Gender Wage Gaps by Education in Spain: Glass Floors vs. Glass Ceilings*

Juan J. Dolado* and Vanesa Llorens**

(*) Dept. of Economics, Universidad Carlos III;
(**)NERA (Madrid)
This draft: 23 October, 2003

Abstract

This paper analyses the gender wage gaps by education throughout the wage distribution in Spain. Quantile regressions are used to estimate the wage returns to the different characteristics at the more relevant percentiles. A correction for the selection bias is included for the group of less educated women. The Oaxaca-Blinder decomposition is then implemented at each quantile in order to estimate the component of the gender gap not explained by differences in characteristics. Our main findings are twofold. On the one hand, when dealing with the group with tertiary education, we find higher discrimination at the top than at the bottom of the distribution, in accord with the conventional "glass ceiling" hypothesis. On the other, for the group with primary and secondary education, the converse result holds, pointing out to the existence of lower wages for women at the bottom of the distribution due to their prospects of lower job stability, a phenomenon that we refer to as "glass floors".

JEL Classification: J16, J71.

Keywords: gender gap, quantile regressions, glass ceilings, glass floors.

---

*Corresponding author: Juan J. Dolado (E: dolado@eco.uc3m.es). We are grateful to Manuel Arellano, Samuel Bentolila, Florentino Felgueroso and Marcel Jansen for useful comments and advice.

1 Introduction

It is a widely documented fact that men earn higher wages than women even after controlling for measurable characteristcs affecting their productivity (see, e.g., Blau and Kahn, 1997). In this respect, there is a large available literature on gender wage differentials and discrimination based on analysing average wages where measures of the so-called *gender wage gaps* (*gender gap* in short) can be interpreted as estimates of discrimination at the mean of the observed distribution of wages. However, the studies on the gender wage gap at other points of the wage distributions are much more scarce.[1] Lately, however, there has been a growing concern about how the gender gap evolves throughout the wage distribution to test whether wage discrimination is greater among high earners or among low earners.

In this paper, following the approach advocated by Albrecht *et al.* (2003) to study gender wage differentials in Sweden, we derive quantile measures of the gender gap in Spain at the end of the 1990s. This is an interesting issue, since Spain, like other Southern-mediterranean countries, has still a much lower female participation than the Nordic countries and therefore patterns of women´s achievements in the labour market may differ markedly from those found for the former countries.[2] As will become clear below, we find it useful to distinguish between workers with higher (tertiary) and

[1] Chamberlain (1994) and Buchinsky (1994, 1995a, 1995b) used quantile regressions to analyze the wage structure in the U.S. An application of this method to Spain can be found in Abadie (1997) and more general applications in Fitzenberger *et al.* (2001). More recently, such estimation techniques have been used to study gender wage discrimination in several former communist countries (Newell and Reilly, 2001), Sweden (Albrecht *et al.*, 2003) and in Spain (García *et al.*, 2001, and Gardeazabal and Ugidos, 2003).)

[2] The Spanish female activity rate (% of population aged 15-64) in 2001 is 51.4% whereas it reaches 73.4 % in Sweden. By educational levels, the corresponding rates are 80.4% and 48.0% for the women with tertiary education and less than tertiary education (84.6% and 68.3% in Sweden), respectively (see OECD, 2002).

lower (primary /secondary) educational attainments in contrast to most of the available studies on this topic. The reason for doing so is that the behaviour of gender gap throughout the distribution differs in an interesting fashion between both groups of workers. Using the 1999 (6th. wave) of the European Community Household Panel (ECHP, henceforth) for workers working more than 15 hours per week in that year, Figure 1a shows the gender gap (measured as the difference in the logged gross hourly wages of male and female workers) in Spain, together with the mean gender gap (solid line)and the trimmed mean.[3]As can be observed, there is a decreasing trend that becomes stable around the 60th percentile and then increases sharply at the higher quantiles. As expected, the gender gap at the mean differs notably from the wage gap at the various percentiles. This non-monotonic evolution, however, stands in sharp contrast to the one found for Sweden where the gap is largest at the top of the wage distribution, giving rise to a *glass celing* phenomenon which Albrecht *et al* (2003) analyze in great detail.

[Figures 1a, b, c about here]

Figures 1b and 1c depict the correponding quantile gender gaps for the two groups of workers described above. As can be observed, while the evolution of the gap is decreasing along the distribution for low-educated workers, it fits well with the glass-ceiling hypothesis for the high-educated ones. Thus, there seems to be a "composition effect" which deserves greater scrutiny. Interestingly, northern and central european countries, such as Denmark and France (Figures 2a and 2b), have an increasing gap for high wages, while in southern european countries, such as Italy and Portugal (Figures 2c and 2d),

[3] We have excluded the extremes of distribution from the graph because their behaviour is erratic. Therefore, trimmed mean represented in the graphs is computed with the observations between the 5th. and 95th. percentiles.

the behaviour is more irregular and resembles the one found for Spain.

[Figures 2a, b, c ,d about here]

Three possible explanations spring to mind regarding these divergent patterns by education:

1. *First*, the OECD (2002) warns that some of these results could be due to measurement errors stemming from the fact that the interviewed persons provide direct information about their own wages, rather than their employers as is the case with matched employer-employee data. If those earning more, mainly men, have a larger propensity to understate their wages, the gap for the higher quantiles would be underestimated. Although this argument could explain the low glass ceilings, it does not explain the pattern found at the bootom of the distribution for the L-group.

2. *Secondly*, in northern and central european countries, female participation in the labor market is much higher than in southern european countries, despite the catching-up process which has taken place during the last two decades. To the extent that less-educated women ´s careers in the labour maket suffer from frequent interruptions in the latter countries, due to societal discrimination in family-duties, employers may use statistical discrimination to lower their wages vis-á-vis more stable men in the lower part of the wage distribution. As their job tenure expands, women become more reliable to employers´s eyes and their wages converge to men´s. In parallel with the glass ceiling phenomenon, we will label this declining pattern as “glass floors”.

3. *Thirdly*, women with tertiary education are bound to be considered much more stable in their jobs, given the human capital investment that they have undertaken. Thus, their wages will be similar to men´s wages at their entry jobs which typically correspond to the lower part of the distribution. As we move along the wage distribution, however, women´s wages fall below mens´s. One explanation for this pattern is known as "dead-end". It argues that women are promoted less frequently because they have jobs with less opportunities of promotion. For example, Polachek (1981) predicts that women choose occupations where the cost of career interruptions is low and the fact that occupational segregation by gender segregation exists in the labor market would validate this argument.[4]. Another explanation, i.e., the "glass-ceiling" phenomenon, explains the fact that women have a lower probability to be promoted to jobs with higher responsabilities even in the case that men and women have ladder jobs and have the same ability distributions. The model by Lazear and Rosen (1990) confers a higher productivity in the household to women, that leads employers to be reluctant to invest in their training on an equal basis with men. Only the more productive women would be promoted.

For these reasons, and trying to achieve a higher degree of homogeneity, we have divided the population into the two groups described above which, for convenience, will be denoted in the sequel as the L (low-educated) group and the H (high-educated) group. Indeed, in the case of Spain, the group of working women is formed by very heterogenous cohorts. Since the 1980s, female participation has upsurged (raising from 33.3% in 1980 until 51.7%

[4]Dolado *et al.* (2003) carry out a comparative study of gender occupational segregation between US and the EU.

in 2001) mainly due to the increase in tertiary education and the decrease in fertility rates.[5]

In the L-group, we are considering women that are often classified by employers as "second-earners" in the households with a higher probability of leaving the labour market. On the contrary, tertiary education plays the role of signaling a stronger commitment with the labor market and therefore, *a priori*, men and women will be treated as equivalent workers. However, women have a lower probability of promotion which increases the wage gap among the top of the distribution. In either group, the gender gap displayed in the previous Figures could be attributed to a lower productivity of women or to a lower return for a given characteristics, usually related to the discrimination component or to unobserved variables. In order to disentangle these two components we will follow the standard decomposition procedure, albeit introducing some modifications since we are analysing gaps at the quantiles instead of at the mean wage. First, we estimate quantile regressions (QR) to obtain the return to the productive characteristics for men an women. Applying an extension of the well-known Oaxaca-Blinder decomposition, we can isolate the two effects that we are looking for. Moreover, given their lower participation, a sample selection correction is introduced for women in the L-group.

There are several techniques for decomposing gender gaps at different percentiles. In this paper, we will follow the methodology proposed by Machado y Mata (2000), which is the one used in Albrecht *et al.* (2003).

Finally, as regards the related literature for Spain, there are two recent papers that apply the QR methodology to the study of wage discrimination

[5] Arellano and Bover (1995) conclude that those are the two main factors explaining the rise in female participation, once their endogeneity is properly dealt with.

using different data sets. Neither of them distinguishes by education. García *et al.* (2001), using the 1991 *Encuesta de Conciencia, Biografía y Estructura de Clase*, control both for the endogeneity of education as well as for the selection of women into the labor market and conclude that the discrimintion component, in absolute and relative terms, is higher in the top of the wage distribution. Gardeazabal and Ugidos (2002), who use the 1995 *Encuesta de Salarios*, also find that the raw gender gap increases along the distribution but, by contrast, find that the discrimination component in relative terms is higher at the bottom of the distribution.

The rest of the paper is organised as follows. In Section 2, we describe two simple theoretical models which are consistent with the stylised facts pertaining to the L and H-groups. Section 3 is devoted to describe the QR methodology and to discuss the data employed and the results of the gender regressions. In Section 4 we present the wage gap decomposition at the mean and at the unconditional quantiles. Finally, Section 5 concludes. Appendix A offers a detailed description of the data while Appendix B discusses the quantile gender gap decomposition when there is a sample selection bias correction.

2 Interpretative models

2.1 L-group

To account for the stylized facts in the quantile evolution of the gender wage gap for the L-group, let us use a simple model in the spirit of Acemoglu and Pischke (1998).

Let us assume that workers are endowed with an ability δ whose c.d.f., $G(\delta)$, is identical for men and women. Low-educated workers need to get specific training to perform a job so that two periods are considered. In the initial period, workers receive training so that firms bear an investement cost leading to a productivity $\gamma_1\delta$ with $0 < \gamma_1 < 1$. At this point firms do not know the worker´s productivity which becomes revealed at the begining of period 2. The training leads to a higher productivity $\gamma_2\delta$ in period 2, such that $\gamma_1 < 1 < \gamma_2$. Workers are certain to work and get trained in period 1, but they only work in period 2 if the wage in that period exceeds the nonmarket alternative value of time, ω, which represents a disutility shock forcing them to quit the job (say for family reasons), as in Acemoglu and Pischke (1998). The ω shock is a a random variable with c.d.f $F(\omega)$ that is revealed to the worker after the wage in period 2 is chosen by the firm. Thus, workers will not quit as long as the wage in period 2, W_2, exceeds the realization of ω, i.e. $W_2 - \omega \geq 0$. Moreover, to stress the monopsony argument emphasised by these authors, neither wage renegotiation nor wage offers from outside firms for job quitters are considered.

The key difference between men and women is that the c.d.f. for men, $F_m(\omega)$, is stochastically dominated by the c.d.f. for women $F_f(\omega)$, namely $F_m(\omega) > F_f(\omega)$ for $\omega > 0$. Through this assumption it is captured the fact that women have higher outside opportunities (at home) than men. To simplify the algebra, and without loss of generality in terms of the qualitative results, we will assume that $dG(.)$ and $dF(.)$ are uniform distributions, such that the density functions $g(\delta) = U[0, \tau]$, $f_m(\omega) = U[0, \varepsilon_m]$ and $f_f(\omega) = U[0, \varepsilon_f]$, with $\varepsilon_f > \varepsilon_m$.

Under the assumptions that the wage in period 2, W_{2i}, is offered before ω is realized, that firms know δ in that period and that no wage renegotation is allowed, they will choose W_{2i} to maximise maximize profits in period 2, Π_2,that is

$$\max_{W_{2i}} \int_0^{W_{2i}} (\gamma_2\delta - W_{2i}) dF_i(\omega) = \max_{W_{2i}} \left[\frac{\gamma_2 \delta W_{2i}}{\varepsilon_i} - \frac{W_{2i}^2}{\varepsilon_i} \right], \ i = f, m, \quad (1)$$

whereby the first-order condition w.r.t. W_{2i} implies that the same wage will be paid in equilibrium to workers of each gender with observed productivity δ, namely $W_2^* = \frac{\gamma_2 \delta}{2}$. Thus, the gender wage gap in period 2 will be zero.

Next, having chosen W_2^*, under a free-entry assumption, firms choose the training wages in period 1, W_{1i}^*, so that there are zero expected profits when hiring, that is

$$\int_0^{\tau} \Pi_2(W_2^*) dG(\delta) + \int_0^{\tau} \gamma_1 \delta dG(\delta) - W_{1i}^* = 0, \quad (2)$$

so that

$$W_{1i}^* = \frac{\gamma_1 \tau}{2} + \frac{\gamma_2^2 \tau^3}{12 \varepsilon_i}. \quad (3)$$

Given the higher quitting probability of women, the gender wage gap in period 1 will be $W_{1m}^* - W_{1f}^* = \frac{\gamma_2^2 \tau^3}{12 \varepsilon_m \varepsilon_f} [\varepsilon_f - \varepsilon_m] > 0$. Insofar as $W_{2i}^* > W_{1i}^*$ which occurs when $(\gamma_2 - \gamma_1)\delta > \frac{\gamma_2^2}{6\varepsilon_i} \tau^2$, the previous result implies that the gender gap will be larger at the bottom of the distribution than at the top if the distribution.

