Las Palmas de Gran Canaria, 12-13 de Julio de 2007. VII JEL.

Gender and potential wage in Europe[°]: a stochastic frontier approach

Mª Ángeles Díaz and Rosario Sánchez[*].

[°*] The authors wish to express their gratitude to the participants of the seminar given at the Employment and Social Security Research Centre in Fudan University of Shanghai in July of 2006. Also for helpful comments, we thank the participants of the seminar given in Las Palmas de Gran Canaria (New Issues in Labour Economics, 6-7 of November of 2006) and the participants of the Symposium of Economic Analysis held in Oviedo in December of 2006. We have obtained the data of the European Community Household Panel Survey from EUROSTAT (ECHP contract nº ECHP/2004/17). The project of Science and Technology SEJ2005-08054 has financed this work.

[*]**Corresponding author:** Rosario Sánchez

University of Valencia, Department of Economic Analysis, Faculty of Economics, Campus dels Tarongers, Av. Dels Tarongers s/n, 46022 Valencia, Spain. Fax: +34 963828249; Tel: +34963828252.

E-mail: rosario.sanchez@uv.es

Abstract:

The aim of this paper is to analyse the gender wage gap in Germany, Denmark, France, Italy, Spain and United Kingdom. We use the stochastic frontier approach to study the wage differentials due to gender, after controlling for human capital variables and other individual characteristics. The data set is the ECHP (European Community Household Panel Data) from 1995 to 2001. Our results indicate significant differences among the countries in the size of the wage gap, that range from Denmark, where the potential wage of women are 3% higher than that of men, to Spain where this difference is about 34% against women.

JEL: J24, J31, J71

Key words: Wages differentials, gender, stochastic frontier, inefficiency, Europe

I. Introduction

Men earn higher wages than women even after controlling for measurable characteristics related to their productivity (Blau and Kahn, 2003). The European Employment Strategy has developed a strong compromise with the objective of gender equality in employment. The national action plans include specific goals from 1997 onwards. However, even though this action plans in Europe in 2000, women's hourly wages represented between 80% and 95% of men's hourly wages (European Commission, 2003, p.10).

In this paper, we propose six hourly wages frontiers for Denmark, France, Germany, Italy, Spain and UK to analyse the factors that contribute to explain the differences in the potential wage of women among these countries.

The purpose of this paper is to provide evidence concerning this wages differential and their determinants. The gender gap could be attributed to a lower productivity of women or to a lower market return for a given characteristics, usually related to the discrimination component or to unobserved variables. Using the frontier approach, we also consider the existence of inefficiency in the transformation of human capital into market earnings and their determinants.

The traditional analysis of discrimination consists in the estimation of an earnings equation by gender, separately, following the work of Oaxaca (1973) and Blinder (1973), assuming that males have a non-discriminatory wage structure. Then, with this method, we decompose the wage gap and we explain the differences in the potential wage due to differences in productivity and unexplained differences

attributed to gender discrimination. This methodology has been criticised in two essential ways. Firstly, the residual wage gap could include omitted variables that difficult its interpretation as discrimination. Secondly, because this method uses reduced form wage equations for the two groups and estimates them separately, considering that both groups are identical. However, there may be unobservable differences between the characteristics of the two groups for which the wage equations cannot account[1]. To avoid these difficulties, different methods have been developing (see Beblo et al, 2003). We use the stochastic frontier approach as alternative method. This econometric method includes a one-sided error to capture the possibility of an inefficient behaviour of the economic units analysed when trying to reach an economic objective. The frontier approach is a methodology usually applied to analyse inefficiency in firms' production where a production frontier shows the maximum amount of output attainable with a given level of inputs. If firms do not obtain this maximum level of output, they are inefficient and we could measure their level of inefficiency by the one-sided error term. Likewise, adapting this methodology to the analysis of wage differentials, the earnings frontier will describe the highest potential income associated with a given stock of human capital. If the worker earns less than its potential wage, that indicate inefficiency in the transformation of human capital variables (schooling, experience, and tenure) into earnings and we could analyse which factors explain these differences between the potential and the effective wage. Moreover, this methodology allows us to know if

[1] Robinson, M and P. V. Wunnava (1989)

identical observable workers can achieve different potential wage due to market discrimination.

There is a growing literature that uses the stochastic frontier approach to estimate earnings functions. Among others, we can find the contributions from Hofler and Polacheck (1985), Herzog, Hofler and Schlottmann (1985), Polachek and Yoon (1987), Robinson and Wunnava (1989), Daneshvary et al. (1992), Hunt-McCool and Warren (1993) and Polachek, and Robst (1998). Robinson and Wunnava (1989) paper is one of the first attempts to measure discrimination using the frontier methodology. Hunt-McCool and Warren (1993) considered the differences in potential earnings between black and white workers associated with labour market discrimination when analysing the extent of inefficiency in the transformation of human capital into earnings. More recently, Dawson et al. (2001) uses the stochastic wage frontier to estimate the relative underpayment of females and men. Garcia et al. (2002) analyse the differences in wages by gender and the existence of gender discrimination in Spain while Lang (2005) estimates an earning frontier to search for systematic differences and discrimination between native and immigrants in Germany.

The stochastic frontier methodology contributes to a better estimation of the wage gap and discrimination in the following aspects. It establishes a relationship between the maximum wage attainable by an individual, given their human capital and other personal characteristics, according to the neo-classical theory instead of considering an average wage obtained by the estimation of a reduced wage equation. Then, the earnings function, represent the relation between the human capital

variables (inputs) and the maximum wage attainable (output) and allows to comparing the wage obtained by a worker with their potential or theoretical wage. In addition, the frontier method not only estimates the determinants of wages, like the traditional wage equation procedure but also gives other interesting information. Firstly, it allows evaluating the individual gaps between the frontier and the obtained wage; and secondly, it is possible to analyse the variables that could explain the wage inefficiency (the distance to the frontier) through the model of inefficiency effects proposed by Batesse and Coelli (1995)

The paper is organised as follows: in section 2 we analyse the stochastic frontier methodology and its application to the earnings equation. Section 3 shows the data and variables. Section 4 provides the discussion of results. Finally, in section 5 we present the concluding remarks.

