this decomposition could give us. Panel A shows the observed values of 90^{th}-10^{th} percentile difference of log wages. It is used here to be able to see the status of wage inequality at any certain point of time. It was actually reported earlier, as Graph.... 90^{th}-10^{th} percentile gap grows starting from about mid-70s. After that, it keeps growing, at first remarkably fast during 80s, then slower recently. Panels B to D give us an idea about the contribution of observable quantities (skills and characteristics), observable prices, and unobservable skills and prices to overall wage inequality[2].

Before going into what the Figure 3.1 says to us, we need to make a few points about its structure. Panels B to D show the decomposition of change in 90^{th}-10^{th} percentile difference for each year compared to the values obtained from a pooled regression. Thus, they are reported as difference from their long-term means. This structure lets

---

2 These results are obtained from application of JMP on the regression of log weekly wages on a quadratic of experience, education dummies for less than high school, high school graduate, some college and college graduate, industry dummies and demographic dummies like married, white, metropolitan area and living in the south. All regressions are 3-year pooled regressions centered on the indicated year except 2005, which is a 2-year pooled regression of 2004 and 2005. Dropped education category is less than high school, dropped industry is agriculture.

us see if a component contributes to wage inequality significantly or not. The farther are the values from 0, the more contribution that component makes to the overall wage inequality for the given period of time.

When we look at Panel B, it is quite clear that the contribution of observed skills and characteristics to the sharp increase in wage inequality during the 80s are very limited. We see that after mid-70s the contribution of the change in observable characteristics to the total change is very limited, it is always close to zero with no clear trend. This corresponds to the changes in composition of education and experience categories. They do not seem to affect wage inequality much after the 80s, either.

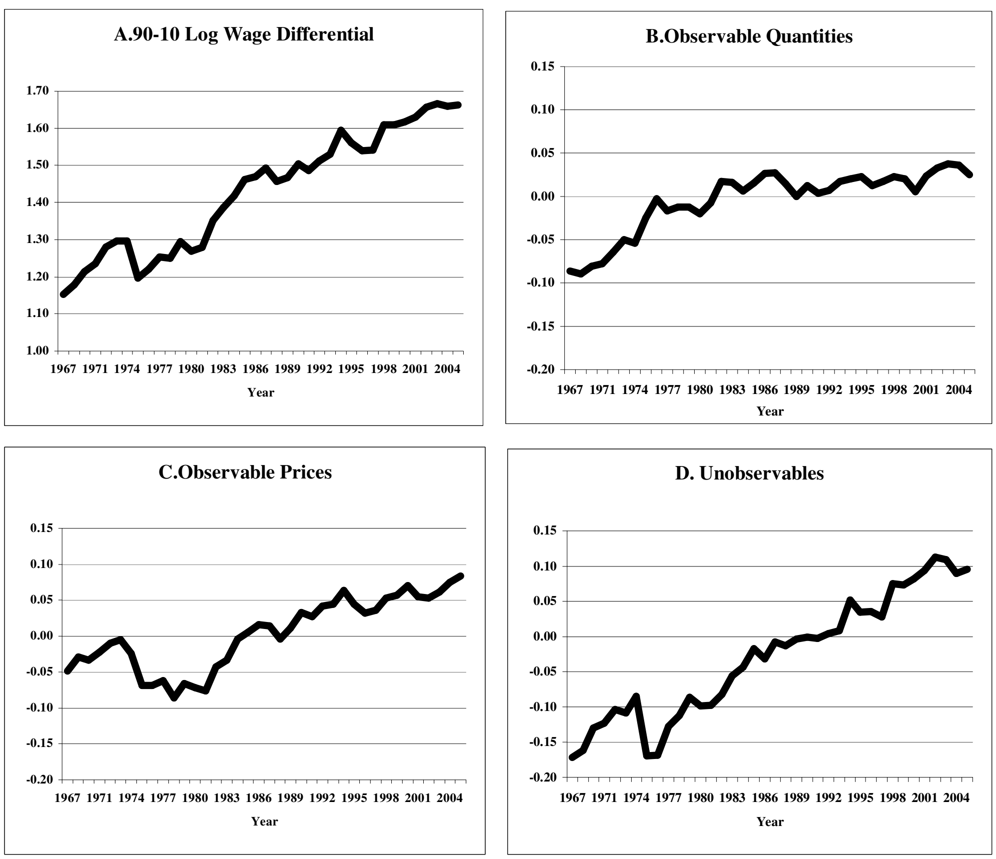


Figure 3.1 Decomposition of 90th-10th Log Wage Differential over the Years

On the other hand, one can see from Panel C that changes in observable skill prices, or returns to skills, have a clearer effect on the overall change, beginning in 80s. This is in accordance with our earlier remarks about changing returns to education and experience during the 80s. Due to both supply and demand factors, the increase in the earnings of people with higher education and more experience result in higher wage inequality. Panel D shows that effect of the change in the composition and prices of unobserved skills looks to be both more noticeable and older than that of the prices of observable skills. It is important to note here that while the impact of observable prices is not very large prior to 80s, the unobservables constitute a large part of inequality from the very beginning of the time period we analyze. This point was made by JMP (1993) as well.

Table 3.1 lets us see the story that lies under Figure 3.1. The coefficients and distributions are all taken from the same regressions that were used to construct Figure 1. However, this table is different from Figure 1 in its setting. Here, we do not decompose the difference from overall means. This time we decompose the change in inequality between two points in time for five-year intervals. The first interval is actually seven years since it would not mean much to create a 2-year interval to decompose the change in wage inequality. It is also important to remind here that all references to a specific year actually mean 3-year data centered around the mentioned year. The table reports JMP decomposition for three measures of inequality: 90^{th}-10^{th}, 90^{th} –50^{th} and 50^{th}-10^{th} percentile differences of log wages. While the first one of these three measures gives us an idea about the overall wage inequality, the second shows a picture of upper half of the wage distribution and the third lets us see the situation in the lower half.

We have already looked into the change in 90^{th}-10^{th} percentile difference a bit. It is evident that there has always been some sort of increase in wage inequality since 1967. While not very remarkable during the 70s, we can see it jumping in the first half of 80s. We also notice that the overall inequality has been growing since the 90s, but at a lower rate than it did during the first half of 80s. When we compare the total effect of observable prices and quantities (they even cancel each other out in some cases) to that of unobservable prices and quantities, we sees that the latter is a deciding factor in the composition of total change. The effect of unobservables is always positive, meaning that they tend to increase the inequality. They account for about 47% of the total change between 1980 and 1985. For all the periods after that, with the exception of 2000-2005, they account for more than half of the total change.

Table 3.1 JMP Decomposition for 5-year Intervals

Year	Percentiles	Total Change	Components		
			Obs.Quant.	Obs.Price	Unobserved
A.1967-1975	90th-10th	0.03	0.06	-0.04	0.01
	90th-50th	0.01	0.02	-0.01	0.00
	50th-10th	0.01	0.04	-0.03	0.01
B.1975-1980	90th-10th	0.04	0.08	-0.13	0.09
	90th-50th	0.01	0.01	-0.06	0.05
	50th-10th	0.04	0.06	-0.07	0.04
C.1980-1985	90th-10th	0.17	0.03	0.06	0.08
	90th-50th	0.06	0.00	0.03	0.03
	50th-10th	0.11	0.03	0.03	0.05
D.1985-1990	90th-10th	0.02	-0.01	0.01	0.03
	90th-50th	0.06	0.02	0.02	0.02
	50th-10th	-0.04	-0.03	-0.01	0.00
E.1990-1995	90th-10th	0.09	0.02	0.02	0.05
	90th-50th	0.07	0.03	0.01	0.03
	50th-10th	0.03	-0.01	0.01	0.02
F.1995-2000	90th-10th	0.06	0.00	0.03	0.03
	90th-50th	0.05	0.01	0.02	0.02
	50th-10th	0.01	-0.01	0.01	0.01
G.2000-2005	90th-10th	0.04	0.02	0.01	0.01
	90th-50th	0.04	0.04	0.00	0.00
	50th-10th	0.00	-0.02	0.01	0.01

The dispersion in the upper half of the wage distribution (90^{th}-50^{th} percentile difference) is quite limited until 80s, then it picks up speed for the rest of the decade, and continues to grow during the nineties as well. It is remarkable that 1990-95 period is where it grows the fastest. The effect of unobservable skills and their prices is quite strong here, as well.

The story in the lower half of the wage distribution is a bit different from that of the upper half. We notice that about two thirds of the increase in overall inequality comes from the lower half of the wage distribution for the period 1980-85. After this point, the change in the dispersion of the lower half of wage distribution seems to be quite limited, especially since the mid-90s.

After 1980, observable characteristics affect inequality in upper and lower half of the wage distribution in opposite ways. In the upper half, they work towards increasing the wage inequality. On the other hand, they work towards stretching it in the lower half.

Different directions of upper and lower half wage inequality is actually a point gaining interest recently, with two explanations offered: high demand and tight labor markets which benefited workers with low skills during the 90s, and that the skill-biased technological change did not affect the lower end of the wage distribution (Autor, Levy and Murnane (2003)). Autor, Katz and Kearney(2005) confirm the second argument.

It is also interesting to note in the table that while observable and unobservable quantities have both increasing effect on the wage inequality for the first two periods, observable prices work in the other direction. We had noted above that most of the change during the period of analysis comes from the changes in unobserved characteristics, and the effect of the change in the composition of observed characteristics was very small after mid-80s. We see this trend here as well: Although the change in composition of observed characteristics constitutes a sizable part of the modest overall inequality increase overall until 1980, starting from 1980-

85 period it is less important than the other two components except the last 5 years. A similar observation can be made for the upper and lower half inequality as well. It actually works to decrease the lower level inequality.

Observed skill prices constitute most of the change for 1985-90 period, and generally constitute a big part of the changes. With only one exception (50^{th}-10^{th} percentile difference for 1985-90), they always work towards increasing inequality.

One has to point out that all 3 of these components have some time periods when they become more important than the others. Thus, there are differences among the timing of the change in inequality.

In Previous Reference, we showed that there has been a shift in the composition of work force between the industries. To test if these shifts had a considerable effect on the wage inequality, we use a simple variance decomposition that has been used often in the literature. The variance of all wages in an industry in year t can be decomposed as follows:

$$\sigma_t^2 = \sum_j s_{jt}\sigma_{jt}^2 + \sum_j s_{jt}(w_{jt} - \overline{w}_t)^2$$

where s_{jt} denotes the share of industry j in the employment of year t, σ_{jt}^2 is the variance of wages in industry j at year t, w_{jt} is the average wage in industry j at year t, σ_t^2 is the overall variance of wages at year t, $\overline{w}_t$ is the average wage in the economy at year t. The first term on the right hand side shows the "within" effect, while the second term shows the "between" effect. Here, the between effect symbolizes the effect of wage changes between industries, and the within effect stands for the effect of wage variance changes within industries. What we want to see is if there is a strong between effect, since it would indicate that the composition of workforce between industries has a significant effect on wage inequality.

Table 3.2 Variance Decomposition of Wages For Industry Composition Effect

		Within Industry			Between Industry		
Period	**Total Chng. in Variance**	**Chng. in Variance**	**Comp. Effect**	**Total**	**Chng. in Wages**	**Comp. Effect**	**Total**
1967-75	0.011	0.004	0.004	0.008	0.002	0.001	0.003
1975-80	0.009	0.006	0.002	0.008	0.002	0.001	0.003
1980-85	0.059	0.052	0.002	0.054	0.004	-0.001	0.003
1985-90	0.020	0.018	0.001	0.019	0.001	0.000	0.001
1990-95	0.046	0.039	0.002	0.042	0.002	0.001	0.003
1995-00	0.041	0.037	0.002	0.039	0.002	0.001	0.003
2000-05	0.016	0.019	0.005	0.024	-0.007	-0.002	-0.009

One can compare the variances of wages in two separate years using this decomposition to see which part affects the change more. Manipulating the difference equations a little bit, we can decompose within and between effects into the effects of changing wages and effects of changing industry composition. We report these values in Table 3.2.

We see that the in each case the change in the variance of wages comes mostly from within industry changes. This is in harmony with our earlier findings. We also notice that the composition effect, that is the effect of change in the industry composition, is small in both within and between industry changes.

3.3 Concluding Remarks

Data indicates that the relatively fast increase during the 80s in overall wage inequality, measured by 90-10 percentile difference of log wages, slowed down in the next decade. However, it still keeps increasing at a lower rate. Although some of the increasing wage inequality can be explained by changes in the supply of observable skills (high school vs. college graduates and higher experienced versus lower experienced) and the skill premia for these skills, a large portion of it comes from unobservable skills, in other words inequality within same observable skill groups. When we look at the change in wage inequality among the industries, we find that within industry wage variance effect surpasses between industry wage changes and composition changes in explaining the change in variance of wages in the economy

Our estimates using the JMP decomposition confirm that the increase in overall inequality was driven mostly by lower-end inequality until mid-80s. Since then, it is mostly upper-end inequality.

As it was reported in JMP (1993), although the skill premia increased since the 1960s, the changes in returns to experience, education and unobserved skills (as computed by residuals) follow very different trends, thus making it very hard to attribute the overall change exclusively to one of them. However, the effect of unobservables is always large. The total effect of changes in observable characteristics have been limited since the 80s. However this seems to be a result of opposite movements in the two halves of the wage distribution. They increase inequality in the upper half and decrease in the lower half. Although the effect of observable prices does not follow quite the same trend, in general we can say that their combined effect has matched that of the unobservables in the upper half since 1980s, sometimes going well beyond. This could be interpreted as evidence that between inequality plays an important role in the surge of the upper half wage inequality in the recent years.

CHAPTER 4

QUANTILE REGRESSION

The JMP decomposition technique provides its results based on Ordinary Least Squares (OLS) regression, that is the conditional mean of the data. Although the method is quite informative as we saw in the previous chapter, focusing only on OLS estimates might be misleading in certain situations when the "mean" loses some of its "meaning". For example when the distribution has thick tails, or not symmetric, that is when the distribution is less like a normal distribution, one needs to take these unusual characteristics into account when doing analysis.

Estimating the model for each quantile using the quantile regression techniques has the benefit of giving a parsimonious description of the entire conditional wage distribution, whatever the shape of it. Then, it can be used to examine the dynamics of wage inequality under a new light. Quantile regression estimates can be used to see the effect of a covariate on within-group wage inequality, as well as seeing the effects of different skill attributes in each quantile (Buchinsky (1994), see Martins and Pereira (2004) and Machado and Mata (2001) for applications).

4.1 The Method

Given $\theta \in [0,1]$, the θth quantile of a random variable Y is a number q_θ such that Pr (Z< q_θ) $\leq \theta \leq$ Pr (Z $\leq q_\theta$). In a more standard expression,

$$q_\theta = \inf(Y : F(Y) \geq \theta).$$

This means that θ % of the cumulative density of Y falls below q_θ. For example, the median can be symbolized as $q_{0.5}$.

One can also think of the quantiles as the solution of a minimization problem: $\min_k \sum_{i=1}^{n} |Y_i - m|$. This minimization obviously leads us to the median. Following a "pinball logic" as it was called in Koenker and Hallock (2000), this approach can easily be modified to obtain other quantiles as well:

$$\min_k \left\{ \sum_{i:Y_i \geq k}^{n} \theta |Y_i - k| + \sum_{i:Y_i \prec k}^{n} (1-\theta) |Y_i - k| \right\}.$$

What we do here is to minimize the weighted sum of absolute deviations to find the θth quantile of the sample distribution. For example, if we want to find the 25th quantile, we have the following:

$$\min_k \left\{ \sum_{i:Y_i \geq k}^{n} 0.25 |Y_i - k| + \sum_{i:Y_i \prec k}^{n} 0.75 |Y_i - k| \right\}$$

It is obvious that we assign more weight to the values below the θth quantile (positive errors) than the ones above it (negative errors). This is how we move the weights around to find different quantiles. This idea is more or less what the quantile regression is centered around.

The standard quantile regression model that is in use today was introduced by Koenker and Bassett (1978). In a wage equation model, we can define the quantile regression setup as:

$$w_i = x_i^{'} \beta_\theta + u_{\theta i} \quad \text{with} \, Q_\theta(w_i | x_i) = x_i^{'} \beta_\theta \, ,$$

where x_i is a vector of exogenous variables, β_θ is a vector of parameters and $Q_\theta(w_i | x_i)$ denotes the θth conditional quantile of w given x. Any given conditional quantile θ can be derived by solving the following problem:

$$\min_{\beta} \frac{1}{n} \left\{ \sum_{\ln w_i \geq x_i \beta} \theta \, | w_i - x_i^{'} \beta_\theta | + \sum_{\ln w_i \prec x_i \beta} (1-\theta) \, | w_i - x_i^{'} \beta_\theta | \right\}$$, which can be written as

$$\min_{\beta} \frac{1}{n} \sum_i \rho_\theta (w_i - x_i^{'} \beta_\theta) = \min_{\beta} \frac{1}{n} \sum_i \rho_\theta (u_{\theta i}),$$

where $\rho_\theta(\varepsilon)$ is the check function defined as

$$\rho_\theta(\varepsilon) = \begin{cases} \theta\varepsilon & ,\varepsilon \geq 0 \\ (\theta - 1)\varepsilon, & \varepsilon < 0 \end{cases}$$

An alternative expression that can be derived from this setup is:

$$\min_{\beta} \frac{1}{n} \sum_i \left[\theta - \tfrac{1}{2} - \tfrac{1}{2} \operatorname{sgn}\left(w_i - x_i^{'} \beta_\theta \right) \right] \left(w_i - x_i^{'} \beta_\theta \right)$$

In either case, increasing θ from 0 to 1, one can trace the whole distribution of w conditional on x The coefficient estimates of quantile regression denote the effects of covariates on the distribution of the regressor at the corresponding quantile, thus giving the user a means to compare distributions. . Tracing the whole distribution of w this way, we get a chance look beyond the conditional mean and see how the effects of covariates change in the tails and other quantiles of interest.

One can see Koenker and Bassett (1978) and Powell (1984, 1986) for the large sample properties of quantile regression estimators.

Since the objective function is not differentiable, it is not possible to use standard optimization methods. It can be solved as a linear programming model. Buchinsky (1998) shows that Generalized Method of Moments can also be applied for estimation. A number of software packages have quantile regression options. In this study we used Stata which is one of the two standard econometric software packages mentioned by Koenker and Hallock(2001) as having functionality for inference, that is acceptable standard errors.

The method of quantile regression can be seen both as an alternative and a complement to the usual methods of linear regression. As it was forcefully proven by Koenker and Bassett (1978), even though the estimator $\hat{\beta}_\theta$ lacks a bit in efficiency compared to the least squares estimator in case of a Gaussian distribution, it is much more efficient and robust for a large array of non-Gaussian situations. Especially for the cases when the conditional distribution of the dependent variable (conditional on covariates) in question has thick tails, is asymmetric, or unimodal, the meaning attributed to the linear regression estimator can be made much stronger with the help of quantile regression estimators which provide better information about the distribution of the variable in question. The quantile regression estimator is robust to outliers.

Although they are derived by two somewhat analogous methods, $\hat{\beta}_\theta$ should be interpreted in a different way than the linear regression estimator. While the latter simply shows the effect of the covariates on the regressor at the conditional mean, the former is the effect of covariates on the specified quantile of the *distribution* of regressor. This nice feature enables us to draw different regression lines for different quantiles and observe their shape changes as well as their scale and location as one goes along the conditional distribution of the regressor.

The benefits of this method in our analysis are obvious. One does not need to go further than our very first figure to see how closely quantile regression is linked with the analysis of wage inequality. In our discussion so far, we often referred to the notions like 90-10 percentile difference, lower and upper halves of the wage distribution. With the application of quantile regression, we get a change to see the effects of covariates on the creation of these statistics. As we will see later, we also compare the JMP estimates with the quantile regression ones in analyzing the wage inequality picture.

4.2 Application to the US Data

We have given details and facts about the US data concerning wage inequality. In this part we will look at the picture from a different angle. Our aim here is first to analyze the quantile regression estimates then to generate a conditional wage distribution from quantile regression estimates and compare it to the empirical one. One can also create counter-factual distributions, keeping the distribution of observed qualities or estimated skill prices between two years, to see the effect of the other (Machado and Mata (2001)).

We ran regressions for 10^{th}, 25^{th}, 50^{th} 75^{th} and 90^{th} quantiles. We used the same covariates as in JMP analysis in our quantile regressions.We evaluate marginal effects of having high school, some college, and college degrees in Table 4.1. The dropped category of education is less than high school. All quantile regressions were run on 3-year pooled data centered on the indicated year. State standard errors are given in parentheses. We also report the OLS estimates for corresponding years to be able to compare the results.

Just looking at a few coefficients, our remarks about the uses of quantile regression in the previous section become clearer. Although the OLS estimates give a good idea about what the data “on average” shows, there are some other stories not told by them. For example, while the OLS suggests that returns to having a high school degree have kept increasing from 1975 until 2005, the 10^{th} quantile workers did not see much of this increase, and the high school diploma actually meant less in 2005 than what it meant in 1980 (a drop from 33.5 to 27.9). Since this trend is offset by the changes in other quantiles, we do not see it in the OLS estimates. It looks like we can split the high school table into two. For the median and higher quantiles, the returns to having a high school degree increase as years pass. On the other hand, as we just mentioned, it goes down for the 10th percentile since 1985. It also peaked for the 25th percentile in the same year, but it remained more or less the same level after.

