ESTADÍSTICA ESPAÑOLA
Vol. 52, núm. 174, 2010, págs. 333 a 366

# Wage Inequality in Spain, 1980-2000: the case of male head-of-household

por
MANUEL A. HIDALGO(1)
Universidad Pablo de Olavide

**RESUMEN**

I use quantile regression to simulate counterfactual densities to decompose Spanish wage inequality evolution over the period 1980-2000 into changes due to observable prices, labour market composition and prices of non-observable worker characteristics. Our empirical results are threefold: first, wage inequality follows a counter-cyclical trend from the mid-eighties onwards; second, changes in both prices and composition play an important role in this evolution; and third, changes in observable prices mirror this behaviour above and below the median, while, non-observable inequality below the median increases from 1985 onwards, following a different trend above it. Finally some tentative explanations are given.

*Key-words*: wage inequality, quantile regression

*JEL codes*: J24, J31

---

(1) I am grateful to seminar participants at 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 their very useful suggestions. I gratefully acknowledge support from project 30.04.01.27.27 (Ensayos sobre Andalucía y Nueva Economía).

## 1. INTRODUCTION

Wage inequality has grown substantially since the 1970s in many countries. In the United States, for example, the $90^{th}/10^{th}$ percentile ratio of male hourly earnings grew by 23.4 log points between 1979 and 2003 (Autor, Katz, and Kearney 2006 & 2008). 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 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 male wage inequality trends in Spain between 1980 and 2000, providing a level of detail similar to that found in analyses for other countries. Specifically, it examines trends in male overall wage inequality, and decomposes them into three different sources: changes in the education and experience wage premium (between-group inequality, BI)(2); the observed skill distribution (composition effects); and changes in wage dispersion among workers with the same levels of experience and education (residual wage inequality, RWI). Thus, the first important contribution of this paper is to provide a comprehensive explanation for the evolution of wage inequality in Spain since records began(3).

The literature on this subject describes a number of approaches to the decomposition of changes in overall wage inequality into between-group variance, composition effects, and residual wage inequality. A seminal study by Juhn, Murphy and Pierce (1993) (JMP hereafter) extended the "mean" Oaxaca-Blinder procedure to include the decomposition of distributions. In applying this approach, JMP first estimated annual 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

(2) There are two reasons for using only these variables. The first is for purposes of comparison with studies of other countries and Spain. The second is that the relevant EPF and ECPF data do not greatly increase the percentage of variance explained. Furthermore, some information, for example, type of contract (fixed-term or indefinite) is unavailable.

(3) For a sensitivity analysis, we expand the workers' characteristics included in the decomposition.

(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 Kearny (2006 and 2008) (hereafter, AKK) extend Machado and Mata's (2005**)** (MM hereafter) quantile decomposition technique. This study relies on the latter approach, because it incorporates both the work of JMP and that of DiNardo, Fortin and Lemieux (1996).

Our analysis uses male data from the Household Budget Surveys for 1980-81 and 1990-91 and the Continuous Household Budget Surveys for 1985-86, 1990-91, 1995-96 and 2000-01. These are the only data sources that contain all the information needed for this study and that cover the entire period of interest. By contrast, other data sources either cover shorter time-periods or lack key information on individual workers. For example the Social Security Records (Muestra Contínua de Vidas Laborales) are, by construction, inadequate for investigation beyond 2004, and the Structural Wage Survey (Encuesta de Estructura Salarial) provides data only for 1995, 2002 and 2006. Also, Social Security Records are tax-based rather than wage-based, and the data are top-cut (Hospido and Bonhomme, 2010). The reason for using only male data is due to the lack of information about women in this survey, simply because some of the data refer only to the head-of-household, who tends to be male and provide most of the information(4).

We find that overall male wage inequality in Spain, measured as the difference between log wages at the $90^{th}$ and $10^{th}$ distribution percentiles, decreased very slightly between 1980 and 2000, and more intensively in the eighties. Our findings for the 1980s are on a par with those of studies published at the time. Using data from the Social Security archives, Arellano, Bentolila and Bover (2001) show that wage dispersion rose during the first half of the 1980s. Abadíe (1997) demonstrated that wage inequality fell slightly throughout the 1980s based on data from the Household Budget Survey for 1980-81 and 1990-91. Our results for the nineties are similar to those obtained by Pijoan and Márquez (2010), who use ECPF data for 1985-1995 to show that wage inequality, measured by the distance between the $10^{th}$ and $90^{th}$ percentiles, fell in the late eighties and rose from 1992 to 1995. They are also in the line of Izquierdo and Lacuesta (2007), who use data from a Structural Wage Survey to explain a decrease in wage inequality in Spain between 1995 and 2002.

Further to these analyses, our contribution is to decompose changes in wage inequality into between-group inequality (the distance between distributions conditional to one observable characteristic), residual inequality (inequality within condi-

(4) See, for example, Abadíe (1997) and Pijoan and Márquez (2010).

tional distributions), and composition effects (changes in wage inequality that arise due to a shift in labour force composition). Between-group inequality behaves much like overall inequality. This is true for the entire 20-year period, 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 late 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 except in the 1990-1995 period. The 50/10 wage gap, on the other hand, increased from the second half of the 1980s up to 1995 and decreased from then to the end of the sample period.

A close look at these figures reveals that between-group inequality evolves similarly in the upper and the lower part of the wage distribution. Thus, both the $50^{th}/10^{th}$ and the $90^{th}/50^{th}$ gaps behave qualitatively like the $90^{th}/10^{th}$ gap with respect to this type of wage inequality. At a quantitative level, however, the changes are more negatively 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 between 1985 and 1995. While the $50^{th}/10^{th}$ gap increased from the second half of the 1980s, the $90^{th}/50^{th}$ gap shows a decrease between 1985 and 1995 and an increase from then onwards. Hence, the overall increase in wage inequality that appears in the upper part of the distribution is mirrored by rising between-group inequality, but not by rising residual inequality. In the lower part of the distribution, in contrast, the behaviour of wage inequality is much like that of residual inequality but different from that of between-group inequality. The different patterns of wage inequality above and below the median, and the observable behavioural differences 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(5). These empirical findings are the second important contribution of this paper. Never before has wage inequality in Spain been linked to the recent economic evolution and cyclical pattern in Spain.

