Selectivity and the gender wage gap decomposition in the presence of a joint decision process

**Marcia M.A. Schafgans*
and
Morton Stelcnery†**

**Discussion paper
No. EM/2006/513
December 2006**

**The Suntory Centre
Suntory and Toyota International Centres for Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel: 020 7955 6679**

---

* Correspondence: Marcia Schafgans, Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, U.K. Email: m.schafgans@lse.ac.uk. I wish to thank Bas Donkers and Stephen Nickell for their support and comments. All remaining errors and omissions are mine.

† In memoriam. 1943-2000. Department of Economics, Concordia University

Abstract

In this paper we revisit the gender decomposition of wages in the presence of selection bias. We show that when labor market participation decisions of couples are not independent, the sample selection corrections used in the literature have been incomplete (incorrect). We derive the appropriate sample selection corrections, based on a reduced form model for the joint participation decisions of both spouses. The influence that husbands' participation decision has on the female participation decision also highlights the importance of using data on both spouses for the analysis of the gender wage gap. Taking account of these issues might influence the outcome of the decomposition analysis and affect the evidence of discrimination. We analyze its potential impact by analyzing the gender earnings differential using Canadian census data.

Keywords: Sample selection model, gender wage differences, Oaxaca wage decomposition, 'discrimination'.

JEL Classifications: C34, J71, D13, J31

1 Introduction

There is a large literature that considers household decisions as a joint decision and bargaining process between husband and wife (Manser and Brown, 1980, McElroy and Horney 1981, and Becker, 1981). Although there is much debate as to what model describes the decision process best, there is agreement on the fact that the spouses decisions are generally interrelated. This is also likely to hold for labor market participation decisions. The literature on the gender wage gap decomposition, however, has neglected this issue, treating the wife's participation decision as independent of that of the husband's.

Since Oaxaca's 1973 influential paper, a lot of empirical research has been devoted to measuring the extend of the gender earnings gap which cannot be explained by wage-related characteristics, also labelled "discrimination". One issue about these decompositions that has received a lot of attention is the assumption made concerning what the level of wages would have been in the absence of discrimination, that is, the nondiscriminatory wage structure.[1] Whereas correcting for the individuals' participation decision has become standard practice in the gender wage decomposition literature (Newman and Oaxaca, 2004), we show that when labor market participation decisions of couples are not independent, the sample selection corrections used have been incomplete (incorrect).

This paper develops the appropriate sample selection corrections when the labor market decisions of spouses are dependent. A reduced form model is used to avoid a strong dependence of the results on a specific bargaining/decision model. The influence that husbands' participation decision has on the female participation decision also highlights the importance of using data on both spouses for the analysis of the gender wage gap.

For the nondiscriminatory wage structure we follow Neumark (1988) and Oaxaca and Ransom (1994). Empirical application of the gender gap decomposition using the Oaxaca-Ransom-Neumark procedure in the presence of sample selectivity (e.g., Mavromaras and Rudolph, 1997) has not accounted for the distinctive and jointly determined participation rules of men and women. A recent discussion of the gender gap decomposition under the pooled nondiscriminatory wage structure (in absence of selectivity corrections) can be found in Fortin (2006). She highlights

[1] Various suggestions have emerged: adopting either the male wage or the female wage structure as the reference wage structure (Oaxaca (1973)); adopting a weighted average of the male and female wage structure (with the proportion of each subgroup in the population used as weights) Cotton (1988); or adopting the pooled wage structure (Neumark, 1988, Oaxaca and Ransom, 1994, Fortin, 2006).

potential problems in the analysis arising from the absence of gender dummies in the pooled regression. Not accounting for distinctive participation rules, just like ignoring a gender dummy in the pooled regression, could lead to the pooled coefficients capturing part of the "between" male and female effects (particularly with large gender differences in explanatory variables and/or participation rules), thereby affecting the outcome of the decomposition analysis and affecting the evidence of discrimination. We illustrate our methodological contribution by analyzing the gender earnings differential using Canadian census data.

The empirical results give strong evidence of correlatedness between the selectivity errors of men and women indicating that unobserved factors positively affecting the husbands' paid employment status affect his wife's paid employment status positively as well, pointing to similarity of partners' unobserved skills. The significant evidence of intrahousehold correlatedness of the unobservables has important consequences concerning the explanatory variables in light of our assumed exogenous regressors assumption. In particular, we note that the definition typically used for unearned income in empirical work "household income minus the individual's earnings" should be avoided due to the resulting endogeneity such a variable would induce (emanating from spousal income). This endogeneity would yield inconsistent parameter estimates.