II. Stochastic frontier and the inefficiency model

We use the Stochastic Frontier Approach to estimate an earnings frontier, adding to the standard earnings equation an asymmetric error term representative of the wage inefficiency. Specifically, we use the model of Battese and Coelli (1995), which is a panel data version of Aigner et al. (1977) approach, in which we estimate the wage inefficiency using the stochastic frontier and explain, simultaneously, this inefficiency by a set of variables. This approach avoids the inconsistency problems of the two-stage approach, when analysing the inefficiency determinants[2].

[2] In a two-stage procedure, firstly, a stochastic frontier function is estimated and the inefficiency scores are obtained under the assumption of independently and identically distributed inefficiency

The Battese and Coelli (1995) model is:

$$Y_{it} = f(X_{it};\beta)\exp(v_{it} - u_{it}) \quad (1)$$

Where X is the set of inputs; β is the set of parameters, v_{it} is a two-sided term representing the random error, assumed to be iid $N(0, \sigma_v^2)$; u_{it} is a non-negative random variable representing the inefficiency, which is assumed to be distributed independently and obtained by truncation at zero of $N(\mu_{it},\sigma_u^2)$. The mean of this distribution is a function of a set of explanatory variables:

$$u_{it} = \sum_i \delta_i Z_{it} + \Psi_{it} \quad (2)$$

Where Z_{it} is a (Mx1) vector of variables that may have effects over individuals' outcomes, δ_i is a (1xM) vector of parameters to be estimated and Ψ_{it} is a random variable defined by the truncation of the normal distribution with zero mean and variance σ_u^2.

The function coefficients (β) and the inefficiency model parameters (δ) are estimated by maximum likelihood method together with the variance parameters: $\sigma^2 = \sigma_v^2+\sigma_u^2$; and $\gamma = \sigma_u^2/\sigma^2$.

---

effects. However, in the second step inefficiency effects are assumed to be a function of some firm-specific variables, which contradicts the assumption of identically distributed inefficiency effects.

The difficulty of transforming individual characteristic in outcomes is measured by the ratio of observed wage over the maximum or potential wage obtainable for an individual (when there are not inefficiency), the efficiency (EF) of an individual *i* in year *t* is[3]:

$$EF = \frac{f(X_{it};\beta)\exp(v_{it} - u_{it})}{f(X_{it};\beta)\exp(v_{it})} = \exp(-u_{it}) \qquad (3)$$

The scores obtained from expression (3) take value one when the individual totally transform its characteristics in earnings and less than one otherwise.

Specification of the stochastic earnings frontier

The earnings frontier describes the highest potential income associated with a given stock of human capital. We adopt a standard semi-logarithmic earnings equation (Mincer (1974)) of the type

$$\ln W^{*}_{it} = \alpha + \beta' X_{it} + v_{it} \qquad (4)$$

Where W^{*} is the potential or theoretical wage and X the set of human capital variables. This equation is stochastic and only measurement errors or

[3] Individual efficiency scores u_i, which are unobservable, can be predicted by the mean or the mode of the conditional distribution of u_i given the value of (v_i-u_i) using the technique suggested by Jondrow et al (1982).

misspecification of the model could affect the entire transformation of the human capital endowment into markets earnings.

However, potential or theoretical wage could differ from realized wage, that is, workers could not be able to transform totally their human capital stock into earnings because informational deficiencies in the labour market. We call the difference between potential and realized wage "wage inefficiency" and it is included in the analysis adding a one-sided inefficiency error term to the earnings function, obtaining a frontier. The stochastic earnings frontier represents an upper bound to the earnings. The observed wage (W) could be lower because of measurement errors or inefficiency in the transformation of human capital into earnings, the called "wage inefficiency".

$$\ln W_{it} = \ln W^{*}_{it} - u_{it} = \alpha + \beta' X_{it} + v_{it} - u_{it} \qquad (5)$$

As we focus on gender wage differentials, we consider discrimination as one of the possible reasons explaining differences in earnings. The existence of discrimination means that the potential earnings for one group are lower than it should be according to their human capital and then, the position of the frontier is affected. In this sense, we follow Hunt-McCool and Warren (1993), Garcia et al. (2002) and Lang (2005) measuring the degree of labour market discrimination as the differences in the potential wages for different groups of workers, not explained by different human capital endowment or personal and job-related characteristics. Then,

we allow for the existence of some kind of discrimination with the introduction of a dummy representing specific group of workers into the wage equation. If these dummies are statistically significant with a negative sign, we could not reject the hypothesis of discrimination.

Moreover, the estimated wage inefficiency is explained by a set of variables that proxies some characteristics of the individuals, incomplete information and other market imperfections.

$$u_{it} = \delta_0 + \sum_{i=1}^{19} \delta_i Z_{it} + \Psi_{it} \qquad (6)$$

Where, Z_{it} represents a set of variables that could have effect in explaining the degree of inefficiency in the transformation of human capital into earnings.

Then, we estimate the earnings function for the whole sample, adding a term of inefficiency, whose mean is a function of a set of inefficiency determinants.

From the estimation, we expect a positive relationship between earnings and human capital endowment, according to the direct link between human capital and labour productivity and a negative relationship between earnings and women in the presence of some kind of market discrimination. Notice that we estimate a common frontier for all the sample groups instead of estimate separate earnings functions for different groups. Therefore, we do not restrict wage inefficiency to a disfavoured group measured when comparing to a reference group full efficient.