Table 4.1 Marginal Effects-Education

Year	High School						Some College						College					
	OLS	10	25	50	75	90	OLS	10	25	50	75	90	OLS	10	25	50	75	90
1968	21.9	24.6	22.6	19.9	18.9	20.8	34.4	33.1	33.0	31.6	33.5	37.8	62.6	58.0	60.0	60.5	63.6	70.2
	(0.4)	*(0.8)*	*(0.5)*	*(0.4)*	*(0.5)*	*(0.6)*	*(0.5)*	*(1.0)*	*(0.7)*	*(0.5)*	*(0.6)*	*(0.8)*	*(0.5)*	*(1.1)*	*(0.7)*	*(0.5)*	*(0.6)*	*(0.8)*
1975	22.6	25.6	23.3	21.0	19.1	19.2	34.3	36.4	33.8	32.1	32.0	32.4	62.5	60.1	60.9	60.1	60.7	64.4
	(0.4)	*(1.0)*	*(0.7)*	*(0.6)*	*(0.7)*	*(0.9)*	*(0.5)*	*(1.3)*	*(0.9)*	*(0.7)*	*(0.8)*	*(1.1)*	*(0.5)*	*(1.3)*	*(0.8)*	*(0.7)*	*(0.8)*	*(1.0)*
1980	25.6	30.3	29.3	25.5	22.1	20.7	37.6	42.5	41.2	36.7	34.1	32.6	63.9	64.0	65.4	62.5	61.8	65.3
	(0.4)	*(0.9)*	*(0.6)*	*(0.5)*	*(0.6)*	*(0.7)*	*(0.5)*	*(1.0)*	*(0.7)*	*(0.6)*	*(0.7)*	*(0.8)*	*(0.5)*	*(1.0)*	*(0.7)*	*(0.6)*	*(0.7)*	*(0.8)*
1985	29.8	33.5	34.7	30.7	26.1	22.7	45.7	48.7	50.5	46.2	41.7	39.1	77.9	77.7	80.9	77.3	74.5	76.4
	(0.5)	*(1.1)*	*(0.7)*	*(0.7)*	*(0.7)*	*(1.0)*	*(0.6)*	*(1.3)*	*(0.8)*	*(0.7)*	*(0.8)*	*(1.1)*	*(0.6)*	*(1.3)*	*(0.8)*	*(0.7)*	*(0.8)*	*(1.1)*
1990	30.0	32.0	31.9	31.5	29.1	25.4	48.1	50.0	50.3	49.8	46.6	42.9	82.5	79.9	83.3	84.1	82.3	83.4
	(0.6)	*(1.2)*	*(0.9)*	*(0.7)*	*(0.7)*	*(1.1)*	*(0.6)*	*(1.4)*	*(0.9)*	*(0.7)*	*(0.8)*	*(1.2)*	*(0.6)*	*(1.4)*	*(0.9)*	*(0.7)*	*(0.8)*	*(1.1)*
1995	31.3	30.3	33.3	33.9	31.5	28.5	48.4	47.1	50.5	50.6	47.8	45.9	88.2	79.4	86.6	89.8	89.2	94.6
	(0.7)	*(1.4)*	*(0.9)*	*(0.8)*	*(0.8)*	*(1.3)*	*(0.7)*	*(1.4)*	*(1.0)*	*(0.9)*	*(0.9)*	*(1.3)*	*(0.7)*	*(1.5)*	*(1.0)*	*(0.9)*	*(0.9)*	*(1.3)*
2000	32.2	30.4	33.9	34.7	32.6	29.4	51.3	48.7	51.8	53.2	51.7	50.1	93.3	82.2	88.5	93.4	96.0	101.5
	(0.7)	*(1.2)*	*(0.9)*	*(0.8)*	*(0.9)*	*(1.3)*	*(0.7)*	*(1.2)*	*(0.9)*	*(0.8)*	*(0.9)*	*(1.4)*	*(0.7)*	*(1.3)*	*(1.0)*	*(0.8)*	*(0.9)*	*(1.4)*
2005	32.0	27.9	32.6	34.0	32.2	31.4	52.4	47.3	53.2	54.8	53.3	51.2	96.7	80.9	90.3	97.2	101.4	105.4
	(0.7)	*(1.5)*	*(1.0)*	*(0.9)*	*(1.0)*	*(1.6)*	*(0.7)*	*(1.5)*	*(1.1)*	*(0.9)*	*(1.1)*	*(1.7)*	*(0.7)*	*(1.6)*	*(1.1)*	*(0.9)*	*(1.1)*	*(1.6)*

All coefficients are multiplied by 100.

The “some college” table shows very similar trends to the high school one except the fact that this time the 10th percentile peaks at 1990, not 1985. For the people with college degrees the life seems to have been much better. Even the 10th percentile see some improvement over the years.

Buchinsky (1994) showed that one can have an idea about within inequality by looking at the differences between relevant conditional quantiles of the same group. Following this logic, Machado and Mata (2001) suggested that one can see an indication of the effect of each group on the overall within inequality by observing the differences of marginal effects of the same group from relevant quantiles. We see the “impact upon dispersion” values which simply show that in Table 4.2. One can look at these numbers as an indication of the contribution of a group to the overall wage inequality due to the inequality it generates within itself, the within inequality. We have 90-10, 90-50 and 50-10 differences. We follow the changes in these values over time to follow within inequality.

Table 4.2 Impact Upon Dispersion-Education

	Impact Upon Dispersion								
	High School			**Some College**			**College**		
	90-10	**90-50**	**50-10**	**90-10**	**90-50**	**50-10**	**90-10**	**90-50**	**50-10**
1968	-3.8	0.9	-4.7	4.7	6.2	-1.5	12.2	9.7	2.5
1975	-6.4	-1.9	-4.6	-4.0	0.3	-4.3	4.4	4.3	0.0
1980	-9.6	-4.8	-4.8	-9.9	-4.0	-5.8	1.3	2.8	-1.5
1985	-10.8	-8.0	-2.8	-9.6	-7.1	-2.6	-1.3	-0.9	-0.4
1990	-6.6	-6.1	-0.5	-7.1	-6.9	-0.2	3.5	-0.7	4.2
1995	-1.8	-5.4	3.6	-1.2	-4.7	3.5	15.2	4.8	10.4
2000	-1.0	-5.2	4.2	1.4	-3.1	4.5	19.4	8.1	11.3
2005	3.4	-2.6	6.1	3.9	-3.6	7.5	24.5	8.2	16.2

All values are differences of indicated quantiles from Table 4.2

We notice that being a high school graduate is a better thing in the 10th quantile than what it is in the 90th quantile, as it can be expected. The only exception to this is in 2005. In a way we can say that the high school group might have seen a decrease in within inequality over the years, and this trend weakens especially in the nineties. The reason for such a picture is probably the ground lost by the high school graduates to the college ones in the upper quantiles. It becomes more obvious that this unusual picture has much more to do with the 90th percentile than the 10th percentile when we look at the 90-50 and 50-10 differences and table 4.1. High school graduates at the 90th percentiles are relatively worse off compared to the other quantiles. On the other hand, there is a trend of increasing within inequality in the lower half of the high school distribution especially after 1990. The some college category follows more or less the same trends as the high school one, once again. However one has to recognise another trend that might be a strong factor in shaping the values of Tables 4.1 and 4.2. As we observed in Figure 2.1, the share of high school graduates among people with no college has been increasing throughout our data set. On the other hand , the ratio of people with some sort of college education to those with high school diploma has been increasing. In other words, the group with less than high school education is disappearing. These lead one to consider the possibility that the higher returns to high school diploma in the lower quantiles than in the upper quantiles might be related to the fact that high school graduates have become a majority in the lower quantile while they have been losing their share in the upper ones.

Apparently the college graduates group is the one that is contributing most to the overall inequality via within inequality. Also, within inequality keeps growing in this group, both in the upper and lower half of the distribution. We also see that the within inequality of the college group is the lowest in the eighties when wage inequality was increasing rapidly. This might have something to do with the high demand for and the stable supply of college graduates during the eighties. The within inequality of college group has been increasing recently.

Table 4.3 College Premium

	College-High School Difference					
	OLS	**10**	**25**	**50**	**75**	**90**
1968	40.7	33.3	37.4	40.6	44.6	49.4
1975	39.9	34.5	37.6	39.1	41.7	45.3
1980	38.3	33.6	36.1	37.0	39.7	44.5
1985	48.1	44.2	46.2	46.6	48.4	53.8
1990	52.5	47.9	51.5	52.6	53.2	58.0
1995	56.9	49.1	53.3	55.9	57.7	66.1
2000	61.1	51.7	54.7	58.8	63.5	72.1
2005	64.6	53.0	57.7	63.2	69.2	74.0

The difference between the coefficients of college and high school graduate dummies

One interesting point here is the so-called college premium, the difference between the returns to college grade and those to high-school diploma. Both OLS and quantile regression results suggest that this value kept increasing steadily from 1980 on. However we observe that the changes become smaller recently. Between 2000 and 2005, the change in all quantiles is modest. This remarks the contribution of the change in inequality caused by education to the slowing down of wage inequality in the last few years.

Table 4.4 reports the marginal effects of experience at 5, 15 and 25 years. Since experience enters the regressions as a quadratic, the marginal effect has been calculated by evaluating the first derivative of it given for 3 values.

Unlike the education case, the quantile regression estimates for experience data generally follow the direction of OLS estimates, to different degrees depending on the quantile. We can say this for all quantiles and for all three experience categories. However, we have a different story here. Figure 4.2 can give us a better feeling of this. It is obvious that having 25 years of work experience was best rewarded during the 80s. We cannot say the same thing for 5 years of experience. Of course one has to keep in mind that the two figures have different scales.

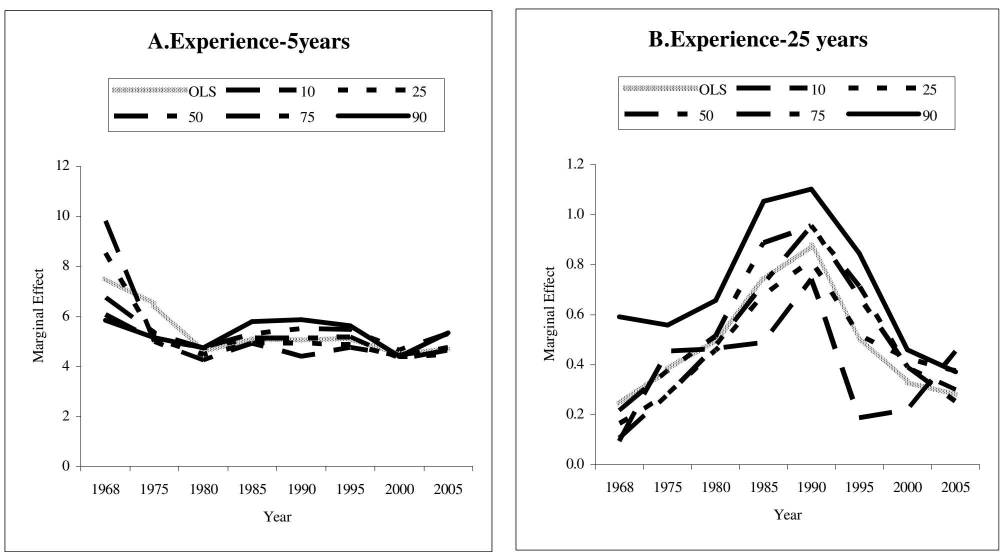
Table 4.4 Marginal Effects-Experience

Marginal Effects of Experience

	5 Years						**15 Years**						**25 Years**					
	OLS	**10**	**25**	**50**	**75**	**90**	**OLS**	**10**	**25**	**50**	**75**	**90**	**OLS**	**10**	**25**	**50**	**75**	**90**
1968	7.53	9.81	8.54	6.76	6.07	5.85	0.85	0.34	0.38	0.84	1.20	1.46	0.24	0.09	0.16	0.10	0.22	0.59
	(0.10)	*(0.22)*	*(0.15)*	*(0.11)*	*(0.12)*	*(0.16)*	*(0.05)*	*(0.09)*	*(0.07)*	*(0.05)*	*(0.06)*	*(0.07)*	*(0.05)*	*(0.10)*	*(0.07)*	*(0.05)*	*(0.06)*	*(0.08)*
1975	6.49	4.97	5.11	5.33	5.07	5.14	1.34	0.99	1.31	1.44	1.58	1.69	0.38	0.45	0.28	0.28	0.37	0.56
	(0.09)	*(0.24)*	*(0.16)*	*(0.13)*	*(0.15)*	*(0.19)*	*(0.05)*	*(0.11)*	*(0.08)*	*(0.06)*	*(0.07)*	*(0.09)*	*(0.05)*	*(0.12)*	*(0.08)*	*(0.07)*	*(0.08)*	*(0.10)*
1980	4.59	4.24	4.47	4.73	4.76	4.74	1.62	1.22	1.47	1.63	1.72	1.84	0.50	0.46	0.46	0.48	0.51	0.66
	(0.09)	*(0.20)*	*(0.13)*	*(0.11)*	*(0.12)*	*(0.15)*	*(0.04)*	*(0.09)*	*(0.06)*	*(0.05)*	*(0.06)*	*(0.07)*	*(0.05)*	*(0.10)*	*(0.07)*	*(0.06)*	*(0.07)*	*(0.08)*
1985	5.09	4.93	4.96	5.15	5.31	5.78	1.80	1.53	1.79	1.83	1.90	1.83	0.74	0.49	0.68	0.72	0.89	1.05
	(0.11)	*(0.24)*	*(0.15)*	*(0.14)*	*(0.16)*	*(0.20)*	*(0.05)*	*(0.11)*	*(0.07)*	*(0.06)*	*(0.07)*	*(0.09)*	*(0.06)*	*(0.12)*	*(0.08)*	*(0.07)*	*(0.08)*	*(0.10)*
1990	5.05	4.40	4.94	5.15	5.50	5.86	1.57	1.33	1.46	1.56	1.78	1.80	0.88	0.74	0.81	0.95	0.96	1.10
	(0.11)	*(0.27)*	*(0.18)*	*(0.14)*	*(0.15)*	*(0.23)*	*(0.05)*	*(0.11)*	*(0.08)*	*(0.06)*	*(0.06)*	*(0.09)*	*(0.06)*	*(0.13)*	*(0.09)*	*(0.07)*	*(0.07)*	*(0.11)*
1995	5.13	4.76	4.89	5.17	5.49	5.63	1.67	1.30	1.47	1.71	1.89	1.95	0.51	0.19	0.52	0.67	0.71	0.84
	(0.14)	*(0.30)*	*(0.20)*	*(0.18)*	*(0.18)*	*(0.28)*	*(0.06)*	*(0.12)*	*(0.08)*	*(0.07)*	*(0.07)*	*(0.11)*	*(0.07)*	*(0.14)*	*(0.09)*	*(0.08)*	*(0.08)*	*(0.12)*
2000	4.46	4.45	4.41	4.31	4.64	4.41	1.39	0.87	1.07	1.41	1.64	1.88	0.33	0.22	0.42	0.39	0.39	0.46
	(0.13)	*(0.26)*	*(0.20)*	*(0.17)*	*(0.19)*	*(0.28)*	*(0.06)*	*(0.11)*	*(0.08)*	*(0.07)*	*(0.08)*	*(0.12)*	*(0.06)*	*(0.11)*	*(0.09)*	*(0.08)*	*(0.08)*	*(0.12)*
2005	4.72	4.75	4.45	4.64	5.32	5.33	1.56	0.81	1.19	1.56	1.72	1.87	0.28	0.45	0.38	0.30	0.25	0.37
	(0.15)	*(0.34)*	*(0.23)*	*(0.20)*	*(0.22)*	*(0.36)*	*(0.06)*	*(0.14)*	*(0.10)*	*(0.08)*	*(0.09)*	*(0.15)*	*(0.06)*	*(0.14)*	*(0.10)*	*(0.08)*	*(0.09)*	*(0.14)*

Since experience enters the regressions as a quadratic, the marginal effect has been calculated by evaluating the first derivative of it given for 3 values. All coefficients are multiplied by 100. Standard errors are in parentheses.

It is also noteworthy from Figure 4.1 and Table 4.4 that the marginal effect of years of experience on wage distribution increases as one goes higher in the quantiles. So, one can assume higher returns to experience in the higher quantiles than the lower ones. This is in accordance with the findings concerning increased wage inequality starting from 70s and going until mid-nineties. At each experience level, additional years of experience gain better returns for workers in higher quantiles than they do for the ones in lower quantiles. At 2000 and 2005, the 10th quantile gains on 25th quantile and the median, thus confirming our earlier findings about slowing or decreasing lower end inequality.



Since experience enters the regressions as a quadratic, the marginal effect has been calculated by evaluating the first derivative of it given for 5 and 25 years.

Figure 4.1 Marginal Effect of Experience

Table 4.5 reports the impact upon dispersion for three fictional years of experience. We notice that the low experience group contributed overall inequality the most during the stretch from 1980 to 1995. Except for 1985 when the within inequality was caused predominantly by upper-level inequality, both tails of the distribution contributed to this effect in a balanced way. However, we see that the upper half inequality gains much more importance in 2000 and 2005 against the lower half which actually stretched, not dispersed. On the other hand, we see that lower half inequality is more important than the upper half inequality from 1990 to 2000. Before 1990, it is the upper half inequality that dominates the picture of within inequality in the 25-year experience group.

Table 4.5 Impact upon Dispersion-Experience

	Impact Upon Dispersion								
	5 Years			**15 Years**			**25 Years**		
	90-10	**90-50**	**50-10**	**90-10**	**90-50**	**50-10**	**90-10**	**90-50**	**50-10**
1968	-3.96	-0.92	-3.05	1.13	0.62	0.50	0.50	0.49	0.01
1975	0.17	-0.20	0.37	0.70	0.25	0.45	0.10	0.28	-0.18
1980	0.50	0.01	0.49	0.61	0.20	0.41	0.19	0.18	0.01
1985	0.85	0.64	0.22	0.29	0.00	0.30	0.57	0.33	0.24
1990	1.46	0.71	0.76	0.47	0.24	0.23	0.36	0.15	0.21
1995	0.87	0.45	0.42	0.65	0.24	0.41	0.66	0.18	0.48
2000	-0.04	0.10	-0.14	1.01	0.47	0.54	0.24	0.07	0.17
2005	0.58	0.70	-0.11	1.06	0.32	0.74	-0.08	0.07	-0.16

The effect of being in an industry on wage distributions is given in Table 4.6 for certain quantiles. All values in this figure are given in the form of deviation from the mean. We note that trade and finance and services always give less than average returns, while mining, manufacturing and transportation, communication and public utilities always give higher returns than average (with the exception of 2004). We can also see the differences of return within the sectors here, depending on the quantile. Looking at this picture, one could probably say that industry differences contribute to wage inequality as well.

Table 4.6 Industries

1968	Mining	Construction	Manuf.	Tr.Com&P.U.	Trade	Fin.&Serv.	Gov.
OLS	8.8	3.4	2.2	3.8	-7.8	-10.2	-0.3
10	9.8	-2.7	4.9	6.1	-10.7	-14.8	7.4
25	9.0	1.6	3.3	5.6	-10.4	-12.4	3.3
50	7.9	4.8	2.6	4.4	-8.7	-10.7	-0.3
75	8.0	7.4	0.5	2.7	-6.1	-8.4	-4.1
90	5.5	8.4	-0.9	0.9	-2.2	-4.1	-7.7
1975	**Mining**	**Construction**	**Manuf.**	**Tr.Com&P.U.**	**Trade**	**Fin.&Serv.**	**Gov.**
OLS	17.2	3.0	0.4	4.8	-10.5	-14.5	-0.4
10	19.1	-2.9	4.0	4.8	-14.5	-16.4	5.9
25	18.6	-0.6	1.8	6.9	-12.7	-16.1	2.1
50	17.2	2.7	-0.2	6.6	-10.9	-15.5	0.2
75	15.2	7.0	-1.6	4.9	-8.8	-13.6	-3.1
90	16.1	9.1	-3.3	2.6	-6.0	-11.2	-7.4
1980	**Mining**	**Construction**	**Manuf.**	**Tr.Com&P.U.**	**Trade**	**Fin.&Serv.**	**Gov.**
OLS	26.6	-1.8	1.7	7.2	-12.9	-15.7	-5.2
10	24.8	-7.1	5.6	8.5	-17.7	-17.6	3.5
25	28.0	-3.8	3.5	8.8	-16.2	-18.5	-1.8
50	26.2	-1.2	2.4	9.0	-13.5	-17.8	-5.1
75	24.9	1.9	0.1	6.4	-9.8	-15.2	-8.3
90	25.3	1.5	-2.1	4.8	-7.1	-9.3	-13.1
1985	**Mining**	**Construction**	**Manuf.**	**Tr.Com&P.U.**	**Trade**	**Fin.&Serv.**	**Gov.**
OLS	24.3	-2.6	3.3	7.4	-13.3	-13.4	-5.6
10	26.2	-6.6	4.7	10.9	-19.9	-17.5	2.2
25	25.7	-5.7	5.0	10.0	-16.7	-15.7	-2.6
50	24.8	-3.1	4.1	8.5	-13.7	-15.3	-5.4
75	22.0	-0.1	2.4	5.8	-10.0	-11.5	-8.7
90	22.5	1.8	-0.1	2.7	-7.5	-6.7	-12.7
1990	**Mining**	**Construction**	**Manuf.**	**Tr.Com&P.U.**	**Trade**	**Fin.&Serv.**	**Gov.**
OLS	23.4	-1.5	3.1	6.4	-15.0	-10.5	-5.9
10	26.7	-5.7	5.4	9.8	-21.7	-15.7	1.3
25	27.9	-3.5	4.0	8.6	-19.3	-14.5	-3.1
50	23.4	-0.2	3.7	7.3	-15.5	-12.4	-6.3
75	19.9	1.3	2.3	4.9	-10.9	-8.3	-9.2
90	19.9	2.3	-0.1	1.9	-8.7	-2.4	-12.9
1995	**Mining**	**Construction**	**Manuf.**	**Tr.Com&P.U.**	**Trade**	**Fin.&Serv.**	**Gov.**
OLS	24.1	-3.0	2.9	6.2	-16.0	-10.5	-3.6
10	29.8	-7.0	3.0	7.5	-22.0	-15.7	4.4
25	28.3	-3.7	2.9	7.3	-20.5	-14.3	0.1
50	23.5	-1.5	2.7	8.0	-16.5	-13.3	-2.8
75	18.9	-1.4	2.7	6.4	-12.0	-8.7	-5.9
90	15.8	-2.1	2.5	4.4	-8.7	-2.6	-9.4
2000	**Mining**	**Construction**	**Manuf.**	**Tr.Com&P.U.**	**Trade**	**Fin.&Serv.**	**Gov.**
OLS	17.3	-2.2	3.0	5.5	-14.2	-5.4	-3.9
10	18.2	-4.3	5.3	7.4	-20.6	-10.3	4.4
25	18.5	-2.9	4.5	7.7	-18.4	-9.8	0.4
50	20.0	-1.8	2.5	6.7	-15.1	-8.7	-3.7
75	17.3	-0.6	1.2	4.6	-11.3	-4.8	-6.4
90	10.4	1.2	-0.1	3.7	-7.7	3.4	-10.8
2005	**Mining**	**Construction**	**Manuf.**	**Tr.Com&P.U.**	**Trade**	**Fin.&Serv.**	**Gov.**
OLS	25.3	-4.9	-0.9	2.5	-9.9	-11.6	-0.5
10	22.5	-6.3	0.4	3.6	-13.4	-17.0	10.2
25	26.7	-6.2	0.9	3.4	-12.1	-15.1	2.5
50	24.2	-4.2	0.5	4.8	-11.1	-13.9	-0.4
75	25.1	-3.4	-2.3	2.2	-9.1	-11.3	-1.2
90	23.0	-2.9	-4.3	2.3	-6.5	-5.4	-6.3

Industry effects are calculated as the difference of regression coefficients from the average industry effect for that quantile for that year.

4.3 Conditional Wage Distribution

Analysis of quantile regression estimates naturally leads us to conditional wage distributions. We see such distributions for our data in Table 4.7. Each value in this table has been created by predicting the wage for the corresponding quantile by means of quantile regression estimates for that quantile and mean values of the covariates for the whole sample.

Table 4.7 Conditional Wage Distribution

Percentil	1968	1975	1980	1985	1990	1995	2000	2005
10	5.410	5.517	5.460	5.414	5.379	5.346	5.432	5.416
25	5.660	5.756	5.724	5.701	5.670	5.653	5.740	5.729
50	5.903	5.993	5.986	5.985	5.964	5.960	6.057	6.056
75	6.129	6.218	6.230	6.254	6.241	6.251	6.367	6.369
90	6.346	6.421	6.461	6.498	6.494	6.536	6.676	6.678

All values are predicted wages for each quantile using quantile regression estimates and average values of covariates.