The empirical results moreover give evidence of a similarity of spouse's unobservable productivity enhancing skills. Moreover, strong evidence is found of correlatedness between the unobservables in the selection and the earnings equation – both individual specific and intrahousehold (between husband and wife) – indicating that sample selectivity corrections are called for. These correlations are found to be negative, indicating that unobservables in the error terms which encourage participation in the wage sector are associated with lower earnings.

Part of the observed wage gap is a direct result of their differential participation decisions – or, the offered wage gap is significantly lower than the observed wage gap, when accounting for selectivity corrections. Taking the distinctive and jointly determined participation rules of men and women into account when obtaining the nondiscriminatory (pooled) wage structure reduces the evidence of discrimination. Moreover, a larger part of the level of "discrimination" should be attributed to a market undervaluation of women than a more favorable market evaluation of observed characteristics for men. Adding additional controls for field of study does not significantly affect the decomposition analysis, but there does appear some reduction in "discrimination" (in the narrower sense of 'unequal pay for equal work') when additional controls for occupation and industry are added.

In Section 2 the model is discussed and the gender wage decomposition is outlined. The data are described in Section 3. The empirical analysis is presented in Section 4, and Section 5 concludes. In the appendix, we provide the derivation of the standard errors of the two-step procedure, elaborate on the pooled earnings regression, provide a glossary of the standard errors of the decomposition terms. Various tables are relayed to the appendix as well.

2 Wage Determination and the Gender Gap

We consider the estimation of standard human capital earnings functions for men and women in the presence of selectivity when labor market participation decisions of couples are not independent. As one may expect the participation and earnings equations for men and women to differ, the model we consider is given by

$$\begin{aligned} d_{ki} &= 1(x'_{ki}\beta_k + \varepsilon_{ki} > 0) \\ y^*_{ki} &= z'_{ki}\theta_k + u_{ki} \\ y_{ki} &= d_{ki}y^*_{ki} \quad \text{for} \quad i = 1, ..., n. \end{aligned} \tag{1}$$

where $k =$ male, female and i denotes a given household. The first equation is the participation equation (where d_{ki} is a dummy variable indicating whether this individual is a wage worker), the individual's latent offered wage (in logarithms) is given by y^*_{ki} and y_{ki} represents his/her observed wage.[2] The variables influencing the decision to participate in the labor market are x_{ki} and z_{ki} are the variables influencing the offered wage. Both x_{ki} and z_{ki} may contain spousal characteristics. Let $x_i = (x_{mi} \cup x_{fi})$ and $z_i = (z_{mi} \cup z_{fi})$. As in the standard selection model, Heckman (1976, 1979), we allow the error terms of the selection and outcome equations of individuals, $(\varepsilon_{ki}, u_{ki})$, to exhibit correlations. In addition, dependencies between husband and wife's error terms are allowed: conditional on (z_i, x_i) the errors have a multivariate normal distribution, given by

$$\left[\begin{array}{c} u_{mi} \\ u_{fi} \\ \varepsilon_{mi} \\ \varepsilon_{fi} \end{array}\right]\Bigg|_{z_i,x_i} \sim N\left(0, \left(\begin{array}{cccc} \sigma^2_{u_m} & \sigma_{u_m u_f} & \sigma_{\varepsilon_m u_m} & \sigma_{\varepsilon_m uf} \\ \sigma_{u_m u_f} & \sigma^2_{u_f} & \sigma_{\varepsilon_m u_m} & \sigma_{\varepsilon_m u_f} \\ \sigma_{\varepsilon_m u_m} & \sigma_{\varepsilon_m u_m} & 1 & \rho_{\varepsilon_m \varepsilon_f} \\ \sigma_{\varepsilon_m uf} & \sigma_{\varepsilon_m u_f} & \rho_{\varepsilon_m \varepsilon_f} & 1 \end{array}\right)\right). \tag{2}$$

For identification purposes, the variances of the selection errors $(\varepsilon_{mi}, \varepsilon_{fi})$ are standardized to equal 1.

[2] An individual chooses to be 'in the sample' of workers ($d_{ki} = 1$), if his/her offered wage exceeds their reservation wage ($y^*_{ki} - y^R_{ki} = x'_{ki}\beta_k + \varepsilon_{ki} > 0$) and $d_{ki} = 0$ otherwise.

