Working papers series

WP ECON 08.08

Wage Inequality in Spain, 1980-2000

Manuel A. Hidalgo (U. Pablo de Olavide)

JEL Classification numbers: J24, J31

Keywords: Wage inequality, counterfactuals, human capital.

Department of Economics

Wage Inequality in Spain 1980-2000

Manuel A. Hidalgo *

15th April 2008

Abstract

We use recent developments in quantile regression to simulate counterfactuals densities that allows to decompose the Spanish wage inequality evolution over the 1980-2000 period between changes due to observable prices, labour market composition and non observable characteristics' prices. Our empirical results are threefold: first of all, the wage inequality decreases during each decade first half and increases during the second ones, second both changes in prices and composition has an important role in this evolution and third, changes in observable prices mirrors this behavior above and below the median while non observable inequality increases since 1985 onwards below the median, with an erratic trend above. Finally some tentative explanations are give that could reasonably explain our findings.

JEL codes: J24, J31

1 Introduction

Wage inequality has grown substantially since the 1970s in many countries. In the United States, for example, the 90/10 percentile ratio of male hourly earnings grew by 23.4 log points between 1979 and 2003 (Autor, Katz, and Kearney 2005a). This ratio increased by about 20 log points in the UK between the early 1980s and late 1990s, with similar increases taking place in Germany between the early 1980s and the mid-1990s. In Canada during the 1990s, the ratio increased by 15 log points. These figures contrast with slight increases or

*U. Pablo de Olavide. I am grateful to seminar participants of the III Jornadas de Economía Aplicada organized at the University of Granada (Área de Fundamentos del Análisis Económico Dpto. Teoria e Historia Economica). I especially thank Antonio Ciccone, Florentino Felgueroso, José Ignacio García, Julian Messina, Sara de la Rica, Jesus Rodriguez and Diego Romero for very useful suggestions. I gratefully acknowledge support of project 30.04.01.27.27 (Ensayos sobre Andalucía y Nueva Economía). Ctra de Utrera, s/n, 41013, Seville, Spain. Email: mhidper@upo.es.

even declines in other countries such as the Netherlands, Sweden or Belgium, as summarized by Acemoglu (2003).

This paper aims to provide a detailed analysis of wage inequality trends in Spain between 1980 and 2000s, providing a level of detail similar to that found in analyses for other countries. Specifically, it will examine the trends in overall wage inequality and decompose these trends into three components, corresponding to three types of inequality: that which responds to changes in the wage premium paid for education and experience (between-group inequality, BI); that which responds to the observed skill distribution (composition effects); and that which responds to changes in wage dispersion among workers with the same levels of experience and education (residual wage inequality, RWI).

The literature on this subject mentions a number of approaches to the decomposition of changes in overall wage inequality into between-group, composition effects, and residual wage inequality. Juhn, Murphy, and Pierce (1993)'s seminal study extended the "mean" Oaxaca-Blinder procedure to include the decomposition of distributions. In applying this approach, Juhn, Murphy and Pierce (1993) first estimated yearly Mincerian wage equations to obtain returns for education and experience, as well as a measure of the residual wage distribution (wage dispersion that cannot be explained by differences in education, experience or other observable worker characteristics). They then used the results to simulate counterfactual densities (the wage distributions that would prevail if some parameters, such as the return to education, were changed but the rest remained constant). Lemieux's (2002, 2006) decomposition of changes into overall wage inequality extends the kernel procedure of DiNardo, Fortin, and Lemieux (1996). An advantage of his approach is that it allows residual wage inequality to be a function of workers' observable characteristics. Finally, Autor, Katz, and Kearney (2005a) extend Machado and Mata's (2005) quantile decomposition technique. We have chosen to rely on the latter approach, because it incorporates the work of both Juhn, Murphy and Pierce and that of DiNardo, Fortin and Lemieux.

Our analysis uses data from the Household Budgets Surveys for 1980-81 and 1990-91 and the Continuous Household Budgets Surveys for 1985-86, 1990-91, 1995-96 and 2000-01. These are the only data sources that contain all of the information needed for this study and that cover the entire period of interest to us here. By contrast, other data sources either cover shorter time periods or lack key information on individual workers.

We find that overall wage inequality in Spain, measured as the difference between log wages at the 90th and 10th distribution percentiles, increased only very slightly between 1980 and 2000. More specifically, we find a decrease in wage inequality during the 1980s, compensated for by an increase during the 1990s. Within each of these decades, wage inequality initially increased and then decreased during the later years of the decade. Given these results, it is interesting to note that GDP grew below trend during the early 1980s and the 1990s and above trend during the final years of both decades. Hence, overall wage inequality behaved countercyclically within each of the two decades.[1] For the 1980s, our findings are on a par with those of earlier studies, published during the 1980s. Arellano, Bentolila, and Bover (2001), for example, used data from the Social Security archives to show that wage dispersion rose during the first half of the 1980s, while Abadie (1997) demonstrated that wage inequality fell slightly throughout the 1980s on the basis of data from the Household Budget Survey for 1980-81 and 1990-91.

We also decompose changes in wage inequality into between-group inequality (the distance between distributions conditional to one observable characteristic), residual inequality (inequality within conditional distributions), and composition effects (changes in wage inequality that arise due as the result of a shift in labor force composition). Between-group inequality behaves much like overall inequality. This is true for the entire 20-year period, as well as for and within each of the two decades encompassed by that period. The composition effect contributes positively to overall wage inequality until the late 1990s. Residual inequality, on the other hand, only begins to increase during the early 1990s.

One advantage of our decomposition method is that it allows us to examine wage inequality trends at different wage distribution points. In particular, we look at wage gaps situated in the lower part of the distribution, between the 50th and the 10th percentiles (where the gap is 50/10), and in the upper part of the distribution between the 90th and the 50th percentiles (where the gap is 90/50). Regarding overall inequality, we find that the 90/50 wage gap behaves qualitatively like the 90/10 gap. The 50/10 wage gap, on the other hand, has been increasing since the second half of the 1980s.

A close look at these figures reveal that between-group inequality evolves similarly in the upper and the lower part of the wage distribution; thus, both the 50/10 and the 90/50 gaps behave qualitatively like the 90/10 gap with respect to between-group inequality. At a quantitative level, however, the changes are more marked in the upper half of the distribution. By contrast, residual inequality behaves rather differently when we look at the upper half instead of the lower half of the distribution. While the 50/10 gap has been increasing since the second half of the 1980s, there is no clear pattern to be found with respect to the 90/50 gap. Hence, in the upper part of the distribution the rising overall wage inequality shown is mirrored by rising between-group inequality, but not by residual inequality. In the lower part of the distribution, on the other hand, wage inequality behaves much like residual inequality but not like between-

[1]For a discussion of the link between business cycle and wage inequality in the US and the U.K. see Dimelis and Alexandra (1999).

group inequality. The different patterns of wage inequality above and below the median, and the behavioral differences that can be observed between the different components of our analysis, suggest that there was no unique driving force behind inequality trends in Spain during the period under study.[2]

The rest of the paper is organized as follows: Section 2 reviews the related literature, Section 3 describes our data, Section 4 gives a preliminary survey of wage inequality between 1980 and 2000, Section 5 explains our decomposition method, Section 6 discusses our main empirical results, and Section 7 concludes.

2 Related Literature

Wage inequality and the decomposition of overall wage inequality into between-group inequality (BI, hereafter), composition effects, and residual wage inequality (RWI) has generated a large body of research to date. Most of the literature uses one of three methods: the full variance accounting method presented by Juhn, Murphy and Pierce (1993); the semi-parametric procedure of DiNardo, Fortin, and Lemieux (1996); or the Mata and Machado (2005) approach as extended by Autor, Katz, and Kearney (2005a).

The full variance accounting decomposition put forth by Juhn, Murphy and Pierce (1993) (JMP) begins with the supposition that wages can be characterized by the canonical Mincer equation

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

where w_{it} denotes the wage logs at time t for individual i, X_{it} is a specific set of individuals and environmental characteristics that may potentially affect wages, β_t is a set of returns or prices that set the value of these characteristics, and u_{it} is a compendium of non-observable characteristics that may potentially affect individual wages.