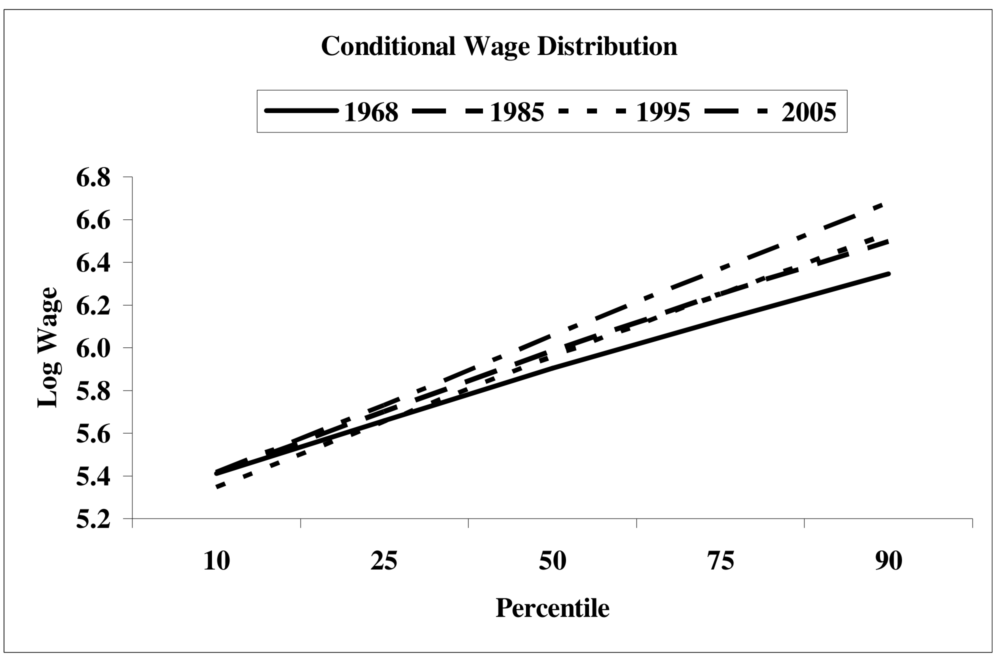


Figure 4.2 Conditional Wage Distributions

Such an approach indirectly means that we create the conditional distribution for an "average" individual, thus it ignores the dispersion created by the differences between individual workers. Figure 4.2 shows how the conditional wage distribution changed over time. The changing slope that marks increasing wage inequality is unmistakable. We notice, however, that the slope change is smaller between 1995 and 2005, confirming the slowdown in wage inequality increase.

Table 4.8 Comparison of Conditional and Empirical Wage Distributions

				A.Dispersion from Conditional Distribution				
Perc.Diff.	**1968**	**1975**	**1980**	**1985**	**1990**	**1995**	**2000**	**2005**
90-10	0.94	0.90	1.00	1.08	1.11	1.19	1.24	1.26
90-50	0.44	0.43	0.47	0.51	0.53	0.58	0.62	0.62
50-10	0.49	0.48	0.53	0.57	0.59	0.61	0.62	0.64
				B.Dispersion from the Data				
Perc.Diff.	**1968**	**1975**	**1980**	**1985**	**1990**	**1995**	**2000**	**2005**
90-10	1.18	1.24	1.28	1.45	1.49	1.57	1.62	1.66
90-50	0.54	0.55	0.56	0.62	0.68	0.75	0.80	0.85
50-10	0.64	0.68	0.72	0.83	0.81	0.82	0.82	0.81
				Changes in Dispersion(Conditional)				
	75-68	**80-75**	**85-80**	**90-85**	**95-90**	**00-95**	**05-00**	
90-10	-0.03	0.10	0.08	0.03	0.07	0.05	0.02	
90-50	-0.01	0.05	0.04	0.02	0.05	0.04	0.00	
50-10	-0.02	0.05	0.04	0.01	0.03	0.01	0.02	
				C.Conditional/Empirical Ratio				
Perc.Diff.	**1968**	**1975**	**1980**	**1985**	**1990**	**1995**	**2000**	**2005**
90-10	0.79	0.73	0.78	0.75	0.75	0.76	0.77	0.76
90-50	0.82	0.78	0.85	0.83	0.78	0.77	0.77	0.73
50-10	0.77	0.69	0.73	0.69	0.73	0.75	0.76	0.79

90-10, 90-50 and 50-10 percentile differences from conditional and empirical log wage distributions. Panel C shows the percentage of empirical distribution measure that is explained by the conditional distribution one.

We see a comparison of conditional and empirical wage distributions in Table 4.8 using the dispersion measures we have become quite familiar with now. It is quite obvious from panels A and B that the dispersion from the conditional distribution is smaller than the one directly from the data. Of course the main reason for this difference is the usage of average values of covariates in the calculation of conditional distribution. Panel C shows the ratio of the dispersion from conditional distribution to the dispersion from empirical data. More than half of the values are within the 0.72-0.77 range. It looks like we can assume that our conditional distribution explains about three fourths of the wage inequality. The rest are attributable to personal characteristics.

Table 4.6 shows us that the conditional wage distribution keeps shifting over time. From Table 4.7 we also see the overall wage inequality measure of 90th-10th percentile difference confirms that the wage inequality begins to increase from the second half of the seventies, continues to rise in the 80s and 90s, and still keeps increasing at a low rate. However, the story between 1975 and 1985 is told a bit differently here. 90-10 measure change between 1975 and 1980 for conditional distribution is more than twice as much as the empirical one. On the other hand, the same statistic for conditional distribution increase less than half as much as the empirical one from 1980 to 1985. For the following years, conditional and empirical distributions more or less match each other in 90-10 measure. If we remember how we set up the conditional distribution, this difference might be attributable to the changes in observable personal characteristics. This kind of approach leads us to think that the effect of changes in personal characteristics were more important for the wage inequality increase between 1980 and 1985 than between 1975 and 1980.

The balance between the lower and upper half wage inequality is more or less maintained until 1990. However, starting from 1995 data, we notice that the inequality in the upper half of the distribution begins to increase faster than the dispersion in the lower half of the distribution. This trend is captured in both conditional and empirical distributions. This was also mentioned in our analysis with JMP results.

The conditional distribution that we have just seen can be used to create another very nice tool to see the different dimensions of wage inequality. The counterfactual distribution approach is certainly not limited to quantile regression analysis. It has been used in some very influential studies of wage inequality (Dinardo, Fortin and Lemieux (1997). However, this is a tool that is best suited for use with quantile regression estimates since they represent the entire distribution, without any assumptions about the shape and properties of it.

In Table 4.9 we see the counterfactual distributions created using 1968 personal characteristics of the workers and quantile regression coefficients for each year. This way we take what we did in creating conditional distributions one step further. Now we compare the conditional distributions for different years assuming that personal characteristic and skill composition is what it was in 1968. In other words we isolate the changes in prices of skills and personal characteristics (represented by the regression estimates) from the changes in demographic characteristics.

Table 4.9 Counterfactual Conditional Distributions-1968 Covariates

	Counterfactual Wage Distribution(1968 averages)							
Percentile	**1968**	**1975**	**1980**	**1985**	**1990**	**1995**	**2000**	**2005**
10	5.410	5.531	5.485	5.415	5.368	5.309	5.380	5.377
25	5.660	5.767	5.742	5.689	5.649	5.603	5.669	5.669
50	5.903	6.001	6.001	5.969	5.931	5.895	5.966	5.972
75	6.129	6.222	6.239	6.234	6.198	6.177	6.259	6.260
90	6.346	6.421	6.463	6.469	6.443	6.453	6.546	6.549
	Dispersion							
Percentile	**1968**	**1975**	**1980**	**1985**	**1990**	**1995**	**2000**	**2005**
90-10	0.94	0.89	0.98	1.05	1.08	1.14	1.17	1.17
90-50	0.44	0.42	0.46	0.50	0.51	0.56	0.58	0.58
50-10	0.49	0.47	0.52	0.55	0.56	0.59	0.59	0.59
	Changes in Dispersion							
Percentile	**75-68**	**80-75**	**85-80**	**90-85**	**95-90**	**00-95**	**05-00**	
90-10	-0.05	0.09	0.08	0.02	0.07	0.02	0.01	
90-50	-0.02	0.04	0.04	0.01	0.05	0.02	0.00	
50-10	-0.02	0.05	0.04	0.01	0.02	0.00	0.01	

Conditional wage distributions are created from 1968 averages of covariates and corresponding year coefficient estimates.

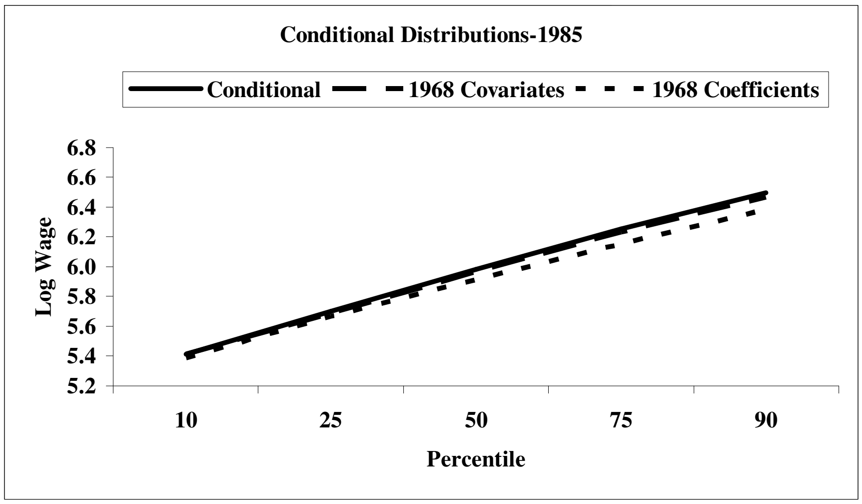
We notice that while the change in dispersion generated by the counterfactual distribution (1968 averages) explains the change in dispersion in the original conditional wage distribution quite reliably until 1985, its explanatory power decreases afterwards. This shows that using the 1968 characteristics does not really change the conditional wage distributions very much until 1985. Comparing the numbers here, one can come to the conclusion that observable characteristics contribute to the change in wage distribution more after 1985.

Table 4.10 Counterfactual Conditional Distributions-1968 Coefficients

	Conditional Wage Distribution(1968 coefficients)							
Percentile	**1968**	**1975**	**1980**	**1985**	**1990**	**1995**	**2000**	**2005**
10	5.410	5.386	5.375	5.392	5.399	5.407	5.405	5.391
25	5.660	5.651	5.648	5.667	5.675	5.686	5.688	5.678
50	5.903	5.898	5.895	5.914	5.922	5.936	5.942	5.935
75	6.129	6.131	6.131	6.154	6.164	6.184	6.194	6.191
90	6.346	6.355	6.356	6.383	6.396	6.422	6.438	6.439
	Dispersion							
	1968	**1975**	**1980**	**1985**	**1990**	**1995**	**2000**	**2005**
90-10	0.936	0.969	0.981	0.991	0.997	1.015	1.033	1.049
90-50	0.443	0.457	0.462	0.469	0.474	0.485	0.496	0.504
50-10	0.493	0.512	0.519	0.522	0.523	0.529	0.537	0.544
	Changes in Dispersion							
	75-68	**80-75**	**85-80**	**90-85**	**95-90**	**00-95**	**05-00**	
90-10	0.033	0.012	0.010	0.005	0.018	0.018	0.016	
90-50	0.014	0.005	0.007	0.005	0.012	0.010	0.009	
50-10	0.019	0.007	0.003	0.001	0.006	0.008	0.007	

Finally, the counterfactual distribution that uses 1968 quantile regression coefficients and corresponding year's average values of covariates is given in Table 4.10. To see how the changes in skill prices (quantile regression coefficients) affect the wage inequality picture, this time we hold the skill prices constant, let the covariates change and create counterfactual distributions for each year.

This time we notice that the last counterfactual distribution does a poor job in matching the changes in dispersion from the conditional log wage distribution. However we see that the difference is most serious during the period 1975 and 1990. After that the difference starts falling.



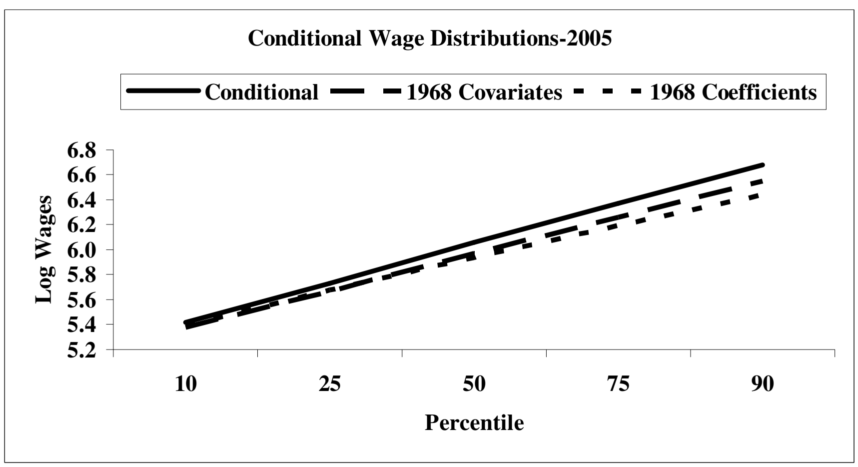


Figure 4.3 Conditional vs. Counterfactual Wage Distributions

One can see a possible reason for these trends in Figure 4.3, where we have all three distributions for years 1985 and 2005. The 1985 figure shows that the conditional distribution and the counterfactual distribution using 1968 covariates are more or less the same. On the other hand, the distribution using 1968 coefficients stands alone, indicating that between 1968 and 1985, changes in skill prices were more serious than the change in skill composition. However, when we come to 2005 we see that changing skill compositions now contribute to the changes in wage distribution as well, even though their effect is still less than the skill prices.

4.4 JMP Decomposition and Quantile Regression

The JMP decomposition is a very powerful and practical method which enables us to identify the magnitude and shape of the effect of unobserved skills on the increase or decrease of wage inequality. This is why it has been in use for the last 15 years despite the facts that many other methods of decomposition have been proposed during the same time period. However, as our analysis with the quantile regression estimates proved, one has to be careful about the shape of the distribution of wage when commenting on JMP results. The method itself depends on OLS estimates which represent the conditional mean. When the distribution of wages is normal, one can comfortably make use of JMP results, since most of the data is centered on the mean and the tails are not strong. However, if the shape is different there might be problems. Since JMP decomposition is done for the changes in the difference of quantiles (90-10, 90-50, 50-10 in our case), such a distribution could limit the meaning of decomposition. Table 7.1 lets us make a comparison of JMP and quantile regression. In Panel A, we see the changes in 90-10 percentile difference that was drawn from conditional distributions in five-year periods. In Panel B, we have the total change value calculated by JMP decomposition for the same time periods.

Generally the 90s, especially 1995-2000 period is when the two tables match the most in both magnitude and the weight of changes. When we check the marginal effects of education and experience from Tables 4.1 and 4.2 for these periods, we see that the change in all percentiles are in the direction of the change in OLS estimates with an exception or two, and of similar magnitude. However, for the 1985-1990 difference, we typically see contradicting patterns in the upper and lower halves. Marginal effects of having a high school degree , for example, follow OLS in the upper half, but goes in the opposite direction in the lower.

Table 4.11 Comparison of JMP and Quantile Regression

A.Quantile Regression (90-10)							
	75-68	80-75	85-80	90-85	95-90	00-95	05-00
90-10	-0.031	0.096	0.083	0.031	0.075	0.054	0.018
90-50	-0.014	0.045	0.039	0.016	0.047	0.044	0.002
50-10	-0.017	0.050	0.044	0.015	0.028	0.011	0.016
B.JMP(90-10)							
	75-68	80-75	85-80	90-85	95-90	00-95	05-00
90th-10th	0.029	0.042	0.172	0.022	0.094	0.057	0.042
90th-50th	0.014	0.006	0.062	0.060	0.067	0.047	0.042
50th-10th	0.015	0.036	0.110	-0.038	0.027	0.010	0.001

Changes in the values of 90^{th}-10^{th} percentiles o

As we mentioned before, Koenker & Bassett (1978) proved that quantile regression estimates perform better than the OLS ones for most non-normal distributions. It is certainly a very handy tool to see how the effects of different covariates change in different parts of the distribution. We have seen time and again in our tables that the OLS estimates sometimes miss important shifts in the tails. For this reason, we feel it is always a good idea to support any analysis of wage inequality with quantile regression estimates. It is possible to get some within inequality measure for a narrowly defined group from quantile regression (our impact upon dispersion tables, for example). However, they lack the practicality and functionality of the JMP method when looking at the shape and magnitude of within inequality in the big picture.

4.5 Concluding Remarks

Quantile regressions on our Mincerian wage equation for 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} quantiles reveal a more detailed picture of the inequality in the US. Since the quantile regressions give a better picture of how the effects of covariates change on different parts of the distribution, they let us see the effects of experience and education on wages in a more informative way.

It appears that college degree is the main educational component that gives way to increased wage inequality. The returns to high school seem to be higher in the lower quantiles than upper quantiles, probably due to the fact that there are more high school graduates in the lower quantiles and less in the upper quantiles compared to earlier years.We have also seen that the increase in difference between college and high school marginal effects slowed down after 2000.

Conditional wage distributions that have been created from the estimates of quantile regressions and average characteristics of workers in a given year are useful for comparison of effects of the changes in composition of observable characteristics and the effects of changes in the prices of observable characteristics. One such distribution showed that observable characteristics affected wage distribution more after 1985 than it did before 1985. Also the earlier findings from JMP decomposition about skill prices affecting wage inequality more than the observable skills has been confirmed by our counterfactual distributions.

Quantile regression and the JMP decomposition have their own advantages compared to the other. While the JMP decomposition gives us the advantage of comparing the relative magnitudes of between and within inequalities as well as an average picture of inequality representing the whole distribution, quantile regression enables us to see what happens at the tails and different parts of the wage distribution as well as created conditional distributions to compare different alternatives. The more normal the wage distribution, the closer the estimates of both methods.

However if the wage distribution is not normal, then one has to support JMP decomposition with quantile regression estimates to avoid serious mistakes.

CHAPTER 5

A CROSS-COUNTRY ANALYSIS OF WAGE INEQUALITY USING LUXEMBOURG INCOME STUDY

The experience of European countries with wage inequality has been different from that of the US with different degrees. It is generally agreed that none of the European countries with the exception of the United Kingdom had the sharp increase in wage inequality that appeared in the US during the 80s. Some even recorded decreases (Gottschalk & Smeeding (1997)). In this chapter we examine the data for a number of European countries to see how their pathc compares to that of the US in more recent years.

Sapir (2006) states that European countries can be classified into four social models: Nordic Countries represent the high levels of government intervention and strong labor unions that compress wages. Anglo-Saxon countries are characterized by weak labor unions, high dispersion of wages and relatively higher rate of low-pay employment. Continental countries have non-employment benefits and strong unions. The Mediterranean countries display a highly compressed wage structure due to collective bargaining in the formal sector. Our group of countries in this study includes Sweden from the first, Netherlands from the second, Germany from the third, and Spain from the last group. In addition to these four social models, one could possibly mention the East European countries, from which we have Hungary. The choice of each country among the groups they belong has been made according to the availability of data we need to apply our analysis. We give the details below.

5.1 Data Set

LIS is a research project that started in 1983 under the joint sponsorship of the government of Luxembourg and by the government of Luxembourg and the Centre for Population, Poverty and Policy Studies (CEPS). A non-profit project, it is mainly funded by research and statistics organizations of 30 member countries which are spread over four continents. Although mainly European, the member roster also includes countries like the USA, Canada, Mexico, Taiwan, Australia and Israel. Participation into the project requires submission of microdata from the participating organization's country that is applicable to LIS system, as well as yearly contribution is made by most members, except for those that are not able to supply such funds (non-contributing countries).[3]

The LIS database has been prepared with a certain objective in mind: Creating a data structure that will enable researchers to make international comparisons on household, person(adult) and child level. The data is collected from household surveys of the member countries and "lissified" to fit into a certain standard data structure which includes "standardized" or "harmonized variables". Standardization corresponds to simply recoding the country information into some standard file structure. Many income variables are this way. Calculation of wage and salary income at person level is a good example of this procedure. Although LIS database reports a single wage and salary income value in the domestic currency for each country, these values come from different questions in each country's domestic household or similar survey. A German person's gross wage and salary income value

---

3 I had the chance to participate LIS Summer Workshop 2006. Held in Luxembourg, this is a a yearly gathering of academicians to work on the LIS data and exchange ideas on their own projects. During my time in Luxembourg I had the opportunity to learn about the applications of LIS data as well as how to access it. Except for some extreme cases, the organization does not let the users do hands-on work with microdata, not even during the Workshop. The users send short programs written for a selection of software packages like Stata, SAS and SPSS and send them to a system called LISSY via e-mail, receiving the output via e-mail from this system in turn. It is a well organized and functioning system, however the process might be a bit slow sometimes due to the nature of the remote access relationship.

is calculated by adding up eight different items for 1989 data. On the other hand, an Italian person's wage and salary income is reported in a single variable from the original data set.

Harmonized variables, on the other hand, keep the country-specific data without recoding it. For example, the eduational level variable (PEDUC) exists in every person data set, provided that the member country supplied any information on educational level. However, the structure of this variable is totally country-specific, and it changes by time and by country.

LIS collects its data in "waves" of five-year intervals. Each wave centers around a certain year and the member countries submit their data from a survey done in a year as close to the center year as possible. Some countries report more than one year around (or including) the reference year. So far there have been 5 waves: years centered around 1980, 1985, 1990, 1995 and 2000. The wave VI, which corresponds to the years centered around 2005, is currently under "lissification".

Even though it is a very impressive attempt at creating reliable international microdata, the LIS data sets have their own problems and shortcomings, which have to be taken into account for an analysis based on them to be meaningful:

a) First of all, it is impossible to perfectly fit the data of every country in a pre-set data structure. One has to make decisions to create comparable data sets, and each decision made as such increases the number of footnotes. Inevitably the data coming from some countries is much more detailed and better classified than the others (US is the obvious example) and it is very hard to avoid overstretching of data to create something that compares to it from a country which simply does not report much. In some cases wage and salary income values are not reported, or simply not available, and in some cases the industry data, standardized, is obviously not correct, or does not involve some industries. Although we have been very careful in our choice of the countries to reflect as much reliable and comparable information as possible, one still has to keep in mind that this is second-hand, recoded data.

b) Although there are 5 waves, the availability of data changes from country to country. While some countries have data only for the last two waves, some others have older information than the first wave. The German data goes as far back as 1973, while it is quite hard to find reliable data older than 1990 for ex-communist countries of Eastern Europe.

c) Especially for the earlier waves, there are holes in most data sets, simply because some questions were not asked in the original survey, or it is not possible to generate the needed information from other survey questions. The most noteworthy example of this problem is the variable representing a person's industry. While some countries use standard 2- and 3-digit coding, others do their own classification, grouping industries in different ways, making it impossible to convert it into the standard system. Also, sometimes it is obvious that the industry data is not reliable. For example 1995 data for Spain can link only 1,822 out of 18,643 persons to an industry in a standard 3-digit set.

d) The data structure of one country sometimes changes over time, due to re-design or different survey, making it hard to do within-country comparisons. Among the variables that are of interest to us, this is mostly a problem with education data.

e) Reliability of data in the harmonized variables is a problem that is also hard to deal with. In some cases the numbers reported just do not make much sense.

f) Some of the critical labor market variables are not well documented in some countries' data sets, making it hard to work with them. The most notable examples are data on educational attainment, number of weeks worked and industry.

