COUNCIL OF
THE EUROPEAN UNION

Brussels, 20 November 2001 (23.11)
(OR. fr)

14099/01
ADD 2

LIMITE

SOC 470

ADDENDUM 2 TO NOTE

from:	Working Party on Social Questions
to:	Permanent Representatives Committee (Part 1)
No. prev. doc.:	14005/01 SOC 451 + ADD 1 + ADD 2
Subject:	**PREPARATION FOR THE COUNCIL (EMPLOYMENT AND SOCIAL POLICY) ON 3 DECEMBER 2001** **Addendum 2 to Draft Council conclusions on the review of Member States' implementation of the Beijing Platform for Action: pay inequalities between women and men**

Delegations will find attached an Addendum 2 to the above draft Council conclusions.

Breakdown of the wage gap between men and women
using the Oaxaca technique

Why?

Indicators 1 to 8 are deficient in that they are insufficient to capture the specific effect of the different factors explaining the gender wage gap. In order to have a better understanding of the specific effects of most of the determinants of the wage gap, it is useful to look at the results obtained from econometric estimates of a wage equation. The coefficients obtained, when statistically significant, measure the specific effect of the variable to which they relate, all other things being equal.

Oaxaca (1973) and Blinder (1973) proposed a breakdown of the gender wage gap so as to distinguish the effect of the different factors that determine the gender wage gap. This consists of calculating, on the basis of separate earnings equation estimates for men and women, on the one hand, the share of the average wage gap that is due to differences in the composition of the female and male labour force (at least the differences identifiable on the basis of the available variables) and on the other hand, the share of the gap due to different remuneration for men and women of characteristics that are nevertheless identical. The first element will take into account the wage inequalities that are caused by the concentration of female employment in a limited number of industries, occupations or types of company, for example the different valuation of female and male occupations (the problem of job classifications and wage scales), the effect of differences in the education level, of professional experience (including the indirect effects of career breaks or periods of unemployment). The second element of the wage gap is caused by a different valuation of identical characteristics: the financial return of university education can be different between men and women, or different values can be attached to years of service or to all other variables used in estimating wage equations.

The information provided by this breakdown technique is particularly useful for implementing policies to reduce the earnings gap between men and women. If it appears that differences in occupation contribute significantly to wage inequalities, measures contributing to a better balance between men and women in different occupations, or measures such as a review of job classifications, will be appropriate. If the effect of seniority or experience is significant, the pertinence of career break systems could be challenged. However, if the differences in the remuneration of identical characteristics are significant, the question why, for instance, the same professional experience is rewarded differently in men and women should be investigated. Consequently, this breakdown allows a better follow-up of the evolution of the elements that constitute the wage inequalities between countries and can suggest areas for priority actions.

In view of the replies to the questionnaire "Indicators of gender-related salary inequalities – Follow-up to the Beijing Conference", calculating an Oaxaca-type indicator for all the countries in the European Union is a realistic option. Several Member States refer to studies in which such indicators have been calculated for their countries (Finland, Denmark, Ireland, France, Belgium, Sweden and the Netherlands). Furthermore, the Oaxaca-type breakdown of wage differences carried out by the Department of Applied Economics of the Free University of Brussels (DULBEA) for all European Union countries on the basis of the ECHP (European Community Household Panel) also bears witness to the feasibility of this option.

Given that an Oaxaca-type breakdown is appropriate and feasible, we propose that one be carried out every four years for all European Union countries on the basis of the ESES (European Structure of Earnings Survey). The choice of that database is justified by the quality of the earnings information and the high degree of differentiation of occupations and sectors of activity it contains. However, the absence of data on the public sector is a major obstacle to the use of ESES. It should be supplemented by administrative data on the public sector, as is the case in Finland and the Netherlands, so that the study can cover both the private and public sectors.

How?

Specifically, we propose that the 15 Member States:

1) Calculate the average wage gap between men and women.

This will enable us, for the first time, to compare wage gaps between men and women in the European Union's countries based on a homogenous database containing gross hourly wages. In addition, by repeating this exercise every four years, we will be able to analyse changes in gender wage gaps in each of the member countries, as well as changes in their relative positions.

2) Break down the average wage gap using the Oaxaca technique.

This will enable us to distinguish between remuneration gaps explained by the heterogeneity of characteristics of men and women and the proportion of the wage gap that is still unexplained after these characteristics have been taken into account.

This proposal's technical details and a concrete example for Belgium are given below.

a) Oaxaca breakdown method

Oaxaca (1973) demonstrated that the differential between the average hourly earnings (as a logarithm) of men and women can be broken down as follows:

$$\overline{\ln(W_h)} - \overline{\ln(W_f)} = \overline{X}_h\ (\hat{\beta}_h - \beta^*) + \overline{X}_f(\beta^* - \hat{\beta}_f) + (\overline{X}_h - \overline{X}_f)\beta^* \qquad (1)$$

where

– The indices h and f designate men and women respectively.

– The terms on the left measure the average value of gross hourly earnings (as logarithms) for men and women.

– The term $\overline{X}$ represents the average values or frequency of occurrence of the different variables (educational level, work experience, etc.) for men and women.

– $\hat{\beta}$ is obtained by estimating a separate earnings equation for men and for women:

$$\ln (W_{h,i}) = \beta_h X_{h,i} + \varepsilon_i \longrightarrow \hat{\beta}_h$$

$$\ln (W_{f,i}) = \beta_f X_{f,i} + \varepsilon_i \longrightarrow \hat{\beta}_f$$

$\hat{\beta}$ measures the yield - in terms of earnings - of the variables contained in the vector $\overline{X}$ for men and women.