III. Data and variables

We use data from the ECPH for Denmark, France, Germany, Italy, Spain and United Kingdom to estimate a stochastic frontier, for each country, to investigate the determinants of wage differentials. We analyse the corresponding unbalanced panel of wage earners currently working 15 or more hours per week, from 1995 to 2001. The samples are of employed people with an age ranged from 25 to 65 years old, which remain in the sample at least three consecutive years and work in the industrial and services sectors.

Table 1 shows the mean and standard deviation of variables that indicates the differences among these countries. Female activity rates are still much higher in northern and central European countries than in most southern European ones, despite the catching-up process in participation that has taken place during the last two decades in the latter countries. The percentages in the sample of female workers vary from United Kingdom, which has the highest percentage, to Spain that is the country with the smaller. We have to mention that the level of female unemployment in United Kingdom is as high as that of men, while in Spain this level is nearly twice of men. The average ages of the individuals of the sample are quite similar around 42 years old. The number of children smaller than 12 years old ranges from 0.62, in Italy, to 0.74 in France. Workers of Denmark seem to be more satisfied with respect their job while the Italian workers seem to be the less satisfied.

[Insert Table 1]

Education and experience are important variables in the classical formulation of wage equations (Mincer, 1974; Willis, 1986). Education is a variable that present difficulties of harmonization on an international level. In the ECHP we have only three levels and do not include any specialities. To solve partially that inconvenience we have included in our estimation the variable *formal training* or education that gives the individual skills needed for their present job. As it appears in Table 1 the highest percentage of people with formal training belong to United Kingdom (84%), Denmark (77%) and Germany (76%), followed with some distance for Spain (58%), France (55%) and Italy (36%). Seniority is a variable traditionally included in the estimation of the wage equations. We have used from the ECPH the variable that picks the number of years that individual has been working with the last employer. In France (18%), Italy (17.99%) and Spain (17.32), we have workers with more seniority in the current job than in Denmark (13.63), Germany (12.99) and United Kingdom (8.43%). Mobility allows workers to obtain a better fit in the labour market, so this variable could be related with efficiency. We have on average a 37% of mobility for Denmark, France and United Kingdom while these percentages are smaller for Germany (22%), Italy (27%) and Spain (29%). In this sample Denmark and United Kingdom have a high proportion of services' sector than the other countries. On the opposite, when we compare the differences in type of contract we obtain that Spain is the country with the smallest percentage of permanent workers. The proportion of individuals that work in private sector is bigger in Spain (71%), United Kingdom (70%) and Germany (70%).

The variables

The dependent variable used for estimation is the logarithm of gross hourly wage.

The explanatory variables of the wage equation are:

Trend: It is the time trend.

Gender: This is a dummy variable that takes value one if the individual is a woman, zero otherwise.

Age: This is a set of four dummy variables

From 25 to 35 years old

From 36 to 45 years old

From 46 to 55 years old (category of reference)

More than 55

Occupation in current job: This is a set of eight dummy variables

Legislator, seniors' officials and managers

Professionals

Technicians and associate professionals

Clerks

Service workers and shop and market sales workers

Craft and related trade workers

Plant and machine operators and assemblers

Elementary occupations (category of reference)

Tenure in Industry: this variable measures the seniority (in years) of individual with the actual employer in any firm that belongs to the industrial sector.

Tenure in Services: this variable measures the seniority (in years) of individual with the actual employer in any firm that belongs to the service sector.

Private sector: This is a dummy variable that takes value one when the individual works in the private sector, zero otherwise.

Formal training: Formal training or education that gives workers the skills needed for their present type of work. This is a dummy variable that take value one when they have this formal training.

Permanent contract: This is a dummy variable that takes value one when the worker has a permanent contract, zero otherwise (fixed-term contract or a non-standard contract)

Education Classification: This a set of three dummy variables

Lower: Less than second stage of secondary education (ISCED 0-2)

Second stage: Second stage of secondary level of education (ISCED 3), (this is the category of reference)

Higher: Recognised third level of education (ISCED 5-7)

The inefficiency model:

Previous unemployment: This is a set of four dummy variables

Short-run unemployment, when the individuals have experienced a period of unemployment ranged from 0 to 11 months before to obtain their present job.

Long-run unemployment (12-24), when the individuals have experienced a period of unemployment ranged from 12 to 24 months before to obtain their present job.

Long-run unemployment (more than 24), when the individuals have experienced more than 24 months of unemployment before to obtain their present job

Not unemployed (category of reference)

Mobility: this is a dummy variable that takes value one when the individual has move to another place, area or country, zero otherwise.

Number of children under 12 years old in the household (taking out baby born): This is a set of four dummy variables

0 children (category of reference)

1 child, takes value one when the household has a child, zero otherwise.

2 children, takes value one when the household has two children, zero otherwise.

3 children or more, takes value one when the household has three children, zero otherwise.

Born: This is a dummy that takes value one when a baby born in the household.

Household income: This is the total monthly income of the household taking out the wage of the individual.

Level of job satisfaction: This is a set of three dummy variables:

Not satisfied, this variable takes value one when the individual has a level of satisfaction that range from 1 to 2.

Medium, this variable takes value one when the individual has a level of satisfaction that range from 3 to 4

Fully satisfied, this variable takes value one when the individual has a level of satisfaction that range from 5 to 6.

Gender: This is a dummy variable that takes value one if the individual is a woman, zero otherwise.

Education Classification: This is a set of three dummy variables.

Lower: Less than second stage of secondary education (ISCED 0-2)

Second stage: Second stage of secondary level of education (ISCED 3), (this is the category of reference)

Higher: Recognised third level of education (ISCED 5-7)

IV. Wages differentials

We have estimated six stochastic frontiers corresponding with the six European countries. From the frontier approach, we obtain the measure of individuals inefficiency compared with the best observations of the sample. The value of the estimates allows us to explain the differences in the inefficiency effects among workers inside each country.