The intuition for this result is quite simple. Since the disutility shock is not known at the time when W_2 is offered, the best that firms can do is to match this outside offer by setting W_2 equal to a fraction of the observed productivity $\gamma_2 \delta$ which, under a uniform distribution, equals $\frac{\gamma_2 \delta}{2}$. Hence, firms will obtain a surplus of $\gamma_2 \delta - W_2^* = \frac{\gamma_2 \delta}{2}$ in period 2 and, given the zero-expected profit condition, they have to pay a wage above $\frac{\gamma_1 \tau}{2}$ in period 1. This explain both why the wage profile is flatter than the productivity profile and why women receive a lower wage in period 1 than men (since they are more likely to quit).

2.2 H-group

As for the H-group, whose gender- gap pattern fits well with the conventional glass-ceiling phenomenon, there are several rationalizations in the literature. Amongst the most popular, there is the one provided by Lazear and Rosen (1990) in a model of job ladders. In their model firms have to choose how to place workers, namely either in a flat ladder (A, with no training), where productivity in both periods is δ or in a promotion ladder (B, with training) where productivities are $\gamma_1\delta$ and $\gamma_2\delta$ in periods 1 and 2, respectively, with the rest of the assumptions given above except that firms are competitive and pay wages in period 2 equal to observed productivities, i.e., $W_2^A = \delta$ and $W_2^B = \gamma_2\delta$. Given women´s larger propensity to quit in period 2, firms choose a more stringent cutoff ability to allocate them to the B job than the one chosen for men. Thus, denoting each cutoff by δ_f^* and δ_m^*, respectively, we have that $\delta_f^* > \delta_m^*$. This result implies that there are women with δ such that $\delta_m^* < \delta < \delta_f^*$ who are not promoted. In other words, to be promoted, a woman must be more productive than a man to compensate for her ex-ante probability of departure and the loss of the training investment. This prediction from the model is well supported by the empirical evidence. For example, Bertrand and Hallock (2000), who analyse the group of high-level executives in US corporations, observe that the main wage differences are due to the fact that women lead smaller firms, they are younger and with less tenure but they emphasize that this result does not rule out the existence of discrimination in terms of gender segregation or promotion. However, in a competitive market, the other key prediction from the model, namely that if men and women have the same underlying ability distribution, then the average wage on females in A jobs should be larger than the average wage of men in that job is at odds with the available evidence (i.e., the glass-ceiling phenomenon). As Lazear and Rosen (1990) note, one way to solve this puzzle is to apply Mincer and Polacheck´s (1974) argument suggesting that different expectations by men and women of labour market participation would result in different ability distributions since women would self-select to relatively low-paid occupations where career interruptions are less penalized. Alternatively, Booth *et al.* (2003) introduce some monopsonistic power by firms and assume that women in highly-paid jobs receive a smaller number of outside offers (due to their "perceived" lower mobility) that the firm might be interested to match in order to retain the worker.[6] In either case, women in good jobs will be lower paid than men in those jobs.

3 Quantile Regressions

3.1 Methodology

Following Koenker and Bassett (1978) and Buchinsky (1998), the model of quantile regression in a wage-equation setting can be described as follows. Let (w_i, x_i) be a random sample, where w_i denotes the logged hourly gross wage of an individual i and x_i is a vector $K \times 1$ of regressors, and let $Q_\theta(w_i|x_i)$ be

[6]In a experimental framework, Gneezy *et al.* (2003) notice that men and women have different attitudes to competing. Men try harder to compete and therefore disproportionately win the top jobs, even when to do the job well does not require an ability to compete. In a similar vein, Babcock and Laschever (2003) notice that male graduates with a master´s degree at Carnegie Mellon University earned starting salaries 7.6% higher than female students, because the latter tend to accept the initial pay offer much more frequently than their male classmates. Sociological explanations based on women wanting opportunities but not a life dominated by work may be behind these attitudes.

θth-order quantile of the conditional distribution of w_i given x_i. Then, under the assumption of a linear specification, the model can be defined as[7]

$$w_i = x_i'\beta_\theta + u_{\theta i} \qquad Q_\theta(w_i|x_i) = x_i'\beta_\theta \tag{4}$$

where the distribution of the error term $u_{\theta i}$, $F_{u_\theta}(\cdot)$, is left unspecified, just assuming that $u_{\theta i}$ satisfies $Q_\theta(u_{\theta i}|x_i) = 0$.

As shown by Koenker and Bassett (1982), the estimator for the vector of coefficients β_θ, *i.e.*, $\hat{\beta}_\theta$, can be obtained as the solution of[8]

$$\min_\beta \frac{1}{n} \left\{ \sum_{i:w_i \geq x_i'\beta} \theta |w_i - x_i'\beta| + \sum_{i:w_i < x_i'\beta} (1-\theta)|w_i - x_i'\beta| \right\}. \tag{5}$$

The minimization problem in (5) can be rewritten as follows

$$\min_\beta \frac{1}{n} \sum_{i=1}^{n} \left(\theta - 1/2 + 1/2 sgn\left(w_i - x_i'\beta\right)\right)\left(w_i - x_i'\beta\right), \tag{6}$$

yielding the following $K \times 1$ vector of first order conditions for (6)

$$\frac{1}{n} \sum_{i=1}^{n} \left(\theta - 1/2 + 1/2 sgn\left(w_i - x_i'\hat{\beta}\right)\right) x_i = 0. \tag{7}$$

This formulation makes clear that it is the sign of the residuals and not their magnitude what matters, implying that quantile regressions are robust to the presence of outliers. The estimated coefficient of the quantile regression $\hat{\beta}_\theta$ is interpreted as the marginal change in the conditional quantile θ due to

[7] If the linear specification were not to be correct, we can always interpret model (4) as the best linear predictor for the conditional quantile.

[8] Although $\hat{\beta}_\theta$ is a consistent estimator for β_θ and asymptotically normal, it is not efficient. An efficient estimator requieres the use of an estimator for the unknown density function $f_{u_\theta}(0|x)$.

a marginal change in the corresponding element of the vector of coefficients on x. Its interest lies in the fact that can be interpreted as rates of return (or market prices) of the productive characteristics at the different points of the wage distribution. As Buchinsky (1998) points out, the conditional quantile offers a full characterization of the conditional wage distribution in the same way that sample quantiles characterize the marginal distribution.

Under some regularity conditions, it can be obtained that

$$\begin{aligned} \sqrt{n}(\hat{\beta}_\theta - \beta_\theta) &\xrightarrow{d} N(0, \Lambda_\theta) \qquad \text{where} \qquad (8) \\ \Lambda_\theta &= \theta(1-\theta)(E[f_{u_\theta}(0|x_i)|x_i x_i'])^{-1} E[x_i x_i'](E[f_{u_\theta}(0|x_i)|x_i x_i'])^{-1}. \end{aligned}$$

A consistent estimator under heteroskedasticity for Λ_θ can be obtained by bootstrap methods in the following manner. [9] First, we consider a sample of size n as the population of interest, so that $\hat{\beta}_\theta$ represents now the population value. Next, we take a sample (with replacement) of size n and we obtain an estimator for $\hat{\beta}_\theta$. Repeating this process B times, the estimated asymptotic variance of $\hat{\beta}_\theta$ is given by

$$\hat{\Lambda}_\theta = \frac{n}{B} \sum_{j=1}^{B} (\hat{\beta}_\theta^{(j)} - \hat{\beta}_\theta)(\hat{\beta}_\theta^{(j)} - \hat{\beta}_\theta)'. \qquad (9)$$

3.2 Data and Results

The data are drawn from the 1999 (6th. wave) of the ECHP which provides information in a harmonized format for the EU countries on earnings, employment, hours of work, education, immigrant condition, civil and health status and other socio-demographic variables. The information is obtained

[9] An alternative consistent estimator under heteroskedasticity for Λ_θ is the kernel estimator (see Koenker and Basset, 1982).

from surveys to a fixed panel of households (70,000 in the EU and around 8,000 in Spain) since 1994. Our sample is restricted to full-time workers working more than 15 hours per week and, as discussed above, we distinguish among two groups by educational attaintments. In the H-group there are 721 men and 558 women whereas the L-group is formed by 1585 men and 626 women. Appendix A contains a detailed description of the variables used in the regression models while Tables 1a and 1b offer summary descriptive statistics of both samples.[10]. As can be observed, the mean gender gaps are around 10% and 23% for the H and L-groups, respectively.[11] High-educated men have slightly more experience than high-educated women (2.1 years), are a bit older (1.8 years) and have a larger share in directives jobs (13 p.p. difference). By contrast, low-educated men are much more experienced than women (4 years) and are quite older (4 years). In both groups women have a larger share in firms with less than 20 employees and work more often in the public sector.

[Tables 1a, b about here]

We have estimated quantile regressions (at the 10th., 25th., 50th., 75th. and 90th.quantiles) where the (logged) gross hourly wage is regressed on different subsets of covariates. Heteroskedastic-robust estimation at the conditional mean has also been undertaken for comparison purposes. As is conventional in the literature on wage equations, the covariates controlled for in each of the two educational groups are: (potential) experience, seniority, civil

[10] Descriptive statistics of women in the L-group who do not work are also reported since they are used to run a probit on participation.

[11] The compared percentiles correspond to the wage distributions of men and women separately. If we were to consider the position of women in the men´s distribution, it is found that 13.6% (3.8%) of women are in the bottom (top) percentile of the distribution for the H-group, while 31% (5.4%) of women are in those percentiles for the L-group.

status, age of children and secondary education (only for the L-group). To consider the demand side of the labour market, regional dummies and size of local council have also been included. We control as well for firm size, immigrant condition, type of contract (permanent or temporary), sector (private or public) and supervisory role. Further, we added 15 occupational dummies which are arguably endogenous, yet they are useful in explaining the gender gap from an "accounting exercise" viewpoint.[12] Finally, labour and non-labour household income have also been used in the probit to control for selectivity bias in the L-group, given the low participation rate of female workers in this group.

Table 2 presents the results of a pooled OLS regression, both at the mean and at the above-mentioned quantiles, for men and women in the H and L-groups, respectively, where a (female) gender dummy captures the extent to which the gender gap remains unexplained after controlling for individual differences in various combinations of characteristics. The returns to these characteristics are restricted to be the same for both genders. The chosen combinations are as follows: (i) *basic controls* (experience and its square, experience interacted with age of children, [13]marital status, regional and size of municipality dummies, and secondary education (only for the L-group); (ii) *extended controls* (basic controls plus immigrant condition, seniority, private or public sector, type of contract, supervisory role and firm size); and (iii) *extended controls plus occupational dummies.* The intercept for the gender dummy is always negative and significant, declining (increasing) in (absolute) value in the L-group up to the 75th quantile (H-group) as we move along the wage distribution, in parallel with the pattern found in Figures 1a and 1b for

[12]Unfortunately, the ECHP does not provide information on parents´education or occupation, which could provide good instruments to correct for endogeneity.

[13]This interaction term aims at capturing the effect of child care on experience.

the raw gender gaps. Typically, the estimated gender intercepts are much lower than the raw gaps for the H-group and the differences increase as we move up the distribution. The reason for this pattern is that the differences in experience in favour of men increase from 0.3 years at the 10th. percentile to almost 7 years at the 90th. percentile. However, when using the largest set of covariates (including occupational dummies) the gender gaps become larger since, as shown in Table 1a, the proportion of women working in the occupations which yield higher wage returns (OC1 to OC8) is larger than the proportion of men. By contrast, the differences between the raw gaps and the estimated ones are much lower for the L-group since the differences in experience in favour of men are much lower (2 or 3 years in the lower percentiles and 1 year in the higher percentiles) and there are no substantive diffences in the shares of women and men (all very small) working in the highly-paid occupations.