– β^* corresponds to the value which would be estimated for each of the coefficients in the absence of discrimination, i.e. if $\hat{\beta}$ were identical for men and women, or else if the individual characteristics of men and women were remunerated in equivalent fashion.

What does this expression mean?

– The left-hand side of the equation (1) measures the average wage differential observed between men and women (as a logarithm).

– The first term on the right represents the advantage enjoyed by men. This constitutes the additional earnings which they receive due to the fact that the different variables contained in the vector $\overline{X}$ have a different impact on men's earnings from the one they would have had in the absence of discrimination.

– The second term on the right reflects the disadvantage suffered by women, i.e. the loss of earnings suffered due to the fact that the different variables contained in the vector $\overline{X}$ have a different impact on women's earnings from the one they would have had in the absence of discrimination.

– The third term on the right measures the part of the wage differential which is attributable to differences in the individual characteristics of men and women.

What is the value of β^*? In other words, what is the value of the regression coefficient vector associated with the non-discriminatory earnings structure, i.e. with a natural state where the wage differentials between men and women are explained entirely by differences in their individual characteristics?

There is no consensus in the literature concerning the non-discriminatory earnings structure to be used (for more details see Oaxaca R. and Ransom M. (1994), "On discrimination and the decomposition of wage differentials", *Journal of Econometrics*, Vol.61, No.1, pp.5-21). We nevertheless propose to take that for men as a reference since: i) they make up the greatest proportion of the employed population, and ii) the distribution of male earnings implicitly constitutes the norm in the law governing equality of earnings ("equal pay for equal work"). This decision implies that $\beta^* = \hat{\beta}_h$.

We can therefore rewrite the equation (1) as follows:

$$\overline{\ln(W_h)} - \overline{\ln(W_f)} = \overline{X}_f(\hat{\beta}_h - \hat{\beta}_f) + (\overline{X}_h - \overline{X}_f)\hat{\beta}_h \qquad (2)$$

where

– The left side of the equation measures the average wage differential observed between men and women (as a logarithm).

– The first term on the right constitutes the price effect, i.e. the unexplained wage differential attributable to "pure" wage discrimination.

– The second term on the right measures the explained wage differential, i.e. attributable to differences in the individual characteristics of men and women.

b) Specification of the earnings equation:

Explained variable:

– Naperian logarithm of gross hourly earnings (in national currency), including paid overtime and premiums for shift, night and weekend work. Bonuses such as the 13th monthly wage and profit-sharing are not included.

Explanatory variables:

– Educational level (6 category variables indicating the highest-level qualifications obtained by the individual).

– Previous work experience (number of years spent on the labour market before obtaining the latest job). This variable must be included in the earnings equation: first order, squared and cubed (for further explanations see Murphy K. and Welch F. (1990), "Empirical age-earnings profiles", *Journal of Labour Economics*, Vol.8, No.2, pp.202-29).

– Years of service in the company (number of years spent with the present employer). This variable must be included in the earnings equation: first order and squared. It is also necessary to include a dual variable whose value is 1 if the individual's years of service are strictly in excess of zero and otherwise 0.

– Occupational activity (category variables, ISCO to 2 digits).

– Naperian logarithm of the number of hours remunerated (including the amount of paid overtime).

– Type of contract (3 category variables).

– Premium for shift, night and/or weekend work (dual variable equal to 1 if the individual has received such a premium and otherwise 0).

– Type of economic and financial control at the establishment (category variables).

– Paid overtime (dual variable equal to 1 if the individual has done paid overtime and otherwise 0).

– Sector of activity (category variables, NACE to 2 digits).

– Naperian logarithm of the size of establishment (ln of the number of workers at the establishment).

c) Example for Belgium

The results shown in Tables 1 and 2 are based on the 1995 Survey on Earnings Structure and Distribution. The specification of the earnings equation is as described in b).

As illustrated in Table 1, the average wage differential between men and women was 25,4% in Belgium in 1995. Approximately 50% of this wage differential is explained by the diversity of individual characteristics for men and women (level of human capital, type of employment contract, sector of activity, sectoral affiliation, number of hours remunerated, size of establishment, etc.). In addition, there is a wage differential of some 13% which equates with "pure" wage discrimination.

Table 1: Oaxaca breakdown of the wage differential between men and women

	Average wage differential between men and women:	Proportion of wage differential:	
		explained ≡ $(\bar{X}_h - \bar{X}_f)\hat{\beta}_h$	not explained ≡ $\bar{X}_f(\hat{\beta}_h - \hat{\beta}_f)$
As logarithms	$\overline{\ln(W_h)} - \overline{\ln(W_f)} = 0{,}226$ [= 6,172 – 5,946]	0,108 (48%)	0.118 (52%)
In BEF	$\frac{W_h - W_f}{W_f} = 0.254$ [= (479 – 382)/382]	0,122 (48%)	0,132 (52%)

Source: Survey on Earnings Structure and Distribution, 1995.

Table 2 shows that differences in human capital (educational level, work experience and years of service), occupational activity and sectoral affiliation account for over 75% of the explained wage differential. Indeed, these differences generate a wage differential of 9,2%. Finally, we find that:

1. Sectoral affiliation and human capital contribute more to explaining pay differentials than does occupational activity.

2. Among the human capital variables, years of service in the company is the main source of earnings inequality between men and women.

Table 2: Contribution to earnings inequality

Variables	As logarithms	In BEF	As a % of the explained total
Human capital:	0,031	0,035	28,7%
- educational level	*0,004*	*0,005*	*3,7%*
- work experience	*0,009*	*0,010*	*8,3%*
- years of service	*0,018*	*0,020*	*16,7%*
Occupational activity	0,017	0,019	15,7%
Sector of activity	0,034	0,038	31,5%
Total	0,082	0,092	75,9%

Source: Survey on Earnings Structure and Distribution, 1995.