The maximum-likelihood estimates of the production frontier parameters, defined in equation (5), given the specification for the inefficiency effects, defined in equation (6), are presented in Table 2. We obtain the estimated coefficient using the computer program FRONTIER 4.1 (Coelli (1995)). At the end of Table 2, we present the tests of the null hypotheses, based on the generalised likelihood ratio (LR) test[4], concerning the relevance of the inefficiency effects.

The variance parameter, γ which lies between zero and one, indicates that technical inefficiency is stochastic and it is relevant to obtain an adequate representation of the data. The value of γ picks up the part of the distance to the frontier explained for the inefficiency. In our estimation, the value of the variance parameter γ range from a value of 0.97 in the case of Denmark to a value of 0.36 in

[4] $LR=-2\{\ln[L(H_0)]-\ln[L(H_1)]\}$, where $L(H_0)$ and $L(H_1)$ are the values of the likelihood function under the null and alternative hypotheses. LR has an approximately chi-square distribution with degrees of freedom equal to the number of restrictions.

the case of Spain. That means that the variance of the inefficiency effects is a significant component of the total error term variance and then, deviations from the potential wage are not only due to random factors.

The first test reported in Table 2 reinforces the relevance of the inefficiency effects in the model. Our results strongly reject the null hypothesis, which considers that the inefficiency effects are not present in the model. Then, the frontier model cannot be reduced to a mean-response wage equation (OLS estimation) to represent accurately the data.

The second test picks the jointly effect of the determinants included in the inefficiency model. We strongly reject the null hypothesis that means that these determinants are not relevant to explain inefficiency.

In this method, we estimate only one wage equation for both men and women for each country. The variables included in these equations determine the potential wage. This is a practise potential wage, obtained from the best observations of the sample. As expected, we obtain that human capital variables had a positive sign in the estimation indicating that people with higher human capital could achieve a higher potential wage. We obtain a negative sign for women in all the countries analysed except for Denmark where the sign is positive. Thus being everything equal to be women reduce the potential wage available respect to men in France, Germany, Italy and Spain. Once controlled for human capital variables, individual and occupational characteristics, the differences between the potential wage for men and women can only be due to discriminatory factors.

The wage equation

Here we define the wage frontier as the maximum wage that can acquire an individual given individual, socio-economic and human capital characteristics. As we mentioned above the estimated coefficients of the wage equations are in Table 2.

[Insert Table 2]

We obtain a positive and significant time trend for France, Italy, Spain and UK while we obtain a negative and significant coefficient for Denmark. The augment of the potential wage during the analysed period range from the 6% of France to the 4% of Italy, Spain and UK.

The **human capital** variables are significant and they have the expected sign. Here we have two sets of variables that pick the effect of education and training in the potential wage of individuals. The ECHP have grouped the years of education in three levels: lower, second stage and higher education, as we have defined in section III. In terms of higher education we can dive the countries in three groups; countries in which women have a lower level of higher education than men (United Kingdom); those in which these levels are equivalent (Denmark and Italy) and those in which women have a higher level than men (France and Spain). From the whole sample we have make a sample selection of working people, more precisely, salaried workers that works 15 or more hours. These selections increased the average level of education for both, men and women. For women the differences in education with respect the whole sample are higher than for men in every country and over all in

Spain (where the proportion of women with higher education rises from 28.8% for the whole sample to 48.3% for the sample restricted to salaried workers).

As we expected, to have primary education reduce the potential wage that could acquire the individual with respect to have secondary education, which is the category of reference in every country with the exception of Denmark where we obtain the opposite sign. When we analyse higher education Denmark is once again eccentric because in this case to have higher education reduce around a six percent the potential wage. In table 1 we can check that United Kingdom (50%) and Italy (13%) have the highest and the lowest percentage of workers with higher education. Therefore, as it was expected the returns of education are bigger in Italy (44%) than in United Kingdom (10%). Therefore, to have university education increases the potential wage in Italy in a 44% while in UK this increment is around 10%.

The tenure of individuals, measured in years, includes two levels that try to proxy the specific training of individuals by sectors of activity (Industry and Services). The coefficients of these variables have a positive and significant sign with the exception of Denmark, which is not significant for the service sector, indicating that tenure increases the potential wage in the Industrial and Services sectors.

With the variable formal training, we measure if an individual have received specific education that gives the skills for its present type of work. The coefficient of this variable is positive and significant for the six countries analyzed. As we can check in Table 1, United Kingdom, Denmark and Germany are the countries with a high percentage of formal training while Italy is the country with the low percentage.

Given everything equal, to have formal training increases the potential wage in a rank of 12.57% for Denmark to around a 5% in Spain.

The **individual characteristic variables** are age and gender. As we expected as higher is the age of individual higher is the potential wage. The relation between age and wages is not lineal. The results indicate that in Denmark, Spain and United Kingdom earnings increase with age until 55 years old. In France, Germany and Italy to be a worker between 36 to 45 years old reduce the potential wage with respect 46 to 55 years old. In addition, in France to be older than 56 increases the potential wage around a six percent while for Germany and Italy to be older than 56 do not make significant differences with respect 46 to 55 years old that is the category of reference.

In absence of discrimination, gender is a variable that should not affect the potential earnings of individuals. However, the sign of this variable in our estimation is negative and significant. That means that to be a woman reduces the potential available earnings related to man. In this type of estimation, the coefficient of this variable measures the extent of the wage discrimination against women. Here we observe differences among the countries. In Denmark, the coefficient is positive and significant meaning that to be women increases the potential wage in a 3%. We obtain the opposite result for France, Germany, Italy, Spain and UK, where the coefficient is negative and significant. Spain is the country where the reduction in the women potential wage is higher (34%), while Italy presents the smallest coefficient (10.6%).

We include **type of contracts** as another source of wages differentials. Here we have two categories: permanent and fixed term contract that is the category of reference. Once controlled for human capital variables, individual and occupational characteristics, the type of contract have a positive effect over potential wage. The permanent contract variable increases the potential earning in every country from a value of 4.1% in Denmark to 13% in Spain.