Although it is important to be careful about the problems that we have listed above, one should not fail to see the enormous amount of data and research possibilities that comes with such a data set. Once we look beyond the footnotes, it is a remarkable opportunity to compare coutries and test theories on a wide range of variables and official data sets from 30 separate countries. LIS has a remarkable amount of information for every country's data sets which warn the user against obvious problems and let one see the details of the variable in question. The web site, www.lisproject.org, gives all the necessary information one needs to know before using the data and gives links to each country's own survey agency. Also, it would be unfair not to note that the usability of data sets improve remarkably over time, the best being waves IV and V.

In our study, we use the person level data files for five countries: Germany[4], Netherlands[5], Sweden[6], Spain[7] and Hungary[8]. This choice is as much by availability of data as by representation.

German data comes from German Social Economic Panel Study (GSOEP). The Dutch data comes from Additional Enquiry on the Use of (Public) Services (AVO-1983), Socio-Economic Panel (SEP-1991, 1994, 1999) and Netherlands European Community Household Panel (NL ECHP) (1994, 1999). As we can see, 1994 and 1999 data is a combination of two surveys. The data on Sweden is drawn from Income Distribution Survey (HINK). Family Expenditure Survey(1990) and Spanish

4 The German data is well documented and very detailed for waves 2 to 5. Wave I data does not give any information on full-time or full-year status, so we did not use it.

5 For 1983(Wave 2) it is not possible to identify full-year workers. Also government employees are not identified for Waves 2 and 3.

6 Industry data is not reliable.

7 Wages are reported as net, not gross. Full-year status cannot be identified for 1990.

8 Net wages are reported. Industry data is either unavailable or unusable for all years we use. No person weights given for 1991. Full-year status cannot be identified for any year.

European Community Household Panel (ES ECHP-1995, 2000) are sources for Spanish data. The Hungarian data is from Household Monitor Survey.

We have limited our sample to male, aged 16-64 years, full-time employed wage and salary workers for whom there is enough data on education to create LIS' standard low, medium and high education dummies about which we give information later. The sample is limited further to include only full-year workers whenever it is available for all years of a country.

Table 5.1. Waves and Sample Size

	Waves				
	I(1980)	II(1985)	III(1990)	IV(1995)	V(2000)
Germany		1984	1989	1994	2000
		(2790)	(2417)	(2811)	(8936)
Netherlands		1983	1991	1994	1999
		(2044)	(1876)	(2266)	(4552)
Sweden			1992	1995	2000
			(4929)	(4840)	(14054)
Spain			1990	1995	2000
			(11214)	(2460)	(4458)
Hungary			1991	1994	1999
			(794)	(651)	(1080)

Reported years are each country's contribution to the related Wave. Numbers in parenthesis are sizes of our wage-earner samples for that country and year.

Table 5.1 shows the years in which data as we want is available for our selected countries and their respective waves. Germany actually has longer historical data than indicated here. The German data goes as far back as 1973. However, we were unable to use them because they failed to supply some of the essential criteria for constructing our sample. The numbers in parentheses are the sizes of wage-earner samples. We notice that the sample size grows for all countries as we go higher on waves, probably an indication of improvement of domestic survey. The considerable drop in the size of Spanish sample from 1990 to 1995 is due to the fact that they are from different surveys, as we explained above. Similarly, the Dutch sample for 1983 is from a different survey from the rest.

In LIS data sets, all monetary variables are reported in domestic currency unit. It is obviously not an ideal situation when you want to make comparisons between countries. A number of issues involving differences in purchasing power and inflation make such comparisons unreliable. Thus it is important to convert the monetary values into a unit of exchange that can be compared. First of all, all wages are converted into the year 2000 value of their local currency, using the GDP deflator. Then these values are converted into US dollars, using the Purchasing Power Parity (PPP) conversion rates.

PPP conversions are useful in the sense that they let us compare income values across the countries in a way that does not overlook the differences between countries in terms of the level of prices. Although it is very convenient to have as a tool, one has to keep in mind that PPPs are actually designed to compare GDPs, not personal income. It is generally believed that in most countries personal income is about 2/3 of the GDP. Thus, although they certainly indicate a good deal about personal income, PPPs also reflect some information that is not related to it, making the comparison a bit biased. Also, PPP is calculated for an aggregate bundle of goods consumed in the country. Thus, it is an average concept. On the other hand, the actual consumption bundles might change depending on where you are at the income distribution.

Our sample includes individuals who earn a gross amount of $110 per week in PPP-converted year 2000 values. The choice of this threshold is because of the need to make it comparable to the American data ($67 per week in year 1982 values). We used the GDP deflators from the IMF website for this conversion. Out of the five countries that we include here, we have gross wage and salary income for three Germany, Netherlands and Sweden, and net wage and salary income for two (Spain and Hungary). Using data from Eurostat on implicit tax rates in countries, we decided to limit the Spanish wages to $78 (29% taxation in 1995) and the Hungarian wages to $63 (42.5% taxation in 1995)) and over instead of $110. However these values have only theoretical meaning, since no one in our wage-earner sample seems to be below $110 in either country.

For each sample statistic and calculation we report here, we used the person weights (except in a case or two where no weights are available). Yearly wages, whether gross (for Germany, Netherlands and Sweden) or net (for Spain and Hungary), were imputed at their reported maximum by 1.33 times the topcoded value. Our unit of analysis is the weekly wage, calculated by dividing yearly wages by 52. The usual way of doing this in the literature is by dividing the yearly wages by the actual weeks worked during the year. However, even though we had information on weeks worked in a year for all years from Germany, we chose not to use it since we do not have this information on regular basis for other countries. Hungary simply does not supply it for any year. Netherlands, Sweden and Spain have this info for some years and they do not have it for others. Using the German data on weeks worked would simply overvalue their weekly wage and salary earnings compared to others.

Since we will be looking into changes of wage inequality, reporting of educational attainment is very important. This is one of the more problematic areas since countries display huge differences in their educational systems. This is well reflected in the educational attainment variable of LIS data set, since it is an harmonized variable. The first important diference between educational systems is that the timing of their primary, secondary and tertiary education do not usually match. Also, the educational path shows big differences among different countries. While some countries tend toward a more general approach in their pre-college education, for example Germany has a very detailed mixture of general and vocational training. Thankfully, LIS supplies a Stata “do” file which breaks the sample into three educational categories to be used for every country in our list except Hungary. We had to create our own dummies for the Hungarian data, taking into account the advice given by the LIS staff in their web page. So, we characterize the educational attainment with three groups for each of the five countries: the low education group includes primary and lower secondary level of education as well as initial vocational training. The medium level education group includes high-school and pre-tertiary level general education and vocational training. High education groupincludes university level and beyond general and specialized vocational education.

Table 5.2. Education of the Wage-Earner Sample

Countries			Waves II	Waves III	Waves IV	Waves V
	Germany	Low	13.1	12.7	10.0	8.6
		Medium	56.4	58.0	60.4	56.9
		High	30.5	29.3	29.6	34.5
	Netherlands	Low	51.8	30.5	28.8	16.8
		Medium	40.3	47.4	47.9	51.3
		High	7.8	22.1	23.3	32.0
	Sweden	Low		27.0	22.7	18.0
		Medium		47.4	49.5	60.0
		High		25.6	27.8	22.0
	Spain	Low		65.5	58.3	53.5
		Medium		15.8	26.1	30.4
		High		18.7	15.6	16.1
	Hungary	Low		21.0	13.9	16.0
		Medium		65.6	68.5	67.0
		High		13.4	17.6	17.0

All values reported here are percentages of the wage-earner sample.

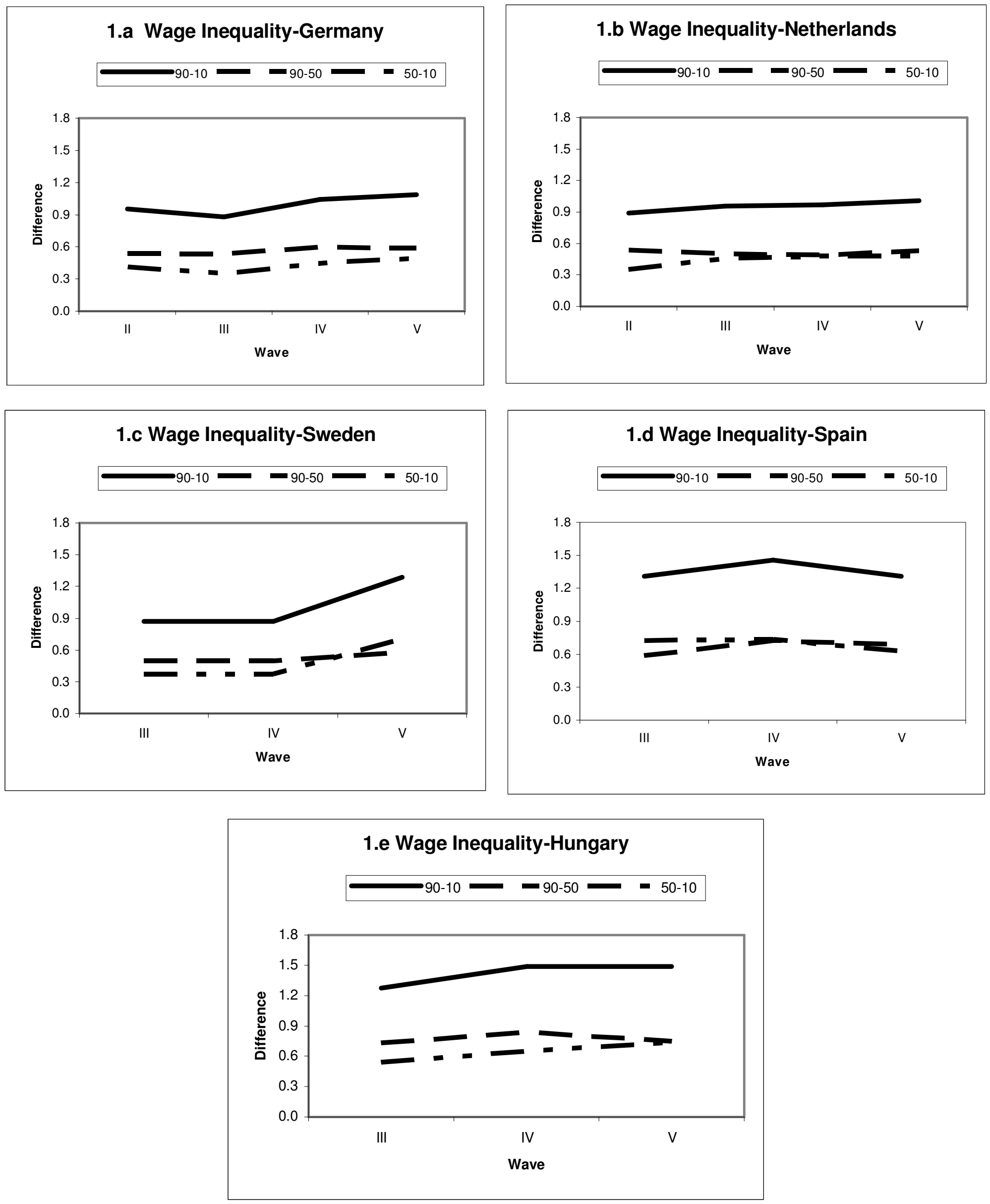
In Table 5.2, we report the distribution of our samples in each country into these three categories, somewhat indicating the supply of different skill levels. The numbers are given in percentage values. In Germany we see that the low education group gets smaller with each wave. While the medium education level group grows in waves III and IV compared to wave II, it returns to more or less the same level as wave II. It is the high level education group that grows up, from 31 percent in wave II to 35 percent in wave V. The German educational system gives a large array of occupational training opportunities at pre-tertiary level, leading to a career. This is probably the main reason behind the high weight of medium level of education in the German sample. As a matter of fact, Germany stands out (together with Hungary, who uses a similar educational system) in the weight of medium level of education.

Sweden and Spain show similar patterns to Germany in the disappearance of low education category. It decreases by a third in Sweden and by a fifth in Spain between waves III and V. In both cases it is the medium level of education category that grows noticeably, with high level of education either remaining the same or decreasing slightly. This probably means that although the overall level of education gets better over the years, either the system does not create enough opportunities to go to college for these newly educated people, or it is not desirable to go to college. This, in return, brings up questions of returns to high level of education which is of very much interest to us.

In Hungary, there are no major changes except the increase in high level education between the 3rd and 4th waves. The share of low level education actually seems to increase in the last wave. One has to be very careful making comments on such variables in a country which recently witnessed drastic changes to its social and economic structure.

The different story here is Netherlands, which demonstrates a sharp decline in low level education and a matching sharp increase in the high level category. Medium level education also increases at a more modest rate. However, we have to keep in mind that we are talking about a span of 16 years here, It is very likely that part of the increase in medium education category is related to the increase in high level education, which is obviously the final target. Evidently, college education has become rather popular in Netherlands. We will see if there is an element of wage inequality anywhere in this story.

5.2 Wage Inequality



Differences between 90^{th}-10^{th}, 90^{th}-50^{th} and 50-10^{th} percentiles of log wage distribution for five countries. Time periods are given in waves as it was explained in Table 5.1.

Figure 5.1 Wage Inequality

As we mentioned earlier, one of the more popular methods of inspecting changes in wage inequality is by looking at the difference between different quantiles of the log wage distribution. We limit our analysis here to three such measures: 90th-10th percentile difference to see the overall change, 90th-50th percentile difference to see the change in the upper half of the wage distribution, and 50th-10th percentile difference to see the change in the lower half.

In Figure 5.1.a we see that Germany records a decline in overall wage inequality in the second half of the 80s. The 90th-50th percentile gap is wider than the 50th-10th percentile gap. Then it changes direction between the third and the fourth waves, that is the in the first half of the 90s. The change comes mainly from the lower half, but the upper level inequality seems to increase slightly as well. As the German overall inequality keeps increasing slightly in the second half of the nineties, we notice that this time the change is mainly fueled by an increase in inequality at the lower half of the log wage distribution. Looking at the general picture, we see that the 50-10 difference increased by 20 percent compared to a 10 percent in 90-50 difference. Also it kept increasing after wave III, while the upper level inequality remained level (it actually decreased a tiny bit).

The overall wage inequality keeps increasing in Netherlands at a very slow pace throughout the years from wave II to V (0.89 to 1.00). We notice serious changes in its composition as well. While the upper level wage inequality (90-50) makes about 60% of the overall change in wave II (0.54/0.89), its share reduces to 51% (0.49/0.97) by wave IV, then picks up a bit to 53% (0.53/1.00) in wave V. On the other hand, wage inequality in the lower half of the wage distribution seems to increase between wages II and IV (50-10 difference increase from 0.36 to 0.48). However, it remains more or less the same between waves IV and V. In short, while the slight increase in overall wage inequality in Netherlands between waves II and V is mainly fueled by a relatively faster increase in lower of the wage distribution (and a decrease in upper-half wage inequality), we see that the tide turns again in wave V and upper half wage inequality picks up pace when lower half wage inequality remains the same. there is a slight change in the composition of it. This is another example of how much one misses just looking at the overall wage inequality values.

Swedish data has a different story to tell. Their wage inequality figures remain stable during the first half of the90s. And this is true for overall, upper half and lower half inequality. In the second half of the 90s, Sweden shows sharp increases in overall and lower level inequality. Although upper half wage inequality also increases, it is the remarkable change in the lower half wage inequality that catches our attention. While the 50-10 difference is 0.37 in wave IV, it jumps up to 0.70 in wave V, a 90% increase. Compared to this, the 16% increase in upper half wage inequality seems quite modest, although it is still higher than what we called "increase in wage inequality" in Netherlands.

Between waves III and V, overall wage inequality in Spain first increases, then decreases back to its level at the beginning. However, its composition changes quite a bit, lower half inequality in Wave V being less than its value in wave III and upper half inequality being higher. It is also remarkable that while the increase in overall inequality from wave III to wave IV came from upper half inequality, the decrease in overall inequality that brought it back to its level in wave III was fueled mostly by a decrease in lower half inequality. This is actually why the composition changed.

The internal dynamics of wage inequality seem to have worked differently for Hungary than they did for Spain. Although Hungary also saw an increase in overall wage inequality between waves III and IV (first half of 90s), the dispersion of wages remained more or less the same in the second half of 90s. However we note that lower half inequality, less than the upper half inequality in wave III, kept increasing and caught it in wave V.

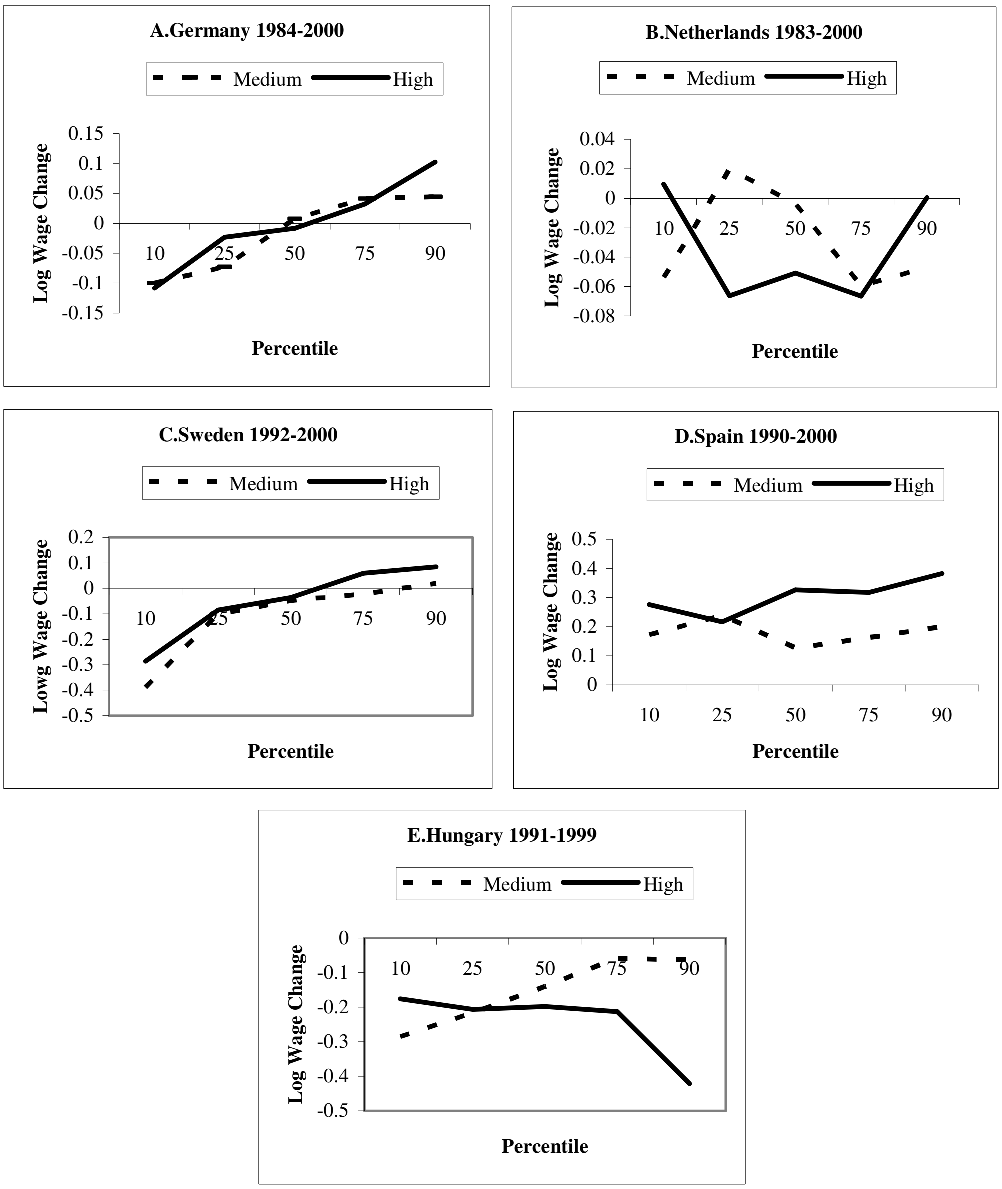
In short, we see an increase in wage overall wage inequality from wave III to IV in all countries except Sweden, and then another increase from wave IV to wave V in Germany, Netherlands and Sweden. In all countries except Spain, lower half inequality becomes more important relative to upper half inequality by wave V, marking increasing dispersion at the lower half of the wage distribution.

When we compare this picture with the US data, we see two important points. First, most of these countries saw increased wage inequality just like the US during 80s and 90s, although to a lesser extent. Second, while the American wage inequality increase came mostly from the upper half during the nineties, all countries in our analysis with the exception of Spain report that there is considerable increase in the lower half of the wage distribution.

5.3 Education and Experience

In figures 2 and 3, we give a picture of how observable characteristics like education and experience affected wage inequality in the countries included in our study.

Figure 5.2 pictures the change in log wages between the beginning and end of the time period we analyze (waves II to V for Germany and Netherlands and waves III to V for other countries) for different percentiles of the wage distribution. Of course these are the change in log real wages, as we explained before. We will try to see if there is any information in these pictures to connect to an increase or decrease of wage inequality in these countries. The high education group always earns more than the medium education group in absolute terms. Thus, if the medium education group gains ground compared to high education group, this signifies a reduction in wage inequality by education. Similarly, if the high education group gains ground, this indicates an increase in wage inequality due to education. Finally, a positive slope in these figures means that higher-percentile earners within the same educational group earn less than the lower-percentile ones, in other words, an increase in within group inequality.



Changes are calculated as the difference between the same percentile of two years. Two separate line for medium and high education groups.

Figure 5.2. Change in Log Wages by Education

It appears that everyone below median of wage distribution earn less in wave V (year 2000) than they did in wave II (1984) in real terms. We do not get a clear indication of change in "between" inequality. However, there is obviously an increase in within inequality. Probably we could even go one step further and say that probably the increase in within inequality is more serious in the high education group.

Both the medium and high level groups of education lose in real terms from wave II(1983) to wave V(2000) in Netherlands with a few exceptions. Perhaps we could conclude that the higher education group is better in the tails and medium education group is better towards the center of the distribution compared to each other. There is no clear indication of change in within inequality here.

Sweden is similar to Germany in the sense that lower half of the wage distribution lost in real terms (in the medium education group, more than the lower half), and there is evidence of increase in within inequality. However, it is different in the sense that there is also an indication of higher between inequality, especially stronger in the tails.

In Spain we notice that both groups got better in real terms between waves III and V. There is also strong evidence of increased between inequality, since high education group gained more than low education group in every point except the 25th percentile, where they are very close. One could probably go ahead and say that there is some evidence of more within inequality in the high education group.

In Hungary everyone lost in real terms. The medium education group gained ground compared to the high education group after 25th percentile, losing less than them. There is increased within inequality in the medium education group, as well.