The rest of the paper is organized as follows: Section 2 reviews the related literature, section 3 describes the data, section 4 gives a preliminary survey of wage

(5) Abadíe (1999) reached the same conclusion for the US and for Spain in the 1980s.

inequality between 1980 and 2000, section 5 explains the decomposition method, section 6 discusses the main empirical results, and section 7 concludes.

## 2. RELATED LITERATURE

Overall wage inequality and its decomposition 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 Machado and Mata (2005) approach as extended by Autor, Katz and Kearney (2006 & 2008).

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

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

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, $\beta_t$ 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, there are three possible sources of change in wage inequality: changes in the distribution of observable characteristics (changes in the distribution of $X_{it}$); changes in the prices of observable skills (changes in $\beta_t$); and changes in the distribution of unobservable characteristics ($u_{it}$). Using this structure, JMP estimate the counterfactual densities that would prevail if any subset of these three components were held 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 due to observables is the result of changes in returns on observable skills (rather than changes in the distribution of observable skills). While finding that residual wage inequality began to increase during the late 1960s, they also find that increasing returns on skills have contributed to higher overall wage inequality only since the early 1980s.

Their preferred explanation for these wage inequality trends is that increasing demand for both observable and unobservable skills has translated into higher skill prices(6).

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)(7). 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, 2006, 2008). As a result, a rise in the number of college degree-holding workers may produce a mechanical composition effect causing overall wage inequality to increase. Such composition effects working through residual wage dispersions demand an explanation, otherwise they might easily be confused with the effects of changing prices on unobservable skills (Lemieux, 2006 and Autor, Katz and Kearny, 2006, 2008). 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(8). He also argues that a great many of the changes in RWI can be explained by institutional factors (rather than changes in skill prices), such as the declining real minimum wage.

Our empirical work relies on the latest decomposition method appearing in the literature on this subject, 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 JMP, and nests the approach proposed by DiNardo, Fortin and Lemieux (1996) and Lemieux (2002, 2006). As applied to U.S. data, the AKK approach confirms the importance of the composition effects emphasized by Lemieux (2006) for the lower part of the wage distribution. In the upper part, meanwhile, rising wage

(6) The important question that this raises is why BI and RWI started rising at different times.

(7) In principle, the approach used by JMP should be able to deal with this issue. However, Lemieux (2002) argues that it does not do so in practice.

(8) The composition effect now encompasses all changes in wage inequality due to changes in skill composition, including those caused by residual wage dispersions re-weighted by skill. RWI refers to changes in wage inequality that are unaccounted for by composition effects and BI.

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

Wage inequality in Spain have been studied by Abadie (1997), Arellano, Bentolila, and Bover (2001), Izquierdo and Lacuesta (2007), and Pijoan and Márquez (2010). Abadie (1997) examines wage inequality trends during the 1980s using quantile regressions. He documents a fall in the return to education during this period, mostly affecting 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 and also within- and between-group wage inequality trends. Arellano, Bentolila, and Bover (2001) use a large Social Security data sample to examine wage inequality trends for the 1980-1987 period. Their analysis focuses on the behaviour 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 reduced inequality, while compositional shifts increased it. Our approach differs from theirs with respect to period of analysis, method, and the special attention our study pays to within-group inequality. Finally, Pijoan and Máquez (2010), using ECPF data for 1985 to 1995 and the European Household Panel Data for 1993 to 1999, document declines in wage inequality during the period 1985 to 1999. The main difference between their analysis and this one is that they conclude by overlapping these two surveys.

## 3. DATA

Let us now review the data sets that enable the analysis of wage inequality trends in Spain.

### *Household surveys*

All of the information necessary for our analysis is supplied by the "Encuesta de Presupuestos Familiares" or EPF (Spanish Household Budget Survey - HBS) for 1980-81 and 1990-91 and its newer counterpart, the quarterly Encuesta Continua de Presupuestos Familiares or ECPF (Continuous Household Budget Survey - CHBS), data sets available from 1985 to 2005. While they provide useful data on wages, education, age, gender, the information refers only to heads of households. This is why this study concentrates on head-of-household data. The absence of information for women restricts the analysis to males only. There are other important limitations. One of them is that hours worked is an unknown factor in all these surveys. This limits the scope of this study to workers on full-time contracts. Further limitations are explained in Appendix A.

Despite the above, alternative sources (some of which are discussed 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, 2000). Appendix A gives more details on these surveys.

### *Other data sources*

There are two other 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 supplies data on about two thousand industrial and service workers for the years 1995, 2002, and, more recently, 2006. While the individual information from this survey is very detailed, the time span covered is too short to be useful here. The (Quarterly) Labour Cost Survey (formerly, the Survey of Wages in Industry and Services) surveys wages for the 1980-2000 period but provides no information on education levels(9).

The Spanish Social Security records are another source of wage data. 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 Continua de Vidas Laborales" from the Social Security records further develops the information used by these authors, covering a panel of workers and an extended number of years and providing more data on individual worker characteristics. This new survey also has information on fiscal variables for 2004 and 2006. The drawback of using these records is that, with the exception of 2004 and 2006, wages are not directly stated and therefore 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 1980-81 and 1990-91 Spanish HBS data for heads of household in full-time employment, and CHBS data for 1985-1986, 1990-1991, 1995-1996 and 2000-2001. The reasons for this choice of data sources are given in Appendix A.

Some important characteristics of these data are reported in Table 1 which shows that individual yearly earnings, in 2000 euros, rose from 13,060.0 in 1980 to

(9) Based on a firm-level sample that may not be representative of Spanish workers.

15,108.0 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. Over the 20-year period, real average wage growth was uneven. The minimal increase in real wages that took place between 1980 and 1985 was followed by a substantial increase between 1985 and 1990. During the 1990s, real wage growth 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 this period, the increase in Spain was much more marked than in most other countries (Acemoglu, 2003).

**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) | 13,060.0 | 13,716.5 | 13,017.9 | 14,129.5 | 14,609.8 | 15,108.0 |
| Sd. Dv. | 1,519.5 | 1,561.0 | 1,401.0 | 1,626.1 | 1,666.8 | 1,863.3 |
| 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 data in the row marked av. wage (c.p. 2000) are average wages for each survey in constant 2000 euros. *St. dv. are* standard deviations. The data in the row marked Wage growth rate are the average growth rates per year between the year heading the column and the previous one. The data in the rows marked *age* and *schooling* are average age and years of schooling for each survey. Lastly, the row marked *women* shows the percentage of women in each sample.