Within this framework, wage inequality changes may come from three sources: changes in the distribution of observable characteristics (changes in the distribution of X_{it}); changes in the prices of observable skills (changes in β_t); and changes in the wage distribution accounted for by unobservable characteristics (u_{it}). With this structure, JMP estimate the counterfactual densities that would prevail if any subset of these three components were held to be constant, finding that residual inequality accounted for a great deal of the overall increase in U.S. wage inequality between 1964 and 1988. They also demonstrate that nearly all of the rising inequality explained by observables is due to changes in returns on observable skills (rather than to changes in the distribution of observable skills). While they find that residual wage inequality began to increase during

[2] Autor, Katz, and Kearney (2005a) reach the same conclusion for the US and Abadie (1997) for Spain in the 1980s.

the late 1960s, they also find that increasing returns on skills have contributed to a higher overall wage inequality only since the early 1980s. Their favorite explanation for these wage inequality trends is that the demand for both observable and unobservable skills has been increasing, and that this demand has translated into higher skill prices.[3]

A shortcoming of this approach is that it does not fully account for the links between observable characteristics and residual wage dispersion, as pointed out by Lemieux (2006).[4] For example, hourly wage dispersion in the U.S. is typically found to be greater for college graduates than for less educated workers (Autor, Katz and Kearny, 2005a, 2005b). As a result, a rise in the number of college degree-holding workers may cause overall wage inequality to increase as the result of a "mechanical" composition effect. Such composition effects working through residual wage dispersions must be accounted for, since they might easily be confused with the effects of changing prices on unobservable skills (Lemieux, 2006 and Autor, Katz and Kearny, 2005a, 2005b). Lemieux's (2002,2006) extension of the approach developed by DiNardo, Fortin and Lemieux (1996) was designed to resolve this problem. Lemieux modelled overall residual wage dispersion as a weighted average of wage dispersions by skill group. Under this model, changes in the distribution of observable skills will also change the balance of wage dispersion by skill group, thereby mechanically changing the overall wage dispersion. Lemieux (2006), who uses this approach in order to examine wage inequality trends in the U.S., finds that composition effects play a greater role and changes in RWI play a smaller role than they do in JMP (composition effect now encompasses all changes in wage inequality caused by changes in skill composition, including those caused by re-weighted residual wage dispersions by skill; RWI are changes in wage inequality that are unaccounted for by composition effects and BI). He also argues that a great many of the changes in RWI can be explained by institutional factors (rather than by changes in skill prices), such as the declining real minimum wage.

Our empirical work will rely on the latest decomposition method appearing in the literature on this subject, Autor, Katz and Kearny's (2005a) (AKK) extension of the Machado and Mata (2005) (MM) quantile decomposition approach, which is explained in detail in Section 5. This method corrects the shortcomings of the original full distribution accounting method developed by Juhn, Murphy, and Pierce (1993), and nests the approach proposed by DiNardo, Fortin and Lemieux (1996) and Lemieux (2002,2006). As applied to U.S. data (Autor, Katz and Kearny, 2005a, 2005b), the AKK approach confirms the importance of the composition effects emphasized by Lemieux (2006) for the lower part of the wage distribution. But in the upper part of the wage distribution,

[3]An important open question is why BI and RWI started rising at different times.

[4]In principle, the approach of JMP should be able to deal with this issue. But Lemieux (2006) argues that it does not do so in practice.

rising wage inequality is found to be almost entirely explained by rising prices for observable skills and by greater RWI (the two main driving forces behind wage inequality emphasized by JMP).

Wage inequality in Spain has been studied by Abadie (1997), Arellano, Bentolila, and Bover (2001) and Izquierdo and Lacuesta (2007). Abadie's analysis examines wage inequality trends during the 1980s using quantile regressions. He documents a fall in the return to education during this period, which mostly affects the lower part of the distribution for younger workers and the upper part for elderly workers. Our approach differs from his in that it allows for a detailed characterization of composition effects as well as for wage inequality trends within and between groups. Arellano, Bentolila, and Bover use a large Social Security data sample to examine wage inequality trends for the 1980-1987 period. Their analysis focuses on the behavior of returns to skill and experience both over time and across sectors. Izquierdo and Lacuesta (2007) use non-parametric techniques to analyze Spanish wage inequality between 1995 and 2002 using the Wage Structure Survey. They show that changes in the return to education and tenure decreased inequality, while changes in composition increased the inequality. Our approach differs from theirs with respect to period of analysis, method, and the special attention paid in our study to within-group inequality.

3 Data

Let us now review the data sets that can be used to analyze wage inequality trends in Spain.

Household surveys. All of the information necessary for our analysis can be found in the "Encuesta de Presupuestos Familiares" or EPF (Household Budget Survey) for 1980-81 and 1990-91 and its newer counterpart, the quarterly Encuesta Continua de Presupuestos Familiares or ECPF (Continuous Household Budget Survey), available from 1985 to 2005 data sets. While they provide useful data on wages, education, age, gender, such information is only available for heads of families. Nevertheless, the alternatives (some of which are explained below) present even greater drawbacks, making these surveys the main source of information for the study of wage inequality trends (Oliver, Raymond, Roig, and Barceinas, 1999). Appendix A gives more details on these surveys.

Other data sources. There are two other of wage surveys in Spain, the Wage Structure Survey and the (Quarterly) Labour Cost Survey, both compiled by the Spanish National Institute of Statistics (*Instituto Nacional de Estadistica* or INE). The first of these provides information for about two thousand industrial and service workers for the years 1995, 2002 and, recently, 2006. While the individual information in this survey is very detailed, the time span covered is too short to be useful here. The (Quarterly) Labour Cost Survey (previously called the Survey of Wages in Industry and Services) surveys wages for the 1980-2000 period but provides no information on education levels.[5]

The Spanish Social Security records provide another source of information on wages. The large sample size and the length of the period covered make this an appealing data set for those analyzing changes in wage inequality over time, (see Arellano, Bentolila and Bover, 2001). Recently, the "Muestra Contínua de Vidas Laborales" from the Social Security records further develops the information used by these authors, covering a panel of workers and an expanded number of years and providing more data on individual worker characteristics. Also, for 2004 and 2006 this new survey has information on fiscal variables. The drawback of using these records is that, with the exception of 2004 and 2006, wages are not directly stated but have to be inferred from social security contributions.

4 Descriptive Statistics and Wage Inequality

Our analysis of wage inequality trends in Spain is based on the data for heads of household who work full-time provided in the EPF surveys for 1980-81 and 1990-91, and in the ECPF for 1985-1986, 1990-1991, 1995-1996 and 2000-2001. The reasons for this selection are described in Appendix A. One drawback of this selection is the smaller female participation as compared with other exercises. For comparative purposes, some of the consequences of this lower participation will be pointed out when we present our results.

Some important characteristics of these data are reported in Table 1. It can be seen that individual yearly earnings, in 2000 pesetas, rose from 2,172,996 in 1980 to 2,513,765 in 2000. This corresponds to a real growth rate of about 15.6% over 20 years. The wage dispersion, measured as the standard deviation of wages, also grew between 1980 and 2000. Within the 20-year period, real average wages grew unevenly. Between 1980 and 1985 real wages increased only minimal; however, they increased substantially between 1985 and 1990. Real wage growth during the 1990s was more even, however. Other Spanish wage statistics for the same period yield similar results (Survey of Wages in Industry and Services or National Accounts).

Table 1 also shows that the average age of workers oscillated between 42 and 43 years during the sample period. Average years of schooling, on the other hand, rose by almost 50 percent between 1980 and 2000 (from 8.63 years to 12.45 years). While schooling rose in nearly every country in the world during

[5]Moreover, sample selection is at the firm level, which implies that the sample may not be representative of Spanish workers.

Table 1: EPF-80/81-90/91 and ECPF 85/86-90/91-95/96-00/01 main characteristics

	EPF		ECPF			
	80/81	90/91	85/86	90/91	95/96	00/01
number	7,027	8,193	2,683	1,965	1,696	2,057
av. wage (c.p. 2000)	2,172,996	2,282,235	2,165,996	2,350,950	2,430,859	2,513,765
sd. dv.	252,818	259,729	233,101	270,562	277,327	310,028
wage growth rate		0.49		1.65	0.67	0.67
age	42.02	42.93	41.58	41.83	41.99	43.60
schooling	8.63	9.28	10.49	11.02	11.69	12.45
women (%)	6.06	9.37	5.14	6.36	9.35	11.74

Notes: The row *av. wage (c.p. 2000)* shows average wages for each surveys at 2000 constant pesetas. *Sd. Dv.* are standard deviations. *Wage growth rate* gives the average growth rate per year between the year corresponding to the column the data and the previous one. The rows *age* and *schooling* present the average age and years of schooling for each survey. Lastly, *women* are the women participation in each sample.

this period, this rise was much more pronounced in Spain than it was in most other countries (Acemoglu 2003). The percentage of female heads of household who work full-time has also grown since the EPF80/81.