[Table 2 about here]

Next, in order to relax the assumption that the returns to the observable characteristics are the same for men and women, results for separate QR equations by gender are presented in Tables 3a (males in H-group) and 3b (females in H-group), and in Tables 4a (males in L-group) and 4b (females in L-group). The reported results correspond to the largest set of regressors.[14] The coefficients on experience for men in the H-group are always larger than the coefficients for women and the gap grows slightltly as we move up the wage distribution, in common with the findings of Albrecht *et al.* (2003) who argue that (potential) experience is a better measure for actual experience in the case of men than for women. We also find that the return from

[14] The qualitative results obtained with the basic and extentended controls are the same, and are available upon request.

performing a supervisory role is larger for men, particularly from the 50th. percentile onwards. Being married has a lower return for women particularly at the bottom of the distribution where it provides a signal to the employers of potential career interruptions. By contrast, working in a firm with more than 20 employees has a larger return for women as is the case of working in the public sector at the 25th.and 50th.percentiles. The presence of strong collective bargaining and affirmative action in the public sector may be behind the latter result. As for the occupational dummies, the results point out that women involved in teaching (OC4, OC6) do better than men and that the difference switches in favour of men at the top quantiles of most of the remaining occupations.[15]

[Tables 3a, b and Tables 4a, b about here]

As regards the L-group, the coefficients on experience and seniority for men are above those for women with the gap decreasing as we move down the distribution. As before, the coefficient on supervisory role is always larger for men as is the return on being married at the lower quantiles. However, having a secondary educational attainment yields a higher return for women as is also the case of working in the public sector or having a permanent contract. Interestingly, women in the top occupations (OC1, OC2) get a higher return than men, in contrast with what happens for the H-group.

Besides the possible endogeneity of choice of occupation, so far we have not considered the problem of selection bias arising from the low participation of women in the L-group. To cater for this problem, we adopt some

[15] The fact that women have larger coefficients than men in some of the occupations (relative to the reference group of unskilled workers) does not imply that they get a higher wage since they may have a lower wage in the reference category. A similar comment pertains to the coefficients on the rest of the dummy variables.

restrictive, albeit simplifying, assumptions leading to the use of Heckman´s lambda approach.[16] As is conventional, we first estimate the inverse of the Mills ´ratio, λ, from a probit equation determining women´s labour market participation. Next, a wage equation is estimated adding λ to the list of regressors in the model, both at the mean and the quantiles.

The participation probit equation includes the following explanatory variables: the presence of grandparents or young children in the household, civil status, household earnings (excluding personal earnings), personal non-labour earnings, age, experience and its square, and the above-mentioned interaction between experience and dependent children. The sample consists of 1628 observations out of which 626 correspond to working women and 1002 to non-working women. Table 5a shows the results of the probit. We find that having children aged below 11 and being married increase the probability of working whereas experience and having secondary education reduce it.[17]Table 5b, in turn, presents the results of the QR for the basic set of covariates, adding λ as an extra regressor. Although the coefficient on λ turns out to be statistically insignificant, the remaining coefficients undergo significant changes which somewhat indicates that selection bias may be present.[18]

[16]A more general estimation methodology for selection models can be found in Newey *et al.*(1990), who propose a semiparametric approach which considerable relaxes the assumptions on the error-term distribution. Buchinsky (1996) generalises this approach to deal with selection-bias correction in quantile regressions.

[17]We are implicitly assuming that a woman willing to work away from home can find a job (see Buchisky, 1995). However, this assumption is fairly restrictive in our case since the 1999 female unemployment rate was 22.4%, seven percentage points higher than the overall unemployment rate.

[18]We also corrected for sample selection the wage equation for men in the L-group. However, the λ term was very small and insignificant, with the remaining coefficients hardly changing.

In sum, the evidence so far points out that returns to observable characteristics differ by gender and that these differences change as we move throughout the distribution. The next step is to investigate how important is discrimination is explaining the gender gap.

[Tables 5a, b about here]

4 Decomposition of the wage gap

4.1 Methodology

A useful way of thinking about the well-known Oaxaca-Blinder decomposition is to compare actual observations with counterfactual ones. In particular, denoting women´s and men´s returns by β^f and β^m and their characteristics by x^f and x^m, respectively, one is interested in knowing the wage that a woman would receive if she were paid according to women´s returns (β^f) but had men´s characteristics (x^m). In a market without discrimination ($\beta^f = \beta^m$), men´s wages would be equal to those fictitious women´s wages as long as they have the same productive characteristics. Therefore, observed wage differences can be attributed to unequal treatment by gender. It should be noted, however, that the discrimination measures based on the mean are not directly applicable to other points of the wage distribution. Indeed, while the decomposition of the mean wage gap is exact, this property is lost when applied to the gender wage gap at quantile θ.

In effect, in the case of the mean, a linear specification implies that

$$w_i = x_i'\beta + u_i \rightarrow E(w_i/x_i) = x_i'\beta, \tag{10}$$

since $\mathrm{E}(u_i/x_i) = 0$. Thus, the properties of OLS estimators ensure that the predicted wage evaluated at the vector of mean characteristics of the sample is exactly the average wage, *i.e.*, $E(w_i) = E(x_i)'\beta$. Hence, the Oaxaca-Blinder decomposition yields

$$E(w^m) - E(w^f) = \left(E(x^m) - E(x^f)\right)' \beta^f + E(x^m)' \left(\beta^m - \beta^f\right), \qquad (11)$$

where the first term measures the differences in the mean wage due to a different endowments of characteristics, whilst the second term captures the differences due to different returns to these characteristics.

However, in QR, expectation of (4) subject to the logged wage being equal to its unconditional quantile of order θ, $w_i = \omega_{\theta i}$, yields

$$\omega_\theta = E(x|w = \omega_\theta)'\beta_\theta + E(u_\theta|w = \omega_\theta),$$

that is, the θ quantile of the (log) wage distribution is equal to its θ conditional quantile evaluated at the vector of mean characteristics of the indiviuals at that quantile, plus the mean value of the error term for this group of individuals. This latter term, in contrast to (11), appears now in the decomposition since (4) implies that $Q_\theta(w_i|x_i) = x_i'\beta_\theta$ but evaluating the *conditional* wage quantile wage function at $E(x|w = \omega_\theta)$ does not yield the *unconditional* quantile ω_θ.

For simplicity, let us denote $\bar{x}_\theta = E(x|w = \omega_\theta)$ and $\bar{u}_\theta = E(u_\theta|w = \omega_\theta)$. Then, an Oaxaca-Blinder decomposition of the gender gap at the θ percentile implies that

$$\omega_\theta^m - \omega_\theta^f = \left(\bar{x}_\theta^m - \bar{x}_\theta^f\right)' \beta_\theta^m + \bar{x}_\theta^{m\prime} \left(\beta_\theta^m - \beta_\theta^f\right) + \left(\bar{u}_\theta^m - \bar{u}_\theta^f\right), \qquad (12)$$

where the third term, $\left(\bar{u}_\theta^m - \bar{u}_\theta^f\right)$, yields the unexplained component.

To eliminate that unexplained term in the decomposition, García *et al.* (2001) consider the gap at a given *conditional* quantile evaluated at the *unconditional* mean of the vector of characteristics, namely

$$Q_\theta(w^m/x^m{=}Ex^m)\text{-}Q_\theta(w^f/x^f{=}Ex^f) = \left(Ex^m - Ex^f\right)' \beta_\theta^m + Ex^{f\prime}\left(\beta_\theta^m - \beta_\theta^f\right). \tag{13}$$

However, as Gardeazabal and Ugidos (2002) point out, (13) suffers from the problem that it weights the contribution of any variable to the gap at the same point, i.e. at the unconditional mean $E(x)$, regardless of which quantile is considered.[19] To correct for this problem, these authors propose an exact decomposition for the difference between *unconditional* quantiles based on evaluating the *conditional* quantiles at a point such that we get the *unconditional* ones. Specifically, within the set of covariates $Z_\theta = \{z \in \mathbf{Z} : Q_\theta(w) = z\acute{}\beta_\theta)$ they choose those points that minimise the distance to $\bar{x}_\theta$. However, their method is burdensome when many covariates are considered, as in our case. For this reason, we follow Albrecht et al. ´s (2003) application of Machado and Mata´s (2000) bootstrap method to implement (12) directly, without attempting to eliminate the unexplained component.[20]

The steps in this procedure can be summarised as follows:

- With the female database, we estimate the coefficient vector β_θ^f at the quantiles of interest.

- From the male database, we take a sample with replacement of size 100 for the vector of characteristics x^m. These individuals are sorted by wage, so we get an observation for each percentile.

[19] Thus, for example, if experience is lower at the bottom than at the top of the distribution, this measure of discrimination would weight its contribution at both ends of the distribution using average experience over the entire sample.

[20] Note that the results Table 5 in Albrecht *et al.* (2003) omit the the unexplained term in (12).

- The previous procedure is repeated 100 times and then we calculate $\bar{x}_\theta = \frac{1}{100}\sum_{j=1}^{100} x^m_{\theta j}$, for $\theta = \{10, 25, 50, 75, 90\}$

Once the vectors of coefficients β^m_θ and β^f_θ and the vector of mean characteristics for each quantile has been estimated, we can proceed to estimate the three components in (12). The whole procedure has been replicated 250 times in order to obtain standard deviations of the contribution of these components.

4.2 Results of the decomposition

Tables 6a and 6b show the results of the Oaxaca- Blinder decomposition for the H and L-groups, respectively, considering the three sets of covariates discussed above. A positive (negative) sign on the $\bar{x}^{m\prime}_\theta \left(\beta^m_\theta - \beta^f_\theta\right)$ term (labeled as *Returns*) implies that the market returns to men´s characteristics are higher (lower) than the returns to women´s characteristics. Likewise, a positive (negative) sign on the $\left(\bar{x}^m_\theta - \bar{x}^f_\theta\right)' \beta^m_\theta$ term (labeled as *Characteristics*) is to be interpreted as the characteristics of men, at men´s returns, being more (less) productive than women´s characteristics.

[Tables 6a, b about here]

For the H-group, Table 6a shows that the *Characteristics* term is mostly positive signifying that men have higher productive characterististics than women. Overall, women have higher education than men.[21]. Yet, considering men and women in the H-group, the former have higher experience, seniority, work in larger firms and hold a larger proportion of managerial jobs. For the largest set of covariates, however, this component is negative a

[21] In the sample, 47.1% of women have tertiary education whereas only 31.3% of men have such an educational attainment.

the 50th and 75th quantiles, due to the fact that women are over-represented in occupations OC2 to OC7. As for the *Returns* term, it is mostly positive as well, pointing out the existence of discrimination against women throughout the distribution. This term, however, is negative for some of the lower quantiles indicating that female graduates concentrated in particular occupations are better rewarded than men. Anyhow, as shown in Figure 3a where the raw and the counterfactual (discrimination) gaps for the three specificactions are depicted, the pattern of discrimination is increasing along the distribution, that is discrimination is higher as the wage increases.

With regard to the L-group, Table 6b displays the corresponding results. The *Characteristics* component is positive in most cases yet, when considering the basic and extended set of covariates, its contribution to the raw gap is smaller than in the H-group. As in that case, female segregation in certain occupations implies that the component is negative. The *Returns* term term is positive and, as Figure 3b shows, discrimination turns out to be decreasing along the distribution, except at the top quantile, and even is above the raw gap in the model including the occupational dummies.[22]

[Figures 3a, b, c about here]

As mentioned above, for the L-group with the basic set of covariates we introduced a correction for sample-selection bias, giving rise to an extra term, λ, in the decomposition (labeled as *Selection*). Appendix B offers a detailed explanation of how to compute the decomposition in QR when the λ term is present. The results appear in the bottom panel of Table 6b and

[22]so This result is not puzzling since the observed gender wage differential does not impose an upper bound on discrimination. If women are more productive than men and receive a lower wage, then discrimination could be above the raw gender gap. The result from the model including the occupational dummies is explained by the concentration of women in certain occupations.

the evolution of the discrimination pattern is displayed in Figure 3c. As can be observed, there is a rise except at the 75th quantile. The fact that a large number of women do not participate in the labour market increases the observed wages of those who work. However, the expected wage considered by women when deciding their labour market participation is lower than for men as a consequence of discrimination. In any case, the decreasing pattern of discrimination found for the L-group remains unaltered.

A brief comment on the size of the unexplained component (labeled as *Residual*) is due. The size of this component is too large in some instances, ranging from -27% to 61% for the H-group and from -27% to 85% for the L-group. In the results of Machado and Mata (2000), the highest proportion reached by this component is 27%, yet their sample size (4,800 observation) is much larger than ours.

Lastly, as regards the contributions to the gender gap along the distribution of the factors considered in the largest set of covariates, they are presented in Figures 4a and 4b for the H and L-group, respectively. In each Figure, the bars show the contribution of the factors split into the part due to the *Characteristics* component (dark bars) and that due to the *Returns* component (light bars). In order to save space, the reported decompositions correspond to the mean and the 25th. and 90th. percentiles. As shown in Figure 4a (H-group) the contribution of experience to the *Returns* component against women turns out to be much higher at the upper than at the lower part of the distribution, in line with the glass-ceiling phenomenon. Women in this group, however, seem to favour from discrimination in the return to seniority, firm size and occupation at the top quantile. By contrast, Figure 4b (L-group) shows how the contribution of the *Returns* component againt women is much higher at the lower part of the distribution than at the top, in accord with the glass- floor hypothesis. It is also interesting to notice that the returns to firm size and occupational composition favour women at the upper quantile while it is unfavourable to them at the lower quantile.