We see the same picture for low experience (10 years or less potential labor market experience) and high experience (20 to 30 years of potential labor market experience) groups in Figure 5.3. In Germany the low experience group earned more wages for all percentiles, and more than the high experience group did. Lower percentiles of the higher experience group actually lost in real terms. We see a similar picture in Hungary in that lower experience group gaining ground compared to higher experience one, but in Hungary only the upper half of the wage distribution of the lower experience group gained higher real wages. Everyone else get lower wages in wave V compared to what they earned in wave III. There is also an evidence of within inequality in the lower experience group.

We have the opposite picture with Netherlands. High experience group recorded higher wage increases in all percentiles. It looks like the lower tail of the low experience group lost ground considerably against the upper tail. Sweden gives more or less the same story, with the exception that everyone lost in real wages. For Spain, we have a different picture than the others. It looks like there should be less within inequality for the low experience group in wave V compared to wave II.

To be able to see how education and experience affect the wages, we ran a Mincerian wage regression of log weekly wages on a quadratic of potential labor market experience[9], a dummy if the person is married, and educational dummies we mentioned before, for medium and high level of education (with low level excluded). The regression did not include industrial dummies, since their reliability is questionable in several cases, as we explained earlier. The coefficients are soundly significant for all years for all variables except experience for Germany in 1989 (wave III), which is only significant at 89% confidence. It is not surprising to get this kind of result for Germany, since the years of education that we use to create the experience value might not reflect fully the diversity of German education system, which could create two individuals with the same number of years in training, but have very different levels of work experience depending on the level of occupational training.

---

9 Calculated as: Age - years of education- 6. It has to be noted that the years of education is generally a recoded value in LIS data. We imputed values to different levels of educational attainment for each country, depending on the educational system. For no-certificate recodes, we attained highest possible number of years of completed education without getting the next certificate.

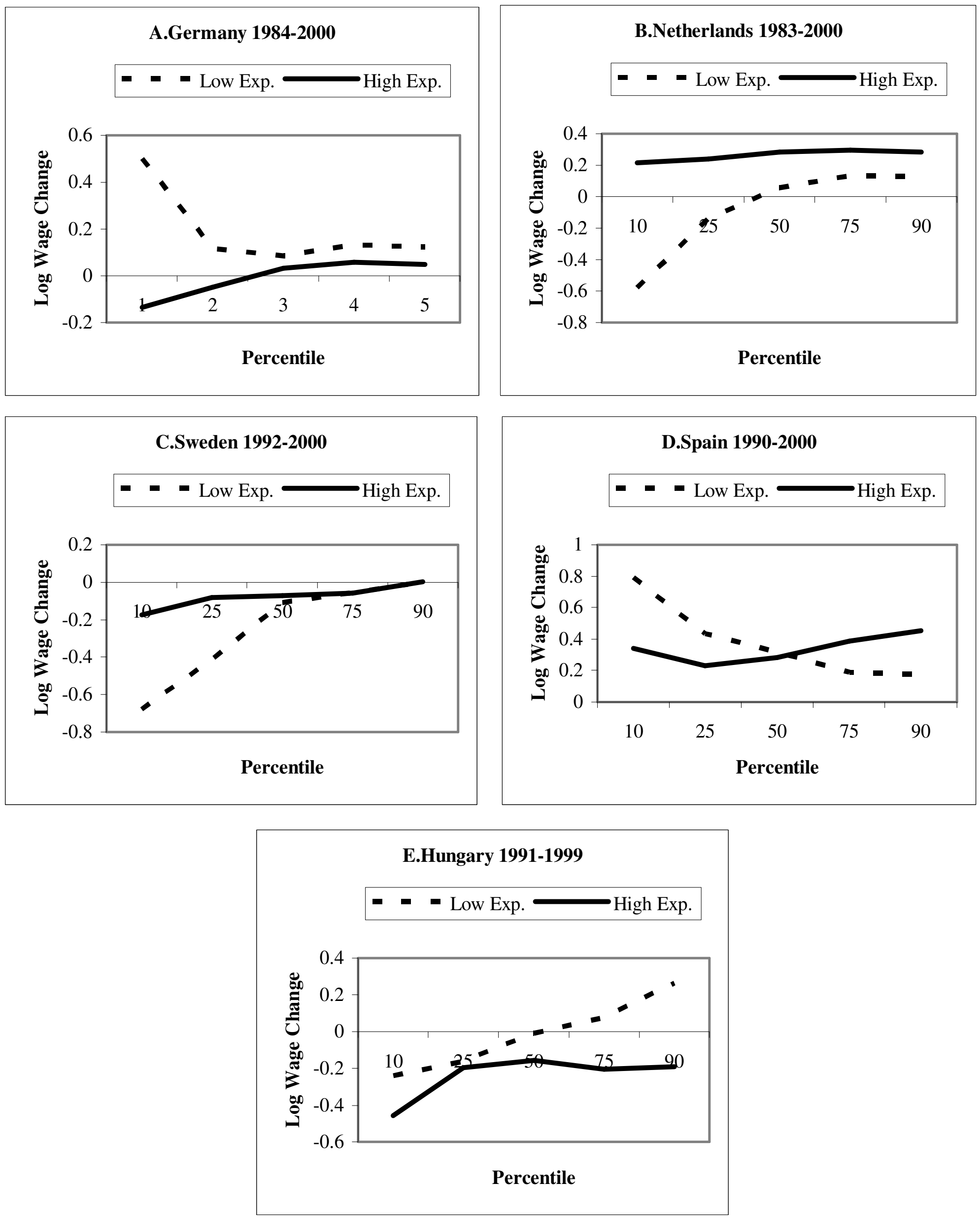


Figure 5.3. Change in Log Wages by Experience

In Table 5.3, we see the regression estimates for education dummies. All values are multiplied by 100. We can observe the effect of having medium and high levels of education on wages here.

Table 5.3 Effect of Education

	Wave	Medium Level	High Level
Germany	II	21.7	53.5
	III	19.2	50.4
	IV	12.3	52.3
	V	10.4	48.6
Netherlands	II	34.8	73.0
	III	24.6	64.3
	IV	20.4	59.4
	V	21.2	60.7
Sweden	III	12.3	38.7
	IV	12.1	37.2
	V	14.1	41.5
Spain	III	39.1	74.7
	IV	40.6	87.1
	V	32.2	72.7
Hungary	III	20.4	73.0
	IV	34.5	107.5
	V	24.5	73.0

In Germany and Netherlands, we see that the effect of having medium level of education decreases during the second half of the 80s and the first half of the nineties. It keeps falling towards the last wave in Germany, while the movement ends at the fourth wave in Netherlands. Spain and Sweden do not observe much change in the first half of the nineties, but they go into different directions in the second half. While medium level (high school) education becomes a bit more favorable in Sweden, it drops drastically in Spain. Hungary sees a sharp increase in the first half of nineties and a following sharp decline in the second half.

There is not much change in the college premium in Germany, the Netherlands and Sweden for any year, except the big fall in Netherlands from wave II to wave III.. On the other hand Spain and Hungary see sharp increases in the first half of nineties followed by sharp decreases to original levels in the second half.

For one thing, we observe that having a college degree, as it should be expected, means an advantage over having a high school equivalent degree in all countries here. Also, the moves in high school premium are matched by parallel moves in college premium in all countries except Germany.In the case of Germany we see that college education becomes more favorable fast. Spain, albeit slower, shows a similar pattern. The other countries do not display considerable changes in this context. Since these coefficients are from an ordinary least squares regression, one has to keep in mind that they do not say anything about the distribution of wages. We will make this point clearer in out discussion about quantile regression.

Since experience enters the regression as a quadratic, we evaluate its marginal effect for a few values of it, namely 5, 15 and 25 years of experience. The marginal effects at these points are reported in Table 5.4. All values have been multiplied by 100.

Table 5.4 Effect of Experience

	Wave	5 years	15 years	25 years
Germany	**II**	10.49	1.70	-0.43
	III	2.79	1.61	0.45
	IV	3.65	1.55	0.50
	V	6.47	1.22	-0.08
Netherlands	**II**	6.96	2.40	0.40
	III	10.84	1.86	0.62
	IV	12.53	2.27	0.39
	V	7.97	2.37	0.75
Sweden	**III**	3.69	1.73	0.67
	IV	4.75	1.25	0.44
	V	8.02	0.90	0.03
Spain	**III**	5.13	2.15	0.89
	IV	6.24	2.32	0.67
	V	4.71	2.19	0.64
Hungary	**III**	5.56	0.84	-0.02
	IV	9.00	0.12	-0.37
	V	3.46	-0.70	0.17

In Germany we see that, after an initial fall, low experience groups begin to recover again. In Netherlands, 5 years experience group gets better until the last wave when they see a fall. In Sweden they see improvement from wave III to wave V. For Spain and Hungary we cannot identify a clear trend.

When we look into the returns to higher experience, we notice that Germany, Sweden and Hungary all see falling values for 15 years of experience. For Netherlands and Spain there is no big change in this group. For the 25 years of experience, the values for Sweden and Spain keep decreasing, while the others are not clear about the direction.

What catches out attention in the experience picture is that while we see different trends for different countries, it is only for 5 years of experience we can identify increasing returns to one additional year of experience (Germany and Sweden). For higher years of experience (15 and 25 years) we either find that either it is less favorable or more or less the same. This could be considered as an indication of higher returns to "unobserved" skills such as the ones related to computer usage, that we talked about before.

5.4 JMP Decomposition

One good way of analyzing these changes in wage inequality is the decomposition method proposed by JMP (1993). As we explained in Chapter 3, the method decomposes the differences between two wage distributions into 3 parts: changes due to the change in quantities of known characteristics (in our case education and experience mostly), changes due to the change in prices of the known characteristics(premia of experience and education) and thirdly, the changes in unobserved prices and quantities.

Tables 5.5.a-e report the results we got from applying this method on our selected countries. The values have been calculated from the same regression we mentioned in section 5.3. As we explained before, T stands for total change, Q for changes in observed skills and characteristics, P for changes in the prices of observed skills and characteristics, and U stands for the changes in the prices and characteristics.

Table 5.5.a JMP Decomposition-Germany

Wave	Percentile Difference	T	Q	P	U
III-II	90th-10th	-0.07	-0.02	-0.01	-0.04
	90th-50th	-0.01	0.01	0.00	-0.02
	50th-10th	-0.07	-0.03	-0.02	-0.02
		T	**Q**	**P**	**U**
IV-III	90th-10th	0.16	0.02	0.03	0.11
	90th-50th	0.06	-0.00	0.02	0.04
	50th-10th	0.10	0.02	0.01	0.06
		T	**Q**	**P**	**U**
V-IV	90th-10th	0.05	0.02	-0.02	0.04
	90th-50th	-0.00	-0.02	-0.02	0.03
	50th-10th	0.05	0.04	-0.00	0.01
		T	**Q**	**P**	**U**
V-II	90th-10th	0.13	0.03	-0.01	0.12
	90th-50th	0.05	-0.01	0.01	0.05
	50th-10th	0.08	0.04	-0.03	0.07

Our model estimates a fall in wage inequality from wave II to wave III in all 3 comparisons (overall, upper half and lower half). Then we se increases in 90-10 difference for the following waves. There is also an increase in wage inequality when we compare wave V with wave II. These estimates are all in harmony with Figure 1.a. In Germany, we see some sizable changes in overall wage inequality (wave IV-III for example). As it was in the US, changes in unexplained characteristics constitute a very large portion of wage inequality for almost all layers, revealing within inequality. However, it seems to be the increase in the dispersion in lower half that makes up for most of the increase or decrease in wage inequality. It seems that in Germany, the between inequality is fueled by changing skill composition in the lower half of the wage distribution, but its main source is the changing skill prices in the upper half. This might be due to the compressed nature of the wages in Germany resulting from institutional structure (Prasad (2004). As we noticed in Figure 1.a, the upper half wage inequality (represented by 90-50 percentile difference) shows either no change or little increase.

Table 5.5.b JMP Decomposition-Netherlands

Wave	**Percentile Difference**	**T**	**Q**	**P**	**U**
III-II	90th-10th	0.07	0.02	-0.02	0.08
	90th-50th	-0.03	-0.04	-0.01	0.02
	50th-10th	0.10	0.06	-0.01	0.06
		T	**Q**	**P**	**U**
IV-III	90th-10th	0.01	0.00	0.02	-0.01
	90th-50th	-0.01	0.00	-0.01	-0.01
	50th-10th	0.02	-0.00	0.03	-0.00
		T	**Q**	**P**	**U**
V-IV	90th-10th	0.04	0.01	0.02	0.02
	90th-50th	0.04	0.02	0.01	0.01
	50th-10th	-0.00	-0.01	0.01	0.00
		T	**Q**	**P**	**U**
V-II	90th-10th	0.12	0.04	-0.01	0.08
	90th-50th	-0.00	-0.02	-0.02	0.04
	50th-10th	0.12	0.07	0.01	0.04

The Dutch data , as we saw in Figure 2, does not show big changes in overall wage inequality from wave to wave, however there is continuing increase which leads to a higher value between waves II and V. There is a remarkable change in composition here. The bulk of wage inequality increase switches from lower half to upper half with each new wave. In this sense we can say that the Dutch experience is a bit similar to that of the US. The unexplained part is important in Netherlands as well, even if it is not as strong as Germany.

Since we have only 3 waves for the remaining countries, it will be good to be careful about making overreaching remarks about them. However, one cannot help but be shocked about the remarkable change in Sweden in a five-year span (wave IV to wave V, or 1995-2000). The change, as we observed before, comes mainly from a jump in the increased dispersion of lower half of the log wage distribution. Unexplained skills and their prices constitute about half of the change.

Table 5.5.c JMP Decomposition-Sweden

Wave	Percentile Difference	T	Q	P	U
IV-III	90th-10th	0.00	-0.03	-0.00	0.03
	90th-50th	0.00	-0.01	-0.00	0.02
	50th-10th	0.00	-0.02	-0.00	0.02
		T	**Q**	**P**	**U**
V-IV	90th-10th	0.41	0.14	0.03	0.24
	90th-50th	0.08	0.01	0.01	0.06
	50th-10th	0.33	0.14	0.02	0.18
		T	**Q**	**P**	**U**
V-III	90th-10th	0.41	0.11	0.03	0.27
	90th-50th	0.08	-0.01	0.00	0.09
	50th-10th	0.33	0.12	0.03	0.18

Table 5.5.d JMP Decomposition-Spain

Wave	Percentile Difference	T	Q	P	U
IV-III	90th-10th	0.15	0.12	-0.02	0.05
	90th-50th	0.14	0.06	0.07	0.00
	50th-10th	0.01	0.06	-0.09	0.05
		T	**Q**	**P**	**U**
V-IV	90th-10th	-0.15	-0.03	-0.09	-0.02
	90th-50th	-0.04	0.01	-0.06	0.01
	50th-10th	-0.11	-0.04	-0.04	-0.03
		T	**Q**	**P**	**U**
V-III	90th-10th	-0.00	0.07	-0.12	0.05
	90th-50th	0.10	0.08	0.00	0.02
	50th-10th	-0.10	-0.01	-0.12	0.04

Table 5.5.d indicates that the unexplained part is important for Spain as well, always pulling the overall inequality value up. In Spain, the upper half of the wage distribution witnesses an increase in wage inequality, while the lower half sees a decrease, and these two moves balance each other, making it look like the overall wage inequality did not change between waves III and V. We see that the unexplained portion explains a bit less than half of the change in overall inequality between waves III and V.

Table 5.5.e shows the Hungarian JMP values. The effect of the unexplained part is very strong in Hungary, perhaps also revealing the side effects of fast economic transition. We notice a remarkable increase in wage inequality in the lower half of wage distribution, while the upper half fluctuates.

Table 5.5.e JMP Decomposition-Hungary

Wave	**Percentile Difference**	**T**	**Q**	**P**	**U**
IV-III	90th-10th	0.21	-0.00	0.13	0.09
	90th-50th	0.11	-0.04	0.12	0.03
	50th-10th	0.11	0.04	0.01	0.06
		T	**Q**	**P**	**U**
V-IV	90th-10th	-0.00	0.03	-0.13	0.09
	90th-50th	-0.09	0.05	-0.15	0.01
	50th-10th	0.09	-0.02	0.03	0.08
		T	**Q**	**P**	**U**
V-III	90th-10th	0.21	0.01	0.01	0.19
	90th-50th	0.02	-0.01	-0.02	0.05
	50th-10th	0.19	0.02	0.04	0.13

5.5 Concluding Remarks

Our analysis of the distribution of log wages in five countries of Europe for the last couple decades with the help of JMP decomposition revealed that even though there are some increases in wage inequality overall in all countries, there is no clear trend except that all five countries showed either an increase or stability between the third and fourth waves, that is the first half of the 90s. There are also different directions of movement in education and experience comparisons. People with high education gained more than the ones with medium education in Sweden and Spain, but not in other countries. Similarly, the gap between high experience and low experience people opened up clearly in Netherlands and Sweden only.

Although the returns to education and experience explain some part of the change in wage inequality, the largest portion is explained by the residuals, in other words by unobserved characteristics. This points to the existence of strong within inequality just like it was in the US.

One thing that catches attention is the increase in lower end inequality, which contradicts with the increase in higher end inequality in the US during the 90s. We observe this change in all five countries when we compare the beginning wave and the end wave. This difference is important in the sense that it means the reasons for change in wage inequality between US and the five countries we have here might be completely different. The main suspect seems to be the institutional differences between the European countries and the US such as the strength of unions and labor market regulations.

CHAPTER 6

QUANTILE REGRESSION ANALYSIS OF THE EUROPEAN DATA

We have run the quantile regressions with the same covariates that were used to fun the OLS regressions that supplied the estimates to apply the JMP method. As we said before, quantile regression estimates are different in nature from the OLS estimates and give us a better understanding of how the effects of covariates might change depending on where we are at the wage distribution.

6.1 Experience

Table 6.1 reports the marginal effects of experience for German data. As in our earlier analysis with the US data, we calculate marginal effects for 5, 15 and 25 years of experience. 5 and 15 years of experience groups both see a drop in their marginal effects on wages from Wave II to Wave III. This is the second half of the 80s. On the other hand, the high-experience group passes from negative to positive marginal effect between these two waves. Combined together, this points to an increase in wage inequality based on experience. The fact that the drop for 15 years of experience is much smaller compared to that of 5 years of experience supports this theory. Then we see the tide turns. Marginal effect of 5 years of experience increases in all quantiles excepth the 90th until wave V, although it never catches up again with the original levels of Wave II. On the other hand, we see decreases for 25 years of experience in all quantiles except for the last two. The 15 years of experience group does not go in a clear direction. In general, we can say that labor market experience has an increasing effect on wage inequality from Wave II to Wave III, then it works toward decreasing wage inequality.

Table 6.1 Marginal Effect of Experience- Germany

Wave	5 Years						15 Years						25 Years					
	OLS	10	25	50	75	90	OLS	10	25	50	75	90	OLS	10	25	50	75	90
II	10.49	18.67	12.75	9.61	3.54	2.75	1.70	1.96	1.65	1.64	1.51	2.04	-0.43	-0.71	-0.56	-0.31	0.10	0.30
	(0.77)	*(1.25)*	*(1.18)*	*(1.04)*	*(0.94)*	*(1.30)*	*(0.20)*	*(0.24)*	*(0.28)*	*(0.24)*	*(0.25)*	*(0.37)*	*(0.13)*	*(0.19)*	*(0.20)*	*(0.16)*	*(0.16)*	*(0.23)*
III	2.79	4.30	2.40	0.22	0.01	2.40	1.61	1.08	1.23	1.41	2.26	1.96	0.45	0.30	0.55	0.64	0.52	0.58
	(0.97)	*(1.92)*	*(1.44)*	*(0.93)*	*(1.31)*	*(1.34)*	*(0.18)*	*(0.34)*	*(0.25)*	*(0.18)*	*(0.28)*	*(0.34)*	*(0.13)*	*(0.26)*	*(0.18)*	*(0.13)*	*(0.20)*	*(0.23)*
IV	3.65	7.76	4.31	1.40	2.71	-0.26	1.55	1.49	1.37	1.32	1.35	2.24	0.50	0.15	0.34	0.54	0.54	0.72
	(1.04)	*(3.06)*	*(1.85)*	*(1.07)*	*(1.50)*	*(2.33)*	*(0.19)*	*(0.66)*	*(0.37)*	*(0.21)*	*(0.34)*	*(0.64)*	*(0.14)*	*(0.37)*	*(0.25)*	*(0.15)*	*(0.24)*	*(0.40)*
V	6.47	10.69	7.09	4.46	3.27	4.45	1.22	0.63	0.93	0.94	1.71	1.93	-0.08	-0.69	-0.28	0.12	0.24	0.42
	(0.67)	*(1.50)*	*(1.05)*	*(0.83)*	*(1.05)*	*(1.64)*	*(0.11)*	*(0.25)*	*(0.19)*	*(0.15)*	*(0.19)*	*(0.30)*	*(0.09)*	*(0.20)*	*(0.14)*	*(0.11)*	*(0.14)*	*(0.21)*

Since experience enters the regressions as a quadratic, the marginal effect has been calculated by evaluating the first derivative of it given for 3 values.Standard errors are given in parentheses. All coefficients and errors are multiplied by 100.

To see how the distributions of each group changed, we go to our impact upon dispersion table. It looks like the five years group has negative impact upon dispersion, and has decreasing within inequality, as we can see from the 90-10 column. The other two experience groups have positive impacts upon dispersion on overall wage inequality (90-10 difference) and increasing within inequality after the second wave. We see that this is mostly fueled by upper half of the distribution for 15 years and the lower half of the distribution for 25 years. Since it looks like the contribution to wage inequality increases as one goes higher in experience years, we could say that experience contributes to wage inequality.

Table 6.2 Impact upon Dispersion-Experience-Germany

Wave	5 Years			15 Years			25 Years		
	90-10	90-50	50-10	90-10	90-50	50-10	90-10	90-50	50-10
II	-15.91	-6.85	-9.06	0.07	0.39	-0.32	1.01	0.61	0.40
III	-1.91	2.18	-4.09	0.89	0.55	0.33	0.28	-0.05	0.34
IV	-8.02	-1.66	-6.36	0.76	0.93	-0.17	0.57	0.19	0.39
V	-6.25	-0.01	-6.23	1.31	1.00	0.31	1.11	0.30	0.81

Differences of indicated quantiles from Table 6.1

Table 6.3 Marginal Effect of Experience- Netherlands

Wave	5 Years						15 Years						25 Years					
	OLS	10	25	50	75	90	OLS	10	25	50	75	90	OLS	10	25	50	75	90
II	6.96	6.55	7.13	6.81	6.05	6.26	2.40	1.99	2.16	2.53	2.64	2.94	0.40	-0.22	0.30	0.58	0.87	1.06
	(0.92)	(1.25)	(0.84)	(1.00)	(0.90)	(1.69)	(0.17)	(0.22)	(0.15)	(0.18)	(0.18)	(0.34)	(0.13)	(0.19)	(0.12)	(0.15)	(0.14)	(0.28)
III	10.84	17.68	12.08	7.88	7.37	6.70	1.86	1.51	1.49	1.74	2.10	2.69	0.62	0.06	0.44	0.46	0.85	1.31
	(0.73)	(1.34)	(0.86)	(0.54)	(0.72)	(1.18)	(0.23)	(0.35)	(0.24)	(0.17)	(0.25)	(0.43)	(0.17)	(0.26)	(0.18)	(0.13)	(0.19)	(0.32)
IV	12.53	20.08	14.34	6.73	7.50	8.22	2.27	2.50	2.03	1.91	2.13	2.51	0.39	-0.10	0.23	0.53	0.75	1.15
	(0.80)	(1.41)	(0.66)	(0.83)	(0.75)	(1.38)	(0.21)	(0.35)	(0.15)	(0.22)	(0.20)	(0.41)	(0.16)	(0.27)	(0.12)	(0.17)	(0.14)	(0.24)
V	7.97	12.25	9.31	6.68	4.91	6.08	2.37	3.36	1.67	1.80	2.13	1.93	0.75	0.13	0.28	0.80	0.96	1.30
	(0.46)	(0.66)	(0.48)	(0.35)	(0.62)	(0.45)	(0.16)	(0.24)	(0.15)	(0.12)	(0.20)	(0.18)	(0.11)	(0.16)	(0.10)	(0.09)	(0.15)	(0.12)

We see the marginal effects of experience for Netherlands in Table 6.3. The 5-year group shows quite a uniform trend in terms of the direction of changes. We see increases in marginal effects for all quantiles, and the OLS, until Wave IV. Between Wave IV and Wave V, there is a decrease. The other two groups do not show clear trends. We could perhaps say that from Wave II to Wave IV, experience works towards less inequality, as the marginal effect of the 5 year experience group is increasing while others do not show a clear move. However we are not in a position to claim this as strongly as we did in the case of Germany, since the other groups show fluctuations on different points of the distribution.