Models of the joint participation decision of husbands and wives that consider their respective participation to be endogenous run into the well known difficulties of coherency (e.g., see Heckman (1978) and Tamer (2003)). This type of endogeneity is not explicitly allowed, nevertheless our model can be viewed as the reduced form of such a model.

Modeling the participation decision simultaneously, indeed, allows us to propose selectivity corrections that take account of the joint decision process faced by men and women. The resulting augmented two-step Heckman estimation proceeds as follows. First, a bivariate probit regression is estimated on the participation decisions of husbands and wives in the wage sector. The log likelihood function is given by

$$\begin{aligned}
&\log L(\beta_m, \beta_f, \rho_{\varepsilon_m\varepsilon_f}; x, z) \\
&= \sum_{\substack{i=1 \\ d_{m_i}=d_{f_i}=1}}^{N} \log \Phi_2(x'_{mi}\beta_m, x'_{fi}\beta_f, \rho_{\varepsilon_m\varepsilon_f}) + \sum_{\substack{i=1 \\ d_{m_i}=1, d_{f_i}=0}}^{N} \log \Phi_2(x'_{mi}\beta_m, -x'_{fi}\beta_f, -\rho_{\varepsilon_m\varepsilon_f}) \\
&+ \sum_{\substack{i=1 \\ d_{m_i}=0, d_{f_i}=1}}^{N} \log \Phi_2(-x'_{mi}\beta_m, x'_{fi}\beta_f, -\rho_{\varepsilon_m\varepsilon_f}) + \sum_{\substack{i=1 \\ d_{m_i}=0, d_{f_i}=0}}^{N} \log \Phi_2(-x'_{mi}\beta_m, -x'_{fi}\beta_f, \rho_{\varepsilon_m\varepsilon_f}),
\end{aligned}$$

where, $\Phi_2(\cdot,\cdot,\rho)$ denotes the standardized bivariate normal cumulative density function with correlation ρ.In the setting where $\rho_{\varepsilon_m\varepsilon_f} = 0$, this reduces to running separate probit regressions for men and women. Second, sample selectivity corrected wage functions are estimated that take into account this joint participation decision, in place of introducing gender specific sample selectivity corrections based solely on the individuals' decision to participate or not. Specifically, the correction we propose involves the introduction of two selectivity correction terms in each wage function. The selectivity correction terms are obtained by evaluating for the male wage function the conditional expectations $E(u_{mi}|d_{mi} = 1, d_{fi} = 1, x_i, z_i)$ and $E(u_{mi}|d_{mi} = 1, d_{fi} = 0, x_i, z_i)$. Each can be written as the linear combination of two selection correction terms, λ_{1i} and λ_{2i}, specifically

$$E(u_{mi}|d_{mi} = 1, d_{fi}, x_i, z_i) = \left(\sigma_{u_m\varepsilon_m}\lambda_1(x'_{mi}\beta_m, x'_{fi}\beta_f) + \sigma_{u_m\varepsilon_f}\lambda_2(x'_{mi}\beta_m, x'_{fi}\beta_f)\right) d_{mi}.$$

where

$$\lambda_1(x'_{mi}\beta_m, x'_{fi}\beta_f) = q_{mi}\frac{\phi(\omega_{mi})\Phi(\frac{\omega_{fi}-\rho^*_{\varepsilon_m\varepsilon_f}\omega_{mi}}{\sqrt{1-\rho^{*2}_{\varepsilon_m\varepsilon_f}}})}{\Phi_2(\omega_{mi},\omega_{fi},\rho^*_{\varepsilon_m\varepsilon_f})} \text{ and}$$

$$\lambda_2(x'_{mi}\beta_m, x'_{fi}\beta_f) = q_{fi}\frac{\phi(\omega_{fi})\Phi(\frac{\omega_{mi}-\rho^*_{\varepsilon_m\varepsilon_f}\omega_{fi}}{\sqrt{1-\rho^{*2}_{\varepsilon_m\varepsilon_f}}})}{\Phi_2(\omega_{mi},\omega_{fi},\rho^*_{\varepsilon_m\varepsilon_f})}.$$