Some trends in wage inequality can be analyzed without a full decomposition approach. Table 2 plots the difference between log wages at the $90^{th}$ and $10^{th}$ percentiles of the distribution (the 90/10 log wag gap), between log wages at the $90^{th}$ and $50^{th}$ percentiles (the 90/50 log wag gap), and between log wages at the $50^{th}$

and 10th percentiles (the 50/10 log wag gap). As the table shows, the 90/10 log wage gap fell by about 7 log points between 1980 and 1990. Between 1985 and 2000, 90/10 wage inequality fell by 2 points. Inequality trends for the 1980s differ qualitatively in the lower and upper halves of the distribution. Looking at the lower tail, one can appreciate that the 50/10 log wage gap fell by 8 points. This contrasts with the no change in the 90/50 gap, indicating that wage inequality was smaller in the lower half of the distribution. During the 1990s, inequality growth was concentrated entirely in the lower tail, with a sharp 27 log-point rise at the 50/10 gap. Inequality in the upper tail decreased during this period.

**Table 2**

MALE WAGE INEQUALITY CHANGES 1980-2000

| | *EPF* | | *ECPF* | | | |
|---|---|---|---|---|---|---|
| | 1980/81 | 1990/91 | 1985/86 | 1990/91 | 1995/96 | 2000/01 |
| 90th-10th | 0.901 | 0.894 | 0.957 | 0.945 | 0.985 | 0.955 |
| 90th-50th | 0.511 | 0.511 | 0.540 | 0.535 | 0.550 | 0.511 |
| 50th-10th | 0.391 | 0.383 | 0.417 | 0.410 | 0.435 | 0.444 |

Note: Each row shows that year's gap in log wages for the percentiles shown in the first column.

Analysis of different sub-periods and points in the wage distribution yields further interesting results. There is a different trend in the upper half of the distribution over the entire twenty-year period. The 90/50 log wage gap increased during the half of nineties, when GDP growth was low, but narrowed during the second half of eighties and nineties, when GDP grew more rapidly(10). 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.

(10) 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.

**Table 3**

BETWEEN-GROUP AND RESIDUAL INEQUALITY

| | *EPF* | | *ECPF* | | | |
|---|---|---|---|---|---|---|
| | 1980/81 | 1990/91 | 1985/86 | 1990/91 | 1995/96 | 2000/01 |
| primary | -0.299*** | -0.276*** | -0.278*** | -0.217*** | -0.212*** | -0.199*** |
| college | 0.279*** | 0.278*** | 0.313*** | 0.296*** | 0.341*** | 0.296*** |
| age | 0.041*** | 0.049*** | 0.059*** | 0.061*** | 0.068*** | 0.039*** |
| agexage | -0.000*** | -0.000*** | -0.001*** | -0.001*** | -0.001*** | -0.000*** |
| constant | 12.566*** | 12.322*** | 12.079*** | 12.048*** | 11.808*** | 12.452*** |
| N | 17,972 | 7,177 | 2,497 | 1,806 | 1,453 | 1,756 |
| $R^2$ | 0.291 | 0.348 | 0.320 | 0.280 | 0.290 | 0.197 |
| Res 90-10 | 0.778 | 0.712 | 0.777 | 0.771 | 0.819 | 0.851 |
| Res 90-50 | 0.423 | 0.398 | 0.419 | 0.399 | 0.422 | 0.440 |
| Res 50-10 | 0.354 | 0.313 | 0.357 | 0.372 | 0.397 | 0.410 |

Note: The first three rows are the standard Mincerian regression coefficients for education (represented by two dummy variables, one for primary education or less; the other for at least a college education) and age (square of age and gender are not shown). The following three rows are, respectively, the non-observed, residual and within-group wage gaps at the 90th, 50th and 10th percentiles. *R*2 shows the goodness of fit; and years of schooling is the return to education based on average years of schooling instead of the dummies in the above Mincerian regression. *** means significance at 1% level.

Wage inequality trends may be driven by changes in between-group or within-group wage dispersion, or by changes in labour force composition. As a preliminary means of isolating between-group inequality changes due to changes in education premium, we estimate Mincerian wage regressions. Table 3 shows different measures of the return to education, all obtained using Mincerian wage equations. All premium/penalties are obtained from a Mincerian regression of wages on age and age squared(11) In particular, the table shows the education wage premium for male college-educated workers relative to those with only a secondary education (col-

(11) These results are derived from a standard Mincer equation that takes the following form:

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

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

lege), and also the wage penalty for having only a primary education relative to having completed secondary education. The results show that wage differentials between primary- and secondary-schooled workers narrowed continuously between 1980 and 2000, forming a trend that should have reduced wage inequality. By contrast, the college/high school wage premium evidenced a more uneven trend, increasing between 1980 and 1985 and between 1990 and 1995, but falling in all the remaining periods. Hence, the college-high school premium appears to follow a countercyclical pattern.

Age/experience contributed to increasing wage inequality between 1980 and 1995. There is a clear increase in the inter-group wage gap due to the return-to-experience effect. During the 1980s is a sharp increase of about 30%. The trend thereafter is roughly similar, but becomes slightly less marked before breaking in 1995. The last five years see a contraction in BI inequality given by age/experience. This results are very similar to that found by Felgueroso, Hidalgo and Jiménez (2010).

Table 3 also shows statistics for residual wage inequality. For now, we simply evaluate changes in 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 (which is fairly standard in Mincerian estimates). The rows in Table 3 reveal a contrast before and after 1990. Despite the increase in the second half of the 1980s', the trend for the decade was negative, due entirely to the considerable decrease in residual wage inequality that took place during the first five years. Thus, the increase in wage inequality since 1985 appears to be due to residual inequality.