[Figures 4a, b about here]

5 Conclusions

In this paper, we have analyzed the evolution of gender wage gaps along the wage distribution in Spain using the 1999 (6th.wave) of the ECHP using a QR framework. Our main finding is that, behind an irregular evolution for the whole sample of individuals, there is distinctive difference between the patterns of the gender gaps when we distinguish by educational attainments (individuals with primary/secondary education, L-group, and with tertiary education, H-group). While for the H-group the gender gap is increasing along the distribution, it happens to be decreasing for the L-group. Further, these patterns remain unchanged when we control for the different observable characteristics which men and women bring to the labour market. Further, while this evolution contrasts with that found for northern and central European countries, where the gender gap is increasing as we move up the distribution, it seems to be similar to that found for southern European countries where, like in Spain, female labour market participation is still low.

Our explanation for these divergent patterns is as follows. Due to the historical low participation of women in the L-group, employers may use statistical discrimination to lower their wages vis-á-vis more stable men in the lower part of the wage distribution since they expect future career interruptions. However, as their job tenure expands, women become more reliable to employers´s eyes and their wages converge to men´s wages with the same characteristics. By contrast, women in the H-group, who have undergone a costly investment in human capital, can be expected to be more stable, since their participation rate is much larger, and therefore are less discriminated at the bottom on the wage distribution. However, for reasons related to their lower job mobility, they suffer from larger gaps at the top of the distribution. Hence, there seems to be a "composition effect" in the overall gender gap, when both groups are lumped together, whereby while there is a "glass floor" floor the L-group, there is a "glass celing" for the H-group.

Using QR we find that the gender components not accounting for observable characteristics or the residual part in the gender wage equation have the the same patterns discussed above, confirming our explanation. Although the quantitative results hinge on different specifications of the set of covariates, and on the use of a correction for sample-section bias in the L-group, the qualitative results remain unaltered.

In general, the weight of the discrimination component in explaining the gender gap is larger than that found in other related papers for the Spanish case (see, García *et al.*, 2001, and Gardeazabal and Ugidos, 2002) which use alternative data sets for earlier years in the 1990s and do not distinguish by education, as we do here. Moreeover, it could be the case that measurement errors in the data collection, as discussed in OCDE (2002) have increased the weight of the *Residual* component in our decompositions. Whatever the case, the trends in the quantile decompositions remain the same irrespectively of the chosen set of covariates, which reinforces our results.

As for future research agenda, we have two extensions in mind. The first one relates to the underlying strategy of the model. We have assumed that that the non-discriminatory wage structure is the male one and, therefore, that all the discrimination stems from an infra-valuation of women´s returns vis-á-vis men´s. Nonetheless, it could be the case that men´s wages will also be affected by discrimination.[23] A more general approach would be to derive a non-discriminatory wage structure from a theoretical model which is neither completely masculine nor femenine (see Neumark, 1988). Several studies show that conclusions on the source of the gender gap hinge crucially on different asumptions on the nature of the non-discriminatory structure. The second one would be to extend the analysis to other EU countries taking advantage of the data harmonization provided by the ECHP.

Appendix

1 A: Definition of variables

The variables are drawn from the 1999 (6th. wave) of theECHP. Our group of interest is composed by wage-earners working full-time and more than 15 hours per week. In this section we include a detailed description of the variables used in the analysis.

Gross hourly wage: The ECHP collects data on average monthly labor income-gross and net-, from salaried workers. Labor income includes salary bonus divided by working months, and overtime. When a worker has more than one job, only the main job income is considered. Weekly hours in the main job are available, including overtime hours. We have set an upper bound of 60 hours to this variable in order to minimize the self-declared bias. This correction affects 2% of men and 0.9% of women from our total sample. Then, gross hourly wage is the monthly gross salary divided by 52/12 and multiplied by the weekly hours worked in the main job.

[23] For instance, there is evidence for the late 1990s that the number of women working in managerial positions who take maternity leaves is three times smaller that the number of men in similar positions taking leaves for stress (see Chichilla, 2003).

Experience:defined as age minus age of first job after leaving full-time schooling. Since data on the age at which individuals left full-time schooling are not available, this information is proxied by 6 plus the minimum number of years necessary to attain the declared educational level (8 for primary education, 12 for secondary and 15 for tertiary).

Exp*Children: interaction between experience and a binary variable that takes a value of 1 when an individual has dependent children (from 0 to 16 years). In the basic set of covariates, we consider separately the case in which children are between 0 and 11 years (**Exp*Children 0-11**) and between 12 and 16 years (**Exp*Children 12-16**).

Level of education: primary or secondary. This dummy variable is only included for the L- group.

Individual characteristics: marital status, immigrant condition, district of residence and district size.

Type of contract: temporary or permanent.

Supervisory role: directive or managing position, supervisor of at least another employee and without responsability for the rest of employees.

Seniority: from 0 to 1 years, from 2 to 5 years, from 6 to 9 years, from 10 to 14 years and above 15 years. Seniority is obtained as the difference between the year of the survey, 1999, and the year of the start of the current job. Seniority is aggregated into five categories and, correspondingly, four dummy variables are included in the regression, since the data is not too accurate.

Firm size: from 1 to 4 employees, from 5 to 19 employees, from 20 to 49 employees, from 50 to 99 employees, from 100 to 499 employees and above 500 employees.

Occupation: fifteen occupational groups have been considered, corresponding to an intermediate level of aggregation of the ISCO-88 (COM) classification. We have distinguished among: Legislators, senior officials and managers (**OC1**); Physical, mathematical, engineering, life science and health professionals (**OC2**), Teaching professionals (**OC3**), Other professionals (**OC4**), Physical, mathematical, engineering, life science and health associate professionals (**OC5**), Teaching and other associate professionals (**OC6**), Clerks (**OC7**), Models, salespersons and demonstrators (**OC8**), Personal and protective services workers (**OC9**), Skilled agricultural and fishery workers (**OC10**), Extraction and building trades workers, other craft and related trades workers (**OC11**), Metal, machinery, precision, handicraft printing and related trades workers (**OC12**), Plant and machinery operators and assemblers (**OC13**), Sales and services elementary occupations (**OC14**) and Agricultural, fishery and related labourers, labourers in mining, construction, manufacturing and transport (**OC15**).

2 B: Decomposition of the gender gap with endogenous selection correction

When considering sample selection bias, the relevant equations become

$$w_i = x_i'\beta + \sigma\lambda_i + u_{\theta i} \qquad \text{for the mean}$$

$$w_i = x_i'\beta_\theta + \sigma_\theta\lambda_i + u_{\theta i} \qquad \text{for the quantiles}$$

where λ is the variable associated to the selection bias which is obtained from the participation equation. For the mean, the decomposition of the gender gap has an analogous expression to equation (11), that is

$$E(w^m) - E(w^f) = \left(E(x^m) - E(x^f)\right)'\beta^f + E(x^m)'\left(\beta^m - \beta^f\right) + \left(\sigma^m E(\lambda^m) - \sigma^f E(\lambda^f)\right), \tag{A.2}$$

where the last term captures the effect of the self-selection.

When extending this decomposition to the case of the wage at the θ quantile, the expression in equation (12) becomes

$$\omega_\theta^m - \omega_\theta^f = \left(\bar{x}_\theta^{m\prime}\beta_\theta^m + \sigma_\theta^m\bar{\lambda}_i^m + \bar{u}_\theta^m\right) - \left(\bar{x}_\theta^{f\prime}\beta_\theta^f + \sigma_\theta^f\bar{\lambda}_i^f + \bar{u}_\theta^f\right) =$$

$$= \left(\bar{x}_\theta^m - \bar{x}_\theta^f\right)'\beta_\theta^m + \bar{x}_\theta^{m\prime}\left(\beta_\theta^m - \beta_\theta^f\right) + \left(\sigma_\theta^m\bar{\lambda}_i^m - \sigma_\theta^f\bar{\lambda}_i^f\right) + \left(\bar{u}_\theta^m - \bar{u}_\theta^f\right), \tag{A.3}$$

where $\bar{\lambda}_\theta = E(\lambda|w = \omega_\theta)$ in common with the notation used in the main text.

Neuman and Oaxaca (1998) underline that it is not obvious how this additional term in equations A.2 y A.3 should be interpreted in the decomposition. We could be interested in splitting this term into the two general components: discrimination and endowments. If this is the case, the way to carry out this additional decomposition depends on the assumptions and the objectives of the analysis. The authors propose various alternatives. In this paper, we have adopted the simplest approximation which consists in considering the gender differences in the selection term as an independent component. Therefore, our decomposition is composed by four terms that are shown in equation A.3, where the male selection term $\bar{\lambda}_i^m$ is zero as no evidence of self-selection for male has been found in our sample.

References

[1] Abadie, A. (1997), "Changes in the Spanish Labor Income Structure during the 1980´s: A Quantile Regression Approach", *Investigaciones Económicas*, 21, 253-272.

[2] Albrecht, J., A. Björklund and S. Vroman (2003), "Is There a Glass Ceiling in Sweden?", *Journal of Labor Economics*, 21, 145-177.

[3] Acemoglu, D. and J.S. Pischke (1998), "Why Do Firms Train: Theory and Evidence", Quarterly Journal of Economics, 113, 79-111.

[4] Arellano, M. and O. Bover (1995), "Female Labour Force Participation in the 1980s: The Case of Spain", *Investigaciones Económicas*, 29, 171-194.

[5] Babcock, L. and S. Laschever (2003), *Women Don't Ask: Negotiation and the Gender Divide*, Princeton University Press.

[6] Bertrand, M. and K. Hallock (2000), "The Gender Gap in Top Corporate Jobs", NBER Working Paper no.7931.

[7] Blau, F. and L. Kahn (2000), "Gender Differences in Pay", *Journal of Economic Perspectives*, 14, 75-99

[8] Booth, A., M. Francesconi and J. Frank (2003), "A Sticky Floors Model of Promotion, Pay and Gender", *European Economic Review*, 47, 295-322.

[9] Buchinsky, M. (1994), "Changes in the U.S. Wage Structure 1963-1987: An Application of Quantile Regression", *Econometrica*, 62, 405-458.

[10] Buchinsky, M.(1998), "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research", *The Journal of Human Resources,* 33, 88-126.

[11] Buchinsky, M. (1996), "Women´s Return to Education in the US: Exploration by Quantile Regression with Nonparametric Sample Selection Correction", Working Paper, Department of Economics, Brown University.

[12] Chamberlain, G. (1994), “Quantile Regression, Censoring and the Structure of Wages”, *Advances in Econometrics, Sixth World Congress* (Volume 1); C. A. Sims (ed.), Cambridge University Press.

[13] Chinchilla, N. (2003), “ Valores Femeninos en la Empresa” (mimeo), IESE Business School, Univeridad de Navarra.

[14] Dolado, J.J., F. Felgueroso and J.F. Jimeno (2003), “Where do Women Work?: Analysing Patterns in Occupational Segregation By Gender”, *Annales d'Economie et de Statistique,* forthcoming.

[15] Fitzberger, B., Koenker, R., and J.A.F. Machado (Eds.) (2001), *Economic Applications of Quantile Regression*, Physica-Verlag.

[16] García, J., P.J. Hernández and A. López-Nicolás (2001), “How Wide is the Gap? An Investigation of Gender Wage Differences using Quantile Regression”, *Empirical Economics*, 26, 149-168.

[17] Gardezabal, J. and A. Ugidos (2002), “A Measure of Gender Wage Discrimination at Quantiles”, (mimeo), Universidad del País Vasco.

[18] Gneezy, U., Niederle L., and A. Rustichini (2003), “Performance in competitive environments: Gender differences,” *Quarterly Journal of Economics*, 128, 1049-1074.

[19] Koenker, R. and G. Bassett (1978), “Regression Quantiles”, *Econometrica*, 46, 33-50.

[20] Koenker, R. and G. Bassett (1982), “ Robust Tests for Heteroskedasticity on Regression Quantiles”, *Econometrica*, 50, 43-61.

[21] Lazear, E. and S. Rosen (1990), “Male-Female Wage Differentials in Job Ladders”, *Journal of Labour Economics*, 8, 106-123.

[22] Machado, J. and J. Mata.(2000), "Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression", mimeo.

[23] Mincer, J. and S. Polacheck (1974), " Family Investment in Human Capital: Earnings of Women", *Journal of Political Economy*, 82, 76-108.

[24] Neuman, S. and R. Oaxaca (1998), "Estimating Labor Market Discrimination with Selectivity Corrected Wage Equations: Methodological Considerations and an Illustration From Israel", Center for Economic Policy Research, Discussion Paper 1915.

[25] Newey, W., J. Powell and J. Walker (1990), "Semiparametric Estimation of Selection Models: Some Empirical Results", *American Economic Association, Papers and Proceedings*, 80,. 324-328.

[26] OECD (2002), "Women at Work: Who Are They and How Are They Faring?", *Employment Outlook*, 63-125.

[27] Polacheck, S. (1981), "Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure", *Review of Economics and Statistics*, 62, 60-69.