The following notation has been used in order to simplify the notation of these selection correction terms:

$$\begin{aligned} q_{ki} &= 2d_{ki} - 1 \\ \omega_{ki} &= q_{ki}\,(x'_{ki}\beta_k) \\ \rho^*_{\varepsilon_m\varepsilon_f} &= q_{mi}q_{fi}\rho_{\varepsilon_m\varepsilon_f}. \end{aligned}$$

The sample selectivity corrected male wage regression, can then be written as

$$y_{mi}d_{mi} = (z'_{mi}\theta_m + \sigma_{u_m\varepsilon_m}\lambda_1(x'_{mi}\beta_m, x'_{fi}\beta_f) + \sigma_{u_m\varepsilon_f}\lambda_2(x'_{mi}\beta_m, x'_{fi}\beta_f) + \nu_{mi})d_{mi}, \tag{3}$$

where $v_{mi} = u_{mi} - d_{fi}E(u_{mi}|d_{mi} = 1, d_{fi} = 1, x_i, z_i) - (1 - d_{fi})E(u_{mi}|d_{mi} = 1, d_{fi} = 0, x_i, z_i)$, and $E(v_{mi}|d_{mi} = 1, x_i, z_i) = 0$ by construction. Note, in absence of any correlation between husbands' and wives' error terms, $\rho_{\varepsilon_m\varepsilon_f} = 0$ and $\rho_{u_m\varepsilon_f} = 0$, the specification reduces to the standard wage equation with the standard inverse Mill's ratio included for the selectivity correction.

Similarly, for women, we obtain

$$y_{fi}d_{fi} = (z'_{fi}\theta_f + \sigma_{u_f\varepsilon_m}\lambda_1(x'_{mi}\beta_m, x'_{fi}\beta_f) + \sigma_{u_f\varepsilon_f}\lambda_2(x'_{mi}\beta_m, x'_{fi}\beta_f) + \nu_{fi})d_{fi}. \tag{4}$$

As in the standard sample selectivity corrected wage regressions, both wage equations (3) and (4) exhibit heteroscedasticity. Moreover, they form a system of seemingly unrelated regressions, due to the correlations between spouses, since $\mathrm{Cov}(v_{mi}, v_{fi}|d_{mi=1}, d_{fi} = 1, x_i, z_i) \neq 0$ in general. Specifically

$$\begin{aligned} \mathrm{Var}(v_{mi}|d_{mi} = 1, d_{fi}, x_i, z_i) &= \sigma^2_{u_m} - \left\{\sigma_{u_m\varepsilon_m}\delta^m_{1i} + \sigma_{u_m\varepsilon_f}\delta^m_{2i}\right\} \\ \mathrm{Var}(v_{fi}|d_{fi} = 1, d_{mi}, x_i, z_i) &= \sigma^2_{u_f} - \left\{\sigma_{u_f\varepsilon_m}\delta^f_{1i} + \sigma_{u_f\varepsilon_f}\delta^f_{2i}\right\} \\ \mathrm{Cov}(v_{mi}v_{fi}|d_{mi} = 1, d_{fi} = 1, x_i, z_i) &= \sigma_{u_m u_f} - \left\{\sigma_{u_f\varepsilon_m}\delta^m_{1i} + \sigma_{u_f\varepsilon_f}\delta^m_{2i}\right\}. \end{aligned} \tag{5}$$

Here

$$\begin{aligned} \delta^k_1(x'_{mi}\beta_m, x'_{fi}\beta_f) &= \sigma_{u_k\varepsilon_m}\left\{\lambda_{1i}(\lambda_{1i} + x'_{mi}\beta_m) + \frac{\rho_{\varepsilon_m\varepsilon_f}}{\sqrt{1-\rho^2_{\varepsilon_m\varepsilon_f}}}\lambda_{3i}\right\} + \sigma_{u_k\varepsilon_f}\left\{\lambda_{1i}\lambda_{2i} - \frac{1}{\sqrt{1-\rho^2_{\varepsilon_m\varepsilon_f}}}\lambda_{3i}\right\} \\ \delta^k_2(x'_{mi}\beta_m, x'_{fi}\beta_f) &= \sigma_{u_k\varepsilon_f}\left\{\lambda_{2i}(\lambda_{2i} + x'_{fi}\beta_f) + \frac{\rho_{\varepsilon_m\varepsilon_f}}{\sqrt{1-\rho^2_{\varepsilon_m\varepsilon_f}}}\lambda_{3i}\right\} + \sigma_{u_k\varepsilon_m}\left\{\lambda_{1i}\lambda_{2i} - \frac{1}{\sqrt{1-\rho^2_{\varepsilon_m\varepsilon_f}}}\lambda_{3i}\right\}, \end{aligned}$$