**Figure 1**

OVERALL WAGE INEQUALITY, COLLEGE AND PRIMARY/HIGH-SCHOOL WAGE PREMIUM AND RESIDUAL INEQUALITIES

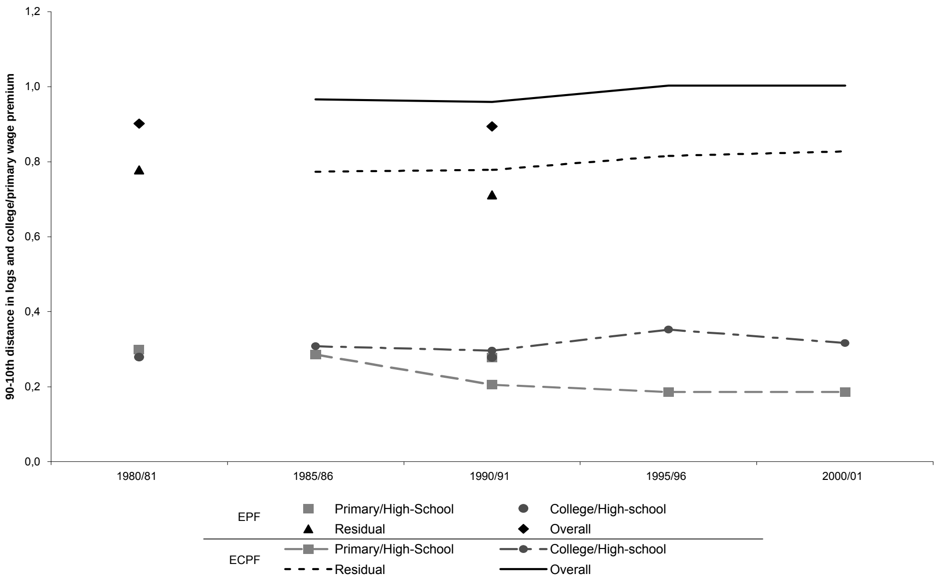



Figure 1 summarizes the trends in 90/10, the college wage premium, the primary wage penalty, and residual inequality (based on data from Table 3). Differences in wage inequality trends between decades are evident. The graph clearly shows that wage inequality fell during the 1980s, increased between 1990 and 1995, and remained constant from then 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, until 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. From 1985 to 1990, changes in wage inequality seem to mirror changes in BI.

The main limitation of our preliminary analysis (based on OLS estimation) is that it does not allow for differences in price evolution or explain the effect of shifts in the labour force composition on wages at different points in the wage distribution.

Moreover, our approach may also 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 technique that can be used to resolve these issues.

## 5. METHODOLOGY

The decomposition technique proposed by MM and extended by 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 $Q_\theta(w_t \mid x_t)$ for $\theta \in (0,1)$ be log wages ( $w_t$ ) at the $\theta^{th}$ 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

$$Q_\theta(w_t \mid x_t) = x_t'\beta_t(\theta). \qquad [2]$$

If $Q_\theta(w_t \mid x_t)$ 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 $\theta$ from a uniform distribution over the open interval (0,1), (ii) estimating the price vector $\beta_t(\theta)$ associated with $\theta$; (iii) calculating $x_t'\beta_t(\theta)$. In general, the specification in (2) must be thought of as an approximation. The accuracy of the approximation, which depends on the particular application,
is easy to check. For the case of Spain, we will show that the resulting approximation is quite accurate.

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

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

where the $\rho_\theta$ is a check function (Koenker and Basset, 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 $\{w_{it} = x'_{it}\hat{\beta}_t(\theta_i)\}_{i=1}^m$, where hats denote estimated values.

This describes the simulation of the wage data for any given $x$, but does not provide the marginal density of $w$, the calculation of which depends also 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 $\theta_i$ from a uniform (0,1) distribution; (iii) obtaining wages as $w_i \equiv x'_i\beta(\theta_i)$ (Machado and Mata, 2005).

The MM conditional quantile decomposition procedure has two important properties. First, like the Oaxaca-Blinder OLS procedure, it separates the observed wage distribution into price and quantity components. However, while the Oaxaca-Blinder procedure characterizes only 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 in research aimed at decomposing wage inequality into composition effects, between-group inequality, and residual inequality. Second, under the assumption that aggregate quantities of skills in the labour market do not affect skill prices (a strong but convenient assumption), the conditional quantile model can be used to simulate the way in which the distribution of wages reacts to shifts in labour force composition or changes in skill prices. For example, to see what wages would have prevailed with the labour force composition for period t $g_t(x)$, and the labour market prices for 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 $\beta(0,5)$ (12).

$$\beta^{w}(\theta) \equiv \beta(\theta) - \beta^{b} \equiv \beta(\theta) - \beta(0,5) \qquad [4]$$

Hence, there are now three ingredients in each wage simulation exercise: (i) estimated within-coefficient vectors $\beta^{w}(\theta)$ for a large number of $\theta$ 's spread over the open unit interval; (ii) an estimated between-coefficient vector $\beta^{b} \equiv \beta(0,5)$; (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 $\beta^{w}(\theta)$ only), between-group inequality (changes in the wage distribution when changing $\beta^{b}$ only), and labour force composition effects (changes in the wage distribution when changing $g(x)$ only).

## RESULTS

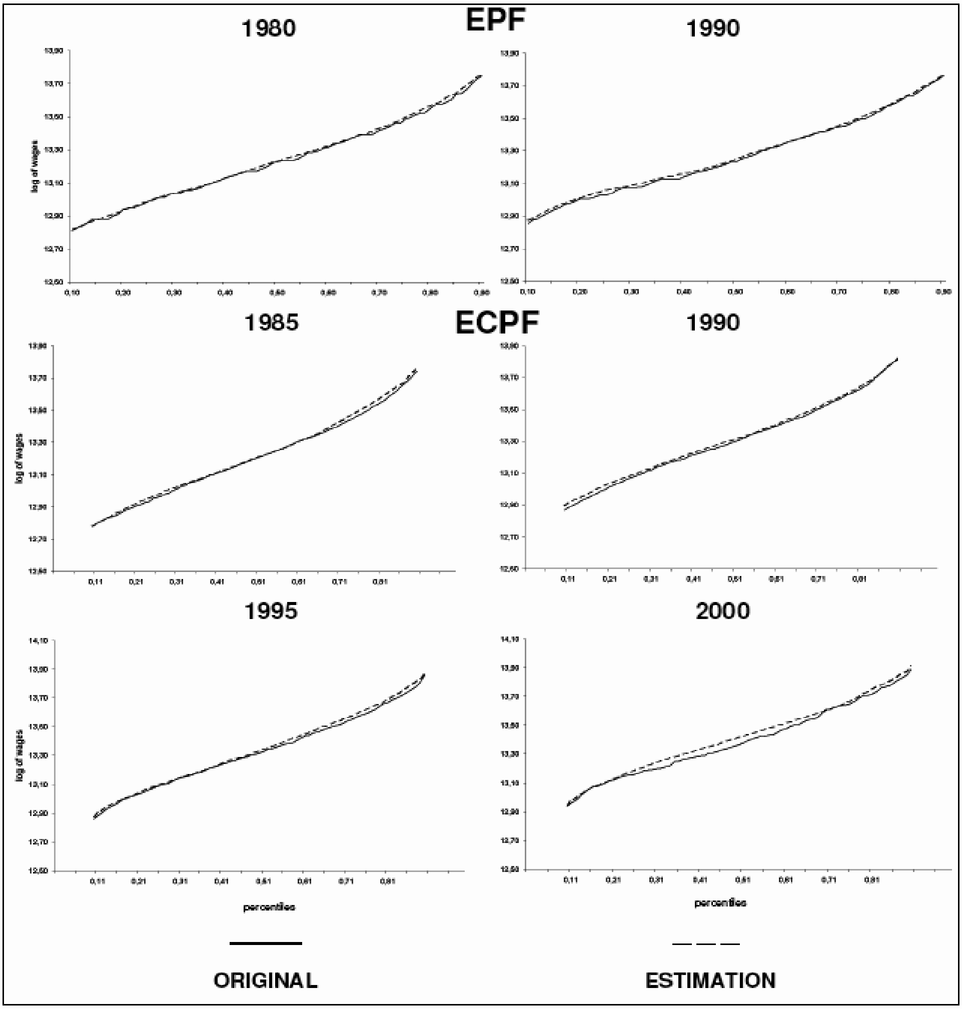
This section describes an exercise to decompose wage inequality change in four steps. First, select four thousand $\theta$ s for each year from a uniform distribution $U(0,1)$. Next, estimate the quantile regressions for each percentile, then implement the counterfactual exercise described in the previous section, and, finally, make the comparison.