Low experience group seems to have a negative impact upon dispersion for Netherlands, with the effect of it always higher in the 10th percentile than in the 90th. However we notice that as one goes higher in experience, the impact upon dispersion values for 90-10 dispersion turn positive, and increase. We notice that this increase is especially related to an increased dispersion in the lower half of the distribution, although the upper half contributes as well. In general, one can say that the within inequality generated by experience groups contributes to increased wage inequality in Netherlands.

Table 6.4 Impact upon Dispersion-Experience-Netherlands

Wave	5 Years			15 Years			25 Years		
	90-10	**90-50**	**50-10**	**90-10**	**90-50**	**50-10**	**90-10**	**90-50**	**50-10**
II	-0.29	-0.56	0.26	0.96	0.41	0.54	1.28	0.48	0.80
III	-10.99	-1.18	-9.80	1.18	0.95	0.23	1.25	0.85	0.40
IV	-11.86	1.48	-13.35	0.01	0.61	-0.59	1.25	0.61	0.64
V	-6.17	-0.60	-5.57	-1.43	0.12	-1.55	1.17	0.50	0.67

Marginal effect of experience for Sweden is reported in Table 6.5. 5-year and 25-year experience groups show opposite movements. While the marginal effect of 5 years of experience increases from Wave III to Wave IV to Wave V, that of 25 years of experience decreases in all quantiles as well as the OLS regression. The 15-year experience group displays a similar picture to 25-year group in the upper half of the distribution, but it fluctuates in the lower tail. All in all, we can say that experience does not contribute to increasing wage inequality, in fact it works in the opposite direction.

Table 6.5 Marginal Effect of Experience- Sweden

	5 Years						15 Years						25 Years					
Wave	**OLS**	**10**	**25**	**50**	**75**	**90**	**OLS**	**10**	**25**	**50**	**75**	**90**	**OLS**	**10**	**25**	**50**	**75**	**90**
III	3.69	4.38	4.05	3.43	3.08	2.87	1.73	1.30	1.03	1.46	2.01	2.43	0.67	0.51	0.32	0.56	0.86	1.19
	(0.43)	*(1.13)*	*(0.68)*	*(0.51)*	*(0.71)*	*(0.85)*	*(0.12)*	*(0.23)*	*(0.16)*	*(0.13)*	*(0.19)*	*(0.25)*	*(0.09)*	*(0.18)*	*(0.11)*	*(0.09)*	*(0.13)*	*(0.18)*
IV	4.75	7.16	4.74	3.98	4.30	3.68	1.25	1.01	1.04	0.98	1.17	1.76	0.44	-0.10	0.21	0.44	0.63	1.16
	(0.52)	*(1.45)*	*(0.72)*	*(0.65)*	*(0.84)*	*(0.90)*	*(0.13)*	*(0.27)*	*(0.15)*	*(0.15)*	*(0.19)*	*(0.24)*	*(0.09)*	*(0.21)*	*(0.11)*	*(0.11)*	*(0.14)*	*(0.17)*
V	8.02	8.23	10.29	7.64	5.85	6.58	0.90	1.68	0.44	0.46	0.98	1.36	0.03	0.25	0.04	-0.07	0.19	0.24
	(0.29)	*(0.92)*	*(0.32)*	*(0.23)*	*(0.22)*	*(0.47)*	*(0.11)*	*(0.38)*	*(0.13)*	*(0.09)*	*(0.09)*	*(0.19)*	*(0.08)*	*(0.24)*	*(0.08)*	*(0.06)*	*(0.06)*	*(0.12)*

Sweden displays a somewhat different picture than Netherlands and Germany about the within inequality generated by experience. Here, it looks like higher experience does not always mean higher within inequality. It is no doubt clear that 15 years of experience creates more dispersion in itself than 5 years of experience, however it is also cause of more dispersion than 25 years of experience. Most of the within inequality we observe in the 15 years of education group comes from the upper half of distribution. So, although the inequality within experience groups seems to contribute to the overall inequality to some degree, one cannot safely say that this contribution increases as one goes higher in experience groups, as we see in Table 6.4.

Table 6.6 Impact upon Dispersion-Experience-Sweden

Wave	5 Years			15 Years			25 Years		
	90-10	**90-50**	**50-10**	**90-10**	**90-50**	**50-10**	**90-10**	**90-50**	**50-10**
III	-1.51	-0.56	-0.95	1.13	0.98	0.16	0.68	0.63	0.05
IV	-3.48	-0.29	-3.19	0.75	0.78	-0.03	1.26	0.72	0.54
V	-1.65	-1.07	-0.58	-0.33	0.89	-1.22	-0.02	0.31	-0.32

The Spanish marginal effects of experience do not give us very clear messages about the wage inequality created by experience in Spain. From Wave III to Wave IV, it looks like all three groups see their effects on the distribution wages increase, with a couple exceptions, namely the lower tails of 15 and 25 years of experience groups. Then from Wave IV to Wave V we see an increase in the marginal overall. This time the upper tails of 15 and 25 years of experience groups do not follow the trend and see their effects increase. In short, it is not possible to observer a clear direction that the experience takes in contributing to wage inequality.

Table 6.7 Marginal Effect of Experience- Spain

Wave	5 Years						15 Years						25 Years					
	OLS	**10**	**25**	**50**	**75**	**90**	**OLS**	**10**	**25**	**50**	**75**	**90**	**OLS**	**10**	**25**	**50**	**75**	**90**
III	5.13	10.07	5.32	3.53	3.17	3.11	2.15	3.18	1.61	1.52	1.69	1.85	0.89	0.61	0.71	0.60	0.93	1.04
	(0.39)	*(0.97)*	*(0.50)*	*(0.29)*	*(0.51)*	*(0.52)*	*(0.13)*	*(0.33)*	*(0.15)*	*(0.09)*	*(0.16)*	*(0.15)*	*(0.10)*	*(0.24)*	*(0.11)*	*(0.07)*	*(0.13)*	*(0.12)*
IV	6.24	12.14	6.92	4.55	4.31	3.75	2.32	3.06	1.94	2.07	2.03	2.71	0.67	-0.22	0.27	0.70	1.01	1.32
	(0.77)	*(1.77)*	*(0.83)*	*(0.66)*	*(0.71)*	*(0.95)*	*(0.26)*	*(0.52)*	*(0.27)*	*(0.23)*	*(0.25)*	*(0.36)*	*(0.20)*	*(0.43)*	*(0.20)*	*(0.17)*	*(0.19)*	*(0.26)*
V	4.71	5.98	4.43	3.91	2.47	1.67	2.19	2.09	1.85	1.55	2.23	2.95	0.64	0.26	0.59	0.63	1.05	1.85
	(0.46)	*(1.04)*	*(0.53)*	*(0.39)*	*(0.49)*	*(0.49)*	*(0.19)*	*(0.38)*	*(0.21)*	*(0.16)*	*(0.21)*	*(0.23)*	*(0.14)*	*(0.29)*	*(0.16)*	*(0.12)*	*(0.15)*	*(0.15)*

Table 6.8 shows that the within inequality among similar experience groups increases as one goes higher in experience in Spain.

Table 6.8 Impact upon Dispersion-Experience-Spain

Wave	5 Years			15 Years			25 Years		
	90-10	90-50	50-10	90-10	90-50	50-10	90-10	90-50	50-10
II	-6.96	-0.42	-6.54	-1.34	0.33	-1.66	0.44	0.44	0.00
III	-8.39	-0.80	-7.59	-0.35	0.64	-0.99	1.54	0.62	0.92
IV	-4.32	-2.24	-2.07	0.86	1.39	-0.53	1.58	1.22	0.37

Lastly, we see the marginal effect of experience for Hungary in Table 6.9. 5 years of experience shows different moves in the two halves of the distribution. In the lower half, the marginal effects first increase, then decrease as one goes from Wave III to Wave V. In the upper half, they increase all the way. The upper tails of 15 and 25 years of experience also show increases all along. The other percentiles do not show clear tendencies. Thus we could say that experience contributes to wage inequality in Hungary, and this contribution is mostly generated by the upper tail.

Table 6.9 Marginal Effect of Experience- Hungary

Wave	5 Years						15 Years						25 Years					
	OLS	10	25	50	75	90	OLS	10	25	50	75	90	OLS	10	25	50	75	90
III	5.56	5.56	3.83	3.24	2.39	7.39	0.84	0.37	0.88	0.35	1.23	0.97	-0.02	0.26	-0.02	-0.11	0.29	0.57
	(1.65)	*(3.06)*	*(1.34)*	*(1.62)*	*(1.70)*	*(2.49)*	*(0.50)*	*(0.84)*	*(0.40)*	*(0.50)*	*(0.57)*	*(0.85)*	*(0.37)*	*(0.61)*	*(0.28)*	*(0.36)*	*(0.40)*	*(0.58)*
IV	9.00	13.85	8.41	5.02	6.95	9.77	0.12	0.99	0.98	0.12	-0.16	-0.44	-0.37	-0.79	-0.56	0.34	0.67	-0.41
	(1.75)	*(2.70)*	*(2.31)*	*(2.01)*	*(2.77)*	*(2.26)*	*(0.63)*	*(0.93)*	*(0.81)*	*(0.70)*	*(0.89)*	*(0.97)*	*(0.51)*	*(0.68)*	*(0.67)*	*(0.57)*	*(0.71)*	*(0.74)*
V	3.46	3.33	2.44	5.39	10.35	12.27	-0.70	-1.03	-0.11	-0.03	-1.51	-1.99	0.17	2.47	0.70	-0.21	-0.10	-1.54
	(2.09)	*(3.35)*	*(1.90)*	*(2.16)*	*(1.36)*	*(2.70)*	*(0.66)*	*(1.10)*	*(0.66)*	*(0.72)*	*(0.40)*	*(0.72)*	*(0.52)*	*(0.82)*	*(0.49)*	*(0.55)*	*(0.33)*	*(0.59)*

Table 6.9 reports the impact upon dispersion values for Hungary. It is quite different from what we have seen in other countries. For one thing, within inequality in the

lower experience group contributes to overall wage inequality, especially in the upper half of the distribution. There is also some contribution of within inequality from other experience groups, but none as strong as the low experience one.

Table 6.10 Impact upon Dispersion-Experience-Hungary

Wave	5 Years			15 Years			25 Years		
	90-10	**90-50**	**50-10**	**90-10**	**90-50**	**50-10**	**90-10**	**90-50**	**50-10**
III	1.83	4.15	-2.32	0.59	0.62	-0.02	0.30	0.68	-0.38
IV	-4.08	4.75	-8.83	-1.43	-0.56	-0.87	0.38	-0.75	1.13
V	8.94	6.88	2.06	-0.95	-1.95	1.00	-4.02	-1.33	-2.68

6.2 Education

The marginal effects of having medium (high school equivalent) and high (college and higher) education on wages are reported in Panel A Table 6.11. For the medium education group, we see a drop in the third and fourth waves compared to the first and second ones. We do not see this sort of drop in the high education level group. Thus, we could probably say that this difference between the medium and high education group is a reason for the increase in wage inequality in Germany between waves III and IV that we observed in Figure 5.1.

A clearer picture of medium versus high level of education is the simple comparison of their marginal effects. In Panel B of Table 6.11 we give this measure. We notice that there is a dramatic increase in high education premium in all percentiles except the 90th from wave III to wave IV. This confirms our observation from the previous table that educational differences contributed to the increase in wage inequality between the third and the fourth waves.

Table 6.11 Marginal Effect of Education-Germany

A.Marginal Effect of Education

Wave	Medium Level						High Level					
	OLS	10	25	50	75	90	OLS	10	25	50	75	90
II	21.67	18.89	18.62	14.78	20.05	23.02	53.47	43.82	44.99	49.58	57.81	58.16
	(2.21)	*(3.12)*	*(3.04)*	*(2.47)*	*(2.60)*	*(3.50)*	*(2.39)*	*(3.34)*	*(3.34)*	*(2.72)*	*(2.81)*	*(3.77)*
III	19.23	13.73	16.72	14.97	21.37	25.11	50.40	26.43	36.59	50.52	56.60	69.29
	(2.08)	*(3.41)*	*(2.43)*	*(1.89)*	*(2.71)*	*(3.15)*	*(2.28)*	*(4.16)*	*(2.84)*	*(2.12)*	*(3.00)*	*(3.64)*
IV	12.25	-6.64	2.14	10.33	14.82	22.18	52.26	24.31	37.07	53.08	60.70	63.18
	(2.43)	*(5.16)*	*(3.67)*	*(2.17)*	*(3.75)*	*(6.62)*	*(2.61)*	*(6.99)*	*(4.47)*	*(2.52)*	*(4.18)*	*(6.87)*
V	10.41	4.77	9.37	6.70	14.91	17.53	48.55	34.06	40.42	42.40	57.96	71.25
	(1.52)	*(3.49)*	*(2.53)*	*(2.13)*	*(2.65)*	*(4.05)*	*(1.59)*	*(3.59)*	*(2.63)*	*(2.23)*	*(2.78)*	*(4.24)*

B. High Education Premium

Wave	OLS	10	25	50	75	90
II	31.81	24.93	26.37	34.80	37.77	35.14
III	31.17	12.70	19.87	35.55	35.23	44.18
IV	40.01	30.95	34.93	42.75	45.88	40.99
V	38.15	29.30	31.05	35.70	43.05	53.73

The impact upon dispersion table for Germany (Table 6.12) shows that both medium and high level of education contribute to wage inequality due to the inequality within themselves, and this contribution has been consistently supported by an increase in upper half wage inequality. The lower half wage inequality within both groups is somewhat shaky, sometimes increasing and sometimes decreasing. However, except for the first wave, it is always positive as well. Thus, we can in general say that education contributes to wage inequality in Germany.

Table 6.12 Impact upon Dispersion-Education-Germany

Wave	Medium Level			High Level		
	90-10	90-50	50-10	90-10	90-50	50-10
II	4.12	8.24	-4.12	14.33	8.58	5.76
III	11.38	10.14	1.24	42.86	18.77	24.09
IV	28.82	11.85	16.97	38.87	10.10	28.77
V	12.76	10.83	1.93	37.19	28.86	8.33

Table 6.13 Marginal Effect of Education-Netherlands

A. Marginal Effect of Education

	Medium Level						High Level					
Wave	**OLS**	**10**	**25**	**50**	**75**	**90**	**OLS**	**10**	**25**	**50**	**75**	**90**
II	34.82	26.21	31.76	35.26	39.36	37.78	73.02	59.11	73.60	79.50	81.28	78.44
	(1.45)	*(1.97)*	*(1.32)*	*(1.60)*	*(1.62)*	*(3.07)*	*(2.58)*	*(3.55)*	*(2.45)*	*(3.02)*	*(2.92)*	*(5.38)*
III	24.59	26.76	20.76	19.76	21.70	25.76	64.33	56.99	55.73	58.00	64.04	66.41
	(1.91)	*(2.97)*	*(2.02)*	*(1.37)*	*(1.91)*	*(3.13)*	*(2.35)*	*(3.70)*	*(2.47)*	*(1.69)*	*(2.37)*	*(3.79)*
IV	20.39	17.51	19.97	19.09	20.17	25.79	59.41	50.36	55.38	53.95	59.65	67.95
	(1.74)	*(3.22)*	*(1.34)*	*(1.86)*	*(1.62)*	*(3.00)*	*(2.09)*	*(3.65)*	*(1.65)*	*(2.22)*	*(1.88)*	*(3.45)*
V	21.23	17.11	18.97	18.08	19.43	23.46	60.66	49.42	54.24	57.04	61.09	67.74
	(1.58)	*(2.24)*	*(1.57)*	*(1.14)*	*(2.10)*	*(1.66)*	*(1.73)*	*(2.51)*	*(1.73)*	*(1.25)*	*(2.30)*	*(1.87)*

B. High Education Premium

Wave	**OLS**	**10**	**25**	**50**	**75**	**90**
II	38.21	32.90	41.84	44.24	41.92	40.66
III	39.74	30.23	34.97	38.24	42.34	40.65
IV	39.02	32.86	35.41	34.86	39.49	42.15
V	39.44	32.31	35.27	38.96	41.66	44.28

The Dutch data shows in Table 6.13 that after an initial drop from wave II to Wave III, the marginal effects of both medium and high level of education do not change much. We also see this from the high education premium table as well. However, the high education premium does not fall between waves II and III at the upper tail, while it falls in the lower one. This is one of the reasons for the mild inequality increases in Netherlands between these two waves. After Wave III, we do not see the evidence of a strong effect of educational differences to increase the wage inequality in Netherlands.

The impact upon dispersion table for Netherlands is given in table 6.14. The within inequality of medium education group is positive except for Wave III. Still, the within inequality in the high education group is both larger and increasing since the third wave. Both groups contribute to dispersion in Dutch wages.

Table 6.14 Impact upon Dispersion-Education-Netherlands

Wave	Medium Level			High Level		
	90-10	90-50	50-10	90-10	90-50	50-10
II	11.57	2.52	9.05	19.33	-1.06	20.39
III	-1.00	6.00	-7.00	9.42	8.41	1.01
IV	8.29	6.70	1.58	17.58	14.00	3.58
V	6.36	5.38	0.98	18.32	10.70	7.62

We do not see big changes in marginal effects of medium or high level of education in Sweden. Table 6.15 shows that while there are some fluctuations in the tails of the distribution, there are no big changes towards increased or decreased inequality. The high education premium decreases in the lower tail, while it remains the same in the upper one. This could be considered a minor move towards lower inequality.

Table 6.15 Marginal Effect of Education-Sweden

A. Marginal Effect of Education

Wave	Medium Level						High Level					
	OLS	10	25	50	75	90	OLS	10	25	50	75	90
II	12.28	10.09	7.89	11.43	16.06	16.20	38.71	31.05	28.03	34.82	46.14	54.89
	(1.21)	*(2.39)*	*(1.52)*	*(1.23)*	*(1.84)*	*(2.52)*	*(1.38)*	*(2.69)*	*(1.72)*	*(1.39)*	*(2.06)*	*(2.75)*
III	12.07	8.91	7.18	10.52	12.20	18.67	37.17	28.05	24.55	32.15	42.50	53.82
	(1.37)	*(3.30)*	*(1.67)*	*(1.63)*	*(2.08)*	*(2.60)*	*(1.53)*	*(3.55)*	*(1.84)*	*(1.78)*	*(2.24)*	*(2.77)*
IV	14.11	16.80	7.46	9.31	16.65	25.93	41.54	31.65	26.20	36.54	51.70	65.75
	(1.14)	*(4.02)*	*(1.31)*	*(0.95)*	*(0.90)*	*(1.88)*	*(1.34)*	*(4.62)*	*(1.53)*	*(1.11)*	*(1.05)*	*(2.17)*

B. High Education Premium

Wave	OLS	10	25	50	75	90
III	26.44	20.97	20.14	23.40	30.08	38.69
IV	25.10	19.13	17.38	21.63	30.30	35.16
V	27.43	14.85	18.74	27.23	35.05	39.82

It is clear from Table 6.16 that both medium and high level of education groups contribute to wage inequality via their within inequality. Besides, their within inequalities increase over time. This is more emphasized in the high level of education group and its upper half. The upper half is also more active in the medium level of education group as well.

Table 6.16 Impact upon Dispersion-Education-Sweden

	Medium Level			High Level		
Wave	**90-10**	**90-50**	**50-10**	**90-10**	**90-50**	**50-10**
III	6.11	4.77	1.34	23.84	20.07	3.77
IV	9.75	8.15	1.60	25.78	21.68	4.10
V	9.13	16.63	-7.49	34.10	29.21	4.89

Table 6.17 Marginal Effect of Education-Spain

A. Marginal Effects of Education

	Medium Level						High Level					
Wave	**OLS**	**10**	**25**	**50**	**75**	**90**	**OLS**	**10**	**25**	**50**	**75**	**90**
III	39.07	40.65	37.69	33.76	30.01	35.07	74.67	74.12	63.23	64.48	66.48	74.52
	(1.55)	*(3.90)*	*(1.84)*	*(1.11)*	*(1.99)*	*(1.92)*	*(1.54)*	*(4.52)*	*(1.92)*	*(1.13)*	*(2.03)*	*(2.01)*
IV	40.57	45.95	36.51	33.98	33.55	35.62	87.07	88.89	78.00	79.79	80.94	99.51
	(2.92)	*(6.18)*	*(2.93)*	*(2.53)*	*(2.85)*	*(4.04)*	*(3.44)*	*(7.34)*	*(3.49)*	*(2.98)*	*(3.33)*	*(4.84)*
V	32.18	29.63	24.26	26.06	28.57	30.86	72.70	69.06	65.42	67.98	76.44	82.82
	(1.95)	*(4.23)*	*(2.19)*	*(1.64)*	*(2.15)*	*(2.25)*	*(2.42)*	*(4.82)*	*(2.68)*	*(2.01)*	*(2.59)*	*(2.71)*

B. High Education Premium

Wave	**OLS**	**10**	**25**	**50**	**75**	**90**
III	35.61	33.47	25.54	30.73	36.46	39.45
IV	46.50	42.94	41.49	45.81	47.39	63.89
V	40.52	39.43	41.16	41.92	47.87	51.96

Table 6.17 shows the marginal effects and high education premium for Spain. The increase in wage inequality that we observed in Figure 5.1 earlier seems to have something to do with educational differences. The medium level education does not see much change in its marginal effects, except for the lowe tail, while the marginal effects of the higher education group increase throughout the distribution between waves III and IV. Then the marginal effects of both groups decrease between waves IV and V, contributing to the decrease in overall wage inequality. We see the same

moves in high education premium values. Especially the changes in the upper tail are noteworthy.

The impact upon dispersion values for Spain are given in Table 6.18. The high education group contributes to overall inequality in all waves. The medium education group contributes only in the last wave. This is the picture of overall within inequality. However, in both groups we see contribution to dispersion from the upper half of the distribution. This effect is offset in the medium education group by the stretching of lower half.