with λ_{3i} defined as

$$\lambda_3(x'_{mi}\beta_m, x'_{fi}\beta_f) = q_{mi}q_{fi}\frac{\phi(\omega_{mi})\phi\left(\frac{\omega_{fi} - \rho^*_{\varepsilon_m\varepsilon_f}\omega_{mi}}{\sqrt{1-\rho^{*2}_{\varepsilon_m\varepsilon_f}}}\right)}{\Phi_2(\omega_{mi}, \omega_{fi}, \rho^*_{\varepsilon_m\varepsilon_f})}.$$

The conditional variance of v_{mi} can be rewritten as,

$$\text{Var}(v_{mi}|d_{mi}=1,d_{fi},x_i,z_i)$$
$$=\sigma^2_{u_m}\left(1-\rho^2_{u_m\varepsilon_m}\left\{\lambda_{1i}(\lambda_{1i}+x'_{mi}\beta_m)+\frac{\rho_{\varepsilon_m\varepsilon_f}}{\sqrt{1-\rho^2_{\varepsilon_m\varepsilon_f}}}\lambda_{3i}\right\}-\rho^2_{u_m\varepsilon_f}\left\{\lambda_{2i}(\lambda_{2i}+x'_{fi}\beta_f)+\frac{\rho_{\varepsilon_m\varepsilon_f}}{\sqrt{1-\rho^2_{\varepsilon_m\varepsilon_f}}}\lambda_{3i}\right\}\right.$$
$$\left.-2\rho_{u_m\varepsilon_m}\rho_{u_m\varepsilon_f}\left\{\lambda_{1i}\lambda_{2i}-\frac{1}{\sqrt{1-\rho^2_{\varepsilon_m\varepsilon_f}}}\lambda_{3i}\right\}\right). \quad (6)$$

This formulation of the conditional variance facilitates our observation that without any correlations between husbands' and wives' errors, that is $\rho_{\varepsilon_m\varepsilon_f}=0$, $\rho_{u_m\varepsilon_f}=0$, and $\lambda_{1i}=\frac{\phi(x'_{mi}\beta_m)}{\Phi(x'_{mi}\beta_m)}$, the simplification to the standard formula pertains, i.e., $\sigma^2_{u_m}\left(1-\rho^2_{u_m\varepsilon_m}\left\{\frac{\phi(x'_{mi}\beta_m)}{\Phi(x'_{mi}\beta_m)}(\frac{\phi(x'_{mi}\beta_m)}{\Phi(x'_{mi}\beta_m)}+x'_{mi}\beta_m)\right\}\right)$.

In matrix notation, our regression model is given by

$$D\begin{pmatrix} y_m \\ y_f \end{pmatrix}=D\left[\begin{pmatrix} W_m & 0 \\ 0 & W_f \end{pmatrix}\begin{pmatrix} \pi_m \\ \pi_f \end{pmatrix}+\begin{pmatrix} v_m \\ v_f \end{pmatrix}\right],$$

where D is a diagonal matrix with elements $(d_{m1},...,d_{mN},d_{f1},...,d_{fN})$, where N equals the number of households, $W_k=[Z_k|\lambda_1|\lambda_2]$ and $\pi_k=\left[\theta'_k,\sigma_{u_k\varepsilon_m},\sigma_{u_k\varepsilon_f}\right]'$ with $k=$ male, female. Correspondingly, we express

$$Var\left[D\begin{pmatrix} v_m \\ v_f \end{pmatrix}|X,Z\right]\equiv Var\left[\begin{pmatrix} D_m & 0 \\ 0 & D_f \end{pmatrix}\begin{pmatrix} v_m \\ v_f \end{pmatrix}|X,Z\right]$$
$$=\begin{bmatrix} D_m\left(\sigma^2_{u_m}I_N-\Delta_m\right) & D_mD_f(\sigma_{u_mu_f}I_N-\Delta_{mf}) \\ D_mD_f(\sigma_{u_mu_f}I_N-\Delta_{mf}) & D_f(\sigma^2_{u_f}I_N-\Delta_f) \end{bmatrix}\equiv V,$$