Prior to the analysis of the results, a test is performed to check the accuracy with which the method used the conditional quantiles. Figure 2 describes both the original and the simulated quantiles. Since the simulated percentiles match the originals, the accuracy of this procedure for Spanish data is almost perfect, except for 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 extent 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 between 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.

(12) See Appendix B for a proof.

**Figure 2**

ORIGINAL VERSUS QR ESTIMATED WAGE DISTRIBUTIONS' PERCENTILES




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

### *Overall Wage Inequality*

Table 4 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 3. For the whole period, the result is a slight decrease. Despite the lack of observations for the 1980-1985 period, inequality might have fallen during the first half of the eighties. This intuition is driven by the 7,2 percent fall in wage inequality between 1980 and 1990 shown by the EPF, when the ECPF shows it to have contracted at a lower rate, 2,2 log points, during the second half of the period under study.

**Table 4**

100 X LOG CHANGES IN OVERALL, BETWEEN-GROUP AND RESIDUAL INEQUALITY

| | *EPF 80/81-90/91* | *ECPF 85-00* | | |
|---|---|---|---|---|
| | | 85-90 | 90-95 | 95-00 |
| Overall | | | | |
| 10-50 | -3.4 | 1.1 | 1.8 | -0.5 |
| 50-90 | -3.7 | -3.3 | -0.1 | -1.8 |
| 10-90 | -7.2 | -2.2 | 1.7 | -2.4 |
| Prices | | | | |
| 10-50 | -5.5 | -0.6 | 1.6 | -1.0 |
| 50-90 | -5.2 | -3.4 | -3.1 | -4.1 |
| 10-90 | -10.7 | -4.1 | -1.5 | -5.2 |
| Quantity | | | | |
| 10-50 | 2.1 | 1.7 | 0.2 | 0.5 |
| 50-90 | 1.5 | 0.1 | 3.0 | 2.3 |
| 10-90 | 3.6 | 1.8 | 3.2 | 2.8 |
| Between | | | | |
| 10-50 | -0.4 | -1.1 | 0.6 | -4.2 |
| 50-90 | -1.5 | -1.2 | 0.1 | -6.3 |
| 10-90 | -1.9 | -2.3 | 0.7 | -10.6 |
| Residual | | | | |
| 10-50 | -5.1 | 0.5 | 1.0 | 3.2 |
| 50-90 | -3.7 | -2.3 | -3.2 | 2.2 |
| 10-90 | -8.8 | -1.7 | -2.1 | 5.4 |

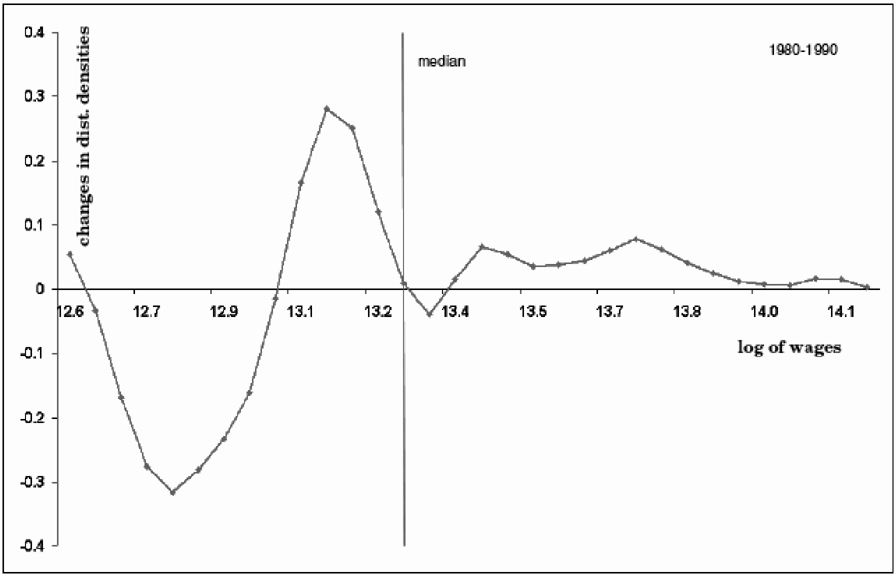
Note: Each value represents changes in log percentile distances. *Overall* defines simulated wage distributions using the MM algorithm. *Prices* represents counterfactual distributions when the only change is in returns to skills (education and experience), *quantity* when the only change is in the distribution of attributes, *between* represents changes between counterfactual densities when the only change is in prices when returns to skills remain the same across all locations in the wage distribution, and *residual* shows the evolution of inequality using counterfactual densities when evaluating only shifts in return distributions.

When examining the two halves of the density rather than the entire distribution, a familiar question arises. As pointed out earlier, the two halves do not always show the same trend. While the upper tail decreased at a higher rate than the lower tail during the eighties and the 1995-2000 periods. Moreover, during the period 1985-1995, inequality increased in the lower tail while decreasing in the upper tail. Thus, the drivers of total wage inequality seem to differ above and below the median(13).