Table 6.18 Impact upon Dispersion-Education-Spain

	Medium Level			**High Level**		
Wave	**90-10**	**90-50**	**50-10**	**90-10**	**90-50**	**50-10**
III	-5.58	1.31	-6.89	0.40	10.04	-9.64
IV	-10.33	1.64	-11.97	10.62	19.72	-9.10
V	1.23	4.80	-3.57	13.76	14.84	-1.08

For the analysis of Hungarian data, we have Table 6.19. The medium education group sees an increase in marginal effects between waves III and IV in all percentiles as well as the OLS. Then the lower half of the distribution and the median lose some of this increase back between waves IV and V. The upper tail does not see this drop. The high education group sees the first increase as well. However, the marginal effects drop back throughout the distribution between waves IV and V. From Panel B of Table 6.19, we see that the high level premium first increases then decreases for all percentiles as well as the OLS. The most serious fall between waves IV and V is in the upper tail, as expected.

Table 6.19 Marginal Effect of Education-Hungary

A. Marginal Effect of Education

Wave	Medium Level						High Level					
	OLS	10	25	50	75	90	OLS	10	25	50	75	90
III	20.42	29.85	19.31	17.98	15.34	15.99	72.99	68.37	72.10	82.10	79.86	97.09
	(4.93)	*(8.56)*	*(3.84)*	*(4.86)*	*(5.82)*	*(9.23)*	*(6.70)*	*(10.92)*	*(5.14)*	*(6.61)*	*(7.81)*	*(12.10)*
IV	34.46	44.89	35.14	27.90	28.28	22.27	107.54	113.93	97.30	101.73	105.92	113.84
	(6.46)	*(9.27)*	*(8.64)*	*(7.20)*	*(9.01)*	*(10.77)*	*(7.85)*	*(11.26)*	*(10.34)*	*(8.77)*	*(10.98)*	*(12.75)*
V	24.48	26.94	24.51	23.29	28.91	22.86	73.01	80.57	75.23	87.25	89.62	79.33
	(6.20)	*(11.45)*	*(6.45)*	*(6.95)*	*(3.77)*	*(5.88)*	*(7.83)*	*(14.65)*	*(8.17)*	*(8.56)*	*(4.81)*	*(8.39)*

B. High Education Premium

Wave	OLS	10	25	50	75	90
III	52.57	38.52	52.79	64.13	64.52	81.10
IV	73.08	69.04	62.16	73.84	77.64	91.57
V	48.53	53.64	50.72	63.97	60.71	56.47

Table 6.20 Impact upon Dispersion-Education-Hungary

	Medium Level			High Level		
Wave	90-10	90-50	50-10	90-10	90-50	50-10
III	-13.85	-1.98	-11.87	28.73	14.99	13.73
IV	-22.62	-5.63	-16.99	-0.09	12.11	-12.20
V	-4.08	-0.43	-3.65	-1.24	-7.92	6.68

Lastly, we see the values that show the impact of education on dispersion in Table 6.20. Although there is strong within inequality among the high education group in wave III, neither group contributes to inequality in the following two waves.

6.3 Conditional and Counterfactual Distributions

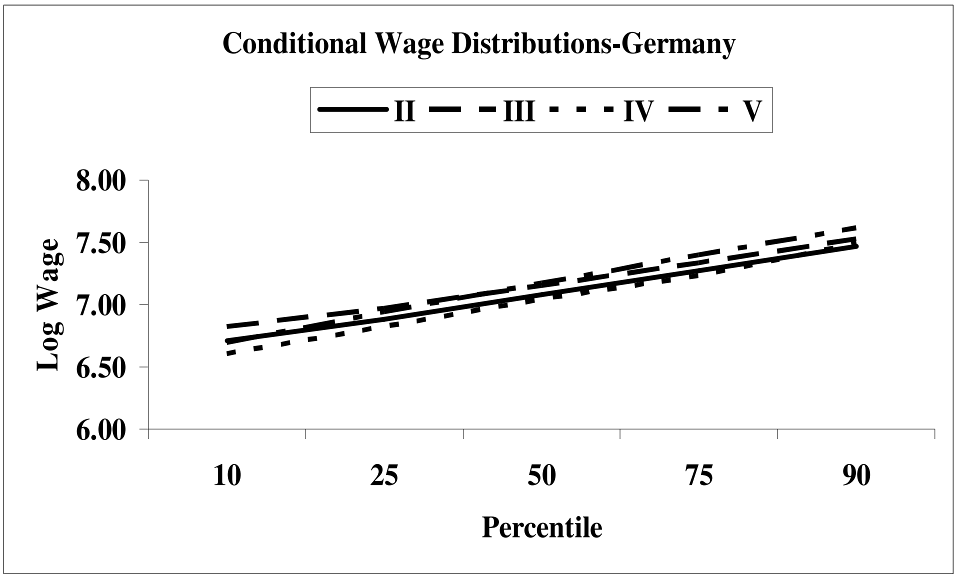
We have the conditional wage distributions for Germany in Table 6.21. These distributions have been prepared following the same logic as we used in Chapter 4. Since we use the average values of covarites for a given wave, they are expected to show less inequality than the empirical distribution.

Table 6.21 Conditional Wage Distribution-Germany

Conditional Wage Distribution				
	Waves			
Percentile	**II**	**III**	**IV**	**V**
10	6.708	6.825	6.604	6.697
25	6.881	6.970	6.823	6.942
50	7.079	7.152	7.045	7.170
75	7.273	7.335	7.231	7.399
90	7.469	7.525	7.494	7.616
Dispersion from Conditional Distribution				
Perc.Diff.	**II**	**III**	**IV**	**V**
90-10	0.76	0.70	0.89	0.92
90-50	0.39	0.37	0.45	0.45
50-10	0.37	0.33	0.44	0.47
Dispersion from Empirical Distribution				
Perc.Diff.	**II**	**III**	**IV**	**V**
90-10	0.96	0.88	1.04	1.09
90-50	0.54	0.53	0.60	0.59
50-10	0.41	0.35	0.44	0.49
Portion of Empirical Dispersion Explained by Conditional				
Perc.Diff.	**II**	**III**	**IV**	**V**
90-10	0.80	0.79	0.85	0.84
90-50	0.72	0.70	0.75	0.75
50-10	0.89	0.94	0.99	0.96
Changes in Dispersion(Conditional)				
Perc.Diff.	**III-II**	**IV-III**	**V-IV**	
90-10	-0.061	0.190	0.030	
90-50	-0.017	0.076	-0.004	
50-10	-0.044	0.114	0.033	

We notice in Table 6.21 that the conditional distribution captures the dispersion measures of 90-10, 90-50 and 50-10 percentile differences quite successfully when compared to the empirical distribution. It is especially on target with the lower half wage inequality measure. As expected, the conditional distribution correctl marks the direction of changes in dispersion over time for Germany.

It is not very straightforward from the table how the conditional distribution evolved over time. For this, we have Figure 6.1. It shows that the conditional distributions for waves II and III have similar slopes. Wave III is a shift up from wave II. Similarly Waves IV and V are alike. V is a shift up from IV. The interesting point here is that the shape of conditional wage distribution changes between the third and the fourth waves. This is a change that favored higher quantiles over lower ones. As well as what we said about education and experience effects, this change helps explaining the nature of increase in wage inequality in Germany in that period.



Roman numbers represent waves of LIS data.

Figure 6.1 Conditional Wage Distributions-Germany

We also have two sets of counterfactual wage distributions for Germany: using the Wave III averages of covariates and each wave's coefficients, and using the Wave II coefficients from our quantile regressions and each wave's averages of covariates. Table 6.22 shows these two distributions. As we explained in Chapter 4, the counterfactual distribution that uses the average values of covariates from wave II helps us isolate the effect of changes in the distribution of covariates. The degree at which this distribution can explain the changes in dispersion from the conditional distribution is an indication of how important are the changes we isolate. It appears that this counterfactual distribution explains the changes in dispersion until wave V quite nicely. However it fails to explain the change in dispersion between waves IV and V. This is an indication that changes in the distribution of covariates became more important at this point.

The second counterfactual distribution does just the opposite. This time we isolate the effects of changes in the prices of personal characteristics and attempt to see how this affects the distribution. It turns out that this distribution explains the changes in overall wage inequality between Waves IV and V better than any other time period. Combining this with our findings from the first counterfactual distribution, we could probably assume that until wave V changes it is the prices of covariates that affects the wage distribution. With wave V the effect of the composition of covariates increases.

Table 6.22 Counterfactual Distributions-Germany

Counterfactual Wage Distribution(Wave III Averages)					Counterfactual Wage Distribution(Wave III Coefficient)				
	Waves					Waves			
Percentile	**II**	**III**	**IV**	**V**	**Percentile**	**II**	**III**	**IV**	**V**
10	6.708	6.827	6.615	6.646	**10**	6.708	6.716	6.736	6.793
25	6.881	6.975	6.828	6.886	**25**	6.881	6.886	6.903	6.959
50	7.079	7.158	7.051	7.119	**50**	7.079	7.081	7.091	7.155
75	7.273	7.340	7.231	7.323	**75**	7.273	7.270	7.281	7.351
90	7.469	7.530	7.496	7.521	**90**	7.469	7.463	7.475	7.548
	Dispersion					Dispersion			
Perc.Diff.	**II**	**III**	**IV**	**V**	**Perc.Diff.**	**II**	**III**	**IV**	**V**
90-10	0.76	0.70	0.88	0.88	**90-10**	0.76	0.75	0.74	0.75
90-50	0.39	0.37	0.44	0.40	**90-50**	0.39	0.38	0.38	0.39
50-10	0.37	0.33	0.44	0.47	**50-10**	0.37	0.37	0.36	0.36
	Changes in Dispersion					Changes in Dispersion			
Perc.Diff.	**III-II**	**IV-III**	**V-IV**		**Perc.Diff.**	**III-II**	**IV-III**	**V-IV**	
90-10	-0.06	0.18	0.00		**90-10**	-0.02	-0.01	0.02	
90-50	-0.02	0.07	-0.04		**90-50**	-0.01	0.00	0.01	
50-10	-0.04	0.10	0.04		**50-10**	-0.01	-0.01	0.01	

The conditional wage distribution of Netherlands is given in Table 6.23. We notice from the comparison of conditional and empirical distributions that the conditional distribution explains the change in dispersion from the data quite accurately, if not as well as the case of Germany. The wage inequality trends of the Dutch data we saw in Figure 5.2 are well captured in the conditional distribution.

Table 6.23 Conditional Wage Distribution-Netherlands

Conditional Wage Distribution				
	Waves			
Percentile	**II**	**III**	**IV**	**V**
10	6.798	6.796	6.859	6.874
25	6.945	6.988	7.044	7.080
50	7.091	7.158	7.226	7.246
75	7.263	7.324	7.383	7.423
90	7.425	7.489	7.571	7.628
Dispersion from Conditional Distribution				
Perc.Diff.	**II**	**III**	**IV**	**V**
90-10	0.63	0.69	0.71	0.75
90-50	0.33	0.33	0.35	0.38
50-10	0.29	0.36	0.37	0.37
Dispersion from Empirical Distribution				
Perc.Diff.	**II**	**III**	**IV**	**V**
90-10	0.89	0.96	0.97	1.01
90-50	0.54	0.50	0.49	0.53
50-10	0.36	0.46	0.48	0.48
Portion of Empirical Dispersion Explained by Conditional				
Perc.Diff.	**II**	**III**	**IV**	**V**
90-10	0.70	0.72	0.74	0.75
90-50	0.62	0.66	0.70	0.72
50-10	0.82	0.79	0.77	0.78
Changes in Dispersion (Conditional)				
Perc.Diff.	**III-II**	**IV-III**	**V-IV**	
90-10	0.066	0.018	0.043	
90-50	-0.004	0.014	0.037	
50-10	0.070	0.004	0.006	

Figure 6.2 shows how the conditional distributions for different waves compare. For the Dutch data, we do not observe the radical slope changes that we saw in German data. The Dutch distribution also shifts, but towards higher wages for all quantiles. The shifts favor higher percentiles a bit more than the lower ones.

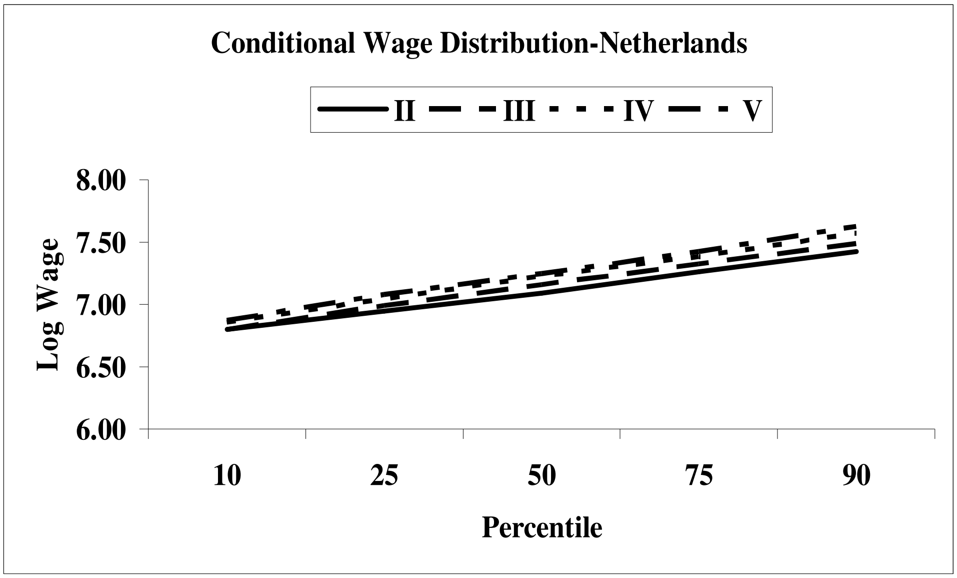


Figure 6.2 Conditional Wage Distributions-Netherlands

Table 6.24 Counterfactual Distributions-Netherlands

Counterfactual Wage Distribution(Wave II Averages)				
	Waves			
Percentile	**II**	**III**	**IV**	**V**
10	6.798	6.781	6.830	6.844
25	6.945	6.962	6.999	7.006
50	7.091	7.128	7.170	7.144
75	7.263	7.286	7.315	7.307
90	7.425	7.465	7.488	7.494
	Dispersion			
Perc.Diff.	**II**	**III**	**IV**	**V**
90-10	0.63	0.68	0.66	0.65
90-50	0.33	0.34	0.32	0.35
50-10	0.29	0.35	0.34	0.30
	Changes in Dispersion			
Perc.Diff.	**III-II**	**IV-III**	**V-IV**	
90-10	0.06	-0.03	-0.01	
90-50	0.00	-0.02	0.03	
50-10	0.05	-0.01	-0.04	

Counterfactual Wage Distribution(Wave II Coefficient)				
	Waves			
Percentile	**II**	**III**	**IV**	**V**
10	6.798	6.838	6.874	6.910
25	6.945	7.004	7.046	7.101
50	7.091	7.160	7.206	7.270
75	7.263	7.323	7.369	7.437
90	7.425	7.461	7.509	7.573
	Dispersion			
Perc.Diff.	**II**	**III**	**IV**	**V**
90-10	0.63	0.62	0.63	0.66
90-50	0.33	0.30	0.30	0.30
50-10	0.29	0.32	0.33	0.36
	Changes in Dispersion			
Perc.Diff.	**III-II**	**IV-III**	**V-IV**	
90-10	0.00	0.01	0.03	
90-50	-0.03	0.00	0.00	
50-10	0.03	0.01	0.03	

The counterfactual distributions for Netherlands are given in Table 6.24. Interestingly, the second counterfactual distribution does a better job of explaining changes in overall dispersion from the conditional distribution from the third wave to the last. The first counterfactual distribution explains about 85% of change in dispersion from Wave II to Wave III but loses its explanatory power after then. Since we use the Wave II averages in the calculation of this distribution, explaining such a big portion of wave II-wave III difference is something that can be expected. We can say that in the later years changes in the composition of personal characteristics become more important, since holding them constant changes the distribution totally.

Table 6.25 Conditional Wage Distribution-Sweden

Conditional Wage Distribution			
	Waves		
Percentile	**III**	**IV**	**V**
10	5.855	5.734	5.547
25	6.024	5.920	5.897
50	6.195	6.085	6.141
75	6.398	6.291	6.380
90	6.603	6.507	6.623

Dispersion from Conditional Distribution			
Perc.Diff.	**III**	**IV**	**V**
90-10	0.75	0.77	1.08
90-50	0.41	0.42	0.48
50-10	0.34	0.35	0.59
Dispersion from Empirical Distribution			
Perc.Diff.	**III**	**IV**	**V**
90-10	0.87	0.87	1.28
90-50	0.50	0.50	0.58
50-10	0.37	0.37	0.70

Portion of Empirical Dispersion Explained by Conditional			
Perc.Diff.	**III**	**IV**	**V**
90-10	0.86	0.89	0.84
90-50	0.82	0.85	0.83
50-10	0.91	0.94	0.84

Changes in Dispersion		
Perc.Diff.	**IV-III**	**V-IV**
90-10	0.025	0.304
90-50	0.014	0.060
50-10	0.011	0.243

The portion of the change in empirical dispersion that is explained by the conditional distribution is quite high for Sweden for all waves. We also notice that the changes in dispersion values represent the increased wage inequality from wave IV to wave V accurately. However, the conditional distributions of Sweden shift in a different way from Germany and Netherlands, as we see in Figure 6.3.

From wave III to wave IV, we see a downward shift, wages for all percentiles decreasing. This is something we have not seen in Germany and Netherlands. Wave V is yet another story, changing the shape of wage distribution so that the upper tail of the wage distribution recovers what it lost in the previous wave. On the other hand, wages of the lower tail drop even further.

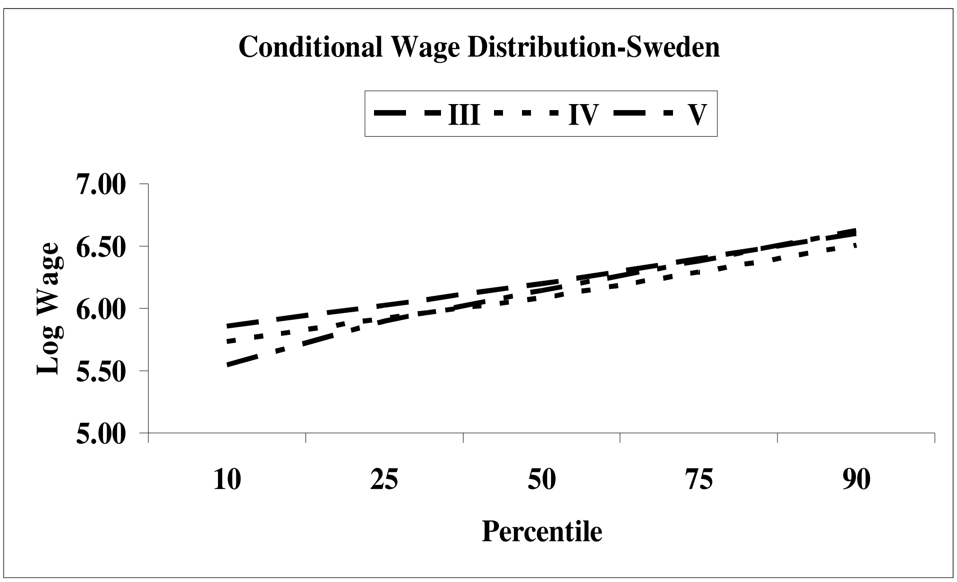


Figure 6.3 Conditional Wage Distributions-Sweden

The counterfactual distributions for Sweden are given in Table 6.26. Although the comparison of the effects of changes in the composition of covariates with that of the prices of them is a bit less meaningful in the case of Sweden, Spain and Hungary due to the shortness of time period, we will still cover that. It appears that the second counterfactual distribution does not have very much explanatory power while the first one explains quite a bit of the change in dispersion. Surprisingly, the first counterfactual distribution does a better job of explaining the change in dispersion from wave IV to wave V than it does with the previous period.

Table 6.26 Counterfactual Distributions-Sweden

Counterfactual Wage Distribution(Wave II Averages			
	Waves		
Percentile	**III**	**IV**	**V**
10	5.855	5.717	5.627
25	6.024	5.906	5.962
50	6.195	6.069	6.201
75	6.398	6.272	6.444
90	6.603	6.484	6.693
	Dispersion		
Perc.Diff.	**III**	**IV**	**V**
90-10	0.75	0.77	1.07
90-50	0.41	0.42	0.49
50-10	0.34	0.35	0.57
	Changes in Dispersion		
Perc.Diff.	**IV-III**	**V-IV**	
90-10	0.02	0.30	
90-50	0.01	0.08	
50-10	0.01	0.22	

Counterfactual Wage Distribution(Wave II Coefficients)			
	Waves		
Percentile	**III**	**IV**	**V**
10	5.855	5.871	5.823
25	6.024	6.038	5.985
50	6.195	6.211	6.153
75	6.398	6.416	6.354
90	6.603	6.624	6.543
	Dispersion		
Perc.Diff.	**III**	**IV**	**V**
90-10	0.75	0.75	0.72
90-50	0.41	0.41	0.39
50-10	0.34	0.34	0.33
	Changes in Dispersion		
Perc.Diff.	**IV-III**	**V-IV**	
90-10	0.00	-0.03	
90-50	0.00	-0.02	
50-10	0.00	-0.01	

The conditional distribution for Spain is reported in Table 6.27. Wages for the Spanish sample seem to be higher than those for other countries. This is probably due to some limitations of the original survey that was given to LIS. Like the earlier countries, the Spanish conditional distribution explains a good deal of wage dispersion from the empirical data, although less so than the Swedish conditional distribution. Changes in dispersion generated from this conditional distribution describes the changes in inequality in Spain adequately.

Table 6.27. Conditional Wage Distribution-Spain

Conditional Wage Distribution

	Waves		
Percentile	**III**	**IV**	**V**
10	10.138	10.294	10.360
25	10.496	10.659	10.708
50	10.705	10.936	10.964
75	10.935	11.185	11.200
90	11.162	11.429	11.467

Dispersion from Conditional Distribution

Perc.Diff.	**III**	**IV**	**V**
90-10	1.02	1.14	1.11
90-50	0.46	0.49	0.50
50-10	0.57	0.64	0.60

Dispersion from Empirical Distribution

Perc.Diff.	**III**	**IV**	**V**
90-10	1.31	1.46	1.31
90-50	0.58	0.72	0.68
50-10	0.72	0.74	0.62

Portion of Empirical Dispersion Explained by Conditional

Perc.Diff.	**III**	**IV**	**V**
90-10	0.78	0.78	0.85
90-50	0.78	0.68	0.74
50-10	0.78	0.87	0.97

Changes in Dispersion(Conditional)

Perc.Diff.	**IV-III**	**V-IV**
90-10	0.112	-0.028
90-50	0.036	0.011
50-10	0.076	-0.038

To see how the conditional distribution evolved in Spain, we look at Figure 6.4. It appears that the conditional distribution jumped up between waves III and IV with minimal change in slope. From wave IV to wave V, the conditional distribution is almost the same, with another and very small shift up. There is no evidence of serious shifts in different parts of the distribution.