To work in the **private sector** increase the potential wage for Danish workers while reduce the potential wage for French, Italian, Spanish and British employees.

With the **occupational variables**, we have controlled the wages differentials due to the differences in occupations. The category of reference is elementary occupations.

Wages differentials and the inefficiency model

We can define wage inefficiency as the distance between the wages that earn a particular group of workers and what they could effectively earn given their observed characteristics. These differences (distance to the frontier) could be explained by the existence of labour market imperfections that makes difficult and expensive for workers the job search process.

Table 2 shows the estimated parameters of the model of inefficiency after those corresponding to the wage equations. Now we will comment the most relevant results obtained in relation to the observed wage inefficiency (a positive sign reflects an increase in the inefficiency) in each of the six European countries analysed.

As we expected individuals that have experienced in previous periods, both **short** and **long run unemployment** are less efficient to be near the average wage

related to their individual, social and educational characteristics. Here we can differentiate between two situations: people that never suffered a dismissal and others that suffered one or more periods of unemployment along of their working life. Possibly, when individuals belonging to the latest group and are hired again the expected lost of human capital skills reduce their effective wage.

One possible explanation is the negative effect that unemployment, overall long run unemployment, has over their reservation wage. Therefore, people that experiment unemployment spells have a higher probability of accepting small wages given a level of human capital endowments. Once again, we obtain different results by countries. From Germany and UK, we obtain positive and significant coefficients for both, short and long run unemployment, with respect do not experience any previous unemployment spell which is the category of reference. In both cases the inefficiency increases with the unemployment spell, meaning that as higher is the period in which the individual has been unemployed higher is the inefficiency to adjust its characteristics to its potential wage. We obtain a negative coefficient for Denmark and Italy but are only significant in the case of Denmark. Therefore, the Danish people that have experienced a period of short-run unemployment reduce inefficiency in obtaining a better wage.

As expected, the variable representing **mobility** is negatively related to inefficiency. As more reluctant to move the worker be, the greater is their distance to the frontier. In this case, Denmark, Germany, Spain and UK have negative and significant coefficient.

The presence in the household of dependants **children** affect differently depending on the countries. Here we have a set of four dummy variables indicating the number of children in the household and the category of reference is the absence of children smaller than twelve years old. In addition, we have included a variable that indicates the born of a baby in this period. The coefficient of this variable is negative and significant for France (one, two and three or more children), Denmark, Germany and Spain (two and three or more children) and UK (for one and two children) what means that individuals with dependants under twelve are closer to the frontier. In addition, the results obtained for baby born in the household for Denmark, France, Germany and UK is negative and significant. These results could be related with the signal that this type of workers (especially men) send to the market in the sense that they are more stable workers. If these characteristics reduce the probability of quit then they could be promoted in a higher proportion than single people could without dependants. From another perspective, family characteristics and education are important factors in employment selection mechanisms. The difficulty here is the endogenous choice: less educated women have fewer job opportunities and then tend to have more children; or women that are more educated wait to have children once they have obtained a stable job. A recent work of Anh and Mira (2002) demonstrate a change of sign in the classic correlation between number of children and access to employment (notably in the Southern countries). This phenomenon can be related to the fall in fertility rates and the rise in levels of education. In addition, it seems that workers of the sample are closer to their wage frontier as higher is the number of children. It could reflect the fact that children are

income and time consuming inside the household and it motivate parents to be more efficient in adjusting their potential wage.

The **household income** is a measure of the non-labour income of the individual we obtain a negative and significant coefficient for all the countries with the exception of Germany where the coefficient is positive and significant. Except for Germany, these results indicate that having non-labour income allows workers to obtain a wage closer to their frontier. In any case, the impact of this variable to the inefficiency model vanishes to zero.

The **job satisfaction** variable measures the subjective relation between wages and job characteristics for workers. The impact of job satisfaction varies among the countries. To be fully satisfied is significant and negatively related with inefficiency with the exception of UK. To be not satisfied is significant and positively related with inefficiency with the exception now of Denmark.

In the wage equation we obtain that to be **woman** reduce the potential wage in all countries except Denmark. However, we obtain a negative and significant coefficient of women in the inefficiency model for French and Spanish women. That means that even women are discriminated in wages they are more efficient in reducing the distant to their wage frontier. Thus, the gender differences reflect females' lower promotion probability (or receive lower wage gains consequent upon promotion) not within job discrimination, see Booth et al (2003) and Arulampalam (2005). As women are less promoted, the range of wages that can achieve is reduced compared with men. A recent study of De la Rica et al, (2007) found that the wage gap in Spain is much flatter than in the Northern countries. They explain the

differences trough the existence of statistical discrimination especially for the group of women of primary and secondary education due to the historical low participation rate of this group. They found a composition effect in the overall gender gap when they lump together this group with the group of tertiary education concluding that there is a glass floor for the group of low education while there is a glass ceiling for the group of higher education. Notice that we obtain evidence in favour of the ceiling glass. From the estimation of the wage equation and the inefficiency model, we obtain a narrowed range of variability of wages for women, due of two facts. The first one, related with the loss of their potential wage obtained in the wage equation and the second because they are more efficient in approaching to their potential wage. For Danish and British women we have a positive and significant coefficient.

The coefficients of **education** are significant and have the expected signs. We have three dummy variables that reflect the level of education of individuals. In the wage equation, we have obtained that there are a positive correlation between education and wages. Here we analyse how efficient are the individuals, with different levels of education, to be more or less close to the frontier. For France, Germany, Italy, Spain and UK we obtain a negative and significant coefficient for lower education while positive and significant for higher education. Once more for Denmark, we obtain the opposite result.

As we expected, the individuals with primary education are closer to their wage frontier. People with primary education have a reduced possibility of job match than the rest of workers with higher education. Thus, these workers are more concentrated around their average wages that correspond with the lowest skilled jobs

in the ranking of jobs and it implies that they are closer to their potential wage. As the level of education increases, the rank of jobs that could occupy the individual augment and it implies a higher variability in wages. That means that with the highest level of education, workers could occupy a wide range of jobs and then we could find more people with wages below their potential wage.