Some trends in wage inequality can be analyzed without a full decomposition approach. Table 2 plots the difference between log wages at the 90th and 10th percentiles of the distribution (the 90/10 log wag gap), between log wages at the 90th and 50th percentiles (the 90/50 log wag gap), and between log wages at the 50th and 10th percentiles (the 50/10 log wag gap). It can be seen that between 1980 and 1990, the 90/10 log wage gap fell by about 15 log points (average real wages at the 90th percentile rose by only 1.8% per year during this period, while at the 10th percentile they registered an increase of 3.3% per year). Between 1985 and 2000, however, 90/10 wage inequality grew by 36 points. Inequality trends for the 1980s differ qualitatively in the lower and upper halves of the distribution. Looking at the lower tail, it can be seen that the 50/10 log wage gap fell by 29 points. By contrast, the 90/50 gap increased by 14 points, indicating greater wage inequality in the upper half of the distribution. During the 1990s, the rise in inequality was concentrated entirely in the lower tail, with a sharp 55 log-point rise rise at the 50/10 gap. Inequality in the upper tail decreased somewhat during this period.

Analyzing different subperiods and parts of the wage distribution yields further interesting results. First, the lower half seems to present increasing inequality between 1985 and 1990. The distance in log wages between the 10th and 50th percentiles falls between 1980 and 1990 but remains constant between 1985 and 1990. This gap also increases between 1990 and 1995, and between 1995 and 2000. However, the upper half shows a different trend for the entire twenty-year period. For each decade, the 90/50 log wage gap increased during the early years when GDP growth was low, but fell during the later years, when

GDP grew more rapidly.[6] Wage inequality above the median thus seems to follow a countercyclical pattern. This result is important enough to merit further attention in the following sections.

Wage inequality trends may be driven by changes in between-group or within-group wage dispersion, or by changes in the composition of the labor force. As a preliminary means of isolating changes in between-group inequality due to changes in education premia, we estimate Mincerian wage regressions. Table 3 shows different measures of the return to education, all obtained using Mincerian wage equations. All premia/penalties are obtained using Mincerian wage regressions that include age, age squared, and gender as additional explanatory variables.[7] In particular, this Table shows education wage premia earned by college-educated workers relative to those with only a secondary education ("college"), and also the wage penalty for primary-schooled workers relative secondary-schooled ones ("primary"). The results show that wage differentials between primary- and secondary-schooled workers narrowed continuously between 1980 and 2000. This tendency should have reduced wage inequality. By contrast, the college/high school wage premium evidenced a more uneven trend. It grew between 1980 and 1985 and also between 1990 and 1995, but fell during the remaining periods. Hence, the college-high school premium seems to follow a counter-cyclical pattern.

The bottom row of Table 3 shows another standard statistic of the return to schooling, the average return to an additional year of schooling.[8] This statistic shows a fall in the return to schooling during the 1980s and a modest increase from 1990 onwards.

Age/experience contributed to increasing wage inequality between 1980 and 1995. The return to experience shows clearly the increase in the inter-group wage gap defined by this characteristic. During the 1980s there was a sharp increase of about 30%. This trend remains roughly similar, but slightly lower, before breaking in 1995. The last five years see a contraction in BI inequality given by age/experience.

[6]Spanish real GDP growth was 1.5%, between 1980 and 1985, 4.5% between 1985 and 1990, 1.3% between 1990 and 1995, and 3.9% between 1995 and 2000.

[7]These results are derived from a standard Mincer equation that takes the following form:

$$w_{it} = \alpha_t + \beta_{1t} p_{it} + \beta_{2t} c_{it} + \gamma_{1t} age_{it} + \gamma_{2t} age_{it}^2 + \rho_t sex_{it} + e_{it}$$

where w_{it} is the log of individual wage, (p_t, c_t) is the vector of educational dummies, $(age_{it}, age_{it}^2, sex_{it})$ is the vector of other wage-influencing variables and $(\alpha_t, \beta_{1t}, \beta_{2t}, \gamma_{1t}, \gamma_{2t}, \rho_t)$ are the returns on wages for these variables.

[8]In this case, instead of introducing dummies to capture the school effect on wages, here we used years of schooling. Then, the estimated mincerian equation is defined by

$$w_{it} = \alpha_t + \beta_t s_{it} + \gamma_{1t} age_{it} + \gamma_{2t} age_{it}^2 + \rho_t sex_{it} + e_{it}$$

where, in this case s_{it} gives the years of schooling of individual i.

Table 2: Wage Inequality Change 1980-2000

	EPF		ECPF			
	1980/81	1990/91	1985/86	1990/91	1995/96	2000/01
90th-10th	0,926	0,911	0,966	0,960	1,003	1,002
90th-50th	0,520	0,534	0,543	0,537	0,541	0,526
50th-10th	0,406	0,377	0,422	0,422	0,462	0,477

Note: Each row represent distances in logs between the wage distribution percentiles represented at the first column.

Table 3: Between-group and residual inequality. Mincerian Wage regressions

	EPF		ECPF			
	1980/81	1990/91	1985/86	1990/91	1995/96	2000/01
primary	-0.311	-0.282	-0.286	-0.223	-0.212	-0.202
	(0.008)	(0.008)	(0.014)	(0.015)	(0.018)	(0.018)
college	0.266	0.283	0.308	0.2961	0.3526	0.316
	(0.011)	(0.011)	(0.021)	(0.023)	(0.024)	(0.02)
age	0.037	0.048	0.055	0.059	0.064	0.035
	(0.002)	(0.003)	(0.004)	(0.005)	(0.006)	(0.007)
res 90-10	0.778	0.714	0.773	0.778	0.816	0.828
res 90-50	0.425	0.394	0.417	0.404	0.419	0.420
res 50-10	0.354	0.319	0.356	0.375	0.398	0.407
R^2	0.29	0.35	0.32	0.280	0.29	0.220
years of sch.	0.063	0.060	0.079	0.064	0.075	0.070
	(0.001)	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)

Note: Each three first rows are the standard Mincerian regression coefficients for education (represented by two dummies, the first if the worker has only primary or less education and second if the worker has at least college education) and age (square of age and gender are not showed). The following three rows are the distance in the 90, 50 and 10 residuals percentiles which evaluates changes in non-observed, residual or within-group inequality. R^2 show the goodness of fit, and years of schooling is the return to education if we use average years of schooling instead of the previous dummies in the Mincerian regression.

Table 3 also shows statistics for residual wage inequality. For now, we simply evaluate changes in the RWI using the distribution of OLS residuals. The analysis of RWI is important because it explains about two thirds of the overall inequality in our data (this number is a quite standard finding using the Mincerian approach). The rows in Table 3 reveal a contrast before and after 1990. The 1980s' trend was negative, despite the increase in the second half of that decade. This was only possible thanks to the considerable decrease in residual wage inequality that took place during the first half of the decade. Thus, the increase in wage inequality since 1985 appears to be due to residual inequality.

Figure 1 summarizes the trends in 90/10, the college wage premium, the primary wage penalty, and residual inequality (the data come from Table 3). Differences in wage inequality trends between decades are evident. The graph shows clearly that wage inequality fell during the 1980s, widened between 1990 and 1995, and remained constant until 2000. The college premium rose until the last period (1995-2000). Changes in the primary wage penalty tended to reduce inequality up to the year 2000. Thus, up to the year 1995, education reduced inequality for less educated workers - those in the lower tail of the wage distribution- and increased inequality in the upper tail.

Overall, residual inequality appears to have run parallel to wage inequality, except during the period between 1985 and 1990 when residual inequality increased while total inequality fell. During 1985-1990, changes in wage inequality seem to mirror changes in BI.

Our analysis so far does not allow us to assess the role of changes in labor force composition for changes in wage inequality. As a simple, preliminary way to examine this issue, we estimated Table 3 again, but holding the labor force composition constant at its 1980 value. To do so, workers in each survey year were classified by level of education (primary, high college and college), age (one-year intervals from 20 to 65 years), and gender, which yielded 276 cells. We then generated counterfactual samples for each survey year by combining average worker characteristics in each cell with 1980 sample weights. Estimating Table 3 using these counterfactual samples yields the results in Table 4. It can be seen that differences with Table 3 are very small, which suggest that labor force composition did not play a major role in Spanish wage inequality trends.