Table 1a. Descriptives statistics. Tertiary education. Spain (1999)

	Men		Women	
	Average	*St. dev*	*Average*	*St. dev*
N. observations	721		558	
Age	37.61	10.34	35.78	9.49
Children 0-11	0.44	0.77	0.28	0.45
Children 12-16	0.22	0.50	0.13	0.33
Age Groups				
17 to 24	0.07	0.26	0.09	0.28
25 to 34	0.37	0.48	0.45	0.50
35 to 44	0.31	0.46	0.28	0.45
≥ 45	0.25	0.43	0.19	0.39
Married	0.62	0.49	0.55	0.50
Immigrant	0.003	0.05	0.01	0.07
Weekly hours	41.22	6.54	38.31	5.94
Gross Hourly wage	1690	1073	1477	719
Ln (gross hourly wage)	7.28	0.55	7.18	0.50
Experience	14.78	10.40	12.68	9.71
Occupation				
OC1	0.05	0.23	0.02	0.13
OC2	0.14	0.35	0.16	0.36
OC3	0.11	0.32	0.25	0.43
OC4	0.09	0.28	0.09	0.29
OC5	0.09	0.29	0.05	0.22
OC6	0.09	0.28	0.13	0.34
OC7	0.10	0.29	0.19	0.39
OC8	0.05	0.21	0.05	0.22
OC9	0.02	0.15	0.03	0.18
OC10	0.01	0.07	0.00	0.00
OC11	0.07	0.25	0.00	0.04
OC12	0.08	0.27	0.00	0.04
OC13	0.07	0.25	0.01	0.12
OC14	0.02	0.13	0.01	0.11
OC15	0.02	0.12	0.01	0.08
Firm Size				
1-4 employees	0.07	0.26	0.10	0.31
5-19 employees	0.21	0.41	0.22	0.41
20-49 employees	0.17	0.38	0.21	0.40
50-99 employees	0.13	0.33	0.10	0.30
100-499 employees	0.18	0.39	0.14	0.35
> 500 employees	0.24	0.43	0.23	0.42
Public sector	0.35	0.48	0.51	0.50
Responsability degree				
Directive	0.16	0.37	0.06	0.24
Supervisor	0.27	0.45	0.25	0.43
Without responsability	0.56	0.50	0.69	0.46
Tenure				
1 year	0.30	0.46	0.32	0.47
2-5 years	0.19	0.39	0.19	0.40
6-9 years	0.10	0.30	0.14	0.35
10-14 years	0.12	0.33	0.10	0.30
≥ 15 years	0.29	0.45	0.25	0.43
Permanent contract	0.80	0.40	0.73	0.44

Table 1b: Descriptive statistics. Non-tertiary education. Spain (1999)

	Men		Working women		Non-working women	
	Average	*St. dev*	*Average*	*St. dev*	*Average*	*St. dev*
N. observations	1585		626		1002	
Age	37.89	11.31	36.19	11.33	33.81	8.39
Children 0-11	0.30	0.46	0.21	0.41	0.53	0.50
Children 12-16	0.16	0.37	0.17	0.38	0.25	0.43
Age Groups						
17 to 24	0.13	0.33	0.18	0.38	0.14	0.35
25 to 34	0.31	0.46	0.32	0.47	0.38	0.49
35 to 44	0.25	0.43	0.23	0.42	0.43	0.50
≥ 45	0.31	0.46	0.27	0.45	0.05	0.21
Married	0.68	0.47	0.53	0.50	0.73	0.44
Immigrant	0.004	0.07	0.01	0.08	0.01	0.09
Secondary ed.	0.29	0.46	0.38	0.49	0.24	0.43
Weekly hours	42.64	6.19	40.36	5.74		
Gross Hourly wage	1037	487	851	484		
Ln (gross hourly wage)	6.86	0.41	6.63	0.47		
Experience	20.77	12.41	16.78	11.85		
Occupation						
OC1	0.02	0.14	0.01	0.11		
OC2	0.00	0.00	0.00	0.00		
OC3	0.00	0.03	0.00	0.04		
OC4	0.00	0.06	0.00	0.07		
OC5	0.02	0.13	0.02	0.14		
OC6	0.05	0.21	0.06	0.24		
OC7	0.07	0.25	0.18	0.38		
OC8	0.08	0.28	0.18	0.38		
OC9	0.04	0.20	0.13	0.33		
OC10	0.02	0.15	0.01	0.10		
OC11	0.22	0.42	0.08	0.27		
OC12	0.10	0.30	0.01	0.10		
OC13	0.20	0.40	0.06	0.24		
OC14	0.05	0.21	0.20	0.40		
OC15	0.13	0.33	0.06	0.24		
Firm Size						
1-4 employees	0.18	0.38	0.22	0.42		
5-19 employees	0.31	0.46	0.25	0.44		
20-49 employees	0.16	0.37	0.16	0.37		
50-99 employees	0.10	0.30	0.11	0.31		
100-499 employees	0.12	0.33	0.14	0.35		
> 500 employees	0.12	0.33	0.11	0.32		
Public sector	0.13	0.34	0.18	0.38		
Responsability degree						
Directive	0.05	0.23	0.03	0.16		
Supervisor	0.16	0.36	0.08	0.28		
Without responsability	0.79	0.41	0.89	0.31		
Tenure						
1 year	0.40	0.49	0.43	0.50		
2-5 years	0.18	0.39	0.19	0.39		
6-9 years	0.08	0.27	0.11	0.31		
10-14 years	0.09	0.29	0.07	0.26		
≥ 15 years	0.25	0.43	0.20	0.40		
Permanent contract	0.64	0.48	0.60	0.49		
Parents 65-75	0.05	0.21	0.06	0.24	0.05	0.22
Parents >76	0.02	0.15	0.05	0.22	0.03	0.17
Households income	1379321	1458148	2207651	1632469	2134753	1365607
Non-labour income	23944	65727	28065	193603	11534	68140

Table 2: Estimated gender gaps under alternative covariates. Spain. 1999
Dependent variable: Ln. Hourly gross wage

H-group (n=1279)	Average	θ=10	θ=25	θ=50	θ=75	θ=90
Observed gender gap	-0.0974***	-0.0801 **	-0.1020***	-0.1054***	-0.1668***	-0.2425***
	(0.019)	(0.039)	(0.032)	(0.038)	(0.036)	(0.031)
Basic controls	-0.0383 *	-0.0397	-0.0376 *	-0.0398 *	-0.0853***	-0.0694***
	(0.021)	(0.025)	(0.023)	(0.024)	(0.021)	(0.024)
Extended controls	-0.0375 *	-0.0461 *	-0.0393 *	-0.0441 **	-0.0868***	-0.0676***
	(0.020)	(0.029)	(0.021)	(0.020)	(0.020)	(0.025)
Extended controls +	-0.0787***	-0.0501	-0.0702 **	-0.0910***	-0.1393***	-0.0773***
occupational dummies	(0.018)	(0.022)	(0.026)	(0.022)	(0.029)	(0.039)

L-group (n=2211)	Average	θ=10	θ=25	θ=50	θ=75	θ=90
Observed gender gap	-0.2273***	-0.3333***	-0.2471***	-0.2031***	-0.1482***	-0.2094***
	(0.021)	(0.035)	(0.023)	(0.019)	(0.032)	(0.041)
Basic controls	-0.2096***	-0.2710***	-0.2230***	-0.1858***	-0.1481***	-0.2014***
	(0.019)	(0.038)	(0.028)	(0.026)	(0.025)	(0.034)
Extended controls	-0.1997***	-0.2175***	-0.2014***	-0.1942***	-0.1553***	-0.2476***
	(0.016)	(0.030)	(0.022)	(0.023)	(0.020)	(0.031)
Extended controls +	-0.2085***	-0.2124***	-0.2032***	-0.1982***	-0.1788***	-0.2002***
occupational dummies	(0.017)	(0.029)	(0.018)	(0.018)	(0.023)	(0.034)

Note: ***, **, * represent significance at 99, 95 and 90% respectively

Table 3a. OLS and QR

H-group. Spain. 1999

Dependent variable : Ln. gross hourly wage

MEN	**Average**	**θ=10**	**θ=25**	**θ=50**	**θ=75**	**θ=90**
Experience	0.020***	0.011	0.016**	0.017**	0.013	0.032***
	(0.006)	(0.008)	(0.007)	(0.007)	(0.009)	(0.010)
Experience^2	-0.0003*	-0.0002	-0.0002	-0.0002	-0.00002	-0.0003
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Exp*Children	0.0025	0.0037	0.0034	0.0027	0.0044	0.0009
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
Immigrant	-0.080	-0.007	-0.137	0.056	-0.160	-0.205
	(0.183)	(0.334)	(0.269)	(0.189)	(0.195)	(0.207)
Public sector	0.085**	0.080	0.094*	0.068	0.070	0.109**
	(0.034)	(0.070)	(0.051)	(0.050)	(0.052)	(0.055)
Permanent contract	0.074*	0.131**	0.141***	0.081	0.034	0.109*
	(0.039)	(0.058)	(0.053)	(0.049)	(0.056)	(0.062)
Responsibility						
Directive	0.320***	0.164*	0.252***	0.298***	0.388***	0.510***
	(0.051)	(0.085)	(0.072)	(0.066)	(0.095)	(0.091)
Supervisor	0.090***	0.005	0.066*	0.107***	0.123***	0.148***
	(0.030)	(0.046)	(0.039)	(0.038)	(0.045)	(0.054)
Tenure						
2-5 years	0.096**	0.177***	0.109**	0.084	0.116**	-0.064
	(0.042)	(0.065)	(0.054)	(0.054)	(0.058)	(0.072)
6-9 years	0.141**	0.073	0.163*	0.176**	0.191**	-0.051
	(0.058)	(0.103)	(0.090)	(0.082)	(0.086)	(0.083)
10-14 years	0.153**	0.167	0.135*	0.174**	0.211***	0.052
	(0.063)	(0.106)	(0.081)	(0.083)	(0.080)	(0.094)
>15 years	0.165***	0.283***	0.187**	0.168**	0.126	-0.135
	(0.062)	(0.109)	(0.075)	(0.078)	(0.085)	(0.116)
Married	0.071*	0.118**	0.030	0.032	0.070	0.049
	(0.039)	(0.056)	(0.049)	(0.045)	(0.061)	(0.072)
Firm size						
5-19 employees	-0.090	0.005	-0.067	-0.087	-0.153	-0.210
	(0.067)	(0.114)	(0.073)	(0.082)	(0.111)	(0.136)
20-49 employees	-0.044	0.076	0.042	0.020	-0.096	-0.232**
	(0.065)	(0.116)	(0.085)	(0.077)	(0.094)	(0.113)
50-99 employees	0.043	0.191	0.124	0.060	-0.031	-0.206*
	(0.065)	(0.124)	(0.086)	(0.083)	(0.105)	(0.124)
100-499 employees	0.088	0.217*	0.165**	0.106	0.009	-0.112
	(0.066)	(0.121)	(0.073)	(0.080)	(0.099)	(0.113)
> 500 employees	0.144**	0.233*	0.216***	0.154*	0.094	-0.008
	(0.065)	(0.125)	(0.081)	(0.080)	(0.096)	(0.119)

Table 3a. (continuation)

MEN	Average	θ=10	θ=25	θ=50	θ=75	θ=90
Occupation						
OC1	0.840***	0.866***	0.657***	0.786***	1.002***	0.983***
	(0.106)	(0.202)	(0.155)	(0.141)	(0.189)	(0.166)
OC2	0.577***	0.548***	0.461***	0.575***	0.626***	0.566***
	(0.082)	(0.207)	(0.148)	(0.107)	(0.113)	(0.133)
OC3	0.646***	0.678***	0.548***	0.626***	0.680***	0.699***
	(0.080)	(0.210)	(0.150)	(0.106)	(0.104)	(0.126)
OC4	0.451***	0.231	0.329	0.504***	0.602***	0.512***
	(0.092)	(0.244)	(0.205)	(0.132)	(0.129)	(0.121)
OC5	0.437***	0.465**	0.319**	0.442***	0.498***	0.434***
	(0.081)	(0.201)	(0.141)	(0.114)	(0.116)	(0.129)
OC6	0.292***	0.344*	0.236	0.319***	0.353***	0.320***
	(0.081)	(0.199)	(0.148)	(0.103)	(0.121)	(0.116)
OC7	0.258***	0.233	0.150	0.232**	0.323**	0.265
	(0.084)	(0.205)	(0.151)	(0.118)	(0.130)	(0.179)
OC8	0.202**	0.238	0.087	0.155	0.319**	0.383**
	(0.090)	(0.217)	(0.147)	(0.129)	(0.148)	(0.192)
OC9	0.191*	0.184	0.186	0.124	0.196	0.100
	(0.099)	(0.235)	(0.161)	(0.125)	(0.142)	(0.225)
OC10	-0.206	-0.221	-0.299	-0.045	-0.120	-0.444
	(0.150)	(0.289)	(0.294)	(0.294)	(0.280)	(0.273)
OC11	0.192**	0.318*	0.136	0.131	0.128	0.027
	(0.081)	(0.191)	(0.136)	(0.101)	(0.124)	(0.142)
OC12	0.242***	0.194	0.165	0.216**	0.278**	0.224
	(0.083)	(0.207)	(0.161)	(0.110)	(0.127)	(0.152)
OC13	0.270***	0.193	0.161	0.215*	0.336**	0.336**
	(0.084)	(0.191)	(0.159)	(0.125)	(0.147)	(0.152)
OC15	0.066	0.224	-0.068	0.127	0.053	-0.056
	(0.099)	(0.242)	(0.193)	(0.137)	(0.122)	(0.151)
Nº Obs.	721	721	721	721	721	721
R^2	0.655	0.402	0.438	0.453	0.449	0.472