**Figure 3**

CHANGES IN WAGE DISTRIBUTIONS

EPF

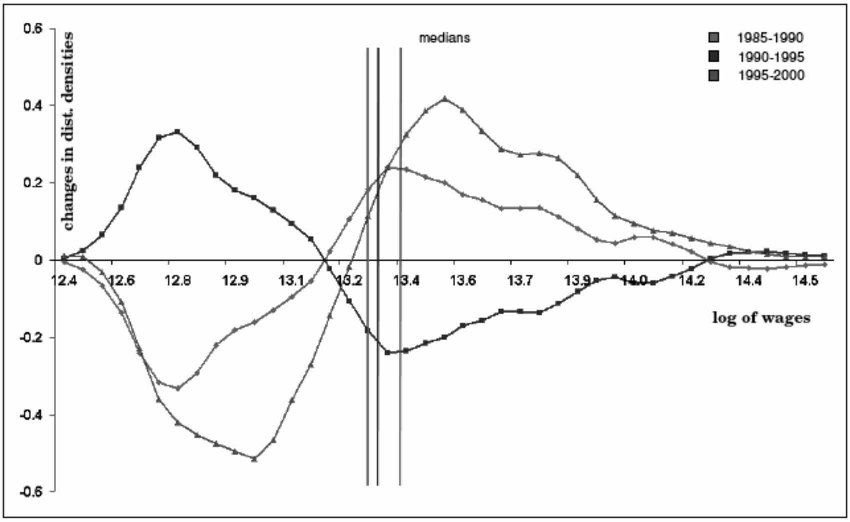



(13) This is also AKK's main finding for the United States.

**Figure 3**

CHANGES IN WAGE DISTRIBUTIONS

ECPF




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(14). 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 close to 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 4. First, the approximately 2,2% decrease in inequality can be explained by the very high concentration of density around the median instead of around the tails. However, the drop in density in the upper half can be explained by the increasing trend of the median, which is not followed by the higher percentiles. Between 1990 and 1995, a dispersion increase occurs that is particularly evident in the lower part of the distribution. Finally, the pattern between 1995 and 2000 mirrors that of the 1980s.

(14) The densities 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.

In view of this information, our first impression is that upper tail inequality adopts a countercyclical pattern from 1985 onwards. This well-known issue finds its parallel in inequality analysis for other countries. Dimelis and Alexandra (1999) for example, find that different US and UK inequality measures show negative correlation with the GDP business cycle (measured as the difference between GDP and its long-term trend). However, 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, the pattern of wage inequality is uneven rather than steady, presenting differences above and below the median. Furthermore, Spanish wage inequality shows a countercyclical pattern shared by 50th-90th inequality since 1985, while 10th-50th inequality increased from 1985 to 1995 and decreased slightly from 1995 onwards. This is our first result.

*Composition and Prices*

To explore the possible causes of this dynamic, we need to 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 10,7 log point decrease in 90th-10th inequality is caused by prices, while quantities play a lesser offsetting role (3,6 log points). The predominance of prices is also evident during the second half of the 1980s. Nevertheless, during the 1990s, the price effect decreases and composition effect increases 90th-10th inequality by -1,5 and 3,2 log points, respectively, until 1995, after which prices resume the dominant role, decreasing inequality by 5,2 log points while the composition effect increases inequality by 2,8 log points. Both effects are important, however. This result is coherent with a major change that took place in the labour force composition in Spain during this period, especially after 1985. This was the increased participation of highly educated workers(15). This change affects 10-50th and 50-90th inequality in the same direction, but is more intensive in the upper tail from 1990 onwards(16).

Despite the compositional factor, the pace marked by total wage inequality repeats the observable price trend. Thus, our data show that prices reduced inequality both during the 1980s and the 1990s, but less intensively during the first halves of those two decades. Nevertheless, the 1980s reduction might have been concen-

(15) As can be seen in Table 1, the increase in average years of schooling that stems from these surveys was around 0.65 in the eighties, while increasing by 0.67 from 1990-1995, and by 0.76 from 1995-2000.

(16) This is coherent with major change due to increasing educational attainment, mainly affecting higher wage levels.

trated in the first half of that period. Except for the 1985-1990 period, the effect of these changes mainly shows up in the lower half of the distribution, which governs the total wage inequality trend. Despite the important role played by the composition effect, the overall evolution of wage inequality mirrors price inequality. These results are coherent with those of previous studies (e.g., Abadíe, 1997 and Izquierdo and Lacuesta, 2007).

Our second result, therefore, is that both price and labour composition played an important role in the evolution of Spanish wage inequality during our study period. Furthermore, while the composition effect is important, always driving overall wage inequality upward, total wage inequality changes mirrored the changes in price inequality during this period. Thus, the distances from the 90$^{th}$ to the 50$^{th}$ percentile, when only price changes decreased, and from the 50$^{th}$ to the 10$^{th}$ percentile present a countercyclical pattern from 1985 onwards.

*Between- and within-group inequality*

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

As shown in the lower panel of Table 4, BI captures a symmetric trend on both sides. Here, 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 features being the increasing trend in the 10$^{th}$ - 50$^{th}$ tail from 1985 to 1995 and the decreasing trend in the 50$^{th}$ - 90$^{th}$ tail in that same period, followed by an increasing trend from them to the end of the study period. In RWI there is no symmetry during any part of the study period. Our third result, therefore, is the difference between the evolution of BI and RWI. Clearly, these two separate forces jointly draw the above-described picture of a price-driven change in inequality. With respect to the reasons for the BI and RWI trends, an initial guess is that they are the result of different causal factors.

*Education and Experience*

Education

Building on previous literature on the effects of returns to education and experience in Spanish wage inequality, a further step is to distinguish the BI for both prices. Table 5 presents the counterfactual densities obtained when only one price changes. For example, dispersion in return to education decreased throughout all of our study period except the 1990-1995 period. This again marks a countercyclical trend throughout the entire distribution, although almost all of the decrease appears in the upper half. The reason for this asymmetry could be the massive incorporation of more highly skilled workers into the labour market during this period. Bearing in mind the results discussed in section 3, the 1980-2000 fall in the high school wage premium relative to other premiums, followed by a fall in the college education wage premium, reduced inequality when only the price of education is entered. Recently, using the Muestra Contínua de Vidas Laborales Felgueroso, Hidalgo and Jiménez (2010) have found that returns to education have fallen since 1996 after a clear increase in 1988 to 1995. Izquierdo and Lacuesta (2007) and Pijoan and Márquez (2010) made the same findings. This exercise reveals that this evidence affects not only mean wage differences between a high- and low-educated workers, but also overall wage distributions.