The main limitation of our preliminary analysis (based on OLS estimation) is that it does not allow for differences in the evolution of prices at different parts of the wage distribution. Similarly, it does not show how changes in the labor force composition affect wages at different parts of the wage distribution. Moreover, our approach may have overestimated the degree to which residual inequality contributes to changes in wage inequality by not fully accounting for links between observable characteristics and residual wage dispersion, as emphasized by Lemieux (2006). The following section explains a quantile regression

Table 4: **Between-group and residual inequality. Minncerian Wage regressions (1980 constant weights)**

	EPF		ECPF			
	1980/81	1990/91	1985/86	1990/91	1995/96	2000/01
Primary	-0.311 (0.008)	-0.261 (0.012)	-0.271 (0.017)	-0.206 (0.020)	-0.186 (0.025)	-0.186 (0.018)
College	0.266 (0.011)	0.295 (0.015)	0.338 (0.032)	0.308 (0.030)	0.357 (0.034)	0.295 (0.02)
Age	0.037 (0.002)	0.051 (0.003)	0.056 (0.005)	0.063 (0.005)	0.072 (0.011)	0.039 (0.009)
Res 90-10	0.778	0.713	0.777	0.782	0.819	0.842
Res 90-50	0.425	0.391	0.423	0.406	0.428	0.424
Res 50-10	0.354	0.322	0.354	0.376	0.390	0.417
R^2	0.29	0.35	0.32	0.280	0.29	0.220

Note: Each three first rows are the standard Mincerian regression coefficients for education (represented by two dummies, the first if the worker has only primary or less education and second if the worker has at least college education) and age (square of age and gender are not showed). The following three rows are the distance in the 90, 50 and 10 residuals percentiles which evaluates changes in non-observed, residual or within-group inequality. R^2 show the goodness of fit, and finally years of schooling is the return to education if we use average years of schooling instead of the previous dummies in the Mincerian regression.
Constant weights implies that each regression are made using 1980 population structure for observable variables.

technique that can be used to resolve these issues.

5 Methodology

The decomposition technique proposed by Machado and Mata (2005,MM) and extended by Autor, Katz and Kearny (2005a,AKK) uses quantile regressions to decompose wage distributions into "price"and "quantity"components. These components are then used to assess the importance of changes in prices and quantities in explaining wage inequality trends using counterfactual analysis. Our exposition of the MM and AKK methodology follows AKK.

Let $Q_\theta(w_t|x_t)$ for $\theta \in (0,1)$ be log wages (w_t) at the θ^{th} quantile of the distribution of wages given the vector of k covariates x_t for year t. Assume that the conditional quantiles can be represented as a linear function of covariates, or more formally, that there are $k \times 1$ vectors of quantile regression coefficients $\beta_t(\theta)$ such that

$$Q_\theta(w_t|x_t) = x_t'\beta_t(\theta). \tag{2}$$

If $Q_\theta(w_t|x_t)$ in (2) is specified correctly, $x_t'\beta_t(\theta)$ provides a full characterization of the distribution of wages given the covariates x_t. The distribution of w_t given x_t can therefore be obtained by (i) repeatedly drawing θ from a uniform distribution on the open interval $(0,1)$; (ii) obtaining the price vector $\beta_t(\theta)$

corresponding to θ; (iii) calculating $x_t'\beta_t(\theta)$. In general, the specification in (2) must be thought of as an approximation. How good the approximation turns out to be is easy to check and depends on the particular application. For the case of Spain, we will show that the resulting approximation is quite accurate.

As is well known, for a given θ, the vector $\beta_t(\theta)$ can be estimated by solving the following minimization problem,

$$\min_{\beta} \quad n^{-1}\sum_{i=1}^{n}\rho_\theta(w_{it} - x_{it}'\beta_t); \tag{3}$$

where the ρ_θ is a "check function" (Koenker and Bassett 1978),

$$\rho_\theta(u) \equiv \begin{cases} \theta u & \text{for } u \geq 0 \\ (\theta - 1)u & \text{for } u < 0; \end{cases}$$

where in this case $u = w_{it} - x_{it}'\beta_t$. This method estimates $\beta_t(\theta)$ consistently under conditions similar to those required for the asymptotic consistency of OLS estimation.

Once $\beta_t(\theta)$ has been estimated for $\theta_i, i = 1, ...m$ with a large value for m spread over the open unit interval, the distribution of wages given x_t is obtained as $\{\hat{w}_{it} = x_{it}'\hat{\beta}_t(\theta_i)\}_{i=1}^{m}$, where hats denotes estimated values.

This describes the simulation of the wage data for any given x, but does not provide the marginal density of w. The marginal density also depends on the distribution of the covariates, which we will denote by $g(x)$. The marginal density of w is obtained by (i) repeatedly drawing rows of data from $g(x)$, x_i; (ii) drawing corresponding θ_i from a uniform (0,1) distribution; (iii) obtaining wages as $\hat{w}_i \equiv x_i'\hat{\beta}(\theta_i)$.

The MM conditional quantile decomposition procedure has two important properties. First, like the Oaxaca-Blinder OLS procedure (Oaxaca, 1973; Blinder, 1973), it separates the observed wage distribution into price and quantity components. But while the Oaxaca-Blinder procedure only characterizes the central tendency of wages (between-group wage differences), the MM approach characterizes both the central tendency of wages and their dispersion (linked to residual wage inequality). This is a key issue if research is aimed at decomposing wage inequality into composition effects, between inequality, and residual inequality. Second, under the assumption that aggregate quantities of skills in the labor market do not affect skill prices (a strong but convenient assumption), the conditional quantile model can be used to simulate how changes in the labor force composition or skill prices affect the distribution of wages. For example, to see what wages would have prevailed with the labor force composition of period t, $g_t(x)$, and labor market prices of period s, $\beta_s(\theta)$, one simply simulates wages using $g_t(x)$ and $\beta_s(\theta)$.

AKK extend the MM approach to a counterfactual analysis of residual wage inequality. Their approach uses the skill price vector at the 50th percentile, $\beta(0.5)$, to characterize changes in between-group inequality. Hence, like the OLS price vector in the Oaxaca-Blinder decomposition, $\beta(0.5)$ is used to estimate the central tendency of the data conditional on x. Within-group inequality is quantified using the difference between the estimated coefficient vector $\beta(\theta)$ and the median coefficient vector $\beta(0.5)$,[9]

$$\beta^w(\theta) \equiv \beta(\theta) - \beta^b \equiv \beta(\theta) - \beta(0.5). \tag{4}$$

Hence, there are now three ingredients in each wage simulation exercise: (i) estimated "within"coefficient vectors $\hat{\beta}^w(\theta)$ for a large number of θs spread over the open unit interval; (ii) an estimated "between"coefficient vector $\hat{\beta}^b \equiv \hat{\beta}(0.5)$; (iii) and the distribution of covariates, $g(x)$. AKK perform counterfactual analysis by changing one of these three elements at a time. This allows them to assess the contribution of residual wage inequality (changes in the wage distribution when changing $\hat{\beta}^w(\theta)$ only), between group inequality (changes in the wage distribution when changing $\hat{\beta}^b$ only), and labor force composition effects (changes in the wage distribution when changing $g(x)$ only).

6 Results

In this section, we perform the following four-step exercise in order to decompose wage inequality change. First, four thousand θ were selected for each year from a uniform distribution $U(0, 1)$. Next, the quantile regressions for each percentile were estimated, the counterfactual exercise described in the previous section was implemented and, lastly the comparison are made.

Before analyzing the results, we have to check whether the method used the conditional quantiles with accuracy. Figure 2 describes both the original as well as the simulated quantiles. Since the simulated percentiles match the originals, the accuracy of this procedure for Spanish data is almost perfect, with the exception of the year 2000. Even for this year, the error seems be restricted to levels (constant at about 3%); thus, it vanishes when we compare the evolution of distances between percentiles. So, to the exent that we are only concerned with log changes, the magnitude of the error is negligible. With the exception of this case, the average error ranges from 0.8% (for the 1989 data) and 1.5% for the 1995 data. Thus, we conclude that the MM algorithm is a suitable tool for decomposing changes in wage distributions.

[9] See Appendix B for a prove.