Note: ***, **, * represent significance at 99, 95 and 90% respectively. Standard deviation in parenthesis. Dummy variables for region and local council size included. Omitted group: wage earners in private sector in less-than-5-employees firms, without responsability, with less-than-1-year tenure, single, and in non-qualified jobs in services and commerce (OC14)

Table 3b. OLS and QR
H-group. Spain. 1999
Dependent variable : Ln. gross hourly wage

WOMEN	Average	θ=10	θ=25	θ=50	θ=75	θ=90
Experience	0.011*	0.002	0.004	0.004	0.019*	0.016
	(0.006)	(0.010)	(0.008)	(0.008)	(0.011)	(0.013)
Experience^2	-0.0001	0.00005	-0.00003	0.000004	-0.0002	-0.0002
	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0003)
Exp*Children	0.0005	0.0020	0.0042**	-0.0009	-0.0004	-0.0007
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
Immigrant	0.008	0.230*	0.089	0.035	0.258	0.024
	(0.128)	(0.128)	(0.143)	(0.224)	(0.311)	(0.331)
Public sector	0.097***	0.097	0.148***	0.100**	0.070	0.103*
	(0.035)	(0.061)	(0.049)	(0.044)	(0.048)	(0.061)
Permanent contract	0.119***	0.074	0.139***	0.155***	0.127*	0.069
	(0.040)	(0.068)	(0.049)	(0.048)	(0.070)	(0.082)
Responsibility						
Directive	0.134**	0.184	0.226**	0.172**	0.138	0.074
	(0.063)	(0.149)	(0.105)	(0.076)	(0.110)	(0.130)
Supervisor	0.049	0.067	0.037	0.036	0.082	0.034
	(0.033)	(0.059)	(0.055)	(0.049)	(0.054)	(0.073)
Tenure						
2-5 years	0.064	0.091	0.044	0.065	0.036	0.083
	(0.039)	(0.088)	(0.056)	(0.049)	(0.065)	(0.077)
6-9 years	0.164***	0.202**	0.124*	0.148**	-0.036	0.081
	(0.052)	(0.092)	(0.068)	(0.069)	(0.086)	(0.114)
10-14 years	0.108*	0.134	0.088	0.099	-0.029	0.180
	(0.063)	(0.117)	(0.091)	(0.074)	(0.113)	(0.117)
>15 years	0.204***	0.299***	0.242***	0.168**	0.018	0.172
	(0.061)	(0.108)	(0.082)	(0.074)	(0.111)	(0.145)
Married	0.038	0.008	-0.033	0.046	0.096**	0.044
	(0.030)	(0.053)	(0.041)	(0.040)	(0.042)	(0.048)
Firm size						
5-19 employees	0.172***	0.215*	0.162**	0.187***	0.081	0.076
	(0.063)	(0.125)	(0.065)	(0.065)	(0.093)	(0.095)
20-49 employees	0.275***	0.412***	0.305***	0.286***	0.168*	0.191*
	(0.063)	(0.137)	(0.078)	(0.081)	(0.096)	(0.103)
50-99 employees	0.286***	0.399**	0.283***	0.364***	0.174	0.225**
	(0.072)	(0.169)	(0.104)	(0.090)	(0.109)	(0.110)
100-499 employees	0.364***	0.498***	0.382***	0.374***	0.305***	0.301***
	(0.064)	(0.113)	(0.076)	(0.084)	(0.092)	(0.094)
> 500 employees	0.316***	0.512***	0.338***	0.349***	0.186**	0.259***
	(0.064)	(0.113)	(0.077)	(0.086)	(0.097)	(0.097)

Table 3b. (continuation)

WOMEN	Average	θ=10	θ=25	θ=50	θ=75	θ=90
Occupation						
OC1	0.818***	0.660***	0.756***	0.838***	0.748***	0.887***
	(0.113)	(0.228)	(0.217)	(0.152)	(0.214)	(0.236)
OC2	0.682***	0.379***	0.546***	0.696***	0.783***	0.834***
	(0.073)	(0.130)	(0.117)	(0.092)	(0.142)	(0.148)
OC3	0.836***	0.714***	0.806***	0.861***	0.925***	0.814***
	(0.069)	(0.134)	(0.115)	(0.095)	(0.130)	(0.151)
OC4	0.764***	0.575***	0.635***	0.719***	0.921***	0.945***
	(0.081)	(0.142)	(0.136)	(0.111)	(0.190)	(0.159)
OC5	0.473***	0.317**	0.441***	0.454***	0.504***	0.376**
	(0.078)	(0.137)	(0.115)	(0.119)	(0.150)	(0.163)
OC6	0.493***	0.240	0.445***	0.425***	0.631***	0.606***
	(0.077)	(0.153)	(0.109)	(0.096)	(0.134)	(0.161)
OC7	0.365***	0.227*	0.288**	0.336***	0.384***	0.352**
	(0.069)	(0.137)	(0.115)	(0.088)	(0.129)	(0.159)
OC8	0.266***	0.226**	0.234**	0.235**	0.268*	0.247
	(0.075)	(0.111)	(0.113)	(0.097)	(0.146)	(0.158)
OC9	0.107	-0.068	0.136	0.103	0.069	0.316
	(0.097)	(0.192)	(0.166)	(0.124)	(0.168)	(0.216)
OC10	0	0	0	0	0	0
OC11	0.364***	0.736**	0.601*	0.426*	0.060	0.017
	(0.105)	(0.375)	(0.310)	(0.225)	(0.151)	(0.171)
OC12	0.420***	0.694*	0.528*	0.476*	0.321	0.129
	(0.098)	(0.380)	(0.303)	(0.256)	(0.203)	(0.161)
OC13	0.244***	0.325	0.240	0.195	0.219	0.243
	(0.092)	(0.232)	(0.163)	(0.127)	(0.173)	(0.214)
OC15	0.431***	0.499***	0.513***	0.404***	0.303*	0.269
	(0.079)	(0.181)	(0.157)	(0.147)	(0.165)	(0.197)
Nº Obs.	558	558	558	558	558	558
R^2	0.655	0.465	0.477	0.472	0.423	0.388

Note: ***, **, * represent significance at 99, 95 and 90% respectively. Standard deviation in parenthesis. Dummy variables for region and local council size included. Omitted group: wage earners in private sector in less-than-5-employees firms, without responsability, with less-than-1-year tenure, single, and in non-qualified jobs in services and commerce (OC14)

Table 4a. OLS and QR

L-group. Spain. 1999

Dependent variable : Ln. gross hourly wage

MEN	**Average**	**θ=10**	**θ=25**	**θ=50**	**θ=75**	**θ=90**
Experience	0.013***	0.017***	0.010***	0.010***	0.008*	0.008
	(0.003)	(0.006)	(0.003)	(0.003)	(0.004)	(0.005)
Experience^2	-0.0002***	-0.0003**	-0.0002***	-0.0002***	-0.0001	-0.0001
	(0.00005)	(0.0001)	(0.00007)	(0.00007)	(0.00009)	(0.0001)
Exp*Children	-0.002**	-0.003*	-0.002	-0.001	-0.001	-0.001
	(0.0008)	(0.001)	(0.001)	(0.0009)	(0.0009)	(0.001)
Secondary ed.	0.060***	0.081**	0.060**	0.055***	0.025	0.023
	(0.020)	(0.040)	(0.027)	(0.019)	(0.024)	(0.035)
Immigrant	-0.143	-0.192	-0.222	-0.244	0.005	0.151
	(0.144)	(0.130)	(0.150)	(0.165)	(0.270)	(0.265)
Public sector	0.020	0.006	0.033	0.032	0.036	0.054
	(0.030)	(0.049)	(0.038)	(0.027)	(0.040)	(0.047)
Permanent contract	0.065***	0.105**	0.061**	0.028	0.037	0.073**
	(0.022)	(0.048)	(0.030)	(0.026)	(0.027)	(0.037)
Responsibility						
Directive	0.161***	0.193***	0.113**	0.166***	0.135**	0.104*
	(0.040)	(0.073)	(0.051)	(0.050)	(0.068)	(0.059)
Supervisor	0.089***	0.083**	0.080***	0.096***	0.092***	0.074**
	(0.023)	(0.035)	(0.028)	(0.031)	(0.032)	(0.035)
Tenure						
2-5 years	0.009	0.021	0.009	0.036	0.017	-0.033
	(0.022)	(0.053)	(0.032)	(0.026)	(0.029)	(0.031)
6-9 years	0.038	0.069	0.038	0.054	0.043	0.035
	(0.031)	(0.054)	(0.040)	(0.038)	(0.051)	(0.057)
10-14 years	0.093***	0.138**	0.095**	0.111***	0.109**	0.117*
	(0.032)	(0.059)	(0.043)	(0.036)	(0.043)	(0.061)
>15 years	0.191***	0.187***	0.207***	0.190***	0.153***	0.148***
	(0.030)	(0.052)	(0.035)	(0.033)	(0.045)	(0.051)
Married	0.079***	0.123***	0.083***	0.070***	0.073**	0.077**
	(0.021)	(0.038)	(0.025)	(0.027)	(0.030)	(0.033)
Firm size						
5-19 employees	0.081***	0.132***	0.078**	0.089***	0.067*	0.082*
	(0.024)	(0.051)	(0.032)	(0.028)	(0.035)	(0.048)
20-49 employees	0.119***	0.104*	0.151***	0.125***	0.102***	0.098**
	(0.026)	(0.060)	(0.038)	(0.032)	(0.034)	(0.045)
50-99 employees	0.110***	0.106	0.147***	0.118***	0.115**	0.079
	(0.033)	(0.067)	(0.047)	(0.033)	(0.046)	(0.051)
100-499 employees	0.239***	0.231***	0.224***	0.248***	0.262***	0.274***
	(0.030)	(0.063)	(0.042)	(0.034)	(0.043)	(0.058)
> 500 employees	0.311***	0.375***	0.283***	0.340***	0.338***	0.344***
	(0.034)	(0.064)	(0.038)	(0.041)	(0.042)	(0.056)

Table 4a. (continuation)

MEN	**Average**	**θ=10**	**θ=25**	**θ=50**	**θ=75**	**θ=90**
Occupation						
OC1	0.305***	0.001	0.172*	0.352**	0.517***	0.460**
	(0.090)	(0.134)	(0.104)	(0.160)	(0.131)	(0.230)
OC2	0	0	0	0	0	0
OC3	0.302***	0.678**	0.383**	0.297**	0.169*	0.062
	(0.052)	(0.306)	(0.196)	(0.137)	(0.097)	(0.086)
OC4	0.318	-0.484	-0.075	0.773	0.694*	0.820***
	(0.319)	(0.494)	(0.614)	(0.585)	(0.400)	(0.374)
OC5	0.269***	0.201	0.221***	0.295***	0.292**	0.380***
	(0.066)	(0.130)	(0.083)	(0.069)	(0.128)	(0.133)
OC6	0.243***	-0.007	0.143*	0.271***	0.342***	0.499***
	(0.055)	(0.105)	(0.074)	(0.074)	(0.091)	(0.089)
OC7	0.198***	0.070	0.096*	0.167***	0.235***	0.511***
	(0.049)	(0.089)	(0.054)	(0.050)	(0.087)	(0.152)
OC8	0.070*	-0.019	0.020	0.118**	0.112*	0.124*
	(0.042)	(0.087)	(0.059)	(0.051)	(0.059)	(0.072)
OC9	-0.008	-0.044	-0.021	0.023	-0.026	0.008
	(0.048)	(0.103)	(0.055)	(0.050)	(0.063)	(0.120)
OC10	-0.082	-0.086	-0.103	-0.057	-0.057	-0.065
	(0.057)	(0.096)	(0.071)	(0.079)	(0.077)	(0.117)
OC11	0.073**	0.046	0.057	0.100**	0.063	0.092
	(0.036)	(0.075)	(0.040)	(0.042)	(0.051)	(0.069)
OC12	0.108***	-0.017	0.111**	0.170***	0.146***	0.144**
	(0.040)	(0.105)	(0.052)	(0.045)	(0.055)	(0.070)
OC13	0.053	-0.021	0.035	0.083*	0.070	0.108
	(0.036)	(0.078)	(0.041)	(0.045)	(0.049)	(0.067)
OC15	-0.044	-0.047	-0.021	-0.017	-0.035	0.001
	(0.040)	(0.091)	(0.042)	(0.044)	(0.051)	(0.074)
Nº Obs.	1585	1585	1585	1585	1585	1585
R^2	0.303	0.237	0.250	0.282	0.327	0.350

Note: ***, **, * represent significance at 99, 95 and 90% respectively. Standard deviation in parenthesis. Dummy variables for region and local council size included. Omitted group: wage earners in private sector in less-than-5-employees firms, without responsability, with less-than-1-year tenure, single, with primary education and in non-qualified jobs in services and commerce (OC14)