V. Concluding remarks

In this paper, we have studied through the stochastic frontier analysis the wage gap in six European countries. Especially we focus in wages differentials due to gender, analysing if the gender gap could be attributed to a lower productivity of women or to a lower market return for a given characteristics, usually related to discrimination. We also consider the existence of inefficiency in the transformation of human capital into market earnings (explained by the existence of labour market imperfections that makes difficult and expensive for workers the job search process) and their determinants.

Using the frontier approach, we propose six hourly wages frontiers for Denmark, France, Germany, Italy, Spain and UK to analyse the factors that contribute to explain the differences in the potential wage of women among these countries.

The earnings frontier will describe the highest potential income associated with a given stock of human capital. This method allows us to consider that there could be differences between the potential and the effective wage due to the existence of inefficiency in the transformation of human capital variables (schooling,

experience, and tenure) into earnings and we could analyse which factors explain these differences. Moreover, this methodology allows us to know if identical observable workers can achieve different potential wage due to market discrimination.

From the estimation, we obtain a positive relationship between earnings and human capital endowment, according to the direct link between human capital and labour productivity and a negative relationship between earnings and gender and type of contract dummies.

Our main result shows that, after controlled for human capital variables and other personal characteristics, in France, Germany, Italy, Spain, and UK the potential wage that can earn women are smaller than that corresponding with men. In absence of discrimination, gender is a variable that should not affect the potential earning of individuals. However, the sign of this variable in our estimation is negative and significant. That means that to be a woman reduces the potential available earnings related to man. In this type of estimation, the coefficient of this variable measures the extent of the wage discrimination against women

From the inefficiency effects model, one of the explanations of this finding is that this wages differential reflects lower promotion probability not within job discrimination. As women are less promoted, the range of wages that can achieve is reduced compared with men so they are more efficient because they have less opportunities and it allows being more concentrated around an average wage.

In addition, we have obtained that a higher level of education increases the potential wage for French, German, Italian, Spanish and British workers while

reduce the wage possibilities for Danish employees. Also, the tenure of individuals, measured in years, have a positive and significant sign with the exception of Denmark

The results obtained with the inefficiency model indicate that there are two groups of education, the first group with primary education and the second group with university studies. Individuals with primary education are closer to the wage frontier because they only can match with lower jobs in the queue of jobs while people with higher education can occupy all rank of jobs available in the labour market. People that are in jobs of a lower category about their skills are far away from their frontier.

VI. References

Aigner, D., Lovell, K., Schmidt, P. (1977), Formulation and estimation of stochastic frontier production function models. Journal of Econometrics, **6**, 21-37

Arulampalam, W., Bototh, A.L, and Bryan, M.L. (2005), Is there a glass ceiling over Europe? Exploring the gender pay gap across the wages distribution, ISER Working Paper, 2005-25, University of Essex, Colchester, UK.

Ahn N. and Mira P. (2002), A note on the changing relations between fertility and female employment rates in developed countries, *Journal of Population Economics*, **15**, 667-682.

Battese, G.E. and Coelli, T.J. (1995), A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics,* **20**, 325-332.

Beblo M., Beninger, D. Heinze, A. and Laisney F. (2003), Methodological issues related to the analysis of gender gaps in employment, earnings and career progression, Raport pour la la Commission européen- DG Emploi et affaires socials.

Blau, F. and Kahn, L.M (2003), Understanding international differences in the gender pay gap, *Journal of labour Economics*, **21**, 106-44.

Blinder, A. (1973), Wage discrimination: reduced form and structural estimates, *Journal of Human Resources,* **8** (4), 436-455

Booth, A. L., Francesconi, M. and Frank J. (2003), A Sticky floors model of promotion, pay and gender, European Economic Review, **47**, 295-322.

Coelli, T., 1996, A Guide to FRONTIER Version 4.1: A computer program for stochastic frontier production and cost function estimation. CEPA Working Paper 96/07, University of New England.

Daneshvary, N., Herzog, Jr. H. W., Hofler, R. A. and Schlottmann, A. M. (1992), Job search and immigrant assimilation: an earnings frontier approach, *Review of Economics and Statistics* **74**, 482-492

Dawson, P., Hinks, T., Watson, D. (2001), German wage underpayment: an investigation into labor market inefficiency and discrimination. *Vierteljahrshefte zur Wirtschaftsforschung* ,**70**, 107-114

De la Rica, S; Dolado, J and Llorens V. (2007), Ceiling or floors?: gender wage gaps by education in Spain. *Journal of Population Economics, forthcoming.*

Garcia, C., Martín, A., Pérez, C. (2002), Using stochastic frontiers to measure gender wage discrimination: The Spanish case. *Universidad de Valladolid*, *Documento de Trabajo FAE* 01/02.

Herzog, Hofler and Schlottmann (1985), Life on the frontier: migrant information, earnings and past mobility, *Review of Economics and Statistics*, **67**. 373-382

Hofler and Polacheck (1985), A new approach for measuring wage ignorance in the labor market, *Journal of Economics and Business,* **37**, 267-276

Hunt-McCool and Warren (1993), Earnings frontiers and labor market efficiency, In Fried, H. (ed.) *The Measurement of Productive Efficiency: Techniques and Applications.* Oxford University Press.

Jondrow, J., Lovell, K., Materov, I.S.and Schmidt, P., (1982), On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics,* **19**, 233-238

Kodde, D. and Palm, F., (1986), Wald criteria for jointly testing equality and inequality restrictions, *Econometrica,* **54-5**, 1243-1248

Lang, G. (2005), The difference between wages and wage potentials: earnings disadvantages of immigrants in Germany, *Journal of Economic Inequality* **3**, 21-42

Lovell, K. (1993), Production frontiers and productive efficiency. In Fried, H. (ed.), *The Measurement of Productive Efficiency: Techniques and Applications.* Oxford University Press.