Several explanations may be explored. For example, 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 are possible. For example, despite the fact that skill bias technical change can be seen to result from unobservable change in the production function, Katz and Murphy (1992) more recently 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 labour is higher than the elasticity of substitution between capital and skilled labour. Thus, in the middle of both decades, there was a decline both in output and in input use. Assuming that capital is a quasi-fixed production factor and that capital-skill complementarity prevails, the demand for unskilled labour reacts more strongly than the demand for skilled labour. When aggregate demand increases, the opposite effect occurs. Thus, it follows that the relative demand for skills is countercyclical due to a positively-sloped relative supply of skilled labour curve, and that the relative wage and employment of skilled labour are also countercyclical.

Supporting this explanation, Hidalgo, O'Kean and Rodríguez (2007) found strong evidence of capital-skill complementarity for Spain between 1980 and 2004, which could explain the countercyclical behaviour of Spanish wage inequality that we have been discussing. While this countercyclical pattern has also been found for other countries(17) the United States evidence only reveals countercyclical differences between skilled and unskilled wages until 1984. The trend turns procyclical in 1984, a shift that has been explained by the concomitant reduction in the degree of capi-

(17) For example Denmark (Skaksen and Sorensen, 2005).

tal-skill complementarities. (Castro and Coen-Pirani, 2005). Nevertheless, since other explanations have been suggested(18), this issue awaits a more in-depth analysis.

Also, Felgueroso and Jiménez (2010) and Felgueroso, Hidalgo and Jiménez (2010) explain that falling returns to education since 1996 may be due to the decrease in average firm tenure and the increase in over-education at the upper end of educational attainment. These two forces explain a large portion of the decrease in returns to education. The fact that the return is higher for higher-educated workers, who are concentrated in the upper tail of the wage distribution, explains the higher decrease in wage inequality among higher wages.

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.

Again Felgueroso, Hidalgo and Jimenez (2010) observe a decrease in return to experience between 1998 and thereafter. This decrease is greater for higher-educated workers, which may explain how wage inequality changes in response to the experience premium.

**Table 5**

100XLOG CHANGES IN EDUCATION AND EXPERIENCE PRICES AND QUANTITY INEQUALITY

| | *EPF 80/81 – 90/91* | *ECPF 85 – 00* | | |
|---|---|---|---|---|
| | | 85-90 | 90-95 | 95-00 |
| Prices | | | | |
| Education | | | | |
| 10-50 | -0.1 | -0.3 | 0.5 | -1.5 |
| 50-90 | -1.8 | -1.3 | 0.4 | -5.7 |
| 10-90 | -1.9 | -1.7 | 0.9 | -7.3 |
| Experience | | | | |
| 10-50 | -0.3 | -0.3 | 0.4 | -1.3 |
| 50-90 | -0.4 | 0.0 | -0.4 | -0.5 |
| 10-90 | -0.8 | -0.3 | 0.0 | -1.9 |

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

*Some tentative explanations for RWI*

Firstly, the evolution of RWI can be explained because some variables are not include in this exercise, the change in residual wage inequality may be due to the change in some other observables characteristics. For example, Davia and Herranz (2004) explain that non-fixed term contracts premium may explain wage inequality changes. Also complete constracts wages versus partial contracts wages may be behind some of this wage inequality change (Pagan, 2007 and Motellón, López-Bazo and El-Attar, 2007).

Secondly, 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 Lemieux, 2001 and Autor, Katz and Kearny, 2005a and 2005b). Second, it fits quite well with Spanish evidence regarding labour market reforms. For example, Peraita (2003) argues that the Spanish labour market was one of the most rigid in the industrialized world during the early 1980s in particular, but market labour 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 1980s might be explained by the increase in the bottom wage percentile due to the appearance of a new institutional wage-setting framework. However, the increase in the bottom half may be explained by the dualization of the labour market following the introduction of temporary contracts since 1984.

Also, 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. These factors would also explain the decrease in inequality 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 involve the generalization of education and the increase in 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 higher-educated. For example, Hidalgo (2009) explains that, after controlling for changes in non-observable skills, RWI decreases from 1995 onwards in Spain.

## CONCLUSIONS

This paper attempts to decompose the change in Spanish male head-of-household 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 (2006 & 2008): changes in between- and within-group inequality and changes in labour composition.

The results for the 1980-2000 period are fourfold. First, instead of the Spanish wage distribution behaves countercyclically the inequality change is negative. Second, both price and labour composition play important roles in the evolution of Spanish wage inequality. However, while the 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 between conditional and observable characteristics in wage distributions, fell during the late 1980s and 1990s but increased during the early years of both decades, while residual wage inequality, which measures the dispersion within the conditional distributions and that caused by non-observable variables, decreased from 1980 to 1995 and increased onwards. Another 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 in each half. Fourth and finally, the observed inequality for education and experience mirrors previous results for Spain and, again, behaves countercyclically in nineties.

There are several possible explanations for these results. While changes in between-group inequality 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 (higher-educated) workers might 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 behaviour, such as changes in Spanish labour institutions in favour of collective bargaining approved between 1980 and 1985, and market reforms since the early 1990s. These intuitions, nevertheless, require deeper exploration.

## APPENDIX A

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

Although these sources do not provide homogeneous wage statistic series, the important and relevant information they supply to this kind of analysis has made them 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 household has worked for more than thirteen hours during the reference week. The only solution to this problem is to use only those workers who report having worked 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 household(19). Therefore this paper, like other Spanish studies (Abadíe, 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. The key differences between these sources centre on the amount of information, number of characteristics reported, the richness of the classifications within each characteristic (with regard, say, 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

(19) For instance, the ECPFs records contain education data for this group only.

over the years and partly because the Spanish legal definition of education changed during the period under study(20).

Table 6 shows education categories based on years of schooling rather than educational attainment, because the former is homogeneous across all years and classifications. The years computed are the same as those given in Vila and Mora (1998).