Overall Wage Inequality

Table 5 shows the log change in 10th-50th, 50th-90th and 10th-90th percentile distances, with the last of these representing what we call overall wage inequality. The first result shown in the top panel of the table refers to the fall in inequality during the eighties and late nineties and the sharp rise in inequality during the early nineties. This result is very similar to that found in section 4. For the whole period, the result is a slight increase. Despite the lack of observations for the 1980-1985 period, inequality might have grown during the first half of the eighties despite the over fall for that decade. This intuition is driven by the 4.7 percent fall in wage inequality between 1980 and 1990 shown by the EPF, when the ECPF shows it to have contracted by a higher rate, 6.7 log points, during the second half of the period under study.[10]

Examining the two halves of the density instead of the entire distribution gives rise to a familiar question. As we pointed out earlier, the two halves do not always share the same trend; thus, total inequality depends on which half prevails. While the upper tail increased during the 1980-1985 and 1990-1995 periods and decreased during the 1985-1990 and 1995-2000 ones, the lower half was highly negative during the eighties (especially between 1980 and 1985) and positive during the nineties. Thus, the causes driving total wage inequality seem to differ above and below the median.[11]

To help clarify the dynamic under discussion, Figure 3 presents changes in wage density distributions. Each line states the increase or decrease in the density of a given wage between two specific years. For example, in the EPF line, the density of a wage equal to 12.6 increased by about 0.05 between 1980 and 1990.[12] The vertical lines represent median levels during the first years of our comparison. Thus, between 1980 and 1990, this figure represents an accumulation of density near but below the median. Between 1985 and 1990, we observe a new low wage concentration beside an increase of density around the median. This explains the results of Table 5. First, the inequality drop of about -6.7% can explained by the huge concentration of density around the median instead of around the tails. However, the drop in density around the upper half can be explained by the median increase, which is not followed by the higher percentiles. Between 1990 and 1995 a dispersion increase can be observed, especially around the lower part. Finally, the pattern between 1995 y

[10]These results are complementary to those found by other Spanish studies, for different periods and using different data: Arellano, Bentolila, and Bover (2001) using Social Security Records for 1981-1987, Abadie (1997) for the eighties as a whole and Izquierdo and Lacuesta (2007) since 1995.

[11]This is also AKK's main finding for the United States.

[12]The density are estimated using the Epanechnikov kernel procedure using simulated wages (Epanechnikov 1969). The width is the one that would minimize the mean integrated squared error if the data were Gaussian.

Table 5: 100 × log changes in Overall, Between Groups and Residual

	EPF 80-90	ECPF 85-00		
		85-90	90-95	95-00
overall				
10-50	-4.6	-1.3	5.1	1.4
50-90	-0.1	-5.3	2.6	-4.4
10-90	-4.7	-6.7	7.7	-2.9
prices				
10-50	-4.4	-3.2	3.7	1.9
50-90	-0.7	-7.2	0.2	-1.8
10-90	-5.1	-10.3	3.9	0.1
quantity				
10-50	-0.2	1.8	1.4	-0.5
50-90	0.5	1.9	2.4	-2.6
10-90	0.3	3.7	3.9	-3.1
between				
10-50	-0.8	-1.0	0.5	-2.1
50-90	-1.8	-3.5	1.6	-2.6
10-90	-2.6	-4.4	2.1	-4.7
residual				
10-50	-3.6	-2.2	3.2	4.0
50-90	1.1	-3.7	-1.4	0.7
10-90	-2.5	-5.9	1.8	4.8

Note: Each value represents changes in log percentiles distances. *Overall* defines simulated wages distributions using MM algorithm. *Prices* represents counterfactuals distributions when only returns to skills (education and experience) change, *quantity* when only the distribution of attributes change, *between* represents changes between counterfactual densities when only princes changes when all of the workers has the same returns to skills no matter the location they are within the distribution of wages, and *residual* shows the inequality evolution using counterfactuals densities when only changes in returns distributions are evaluated.

2000 mirrors that to be found for the 1980s.

In view of this information, our first impression is that inequality adopts a countercyclical pattern. This is a well known issue that finds its parallel in inequality analysis for other countries. For example, Dimelis and Alexandra (1999) find that different US and UK inequality measures show a negative correlation with the GDP business cycle (measured as the difference between GDP and its long-term trend). But they also find a positive correlation for some countries, such as Greece, and mixed results for others, such as Italy. This initial information therefore shows that the Spanish case is closer to that of the US and the UK.

In summary, during the period under analysis wage inequality showed an uneven rather than a steady pattern, presenting differences above and below the median. Furthermore, Spanish wage inequality shows a countercyclical pattern shared with 50th-90th inequality, while 10th-50th inequality has increased since 1985. This is our first result.

Composition and Prices

To explore the possible causes of this dynamic, we must isolate the effects of both price and composition on wage inequality. At a first glance, it is not clear which of these two effects predominates. During the 1980s, for example, the 5.1 log point decrease in 90th-10th inequality is caused by prices, while quantities play a minor offsetting role (0.3 log points). The predominance of prices is also evident during the second half of the 1980s. Nevertheless, during the 1990s prices and composition jointly increase the 90th-10th inequality by 3.9 log points until 1995, while composition plays the main role thereafter, decreasing inequality by 3.1 log points while the price trend remains steady. Both effects are important, however. This result is coherent with a major change in the labor force composition in Spain during this period, and especially after 1985: the increased participation of both highly educated workers and women.[13] This change affects 10-50th and 50-90th inequality to the same degree. For all of these reasons, therefore, the composition effect in wage inequality change must not be negligible.

Despite the compositional factor, the pace marked by total wage inequality repeats observable price movements until the last period, when the shift in inequality was almost zero. Thus, our data shows that prices reduced inequality during the 1980s and increased it during the 1990. Nevertheless, the reduction during the 1980s might have been concentrated in the second half of that period. Except for the 1985-1990 period, the effect of these changes mainly shows up in the lower half, which governs the total wage inequality trend. Despite the important role played by composition effect, the overall evolution of wage inequality mirrors price inequality. Thus, inequality caused by price changes shows an increased trend for the 10th-50th percentile distance and a countercicylcal one for the 50th-90th one.

These results are coherent with those of previous studies (e.g., Abadíe, 1997; Arellano, Bentolila and Bover, 2001 and Izquierdo and Lacuesta, 2007) but not with those of Izquierdo and Lacuesta (2007) with regard to the composition effect. There are two reasons for this discrepancy. First the period of study analyzed by the latter (1995-2000) differs from our own. Second, we are restricted to using data from heads of family. In the EPF and ECPF surveys, a woman is classified as the head of her family when no other male family member is em-

[13]As it can be sawn in Table 1, the increase in average years of schooling that stems from these surveys was around 0.65 in eighties, while during 1990-1995 it increased 0.67 and 1995-2000 by 0.76.

ployed. Thus, although our data reflects the fact that ten percent of the workers considered for this exercise were women, our analysis shows a quite small level of female participation and also a gap due to market participation. Thus, our results differ because our data –unlike that of Izquierdo and Lacuesta (2007)– does not completely reflect shifts in female participation taking place during the period under study, nor does it show the dramatic recent rise in the latter. As they argue, women's participation in the job market was the main driving force behind wage inequality during this period, and our study does not account for this trend.

Our second result, therefore, is that both price and labor composition played an important role in the evolution of Spanish wage inequality during our study period. Furthermore, while composition effect is important, overall wage inequality changes mirrored the changes in price inequality during this period. Thus, the 10th-50th distance when only prices changes increased after 1985, and the 50th-90th distance shows a countercyclycal pattern.

Between and within group inequality

Next, price inequality is decomposed between the BI and RWI change in inequality. Note that BI shows inequality due to the distance between the conditionals distributions, while RWI represents the dispersion within the conditional distributions.

As shown in the lower panel of Table 5 BI captures a symmetric trend on both sides. Here again, the trend suggests that prices reduced inequality during the late 1980s and 1990s but increased it during the early years of both decades. Nevertheless, analysis of RWI reveals a completely different picture, with the most salient feature being the increasing trend. From a fall during the early 1980s to a rise during the late 1990s, the trend is clearly one of steady growth. In RWI there is no symmetry, with most of the change clustering in the lower half with the exception of the 1985-1990 period. Our third result, therefore, is the difference between the inequality evolution at BI and RWI. Clearly, these two different forces jointly draw the picture described previously regarding a price-driven change in inequality. How can we explain the BI and RWI trends? An initial guess is that each of these trends must respond to different causal factors.

Education and Experience

Due to the existence of previous literature that investigates the effects of a return to education and experience in Spanish wage inequality, one further step is to distinguish the BI for both prices. Table 6 presents the counterfactual densities obtained when only one price changes. For example, education became less dispersed throughout our study period, with the exception of the years between 1990 and 1995. Again, this is a countercyclical trend, but one located between the 50th-90th percentiles, so that almost all of the decrease can be found in the upper half. The reason for this asymmetry might be the massive incorporation of more highly skilled workers into the labour market during this period. Bearing in mind the results discussed in section 4, between 1980 and 2000 the fall in the high school wage premium relative to other premiums, followed by a fall in the college wage premium, reduced inequality when only the price of education is accounted. Nevertheless, the lower tail shows a different movement where growth is prevalent. Again, weaker institutions might be the reason for this trend. The period 1980-1985 is of special interest. In this case, comparison of the results for the decade as a whole and for the 1985-1990 period would reveal an increase in BI in education. This result is similar to that found by Arellano, Bentolila, and Bover (2001) (for the first half of the 1980s) and to that found by Abadie (1997) (for the 1980s) and Izquierdo and Lacuesta (2007) (for the second half of 1990s). Experience presents a countercyclical and completely asymmetric trend on both sides of median. The greater weight of less experienced workers within the lower tail of the wage distribution and the greater weight of more experienced workers within the upper tail would explain the different effects. Again, these results mirror those obtained in other Spanish studies for the same periods.