Table 4b. OLS and QR
L-group. Spain. 1999
Dependent variable : Ln. gross hourly wage

WOMEN	**Average**	**θ=10**	**θ=25**	**θ=50**	**θ=75**	**θ=90**
Experience	0.007*	-0.0001	0.009	0.008	0.006	0.006
	(0.004)	(0.010)	(0.006)	(0.005)	(0.006)	(0.007)
Experience^2	-0.0001	0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	(0.00009)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Exp*Children	-0.002	0.00004	-0.004**	-0.003*	0.002	0.002
	(0.001)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)
Secondary ed.	0.113***	0.103*	0.077*	0.069	0.096**	0.155
	(0.033)	(0.059)	(0.043)	(0.043)	(0.041)	(0.060)
Immigrant	-0.479***	-0.249*	-0.358***	-0.528***	-0.584***	-0.868
	(0.073)	(0.140)	(0.112)	(0.148)	(0.144)	(0.173)
Public sector	0.108***	0.150**	0.134**	0.066	0.067	0.066
	(0.039)	(0.073)	(0.066)	(0.051)	(0.060)	(0.072)
Permanent contract	0.121***	0.194***	0.160***	0.111**	0.029	0.120
	(0.035)	(0.062)	(0.060)	(0.048)	(0.041)	(0.051)
Responsibility						
Directive	-0.050	0.018	0.017	-0.150	0.079	0.060
	(0.129)	(0.210)	(0.138)	(0.127)	(0.184)	(0.176)
Supervisor	0.075*	0.081	0.094	0.096**	0.070	0.067
	(0.045)	(0.088)	(0.061)	(0.049)	(0.066)	(0.078)
Tenure						
2-5 years	0.072*	0.027	0.046	0.096**	0.050	0.038
	(0.039)	(0.066)	(0.056)	(0.039)	(0.050)	(0.058)
6-9 years	0.138***	0.187*	0.133*	0.163***	0.205***	0.135
	(0.047)	(0.109)	(0.076)	(0.060)	(0.063)	(0.065)
10-14 years	0.165***	0.167*	0.065	0.149**	0.174*	0.321
	(0.058)	(0.092)	(0.075)	(0.076)	(0.095)	(0.158)
>15 years	0.290***	0.212**	0.199**	0.289***	0.351***	0.350
	(0.049)	(0.107)	(0.081)	(0.065)	(0.062)	(0.079)
Married	0.065**	0.040	0.071*	0.093***	0.022	0.086
	(0.027)	(0.050)	(0.043)	(0.032)	(0.033)	(0.044)
Firm size						
5-19 employees	0.122***	0.137*	0.162***	0.145***	0.097*	0.063
	(0.040)	(0.074)	(0.062)	(0.056)	(0.056)	(0.060)
20-49 employees	0.255***	0.360***	0.270***	0.258***	0.177**	0.142
	(0.044)	(0.086)	(0.062)	(0.058)	(0.069)	(0.070)
50-99 employees	0.211***	0.286***	0.254***	0.191***	0.134*	0.034
	(0.053)	(0.076)	(0.069)	(0.066)	(0.073)	(0.087)
100-499 employees	0.266***	0.313***	0.293***	0.290***	0.213***	0.182
	(0.048)	(0.076)	(0.068)	(0.058)	(0.057)	(0.066)
> 500 employees	0.324***	0.281***	0.329***	0.392***	0.347***	0.194
	(0.060)	(0.092)	(0.082)	(0.075)	(0.076)	(0.082)

Table 4b. (continuation)

WOMEN	Average	θ=10	θ=25	θ=50	θ=75	θ=90
Occupation						
OC1	0.917***	0.464	0.410	0.757*	1.078**	1.549***
	(0.281)	(0.330)	(0.392)	(0.449)	(0.522)	(0.438)
OC2	0	0	0	0	0	0
OC3	1.139***	1.368*	1.353**	1.285**	0.873*	0.824**
	(0.077)	(0.712)	(0.636)	(0.582)	(0.463)	(0.398)
OC4	0.527***	0.577*	0.384	0.637**	0.748**	0.619*
	(0.174)	(0.312)	(0.288)	(0.277)	(0.301)	(0.352)
OC5	0.112	0.157	0.163	0.171	0.001	-0.053
	(0.084)	(0.155)	(0.116)	(0.109)	(0.128)	(0.125)
OC6	0.299***	0.152	0.243*	0.378***	0.344***	0.501***
	(0.078)	(0.143)	(0.124)	(0.101)	(0.132)	(0.168)
OC7	0.237***	0.283***	0.214***	0.254***	0.211***	0.222**
	(0.048)	(0.080)	(0.055)	(0.061)	(0.065)	(0.090)
OC8	0.081*	0.040	0.009	0.111**	0.034	0.203**
	(0.046)	(0.081)	(0.060)	(0.054)	(0.060)	(0.082)
OC9	0.081*	0.152*	0.082	0.080	0.039	0.051
	(0.047)	(0.091)	(0.075)	(0.061)	(0.063)	(0.073)
OC10	-0.402***	-0.383	-0.406**	-0.422***	-0.446***	-0.501***
	(0.090)	(0.257)	(0.198)	(0.120)	(0.085)	(0.120)
OC11	0.020	0.001	-0.011	0.016	0.007	-0.001
	(0.052)	(0.109)	(0.071)	(0.067)	(0.071)	(0.071)
OC12	0.144	0.303*	0.146	-0.046	0.142	0.269
	(0.136)	(0.178)	(0.151)	(0.231)	(0.299)	(0.268)
OC13	0.049	0.152	0.018	0.003	0.039	0.056
	(0.055)	(0.101)	(0.078)	(0.085)	(0.073)	(0.097)
OC15	-0.035	-0.243*	-0.056	-0.042	0.007	-0.065
	(0.058)	(0.130)	(0.088)	(0.072)	(0.067)	(0.086)
Nº Obs.	626	626	626	626	626	626
R^2	0.308	0.372	0.372	0.385	0.421	0.462

Note: ***, **, * represent significance at 99, 95 and 90% respectively. Standard deviation in parenthesis. Dummy variables for region and local council size included. Omitted group: wage earners in private sector in less-than-5-employees firms, without responsability, with less-than-1-year tenure, single, with primary education and in non-qualified jobs in services and commerce (OC14)

Table 5a. Participation Probit Equation
L-group. Spain. 1999.

Parents 65-75	-0.120
	(0.157)
Parents >76	-0.032
	(0.176)
Children 0-11	-0.587***
	(0.143)
Children 12-16	-0.078
	(0.198)
Married	-0.367***
	(0.090)
Household income/100	-0.0006
	(0.00002)
Household n.l. income/100	0.04
	(0.0004)
Secondary ed.	0.364***
	(0.077)
Age	-0.004
	(0.007)
Experience	0.068***
	(0.011)
Experience^2	-0.001***
	(0.0003)
Exp*Children 0-11	-0.017**
	(0.008)
Exp*Children 12-16	-0.012
	(0.009)
Nº Obs.	1628
Pseudo R^2	0.169

Note: ***, **, * represent significance at 99, 95 and 90% respectively. Dummy variables for region and local council size included

Table 5b. OLS and QR
L-group (with selection-bias correction). Spain. 1999
Dependent variable: Ln. gross hourly wage

WOMEN	**Average**	**θ=10**	**θ=25**	**θ=50**	**θ=75**	**θ=90**
Experience	0.033^{***}	0.036^{*}	0.038^{***}	0.051^{***}	0.021	0.041^{**}
	(0.011)	(0.020)	(0.013)	(0.014)	(0.013)	(0.019)
Experience^2	0.000^{***}	-0.001	-0.001^{***}	-0.001^{***}	-0.0003	-0.001^{**}
	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0003)
Exp*Children 0-11	-0.002	-0.012	-0.008	-0.011	0.006	-0.007
	(0.008)	(0.015)	(0.009)	(0.010)	(0.009)	(0.014)
Exp*Children 12-16	-0.003	-0.008^{*}	-0.006^{*}	-0.006^{*}	-0.002	-0.001
	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.005)
Secondary ed.	0.353^{***}	0.249^{*}	0.310^{***}	0.430^{***}	0.291^{***}	0.425^{***}
	(0.070)	(0.149)	(0.089)	(0.077)	(0.075)	(0.107)
Married	0.028	-0.054	-0.021	-0.046	0.149^{*}	0.011
	(0.075)	(0.146)	(0.082)	(0.071)	(0.078)	(0.122)
λ	0.138	0.404	0.225	0.370	-0.128	0.286
	(0.239)	(0.443)	(0.253)	(0.248)	(0.275)	(0.420)
Nº obs.	626	626	626	626	626	626
R^2	0.3946	0.1313	0.1477	0.1956	0.2355	0.2512

Note: ***, **, * represent significance at 99, 95 and 90% respectively. Standard deviations in parenthesis. Dummy variables for region and local council size included

Table 6a. Gender wage gag decomposition
H-group. Spain. 1999.

	Mean	**θ=10**	**θ=25**	**θ=50**	**θ=75**	**θ=90**
Observed Gap	**9.74**	**8.00**	**10.20**	**10.54**	**16.68**	**24.25**
Basic Set						
Returns	**3.75**	**-1.10**	**-0.72**	**0.41**	**5.82**	**22.76**
		(1.40)	(1.27)	(1.14)	(1.59)	(1.77)
%	38.53	-13.76	-7.10	3.90	34.87	93.88
Characteristics	**5.99**	**4.29**	**6.05**	**9.62**	**5.40**	**3.86**
		(3.06)	(3.62)	(4.1)	(3.09)	(2.61)
%	61.47	53.63	59.33	91.33	32.37	15.91
Residual	**0.00**	**4.82**	**4.87**	**0.50**	**5.46**	**-2.37**
%	0.00	60.25	47.76	4.76	32.73	-9.79
Extended Set						
Returns	**5.46**	**4.08**	**3.54**	**4.53**	**9.12**	**13.57**
		(1.81)	(1.72)	(1.55)	(1.66)	(3.53)
%	56.00	50.96	34.74	42.97	54.65	55.95
Characteristics	**4.29**	**-0.98**	**6.01**	**4.07**	**1.12**	**11.10**
		(3.27)	(3.65)	(4.33)	(3.25)	(3.19)
%	44.00	-12.20	58.96	38.62	6.72	45.79
Residual	**0.00**	**4.90**	**0.64**	**1.94**	**6.44**	**-0.42**
%	0.00	61.24	6.30	18.41	38.63	-1.74
Extended set+occ. dum						
Returns	**11.03**	**-5.37**	**0.53**	**10.13**	**18.53**	**23.81**
		(2.86)	(1.96)	(1.77)	(2.52)	(3.16)
%	113.22	-67.13	5.18	96.13	111.09	98.17
Characteristics	**-1.29**	**8.78**	**8.87**	**-3.17**	**-9.21**	**7.01**
		(4.34)	(3.48)	(3.66)	(3.74)	(3.13)
%	-13.22	109.75	86.99	-30.07	-55.21	28.91
Residual	**0.00**	**4.59**	**0.80**	**3.58**	**7.36**	**-6.57**
%	0.00	57.37	7.83	33.94	44.12	-27.09

Note: Standard deviation in parenthesis. The s.d. has been obtained through 250 replications of the decomposition

Table 6b. Gender wage gag decomposition
L-group. Spain. 1999.

	Mean	θ=10	θ=25	θ=50	θ=75	θ=90
Observed Gap	**22.73**	**33.33**	**24.71**	**20.31**	**14.82**	**20.94**
Basic Set						
Returns	**21.00**	**28.48**	**26.24**	**21.44**	**17.08**	**17.13**
		(1.51)	(0.94)	(0.84)	(0.99)	(1.97)
%	92.39	85.46	106.22	105.55	115.31	81.81
Characteristics	**1.73**	**-0.10**	**1.26**	**0.07**	**1.35**	**-1.71**
		(2.60)	(2.81)	(2.75)	(2.74)	(2.95)
%	7.61	-0.29	5.10	0.34	9.09	-8.17
Residual	**0.00**	**4.94**	**-2.80**	**-1.20**	**-3.61**	**5.52**
%	0.00	14.83	-11.31	-5.89	-24.39	26.36
Extended Set						
Returns	**22.02**	**27.71**	**26.14**	**22.11**	**20.66**	**4.02**
		(1.75)	(1.28)	(1.22)	(1.65)	(3.09)
%	96.87	83.14	105.78	108.85	139.46	19.20
Characteristics	**0.71**	**5.94**	**0.29**	**1.79**	**-5.69**	**-1.69**
		(3.15)	(3.03)	(3.51)	(4.04)	(5.42)
%	3.14	17.82	1.16	8.82	-38.41	-8.07
Residual	**0.00**	**-0.32**	**-1.72**	**-3.59**	**-0.16**	**18.61**
%	0.00	-0.96	-6.95	-17.67	-1.06	88.98
Extended set+occ. dum						
Returns	**24.65**	**34.96**	**29.05**	**26.93**	**19.74**	**12.33**
		(2.27)	(1.61)	(1.52)	(1.56)	(3.26)
%	108.43	104.90	117.56	132.57	133.24	58.88
Characteristics	**-1.92**	**0.81**	**-0.33**	**-1.99**	**-3.06**	**-9.30**
		(3.72)	(3.31)	(3.72)	(3.88)	(5.95)
%	-8.43	2.43	-1.32	-9.80	-20.67	-44.40
Residual	**0.00**	**-2.44**	**-4.01**	**-4.63**	**-1.86**	**17.91**
%	0.00	-7.33	-16.24	-22.78	-12.57	85.53
Basic set+selection bias correction						
Returns	**35.12**	**66.71**	**46.37**	**54.80**	**6.07**	**41.30**
%	154.51	200.15	187.69	269.83	40.96	19.23
Characteristics	**-1.13**	**-1.44**	**-6.43**	**-0.76**	**4.03**	**-13.01**
%	-4.97	-4.31	-26.02	-3.73	27.20	-62.13
Selection	**-11.26**	**-39.11**	**-18.84**	**-31.51**	**8.67**	**-18.46**
%	-49.52	-117.35	-76.24	-155.17	58.51	-88.15
Residual	**0.00**	**7.17**	**3.60**	**-2.22**	**-3.95**	**11.11**
%	0.00	21.51	14.57	-10.94	-26.66	53.05

Note: Standard deviation in parenthesis. The s.d. has been obtained through 250 replications of the decomposition

Figure 1: Gender wage gap. Spain.1999.