To conclude, despite their limitations, these surveys possess many redeeming features that make them our best choice for data on the distribution of Spanish wage inequality between 1980 and 2000. Briefly, two main criteria may be cited in support of their use: first, the lack of 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 household 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 household is the only wage-earner. 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 were due to erroneous replies. Thus, when reported annual wages fell below the legal minimum for a particular year, it was necessary to correct for zero-censoring. Table 7 shows the minimum annual wage value in 1980 constant prices, which is 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 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 by using all four quarters in the calendar context.

(20) The education act of 1970 established a school leaving age of fourteen. In 1990, it was raised to sixteen by an educational reform act (the LOGSE), which changed school grades as the 1970 act had.

**Table 6**

SCHOOLING GROUPS. INDIVIDUAL DATA

| | *EPF* | | *ECPF* | |
|---|---|---|---|---|
| | 80/81 | 90/91 | 85/95 | 00/01 |
| Illetary-Primary | 1-3 | 1-4 | 1-4 | 1-2 |
| Low Secondary | 4 | 5 & 7 | 4 | 3 |
| Acad. & Voc. Upper Secondary | 5-6 | 6 & 8 | 5 | 4-5 |
| Higher Education (Short cycle) | 7 | 9 | 6 | 6 |
| Higher Education (Long Cycle) | 8 | 10 | 7 | 7 |

**Table 7**

CENSORED DATA

| | EPF | | ECPF | | |
|---|---|---|---|---|---|
| | 1980 | 1990 | 1985 | 1990 | 1995 |
| Censored wage (euros) | 1,882.1 | 1,728.3 | 1,756.6 | 1,728.8 | 1,677.2 |
| Percentage (%) | 9.44 | 3.16 | 8.18 | 3.83 | 4.47 |

Source: INE and own elaboration.

## APPENDIX B

This Appendix describes the expression (within) that captures within-group inequality. The key is that quantile prices exhibit 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 Basset, 1982). In this case, the quantile coefficients show changes in the percentiles. More specifically, suppose that the 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 $\varepsilon_{it}$ 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

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

Suppose, for the sake of simplicity, that $\sigma(x_t) = x_t$. Then, for the case of (quantile), 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 $\varepsilon$ and represents the value of the percentile $\tau$. 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 (quantile). 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 (betas). 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.

## REFERENCES

ABADÍE,A.(1997): «Changes in Spanish Labor Income Structure During the 1980's: A Quantile Regression Approach,» *Investigaciones Económicas*, XXI(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 (2006):«Rising Wage Inequality: The Role of composition and Prices», *NBER Working Papers 11628*, National Bureau of Economic Research, Inc.

AUTOR,D.H., L.F. KATZ, AND M.S. KEARNEY (2008): «Trends in U. S. Wage Inequality: Revising the Revisionists», *The Review of Economics and Statistics*, nº 2, vol. 90, pp 300-323.

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. 31

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.

DAVIA, A AND V. HERRANZ (2004):«Temporary employment and segmentation in the Spanish Labour Market: An Empirical Analysis through the study of wage differentials». *Spanish Economic Review*, vol. 6, nº4, pp. 291-318.

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), 1001-44.

EPANECHNIKOV,V. (1969): «*Theory of probability and Its Applications*chap». *Non-parametric Estimation of a Multivariate Probability Density, pp.* 153-158, no. 14.

FELGUEROSO, F. AND S. JIMÉNEZ (2009): «The "New Growth Model". How and with Whom?» DT FEDEA 39-2009.

FELGUEROS0, F. HIDALGO, M. AND S. JIMÉNEZ (2010): «Explaining the fall of skill premium in Spain», Mimeo.

HIDALGO, M.A., J. O'KEAN, AND J. RODRÍGUEZ (2007):«Demanda de Trabajo y Tecnologías de la Información: Evidencia para España», *Working Paper nº 08*.12. Universidad Pablo de Olavide.

HOSPIDO, L. AND C. BOHOMME (2010): "Using Social Security Data to Estimate Earnings Inequality", Mimeo.

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), 410-42.

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 Stata Papers 40.

KOENKER, R., AND G. BASSETT (1978): «Regression quantiles», *Econometrica*, (46), 33-50.

KOENKER,R., AND G. BASSETT (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):«Descomposing Changes in Wage Distributions: A United Approach», *The 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.

MOTELLÓN,, E, LÓPEZ-BAZO, E. AND M. EL-ATTAR (2007): «Cambios en la distribución salarial en España, 1995-2002. Efectos a través del tipo de contrato». Doc Trab. IREA.

MACHADO,J.A.,ANDJ.MATA(2005): «Counterfactual Decomposition of Changes in Wages Distributions Using Quantiles Regression», *Journal of Applied Econometrics*, (20), 445-465.

OLIVIER, J. RAYMOND, J.L., ROIG, J.L. AND BARCEINAS, F (2000):«Los rendimientos de la educación y la inserción laboral en España», *Papeles de Economía Española, 2000;* (86), pp. 128-148.

PAGÁN, R. (2007): «Diferencias salariales entre el empleo a tiempo completo y parcial», *Revista de Economía Aplicada* nº 43 (vol. XV), pp. 5 – 47.

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

PIJOAN, J. AND V. SÁNCHEZ-MARCOS (2010): «Spain is Different: Falling Trends of Inequality», *CEPR Discussion Paper*, 7489.

POURPOURIDES,P.M.(2007): «Implicit Contracts, Wages and Wage Inequality over the Business Cycle», *Cardi® Economics Working Papers* E2007/19, Cardiff University, Cardiff Business School, Economics Section.

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*. 33

# DESIGUALDAD SALARIAL EN ESPAÑA, 1980-2000: EL CASO DE LOS CABEZAS DE FAMILIA


## RESUMEN

Utilizando regresiones cuantílicas, se descompone la evolución de la desigualdad salarial en España para el período 1980-2000 entre cambios en los precios, en la composición del mercado de trabajo y entre las características no observables de los trabajadores. Los principales resultados son tres: primero, la desigualdad salarial sigue un proceso contra-cíclico desde mitad de los ochenta; segundo, los precios y la composición juegan un importante papel en esta evolución; y tercero, mientras que los cambios en la desigualdad motivado por los precios es similar a ambos lados de la mediana, los cambios en la desigualdad motivada por cambios en las características no observables se incrementa en la parte "baja" de la distribución de salarios desde 1985, mientras no crece en la parte alta. Por último, algunas explicaciones son ofrecidas ante estos resultados.



*Palabras clave:* desigualdad salarial, regresión cuantílica

*Clasificación JEL*: J24, J31