Table 6: 100 $\times$ log changes in Education and Experience Prices and Quantity Inequality

	EPF 80-90	**ECPF 85-00**		
Prices		85-90	90-95	95-00
Education				
10-50	0.0	1.7	-1.7	1.7
50-90	-2.5	-5.4	2.9	-4.2
10-90	-2.5	-3.7	1.2	-2.5
Experience				
10-50	1.3	-2.7	2.2	-3.8
50-90	-1.5	1.9	-1.3	1.6
10-90	-0.2	-0.7	0.9	-2.1

The range of variation not only between BI, RWI and composition effects, but also within each simulated counterfactual density, suggests that changes in wage inequality were caused by more than one factor. The hidden factors underlying changes in earnings dispersions differ for different wage distribution areas, and in the intensity of their effects. Nevertheless, it is necessary to tackle this result with some possible explanations that could be developed in future research.

Some tentative explanations

Note that our results for BI once again support the idea of countercyclical inequality. In years of low GDP growth, the distance between the returns paid to skilled workers increases, whereas in years of high GDP growth the distance between conditional dispersions contracts. A number of theoretical explanations for this result can be given. For example, despite that the fact the skill bias technical change can be seen to result from unobservable change in the production function (Katz and Murphy, 1992), more recently Krusell, Ohanian, Rios-Rull, and Violante (2000) associate technological change with capital equipment increase. The important assumption is that of capital-skill complementarity, which means that the elasticity of substitution between capital and unskilled labor is higher than the elasticity of substitution between capital and skilled labor. Thus, in the middle of both decades, both output and the use of inputs declined. Assuming that capital is a quasi-fixed production factor and that capital-skill complementarity prevails, the demand for unskilled labor reacts more than the demand for skilled labor. When aggregate demand increases, the opposite effect occurs. Thus, it follows that the relative demand for skills is countercyclical due to a relative supply curve of skilled labor with a positive slope, and that the relative wage and employment of skilled labor are also countercyclical.

Supporting this explanation, Hidalgo, O'Kean and Rodríguez (2008) have found strong capital-skill complementarity evidence for Spain between 1980 and 2004, which could explain the Spanish countercyclical property that we have been discussing. While this countercyclicality has also been found for other countries,[14] for the United States there is only evidence of countercyclical difference in wages earned by skilled and unskilled workers until 1984. The trend turns procyclical in 1984, a shift that has been explained by the concomittant reduction in the degree of capital-skill complementarity. (Castro and Coen-Pirani 2005).[15] Nevertheless, since other explanations have been suggested,[16] this issue awaits a more in-depth analysis.

[14]For example Denmark (Skaksen and Sorensen 2005).

[15]The cyclicality of the skill premium has been largely ignored in the literature since Reder (1955), who was the first who formally investigated the movements of wage differentials between skilled and unskilled workers over the business cycle. Since then, Keane and Prasad (1993), Young (2003) or Lindquist (2004) has analyzed this issue.

[16]For example, firm-specific human capital of skilled labor (see e.g., Becker, 1964), and hiring costs that are higher for skilled labor than unskilled labor (see e.g., Bentolila and Bertola, 1990) or implicit contracts literature (see e.g., Pourporides, 2007).

The evolution of RWI can be explained by institutional factors, for a number of reasons. First, this feature has tended to be considered as an institutional effect (i.e. DiNardo, Fortin and Lemieux, 1996, Card and Lemiuex, 2001 and Autor, Katz and Kearny, 2005a and 2005b). Second, it fits quite well with Spanish evidence regarding labor markets reforms. For example, Peraita (2003) argues that the Spanish labor market was one of the most rigid in the industrialized world during the early 1980s in particular, but market labor reforms in 1984, 1992, 1993 and especially in 1994 induced a flexibilization process. Third, labour institution reforms since 1984 and the fact that Spain's growth structure since 1996 has been skewed towards low-skilled labour sectors would explain the wide dispersion found, especially, in the lower half of the distribution. For example, the inequality decrease described in the lower half of RWI during the early 1980s could be explained by the increase in the bottom wage percentile due to the appearance of a new institutional wage-setting framework. Peraita (2003) explains, by contrast, that Spanish labour market reforms should increase wage inequality among lower-paid workers, especially from 1994 onwards. To summarize, it seems that the possible increase in relative supply due to unemployment and in skill bias technical change due to capital-skill complementarity (BI) and, more particularly, to weaker labour institutions, (RWI), may jointly explain the increase in inequality during the early 1990s and thus explain the overall increase in wage inequality during this period. These factors would also explain the lack of movement in the upper tail: the lesser impact of unemployment effects in these cohorts and the limited labour reforms which had no effect on indefinite labour contract privileges might explain this pattern, in which the labour market appears to dominate the wage inequality trend.

There is room for alternative explanations, however. For example, there may have been an increase in unobservable heterogeneity, which would have increased RWI. A natural explanation might be linked to the generalization of education and the increase in number of scholarships granted since the 1980s, which pushed a greater number of students with non-observable skills into higher education. This scenario might have increased the non-observable skills within each group of workers, especially the more educated ones.

This last hypothesis can be easily tested using counterfactual densities to simulate the densities of groups of workers with, say, 16 years and zero years of schooling, and exactly the same experience, gender and other characteristics. Under this hypothesis it should be expected that the conditional RWI of the more educated workers would mirror the overall RWI trend, especially in the lower tail. People with less skills and more education would be located in the lower tail of the conditional "more educated"distribution. Thus its dispersion must increase through time. Then, once the "more and less educated"conditional counterfactual distributions are simulated, it is possible to calculate the percentile distances. The results for the 1985-2000 period are shown in table 7.

Table 7: 100 $\times$ log changes in percentiles distances for 16 and 0 years of schooling counterfactual densities

	85-90	90-95	95-00
16 years sch.			
10-50	5.8	-5.3	12.9
50-90	-18.8	6.8	-9.6
10-90	-13.0	1.5	3.3
0 years sch.			
10-50	-0.1	6.9	13.8
50-90	-6.9	-4.6	-7.6
10-90	-7.0	2.4	6.2

In these cases, the less educated workers' RWI appears to be closer to overall RWI than to that of more educated workers. In this last case, the 10th-50th distance grows between 1985 and1990 and falls between 1990 and 1995, while the 10th-50th distance in the less educated conditional distribution follows the same trend as the total RWI lower tail dispersion. Thus, the increase in heterogeneity among more educated workers due to easier access to education is not enough to explain the lower tail trend of its conditional residual wage distribution and the resultant increase in total RWI. Rather, institutions seem to represent a key driving force behind the change in the residual inequality for lower wages.

7 Conclusions

This paper attempts to decompose the change in Spanish wage inequality into three components, using a counterfactual analysis based on the quantile regressions developed by Machado and Mata (2005) and extended by Autor, Katz and Kearney (2005a and 2005b): changes in between- and within-group inequality and changes in labor composition.

The results for the 1980-2000 period are fourfold. First, because the Spanish wage behaves countercyclically, the inequality change is almost zero but positive. Second, both price and labor composition play important roles in the evolution of Spanish wage inequality; however, while composition effect is important, the changes in overall wage inequality mirror the changes in price inequality. Third, the between-group inequality, which measures the distances in conditional to observable characteristics in wage distributions, follows a countercyclical trend, while the residual wage inequality, which measures the dispersion within the conditional distributions and that caused by non-observable variables, increased since 1985 onwards. Also, an important result is that while the between-group inequality shows a symmetric pattern above and below the median, the residual inequality tells a different story for each half. Forth and finally, the inequality found for education and experience mirrors previous results for Spain and, again, behaves countercyclically.

There are several possible explanations of these results. While for between-group inequality changes in the supply and demand of different worker cohorts may explain price changes, and therefore changes in wage distribution, institutions may be behind the increased dispersion in the lower tail of the residual wage inequality. For example, countercyclical changes in relative demand for skilled (more educated) workers could be explained by skill-biased technological change, and especially by the degree of complementarity between this factor and capital. However, there are other possible explanations for this behavior, such as changes in Spanish labour institutions in favor of collective bargaining improved between 1980 and 1985, and market reforms since the early 1990s. Nevertheless, these intuitions must be explored in greater depth.