Mincer, J. (1974) *Schooling, Experience and Earnings*, New York: National Bureau of Economic Research

Oaxaca, R. (1973), Male-female wage differentials in urban labour markets. *International Economic Review,* **14**,693-709

Polachek and Yoon (1987), A two-tiered earnings frontier estimation of employer and employee information in the labour market*. Review of Economics and Statistics* **69 (2)**, 296-302.

Polachek, and Robst (1998), Employee labor market information: comparing direct world of work measures of workers' knowledge to stochastic frontier estimates. *Labour Economics*, **5**, 231-242

Robinson and Wunnava (1989), Measuring direct discrimination in labor markets using a frontier approach: evidence from CPS female earnings data. *Southern Economic Journal*, **56 (1)**, 212-216

Willis, R. (1986), Wage determinants: a survey and reinterpretation of Human capital earnings functions, *Handbook of Labor Economics* vol. 3B, O. Ashenfelter and D. Card (eds.) North-Holland, Amsterdam.

Table1: DESCRIPTIVE STATISTICS (MEAN AND STANDARD DEVIATION)												
Variables	Denmark		France		Germany		Italy		Spain		United Kingdom	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Women	0.37	0.48	0.23	0.42	0.30	0.46	0.28	0.45	0.18	0.38	0.41	0.49
Age	43.75	9.47	41.90	8.79	42.46	9.68	41.69	9.78	42.85	9.39	41.83	9.70
Number of children	0.68	0.98	0.74	0.95	0.64	1.34	0.62	0.84	0.66	0.84	0.62	0.98
Job satisfaction	4.92	0.93	4.40	1.04	------	------	4.05	1.22	4.34	1.26	4.26	1.31
Higher education	0.37	0.48	0.30	0.46	0.29	0.45	0.13	0.33	0.32	0.47	0.50	0.50
Formal training	0.77	0.42	0.55	0.50	0.76	0.43	0.36	0.48	0.58	0.49	0.84	0.37
Seniority	13.63	8.54	18.01	10.03	12.99	7.49	17.99	10.88	17.32	10.67	8.43	6.18
Services	0.72	0.44	0.66	0.47	0.54	0.50	0.67	0.47	0.60	0.49	0.71	0.45
Mobility	0.37	0.48	0.37	0.48	0.22	0.42	0.27	0.44	0.29	0.45	0.37	0.48
Permanent contract	0.95	0.20	0.96	0.20	0.95	0.23	0.93	0.26	0.81	0.39	0.96	0.19
Private sector	0.58	0.49	0.64	0.48	0.70	0.46	0.58	0.49	0.71	0.45	0.70	0.46

Note: Germany do not report in the enquire data for the variable PK001 that measure the job satisfaction.

Table 2- Stochastic Frontier Estimation: The Wage Equation

Variables	Denmark	France	Germany	Italy	Spain	UK
Constant	4.757 (16.17)	5.203 (22.81)	3.135 (14.91)	2.543 (95.85)	7.025 (11.06)	1.649 (54.83)
Time trend						
Trend	-0.0105 (-73.05)	0.0594 (32.49)	0.00005 (1.229)	0.0370 (19.18)	0.0371 (15.32)	0.0415 (20.34)
Gender (Category of reference: man)						
Woman	0.0307 (10.86)	-0.1415 (-14.54)	-0.1907 (17.87)	-0.1062 (-7.75)	-0.3483 (-12.81)	-0.2321 (-23.88)
Age (Category of reference: 46 to 55 years old)						
25 to 35 years old	-0.1487 (-13.49)	-0.1086 (-11.03)	-0.0399 (-4.54)	-0.0994 (-8.22)	-0.1325 (-9.07)	-0.0204 (-2.01)
36 to 45 years old	-0.0844 (-5.99)	-0.0214 (-2.68)	-0.2771 (-3.468)	-0.0210 (-2.03)	-0.0383 (-3.17)	0.0585 (5.97)
More than 56	-0.0273 (-2.34)	0.0582 (3.88)	0.0098 (0.983)	-0.0044 (-0.329)	-0.0602 (-3.84)	-0.0485 (-3.54)
Occupation in current job (Category of reference: elementary occupations)						
Legislators, seniors officials and managers	0.0592 (3.70)	0.6203 (33.06)	0.4024 (22.48)	0.5360 (18.61)	0.6881 (23.01)	0.6435 (36.65)
Professionals	0.4594 (21.52)	0.5998 (35.25)	0.4186 (27.19)	0.3715 (19.06)	0.4751 (20.50)	0.6168 (33.68)
Technicians and associate professionals	0.3151 (17.55)	0.3129 (21.52)	0.2375 (17.02)	0.2309 (13.46)	0.2989 (14.86)	0.5126 (28.39)
Clerks	0.2348 (13.69)	0.1383 (9.18)	0.2009 (13.77)	0.1798 (12.32)	0.2186 (10.74)	0.2990 (17.14)
Service workers and shop and market sales workers	0.1218 (6.37)	0.0489 (3.046)	-0.0543 (-3.34)	0.0679 (4.10)	0.0568 (2.74)	0.1000 (5.51)