Figure 1a. Total population

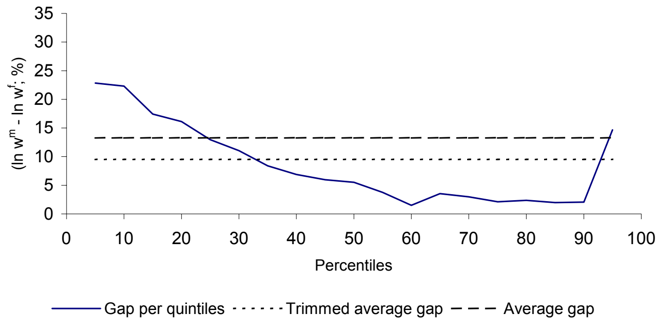


Figure 1b. Tertiary education

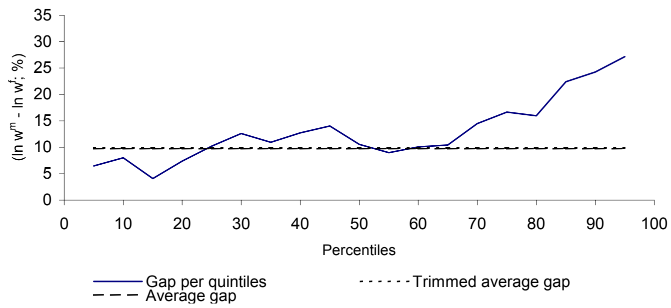
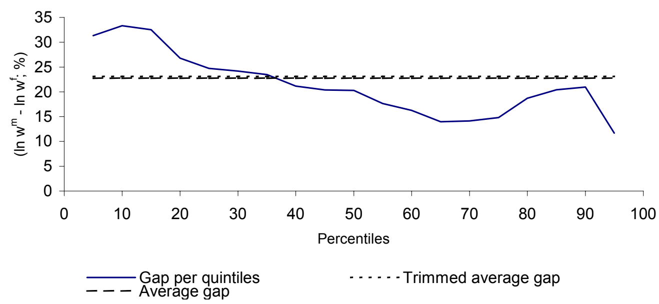


Figure 1c. Less-than-tertiary education



Note: European Community Household Panel (ECHP)

Figure 2: Gender wage gap. European countries. 1999.

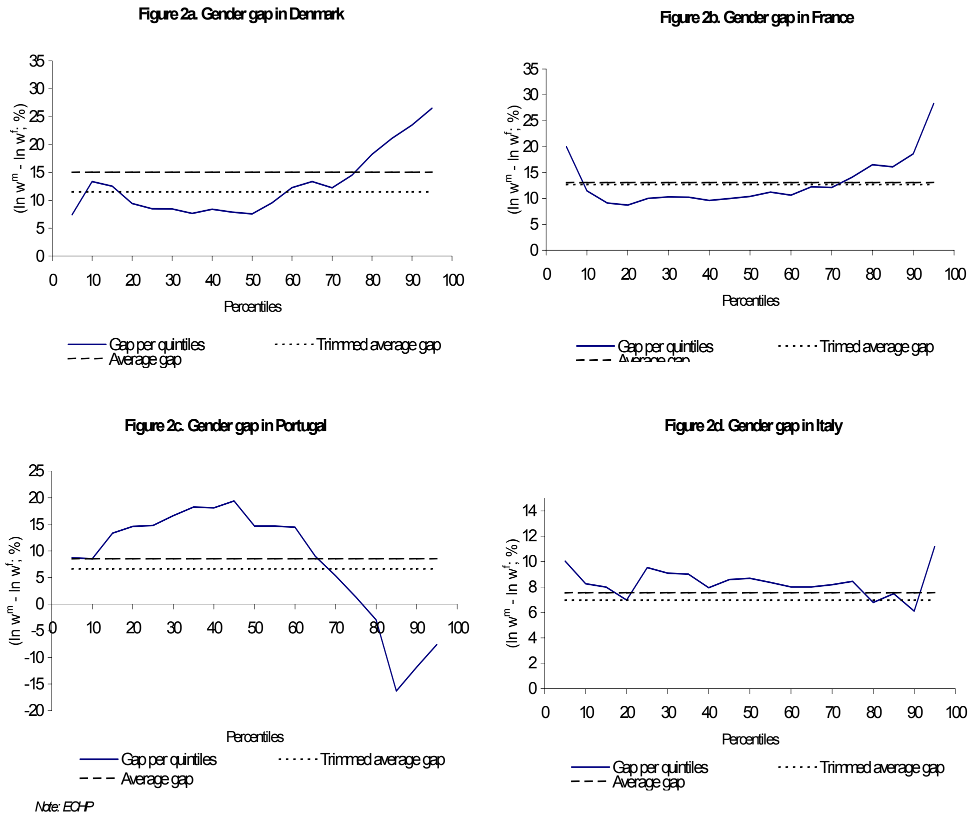


Figure 3: Gender gap (Observed and Counterfactual). H-group. Spain. 1999.

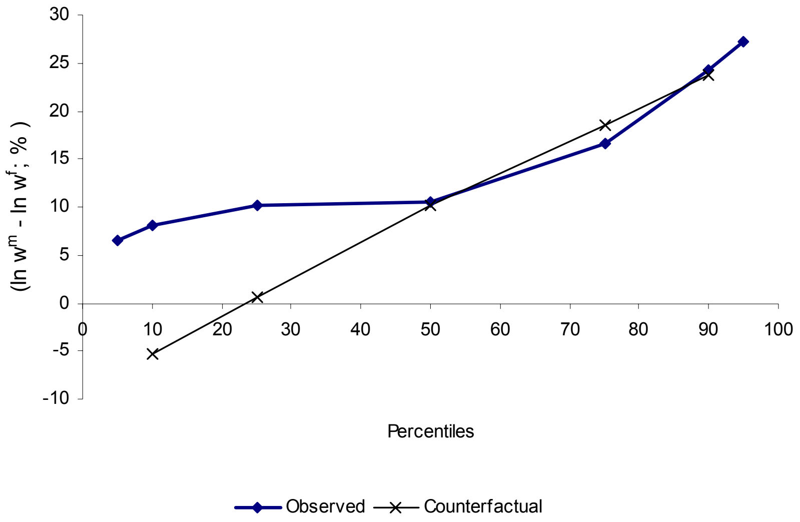


Figure 4: Gender gap (Observed and Counterfactual). L-group. Spain. 1999.

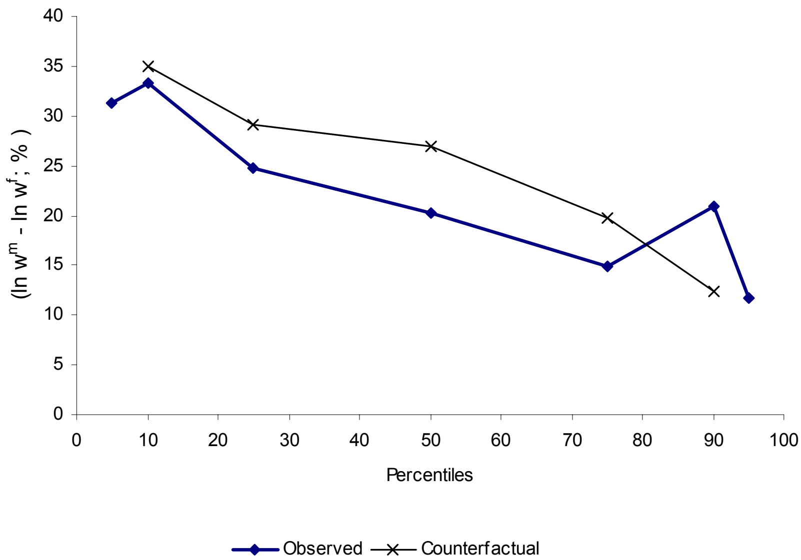


Figure 5: Gender gap (Observed and Counterfactual, sample-selection correction). L-group. Spain. 1999.

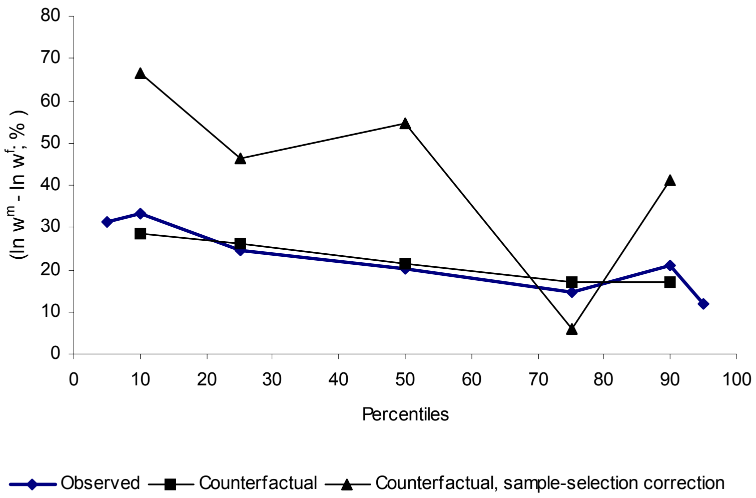


Figure 6: Contribution to factors to the gender gap. H-gourp. Spain. 1999.

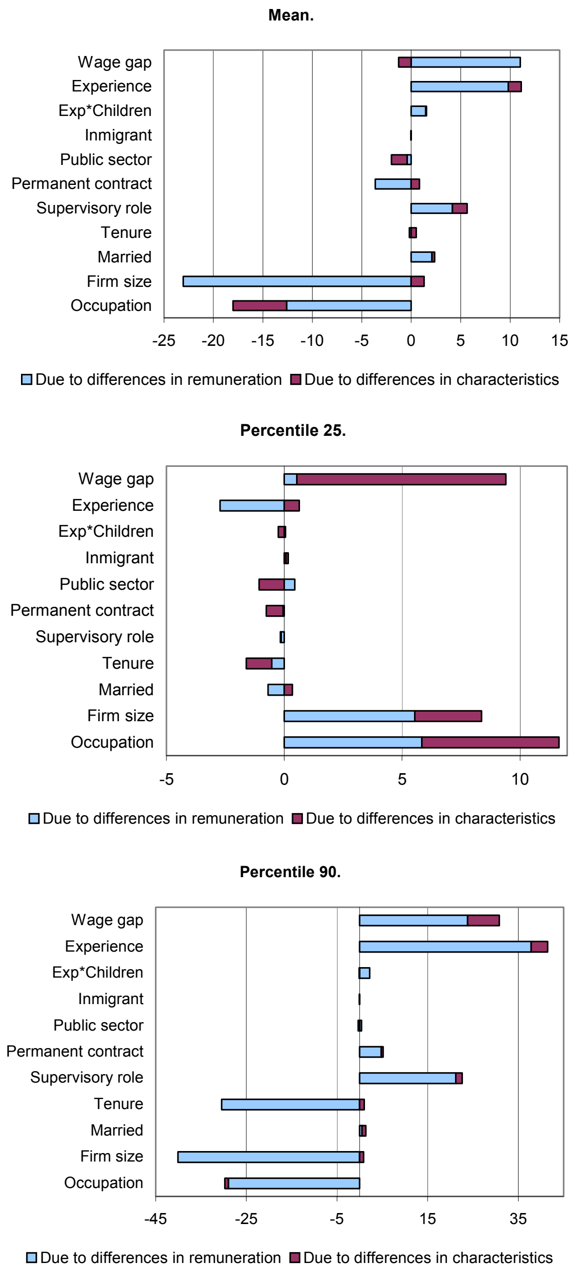


Figure 7: Contribution of factors to the gender gap. L-group. Spain. 1999.