References

ABADIE, A. (1997): "Changes in Spanish labor income structure during the 1980's: a quantile regression aproach," *Investigaciones Economicas*, 21(2), 253–272.

ACEMOGLU, D. (2003): "Cross-Country Inequality Trends," *Economic Journal*, 113(485), 121–149.

ARELLANO, M., S. BENTOLILA, AND O. BOVER (2001): "The Distribution of Earnings in Spain During the 1980s: The Effects of Skill, Unemployment and Union Power," CEPR Discussion Papers 2770, C.E.P.R. Discussion Papers.

AUTOR, D. H., L. F. KATZ, AND M. S. KEARNEY (2005a): "Rising Wage Inequality: The Role of Composition and Prices," Harvard Institute of Economic Research Working Papers 2096, Harvard - Institute of Economic Research.

——— (2005b): "Trends in U.S. Wage Inequality: Re-Assessing the Revisionists," NBER Working Papers 11627, National Bureau of Economic Research, Inc.

BECKER, G. (1964): *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. Columbia University Press, New York.

BENTOLILA, S., AND G. BERTOLA (1990): "Firing Costs and Labour Demand: How Bad Is Eurosclerosis?," *Review of Economic Studies*, 57(3), 381–402.

BLINDER, A. (1973): "Wage Discrimination: Reduced Forms and Structural Estimation," *Journal of Human Resources*, 8(4), 436–455.

CARD, D., AND T. LEMIEUX (2001): "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, 2(116), 705–746.

CASTRO, R., AND D. COEN-PIRANI (2005): "Why Have Aggregate Skilled Hours Become So Cyclical Since the Mid-1980's?," Cahiers de recherche 2005-19, Universite de Montreal, Departement de sciences economiques.

DIMELIS, S., AND L. ALEXANDRA (1999): "Inequality and Business Cycles in the U.S. and European Union Countries," *International Advances in Economic Research*, 5(3), 321–338.

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), 1001–1044.

EPANECHNIKOV, V. A. (1969): "Non-Parametric Estimation of a Multivariate Probability Density," *Theory of Probability and its Applications*, 14(1), 153–158.

HIDALGO, M. A., J. O'KEAN, AND J. RODRÍGUEZ (2008): "Demanda de Trabajo y Tecnologías de la Información: Evidencia para España," Mimeograph. Universidad Pablo de Olavide.

IZQUIERDO, M., AND A. LACUESTA (2007): "Wage inequality in Spain - recent developments," Working Paper Series 781, European Central Bank.

JUHN, C., K. M. 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 K. M. MURPHY (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107(1), 35–78.

KEANE, M., AND E. PRASAD (1993): "Skills Levels and the Cyclical Variability of Employment, Hours and Wages," Discussion Paper 4, IMF Staff Papers 40.

KOENKER, R. W., AND J. BASSETT, GILBERT (1978): "Regression Quantiles," *Econometrica*, 46(1), 33–50.

——— (1982): "Robust Tests for Heteroscedasticity Based on Regression Quantiles," *Econometrica*, 50(1), 43–61.

KRUSELL, P., L. OHANIAN, V. RIOS-RULL, AND G. VIOLANTE (2000): "Capital Skill Complementary and Inequality: A Macroeconomic Analysis," *Econometrica*, 68(5), 1029–53.

LEMIEUX, T. (2002): "Decomposing changes in wage distributions: a unified approach," *Canadian Journal of Economics*, 35(4), 646–688.

LEMIEUX, T. (2006): "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?," *American Economic Review*, 96(3), 461–498.

LINDQUIST, M. J. (2004): "Capital-Skill Complementarity and Inequality Over the Business Cycle," *Review of Economic Dynamics*, 7(3), 519–540.

MATA, J., AND J. A. F. MACHADO (2005): "Counterfactual decomposition of changes in wage distributions using quantile regression," *Journal of Applied Econometrics*, 20(4), 445–465.

Oaxaca, R. (1973): "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14(3), 693–709.

Oliver, J., J. Raymond, J. Roig, and F. Barceinas (1999): "Returns to Human Capital in Spain: A Survey of the Evidence," in *Returns to human capital in Europe: A literature review*, chap. 13, pp. 280–298. R. Asplund and P.Pereira, Helsinki.

Peraita, C. (2003): "The Importance of Institutions and Skill-Based Technical Change in Wage Decompression: Evidence for Spain," V Jornadas de Economía Laboral, Reus.

Pourpourides, P. M. (2007): "Implicit Contracts, Wages and Wage Inequality over the Business Cycle," Cardiff Economics Working Papers E2007/19, Cardiff University, Cardiff Business School, Economics Section.

Reder, M. (1955): "The Theory of Occupational Wage Di®erentials," *American Economic Review*, 45(5), 833–852.

Skaksen, J., and A. Sorensen (2005): "Capital-Skill Complementarity and Rigid Relative Wages: Inference from the Business Cycle," *Contributions to Macroeconomics*, 5(1), 1145–1145.

Vila, L., and J. Mora (1998): "Changing Returns to Education in Spain During the 1980s," *Economics of Education Review*, 17(2), 173–178.

Young, R. (2003): "The Wage Premium: A Puzzle," Florida State University.

A The EPF 80-81/90-91 and ECPFs 1985-2004

Although these sources do not provide homogeneous wage statistic series, they give important information that is relevant to this kind of analysis, for which they have become the main current source of statistics. However, a number of problems need to be considered.

First, the EPFs from 1980-81 and 1990-91 provide a broad spectrum of data for about 20,000 families. However, the quarterly surveys use a smaller sample size, since their main objective is to offer a short-term analysis of consumption, rather than consumption structure. At any rate, since 1997 the sample has doubled. This problem may be resolved by using two-year samples to improve sample size, since the ECPFs poll the same family for six quarters, changing one sixth of this sample each quarter.

Second, wages are not immediately determined by recorded earnings data. The main problem is that there is no information on hours worked or similar criteria unless the head of family has worked for more than thirteen hours during the reference week. The only solution to this problem is to use only those workers who reported working more than 13 hours and to assume that they worked full time. This assumption will no doubt introduce some measurement errors.

Third, despite the vast amount of information available, complete information is only available for the head of family.[17] Therefore this paper, like other Spanish studies (Abadie 1997), works explicitly with this selection.

The fourth problem arises from the use of two similar but somewhat different sources, the EPF and the ECPF. Key differences between these sources center on the amount of information, number of characteristics reported, the richness of the classifications within each characteristic (for example, with regard to education) and others. Any contrast between these two sources must therefore derive from the effect of these differences.

The last problem stems from the heterogeneity of the definitions and the classifications of variables used. For example, different educational level classifications appear each year, partly because current surveys have modified their definitions over the years and partly because the Spanish legal definition of education changed during the period under study.[18] Table 8 shows the different educational groups. To solve this problem, we used years of schooling instead of educational level, because the former is homogeneous across all years and classifications. The years imputed are the same as those given in Vila and Mora (1998).

[17]For instance, the ECPFs records contain information about education only for this group

[18]In the early nineties the education law changed from an earlier one passed in 1970, which introduced compulsory schooling in Spain up to the age of fourteen. In 1990, the LOGSE extended compulsory schooling to the age of sixteen and changed the grades as the 1970 law had.

Table 8: Schooling groups. Individual Data

	epf 80/81	epf 90/91	ecpf 85/95	ecpf 00/01
Illetary-Primary	1-3	1-4	1-4	1-2
Low Secondary	4	5 & 7	4	3
Acad. & Voc. Upper Sec.	5-6	6 & 8	5	4-5
Higher Educ. (Short cycle)	7	9	6	6
Higher Educ. (Long cycle)	8	10	7	7

To conclude, despite the limitations of working with these surveys, they possess many redeeming features which make them our best choice for data on the distribution of Spanish wage inequality between 1980 and 2000. To summarize, two main criteria may be cited in support of their use: first, the lack of no better alternatives, second, their usefulness as a basis for comparative analysis. Once the surveys were selected for each year, they were then refined. Only data for heads-of-family working over thirteen hours per week were considered, after eliminating all self-employed wage-earners. Because ECPF surveys are restricted to household income data, this study focused exclusively on households where the head of family is the only worker. Second, groups of workers with unrepresentative characteristics were eliminated. For instance, education and experience cohorts were defined by five-year segments, and cohorts with fewer than fifteen records were deleted. Third, it was assumed that wages reported as below the legal minimum were either earned by part-time workers or individuals who had been unemployed for at least one year, or represented erroneous replies. Thus, the records were modified when annual reported wages fell below the legal minimum for that year. In this case, a zero was imposed for censored analysis. Table 9 shows the minimum wage value in 1980 constant prices, which will be the censor value and the percentage of records modified to zero. Fourth, calendar effects are taken into account in the ECPF information. The quarterly nature of these surveys implies that the wages reported might be influenced by the quarter in which they are given. To eliminate this effect, families with wages for all six quarters were taken first, and the calendar effect was analyzed in these cases. Then a wage-level factor was obtained for each quarter. Once these factors were obtained, all the workers' wages were "deflated". Finally, the average wage was taken and multiplied by four, to give the yearly wage for all of the workers, regardless of which quarter they had worked. In this case, the six quarter samples were combined to increase (double) the ECPF sample size, although the same result could probably be obtained using all four quarters in the calendar context.