Craft and related trade workers	0.0578 (2.77)	0.1101 (7.49)	0.0331 (2.49)	0.0460 (3.23)	0.1018 (5.86)	0.2878 (15.30)
Plant and machine operators and assemblers	0.0369 (1.75)	0.1051 (6.92)	0.0537 (4.03)	0.0999 (6.26)	0.0693 (3.76)	0.1759 (9.24)
Seniority in the economic sector of activity						
Tenure in services	0.0023 (1.20)	0.0080 (19.52)	0.0031 (5.95)	0.0052 (12.15)	0.0086 (13.21)	0.0020 (2.81)
Tenure in industry	0.0022 (3.42)	0.0106 (23.27)	0.0085 (17.05)	0.0040 (8.74)	0.0108 (16.34)	0.0079 (9.07)
Current job (category of reference: public sector)						
Private sector	0.0030 (3.50)	-0.1096 (-14.19)	0.0133 (1.76)	-0.0369 (-4.11)	-0.0749 (-6.55)	-0.0972 (-11.13)
Have you had formal training or education that has given you skills needed for your present type of work? (Category of Reference: No)						
Formal training	0.1257 (11.59)	0.0719 (10.66)	0.1095 (13.64)	0.0694 (8.34)	0.0495 (4.95)	0.1251 (12.16)
Type of contract (Category of reference: fixed-term contact)						
Permanent	0.0411 (3.35)	0.1152 (7.14)	0.0578 (4.42)	0.1080 (7.49)	0.1295 (9.98)	0.0548 (2.95)
Education (Category of reference: second stage)						
Lower	0.0402 (1.91)	-0.1145 (-12.59)	-0.0501 (-4.18)	-0.0876 (-6.13)	-0.2771 (-8.70)	-0.0564 (-4.49)
Higher	-0.0584 (-2.89)	0.1952 (19.35)	0.1545 (12.15)	0.4445 (8.52)	0.2346 (6.54)	0.0976 (7.96)
Inefficiency Model						
Constant	-4.530 (-6.07)	-0.0500 (-16.30)	-1.758 (-12.05)	-0.4703 (1.94)	0.3503 (4.08)	-0.0628 (-13.58)
Previous unemployment (Category of reference: has not experienced any previous unemployment)						
Short-run unemployment	-1.369 (8.37)	0.1567 (4.97)	0.529 (13.10)	-0.0092 (-0.272)	0.0679 (2.25)	0.1465 (2.13)
Long-run unemployment (from 12 to 24 months)	1.0368 (6.05)	0.5751 (5.78)	0.7341 (13.15)	0.1802 (2.96)	0.0656 (1.44)	2.109 (9.83)

Long-run unemployment (more than 24)	-0.7593 (1.63)	0.3431 (3.90)	1.126 (16.52)	0.1349 (2.64)	0.0652 (1.36)	2.487 (11.78)
Mobility to another place, area or country (Category of reference immobility)						
Mobility	-0.8235 (-4.69)	-0.0285 (-0.963)	-0.1515 (-6.04)	0.0432 (1.61)	-0.0893 (-3.03)	-0.5281 (-14.66)
Number of children under 12 years old in the household (category of reference: no child aged under 12)						
1 child	0.2595 (4.261)	-0.9881 (-20.20)	-0.0102 (-0.419)	-0.0211 (-0.79)	-0.0049 (-0.180)	-0.5824 (-11.56)
2 children	-0.6447 (5.67)	-1.119 (-29.21)	-0.4060 (-9.27)	-0.1199 (-1.96)	-0.1186 (-3.01)	-0.5635 (-10.56)
3 or more children	-0.4118 (-4.79)	-1.0346 (-14.68)	-1.1018 (-10.72)	0.1528 (1.89)	-0.2293 (-2.04)	0.8785 (10.72)
Born (category of reference: no born)						
Born	-1.1292 (-4.15)	-1.1585 (-32.12)	-1.1024 (-9.34)	0.0770 (1.50)	-0.0448 (-0.756)	-1.226 (-8.91)
Household income (taking out the own wage)						
Household income	-0.0003 (-5.83)	-0.0001 (-16.18)	0.0002 (5.58)	-0.0002 (-2.37)	-0.0002 (-3.46)	-0.0001 (-14.47)
Level of job satisfaction (category of reference: medium)						
Not satisfied	-0.2851 (9.56)	1.058 (10.97)	_____	0.4210 (3.82)	0.1922 (4.13)	0.3177 (7.73)
Fully satisfied	-1.9948 (-4.791)	-0.4063 (-16.43)	_____	-0.1653 (-3.45)	-0.0767 (-3.43)	0.4475 (11.87)
Gender (Category of reference: man)						
Woman	0.1367 (2.94)	-0.4429 (-7.83)	-0.0200 (-0.552)	0.0750 (1.73)	-0.3865 (-4.12)	0.8623 (9.65)
Education (Category of reference: second stage)						
Lower	-0.3601 (-6.481)	-1.0456 (-19.51)	-1.8876 (-11.33)	-0.2077 (-4.17)	-0.3138 (-4.17)	1.728 (21.59)

Higher	0.9775 (6.26)	2.198 (27.14)	0.7062 (14.38)_	0.7331 (6.03)	0.2007 (3.82)	0.5435 (8.29)
Variance Parameter						
σ^2	1.260 (7.48)	1.066 (17.85)	0.5753 (18.69)	0.1534 (5.23)	0.1504 (13.55)	1.117 (14.37)
γ	0.97 (31.86)	0.9316 (21.66)	0.9021 (14.27)	0.6698 (10.10)	0.3640 (7.53)	0.9063 (12.40)
Generalised likelihood-ratio (LR) tests of null hypotheses [(a)]						
Null hypothesis, H_0	LR Test	LR Test	LR Test	LR Test	LR Test	LR Test
H_0: $\gamma=\delta_0=.....=\delta_{14}=0$; critical value (23.069)[b]	804.87	940.81	1507.53	258.27	284.67	386.55
H_0: $\delta_1=.....=\delta_{14}=0$; critical value (29.1)	55.82	184.69	279.36	186.48	235.79	79.42

[(a)] The test statistics have a χ^2 distribution with degrees of freedom equal to the difference between the parameters involved in the null and alternative hypothesis. The T─ student statistic in brackets.

[(b)] As γ takes values between 0 and 1, in H_0: $\gamma=\delta_0=.....=\delta_{14}=0$ the statistic is distributed according to mixed χ^2 whose critical value is obtained from Kodde and Palm (1986).