The counterfactual distributions for Spain can be seen in Table 6.28. It looks from the table that both composition of covariates and their prices are important in explaining the changes in dispersion in Spain.

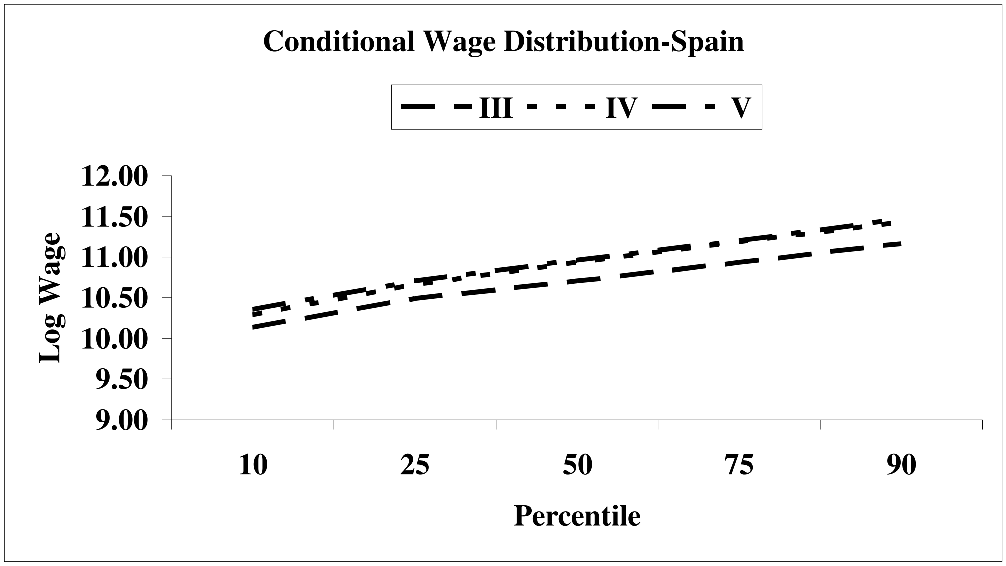


Figure 6.4 Conditional Wage Distributions-Spain

Table 6.28 Counterfactual Distributions-Spain

Counterfactual Wage Distribution(Wave III Averages)			
	Waves		
Percentile	III	IV	V
10	10.138	10.312	10.347
25	10.496	10.671	10.741
50	10.705	10.950	10.994
75	10.935	11.199	11.224
90	11.162	11.442	11.498
	Dispersion		
Perc.Diff.	III	IV	V
90-10	1.02	1.13	1.15
90-50	0.46	0.49	0.50
50-10	0.57	0.64	0.65
	Changes in Dispersion		
Perc.Diff.	IV-III	V-IV	
90-10	0.11	0.02	
90-50	0.03	0.01	
50-10	0.07	0.01	

Counterfactual Wage Distribution(Wave III Coefficients)			
	Waves		
Percentile	III	IV	V
10	10.138	10.138	10.074
25	10.496	10.508	10.471
50	10.705	10.714	10.686
75	10.935	10.930	10.907
90	11.162	11.156	11.134
	Dispersion		
Perc.Diff.	III	IV	V
90-10	1.02	1.02	1.06
90-50	0.46	0.44	0.45
50-10	0.57	0.58	0.61
	Changes in Dispersion		
Perc.Diff.	IV-III	V-IV	
90-10	-0.01	0.04	
90-50	-0.01	0.01	
50-10	0.01	0.04	

Hungarian conditional wage distribution is given in Table 6.29. The conditional distributions give very close dispersion measures to those of the empirical distribution. They also capture the increase in wage inequality from wave III to wave IV as well as anothe slight increase from wave IV to wave V.

Table 6.29 Conditional Wage Distribution-Hungary

Conditional Wage Distribution			
	Waves		
Percentile	**III**	**IV**	**V**
10	8.946	8.752	8.698
25	9.198	9.111	9.056
50	9.468	9.393	9.376
75	9.745	9.691	9.703
90	10.067	10.025	10.033
Dispersion from Conditional Distribution			
Perc.Diff.	**III**	**IV**	**V**
90-10	1.12	1.27	1.33
90-50	0.60	0.63	0.66
50-10	0.52	0.64	0.68
Dispersion from Empirical Distribution			
Perc.Diff.	**III**	**IV**	**V**
90-10	1.28	1.49	1.49
90-50	0.73	0.84	0.75
50-10	0.54	0.65	0.74
Portion of Empirical Dispersion Explained by Conditional			
Perc.Diff.	**III**	**IV**	**V**
90-10	0.88	0.85	0.90
90-50	0.82	0.75	0.88
50-10	0.96	0.99	0.92
Changes in Dispersion			
Perc.Diff.	**IV-III**	**V-IV**	
90-10	0.153	0.062	
90-50	0.034	0.024	

Figure 6.5 shows us the Hungarian conditional distributions as a group. Interestingly enough, the wage inequality increase in Hungary is not reflected by the wages of higher percentiles increasing more than the lower ones. It is marked by a decrease in the wages of lower percentiles while the wages of the upper tail remained more or less the same.

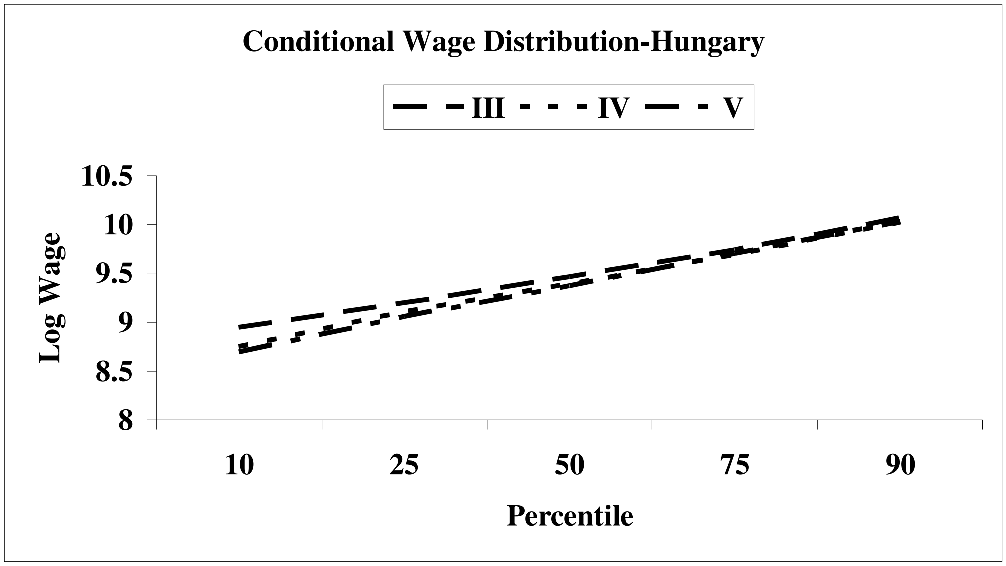


Figure 6.5 Conditional Wage Distributions-Hungary

Table 6.30 Counterfactual Wage Distributions-Hungary

Counterfactual Wage Distribution(Wave III Averages)				Counterfactual Wage Distribution(Wave III Coefficient)			
	Waves				Waves		
Percentile	III	IV	V	**Percentile**	III	IV	V
10	8.946	8.730	8.539	**10**	8.946	8.957	8.968
25	9.198	9.074	8.991	**25**	9.198	9.217	9.226
50	9.468	9.343	9.319	**50**	9.468	9.491	9.500
75	9.745	9.637	9.635	**75**	9.745	9.773	9.783
90	10.067	9.989	9.990	**90**	10.067	10.096	10.118
	Dispersion				Dispersion		
Perc.Diff.	III	IV	V	**Perc.Diff.**	III	IV	V
90-10	1.12	1.26	1.45	**90-10**	1.12	1.14	1.15
90-50	0.60	0.65	0.67	**90-50**	0.60	0.61	0.62
50-10	0.52	0.61	0.78	**50-10**	0.52	0.53	0.53
	Changes in Dispersion				Changes in Dispersion		
Perc.Diff.	IV-III	V-IV		**Perc.Diff.**	IV-III	V-IV	
90-10	0.14	0.19		**90-10**	0.02	0.01	
90-50	0.05	0.02		**90-50**	0.01	0.01	
50-10	0.09	0.17		**50-10**	0.01	0.00	

We can see the counterfactual distributions for Hungary in Table 6.30. Both distributions do a poor job of explaining changes in dispersion from the empirical data. Thus we say that changes in both the composition of personal characteristics and their prices contributed to inequality in Hungary.

6.4 Concluding Remarks

Quantile regression analysis of five European countries reveals, if not anything else, how diverse the wage inequality experience can be among a few countries on the same continent. Due to the nature of quantile regression analysis, we have been able to look into different aspects of wage inequality throughout the wage distribution for these countries. In many cases it is possible to group a few of them under a group, however a general grouping does not seem to be possible since the composition of groups keep changing. Wage inequality experience within a country seem to be able to change direction from wave to wave as well.

Still, there are generalizations one can make out of these different results. For one thing, within inequality seems to be strong to a degree in all countries, and part of it is created by wage dispersion within the same education and experience groups. Secondly, our anaysis with counterfactual distributions lets us divide the countries into two groups in terms of the relative effects of changes in the composition of observable characteristics (covariates) and their prices (quantile regression estimates. In Germany, Netherlands and Sweden it is mostly the changes in prices of covariates that increase the wage inequality, while for Spain and Hungary, both the changes in prices and composition of covariates are effective in creating wage inequality. The fact that this is not obvious from our earlier analysis with JMP might be a good example of the usefulness of quantile regression.

CHAPTER 7

CONCLUSION

The US data shows that wage inequality has been increasing with pauses and jumps in the last three decades. The fast increase during the 80s has been documented in the literature, but we observe a sharp increase in wage inequality in the second half of the nineties as well. There are signs of stability after the turn of the century, but it is still too early to tell if this will be persistent. While the wage inequality in the lower half of the wage distribution was more important in explaining the increase in wage inequality from 70s to early nineties, the upper half of the wage distribution has been gaining more weight since then. Some part of this increase in the wage inequality can be explained by educational differences. For example the college graduates recorded huge gains compared to other groups between 1980 and 1990 throughout the wage distribution. Also, the gains of college graduates are more underlined as one goes higher in experience. The biggest portion of wage inequality is explained by inequality within narrowly defined groups. We observe higher wage inequality within groups of people with similar education and experience levels.

The application of JMP method to the US data revealed that the effect of changes in observable skills on wage inequality has been quite limited since the beginning of 80s. On the other hand, the effect of the prices of these skills (represented by regression estimates) on the year-by year increase of wage inequality has been increasing in the same time period. We also found that the effect of the changes in the composition and prices of unobservable characteristics and skills dominates these two effects combined in causing greater inequality. Inter-industry wage dispersion does not seem to be much of a factor compared to the within-industry wage inequality. The quantile regression results improved our understanding of the behavior of wage inequality in the US by proving that in many cases the marginal effects of covariates differ wildly in the tails from those of OLS. There seems to be sizable wage inequality between college graduates and others, as well as among the college graduates. While the between inequality has been slowing down in the recent years, the within inequality of the college group has been increasing. Experience seems to contribute to both within and between inequalities, but to a lesser degree than education.

The European experience gives us hints on the impossibility of talking about a single European labor market. Countries have very different experiences of wage inequality, which is not surprising if one thinks about the differences of labor market institutions. One thing that is common in all five countries is that they experience some sort of increase in wage inequality between the third and the fourth waves, or at least not experience a decrease (Sweden). This is roughly the first half of the nineties, which is actually one of the time periods that the wage inequality in the US did not increase as fast as it did in other periods. Then, in the second half of the nineties, some record increases in wage inequality (Germany and Sweden, most notably), some see a decrease (Spain) and some stay more or less the same (Hungary and Netherlands). The details of wage inequality are all country-specific with no clear direction that these countries take. However, one could probably say that the theories of European countries not experiencing wage inequality due to their highly regulated labor markets and strong labor unions might not be as strong as they used to be, at least for some countries. Sweden, for example, displays a remarkable increase in the overall wage inequality that bears a heavy within inequality effect in it (also fueled by educational and experience differences to a lesser degree) which concentrated on the lower half of the wage distribution. We also see milder increases in wage inequality in Germany for the whole decade of nineties.

One thing that contradicts in most European countries that we analyzed is that the weight of lower half wage inequality has been increasing since the beginning of 90s in all countries with different rates. The increase is more pronounced in the second half of the decade. This is another difference from the US data in which we saw that the effect of the upper half of wage distribution has been growing in the same period. We have clues that the experience of these countries before the 90s might also be different from the US, since they all start with higher levels of upper half wage inequality than the lower half one. On the other hand, the upper half wage inequality of the US, although it had been closing the gap throughout the nineties, did not catch up with the lower half until the turn of the century.

One has to recognize the importance of this difference since it shows that the reasons for change in wage inequality during the 90s might be different between the two sides of Atlantic Ocean. The change in the American side is probably related to the fact that the skill differences caused by SBTC have reached such a point that the employers are even differentiating between people with higher skills, stretching the wage gap there. On the other hand, the European side has not been able to respond to SBTC effectively due to its institutional rigidities.

REFERENCES

Acemoglu, D. (2002). ”Technical Change, Inequality and the Labor Market”, *Journal of Economic Literature*, 40:7-72.

Acemoglu, D. (2003). ”Cross-Country Inequality Trends”, *The Economic Journal*, 113:F121-F149.

Atack, Bateman and Margo (2004).“Skill Intensity and Rising Wage Dispersion in Nineteenth-Century American Manufacturing”, *The Journal of Economic History* (2004), 64: 172-192

Attanasio, O., P.K. Goldberg, N. Pavcnik (2004). ”Trade Reforms and Wage Inequality in Colombia”, *Journal of Development Economics*, 74(2004) 331-366.

Autor, D.H., L.F. Katz, and M.S. Kearney (2005). ”Trends in U.S. Wage Inequality: Re-Assessing the Revisionist”, NBER Working Paper N.11627.

Autor, D.H., L.F. Katz, and A.B. Krueger (1998). ”Computing Inequality: Have Computers Changed the Labor Market?”, *Quarterly Journal of Economics*, 113 (4): 1169-1213

Autor, D.H., F.Levy, and R.J. Murnane, (2003). ”The Skill Content of Recent Technological Change: An Empirical Investigation”, Quarterly Journal of Economics, 118 (11): 1279-1333

Baker, D., 2001, “Defusing the Baby Boomer Time Bomb: Projections of After-Tax Income in the Twenty-First Century”, International Journal of Health Services, 31(2): 239-278.

Berman, E., J. Bound, and Z. Griliches (1994). “Changes in the Demand for Skilled Labor Within United States Manufacturing: Evidence From the Annual Survey of Manufactures”, *Quarterly Journal of Economics*, 109(2): 367-397

Beyer, H., F. Rojas, and R. Vergara (1999). “Trade Liberalization and Wage Inequality”, *Journal of Development Economics*, 59: 103-123.

Blau, F.D. and L.M. Kahn (1994). “Rising Wage Inequality and the US Gender Gap”, *The American Economic Review*, 84(2):23-28.

Blau, F., L. M. Kahn (1996). “International Differences in Male Wage Inequality: Institutions versus Market Forces”, *The Journal of Political Economy*, 104(4): 791-837.

Bound, J. and G. Johnson (1992). “Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations”, *American Economic Review*, 83: 371-392.

Brainerd, E. (1998). ”Winners and Losers in Russia’s Economic Transition”, *American Economic Review*, 88(5): 1094-1116.

Buchinsky, M. (1994). "Changes in the U.S. Wage Structure1963-1987: Application of Quantile Regression", *Econometrica*, 62(2):405-458.

Buchinsky, M. (1998). “Recent Advances in Quantile Regression Models: A Practical Guide for Empirical Research”, *Journal of Human Resources*, 33:88-126.

Card, D. (2001). ”The Effect of Unions on Wage Inequality in the US Labor Market”, *Industrial and Labor Relations Review*, 54 (2): 296-315

Card, D. and J.E. Dinardo (2002). ”Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles”, *Journal of Labor Economics*, 20(4):733-783.

Dinardo, J., N.M. Fortin, and T. Lemieux (1996)."Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach”. Econometrica, 64(5):1001-1004.

Ebbinghaus, B. and B. Kittel (2005). “European Rigidity versus American Flexibility? The Institutional Adaptability of Collective Bargaining”, *Work and Occupations*, 32(2):163-195.

Fan, C.S. and K. Cheung (2004).”Trade and Wage Inequality: The Hong Kong Case”, *Pacific Economic Review*, 9(2):131-142

Feliciano, Z. M. (2001). “Workers and Trade Liberalization: The Impact of Trade Reforms in Mexico on Wages and Employment”, *Industrial and :Labor Relations Review*, 55(1):95-115.

Galiani, S. and P. Sanguinetti (2003). “The Impact of Trade Liberalization on Wage Inequality: Evidence from Argentina”, *Journal of Development Economics*, 72(2): 497-513.

Goldin, C. nd R.A.Margo (1992). “The Great Compression: The Wage Structure in the United States at Mid- Century”, *Quarterly Journal of Economics*, 107(1):1-34.

Gottschalk, P and T. Smeeding (1997). ”Cross-National Comparisons of Earnings and Income Inequality”, *Journal of Economic Literature*, 35 (2):633-687.

Green, F., A. Dickerson, and J.S. Arbache, (2001). “A Picture of Wage Inequality and the Allocation of Labor Through a Period of Trade Liberalization: The Case of Brazil”. *World Development*, 29(11):1923-1939.

Gregg, P. and A. Manning (1997). ”Skill-Biased Change, Unemployment and Wage Inequality”, *European Economic Review*, 41:1173-1200.

Juhn, C., K. Murphy, and B. Pierce (1991). ”Accounting for the Slowdown in Black-White Wage Convergence”, in *Workers and Their Wages,* ed. by M. Kosters, Washington D.C.: American Enterprise Institute Press, 107-143.

Juhn, C., K. Murphy, and B. Pierce (1993). ”Wage Inequality and the Rise in Returns to Skill”, *Journal of Political Economy*, 101(3): 410-442.

Katz, L.F. and D.H. Autor, (1999). ”Changes in the Wage Structure and Earnings Inequality”, *Handbook of Labor Economics*, Vol.3A.

Katz, L.F. and K. Murphy (1992), “Changes in Relative Wages, 1963-1987: Supply and Demand Factors”, *Quarterly Journal of Economics*, 107(1):35-78.

Kijima, Y. (2005). “Why did Wage Inequality Increase? Evidence from Urban India”, *Journal of Development Economics*, 81(1):97-117.

Koenker, R. And G.Bassett (1978). ”Regression Quantiles”, *Econometrica*, 46(1): 33-50.

Koenker, R. And K.F.Hallock (2000). “Quantile Regression: An Introduction”, University of Illinois Working Paper, December.

Koenker, R. And K.F.Hallock (2001). “Quantile Regression”, *Journal of Economic Perspectives,* 15(4):143-156.

Leamer, E.E. (1992. ”Wage Effects of a U.S.-Mexican Free Trade Arrangement”, NBER Working Paper, No.3991.

Lee, D.S. (1999). “Wage Inequality in the US during the 1980s: Rising Dispersion or Falling Minimum Wage?”,*Quarterly Journal of Economics*, 114(3):941-1023.

Levy, F. and R. Murnane (1992). “US Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations”, Journal of Economic Literature, 30:1331-1381.

Lin, C.A. and P.Orazem, (2004). ”A Reemxamination of the Time Path of Wage Differentials in Taiwan” ,*Review of Development Economics*, 8(2):295-308.

Machado, J.A.F and J. Mata (2001). “Earning Functions in Portugal 1982-84: Evidence from Quantile Regression”, *Empirical Economics*, 26:115-134.

Margo, R.A (1995). “Explaining Black-White Wage Convergence, 1940-1950: The Role of the Great Compression”, *Industrial and Labor Relations Review*, 48(3): 470-481.

Martins, P.S., Pereira, P.T. (2004). ”Does Education Reduce Wage Inequality? Quantile Regression Evidence From 16 Countries”, *Labour Economics*, 11:355-371

Mincer, J.(1974). “Schooling, Experience and Earnings”, New York: Columbia University Press.

Morris, M.A., and B. Western (1999). “Inequality in Earnings at the Close of the Twentieth Century”, *Annual Review of Sociology*, 25 (1): 623-657.

Murphy, K.M. and F. Welch (1992). ”The Structure of Wages”, *Quarterly Journal of Economics*, 107: 285-326.

Nickel, S. and B. Bell (1996). ”The Collapse in Demand for the Unskilled and Unemployment across the OECD”,*Oxford Review of Economic Policy*, 11: 40-62.

Oaxaca, R. (1973). “Male-Female Wage Differentials in Urban Labor Markets*”, International Economic Review*, 14: 693-709.

Park, J.H. (1994). “Estimation of Sheepskin Effects and Returns to Schooling Using the Old and the New CPS Measures of Educational Attainment”, Princeton University International Relations Section Working Paper, No.338.

Prasad, E. (2004). "The Unbearable Stability of the German Wage Structure: Evidence and Interpretation", IMF Staff Papers, 51(2):354-385.

Sapir, A. (2006). "Globalization and the Reform of European Social Models", Journal of Common Market Studies, 44(2):

Siebert, H. (1997). "Labor Market Rigidities: At the Root of Unemployment in Europe", *Journal of Economic Perspectives*, 11(3):37-54.

Skoufias, E. and A. Suryahadi (2002). "A Cohort Analysis of Wages in Indonesia", *Applied Economics*, 34(13): 1703-1710.

Tansel, A. And F. Bircan (2006). "Education and Wage Inequality in Turkey, 1994-2002: A Quantile Regression Analysis", Paper presented at the International Conference of Turkish Economists Association (TEK) in Ankara, 11-13 September, 2006.

Teulings, C.N (2003). "The Contribution of Minimum Wages to Increasing Wage Inequality", *Economic Journal*, 113(490): 801-834.

Tsou, Meng-Wen (2002). "Wage Differentials in Taiwanese Manufacturing, 1982-1997", *Asian Economic Journal*, 16 (4).

Wilthagen, T. and F. Tros (2004). "The concept of 'flexicurity': A new approach to regulating employment and labour markets", *Transfer – European Review of Labour and Research*, 10(2):166-187.

Wood, A. (1995). "How trade hurt unskilled workers", *Journal of Economic Perspectives*, 9(3): 57-80.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Yağanoğlu, Nazmi Yükselen
Nationality: Turkish (TC)
Date and Place of Birth: 9 October 1970 , Sivas
Marital Status: Single
Phone: +90 312 447 42 84 – 0535 639 3051(Mobile)
email: nazmi@metu.edu.tr

EDUCATION

Degree	Institution	Year of Graduation
MS	Boston University Economics	1998
BS	METU Economics	1992
High School	Kayseri Lisesi, Kayseri	1987

WORK EXPERIENCE

Year	Place	Enrollment
November 2000-present	METU Department of Economics	Research Assistant
June 2000-November 2000	Çanakkale 18 Mart University Department of Economics	Research Assistant
1992-1996	T.C.Ziraat Bankası	Expert (Corporate Banking)

FOREIGN LANGUAGES

Advanced English