Table 9: Censored data

	EPF		ECPF		
	1,980	1,990	1,985	1,990	1,995
Censored wage	313,148	287,565	292,280	287,565	279,055
Percentage (%)	9.44	3.16	8.18	3.83	4.47

B Heteroscedasticity and Quantile Regression

This Appendix describes the expression (4) that captures within-group inequality. The key is that quantile prices incorporate the heteroscedasticity as they change through the percentile estimates that they produce.

For the case in hand, different coefficients were estimated for different $\tau \in (0,1)$, which differ from each other whenever conditional wage dispersion depends on the covariate values. In others words, it has been proved that quantile regressions serve to analyze the heteroscedasticity in errors (Koenker and Bassett 1982). In this case, the quantile coefficients show changes in the percentiles. More specifically, suppose that wage equation

$$w_{it} = \beta_t x_{it} + e_{it}$$

is a so-called location-scale model, where $e_{it} = \sigma(x_t)\varepsilon_{it}$, $\sigma(x_t)$ is some function of x_t and ε_{it} is a normal iid error term with continuous and positive density $f(\varepsilon_t)$ and a distribution given by $F(\varepsilon_t)$. In this case, the conditional quantile for year t is given by

$$Q_\tau(w_t|x_t) = x_t'\beta_t + \sigma(x_t)F_{\varepsilon_t}^{-1}(\tau).$$

Suppose, for the sake of simplicity, that $\sigma(x_t) = x_t$. Then for the case of (2) the vector of coefficients estimated is given by

$$\beta_t(\tau) = \beta_t + F_{\varepsilon_t}^{-1}(\tau).$$

Here, $F_{\varepsilon_t}^{-1}(\tau)$ is the inverse of the cumulative distribution function of ε and represents the value of the percentile τ. A test of heteroscedasticity would therefore be to check the null hypothesis $\beta_t(\tau) = \beta_t(\theta)$ for $\tau \neq \theta$. This does not imply that estimations are not consistent. If $\sigma(x_t)$ cannot be fitted to a linear expression, the intuition will be the same, but $\beta_t(\tau)$ will require a more complex expression.

To clarify matters, let us suppose that the quantile model is given by (2). Then, it is easy to see that

$$\beta_t(\tau) = \beta_t + \Psi(x_t; F_{\varepsilon_t}^{-1}(\tau));$$

which is a generalization of (B). In this case, the within-group coefficients are

$$\beta_t^w(\tau) = \beta_t(\tau) - \beta_t(0.5) = \Psi(x_t; F_{\varepsilon_t}^{-1}(\tau)) - \Psi(x_t; F_{\varepsilon_t}^{-1}(0.5)).$$

This expression implies that x_t induces the heteroscedasticity in the $\beta_t^w(\tau)$ estimation. But that is another story, because the within-group estimated counterfactual density changes $Q_\tau(\beta_1^b, \beta_1^w(\tau); g(x;0)) - Q_\tau(\beta_1^b, \beta_0^w(\tau); g(x;0)))$ use a constant weight, $g(x;0)$, applied to each conditional density. So, residual counterfactual densities use the same weighted rule over time, and composition effects are correctly removed from RWI.

Figure 1: Wage inequality, college and primary/high-school wage premium and residual inequality

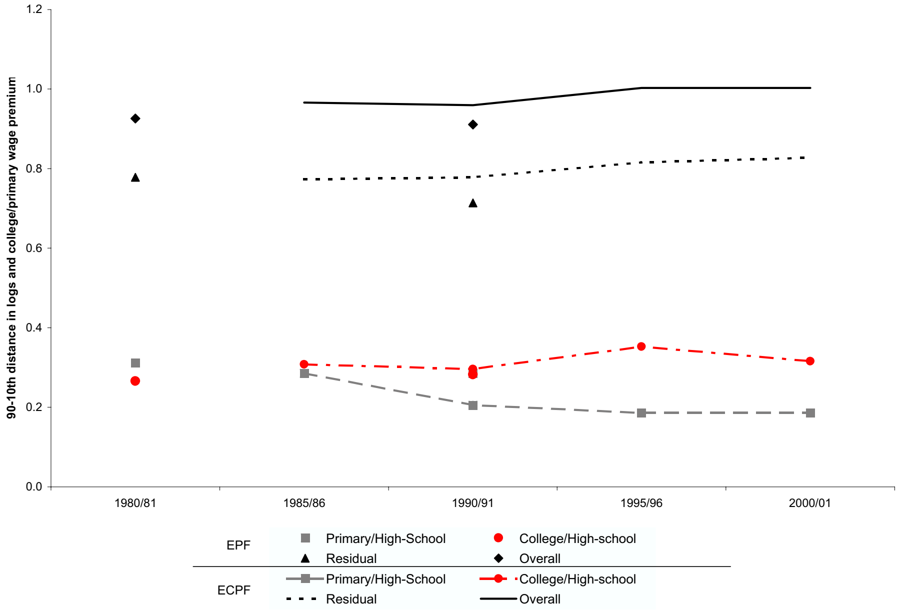
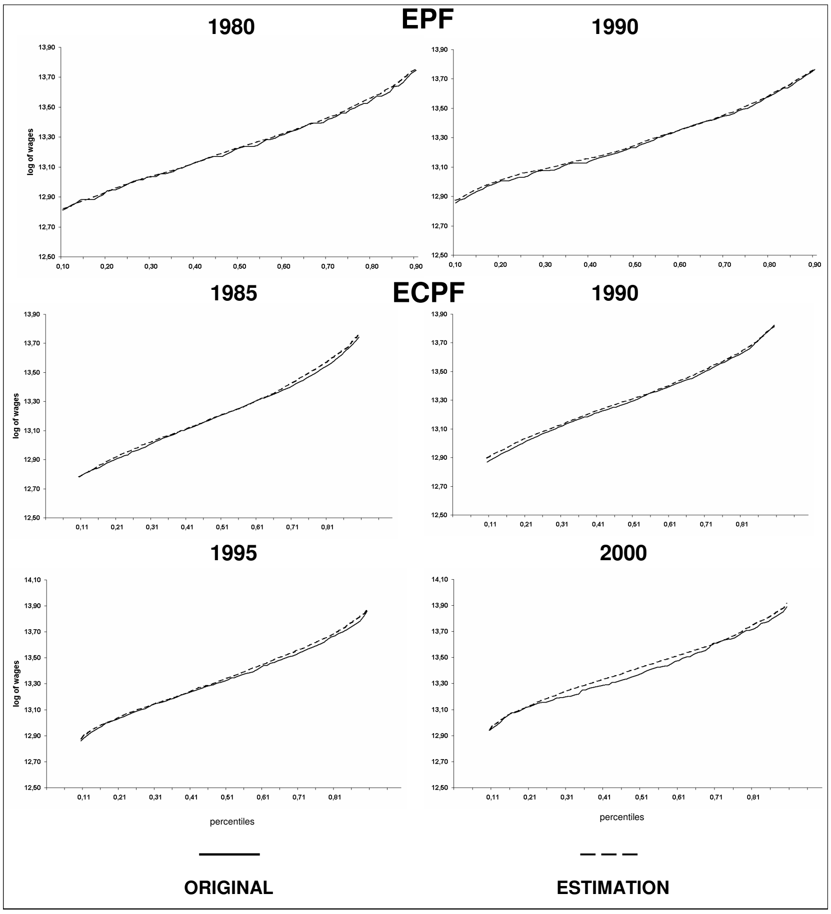


Figure 2: Original versus QR Estimated Wage Distributions' Percentiles.



Solid lines represent original wage distributions percentiles. Dotted lines show the simulated wage distribution percentiles using the Machado and Mata (2005) algorithm.

Figure 3: Changes in Wage Distributions