where Δ_m is a diagonal matrix with $\sigma_{u_m\varepsilon_m}\delta^m_{1i}+\sigma_{u_m\varepsilon_f}\delta^m_{2i}$ on its diagonal, Δ_f is a diagonal matrix with $\sigma_{u_m\varepsilon_m}\delta^f_{1i}+\sigma_{u_m\varepsilon_f}\delta^f_{2i}$ on its diagonal, and Δ_{mf} is a diagonal matrix with $\sigma_{u_f\varepsilon_m}\delta^m_{1i}+\sigma_{u_f\varepsilon_f}\delta^m_{2i}$ (or equivalently $\sigma_{u_m\varepsilon_m}\delta^f_{1i}+\sigma_{u_m\varepsilon_f}\delta^f_{2i}$) on its diagonal. Below, we use the following decomposition of Δ_m and Δ_f :

$$\Delta_k=\sigma_{u_k\varepsilon_m}\Delta^k_1+\sigma_{u_k\varepsilon_f}\Delta^k_2 \text{ for } k=\text{male,female.}$$

We estimate the male and female wage equations (3) and (4) by ordinary least squares (equation by equation) using consistent estimates of our selectivity correction terms (as Heckman, 1976). That is, we regress,

$$D_ky_k=D_k\left(\hat{W}_k\pi_k+v_k+\eta_k\right), \text{ for } k=\text{male,female,}$$

where $\hat{W}_k = \left[Z_k | \hat{\lambda}_1 | \hat{\lambda}_2 \right]$ and $\eta_k = \sigma_{u_k \varepsilon_m}(\lambda_1 - \hat{\lambda}_1) + \sigma_{u_k \varepsilon_f}(\lambda_2 - \hat{\lambda}_2)$.[3] Our parameter estimates are therefore given by

$$\begin{pmatrix} \hat{\pi}_m \\ \hat{\pi}_f \end{pmatrix} = \begin{pmatrix} \left(\hat{W}_m' D_m \hat{W}_m \right)^{-1} \hat{W}_m' D_m y_m \\ \left(\hat{W}_f' D_f \hat{W}_f \right)^{-1} \hat{W}_f' D_f y_f \end{pmatrix}$$

The standard errors of these parameter estimates are corrected both for the non-scalar covariance matrix (heteroscedasticity and serial correlation) and for the fact that we use consistently estimated selectivity correction terms. In the Appendix, we derive the asymptotic distribution

$$\left. \begin{pmatrix} \hat{\pi}_m \\ \hat{\pi}_f \end{pmatrix} \right|_{X,Z} \overset{a}{\to} N \left(\begin{pmatrix} \pi_m \\ \pi_f \end{pmatrix}, \begin{pmatrix} \left(W_m' D_m W_m \right)^{-1} W_m' D_m \\ \left(W_f' D_f W_f \right)^{-1} W_f' D_f \end{pmatrix} (V + R) \begin{pmatrix} \left(W_m' D_m W_m \right)^{-1} W_m' D_m \\ \left(W_f' D_f W_f \right)^{-1} W_f' D_f \end{pmatrix}' \right).$$

V is the covariance matrix, given above, and

$$R = \begin{pmatrix} \Delta_1^m X_m & \Delta_2^m X_f & \Delta_3^m \\ \Delta_1^f X_m & \Delta_2^f X_f & \Delta_3^f \end{pmatrix} I(\beta_m, \beta_f, \rho_{\varepsilon_m \varepsilon_f})^{-1} \begin{pmatrix} \Delta_1^m X_m & \Delta_2^m X_f & \Delta_3^m \\ \Delta_1^f X_m & \Delta_2^f X_f & \Delta_3^f \end{pmatrix}',$$

where $I(\beta_m, \beta_f, \rho_{\varepsilon_m \varepsilon_f})^{-1}$ denotes the inverse of the information matrix from our first stage bivariate probit model, $X_k = \begin{bmatrix} x_{k1}' \\ \vdots \\ x_{kn}' \end{bmatrix}$ (the matrix of explanatory variables related to the decision to participate), Δ_1^k and Δ_2^k are defined above, and Δ_3^k is a diagonal matrix with δ_{3i}^k on its diagonal. Here δ_{3i}^k is defined as

$$\delta_3^k(x_{mi}'\beta_m, x_{fi}'\beta_f) = \sigma_{u_k \varepsilon_m} \left\{ \frac{1}{\sqrt{1-\rho_{\varepsilon_m \varepsilon_f}^2}} \lambda_{3i} (\lambda_{1i} + \frac{x_{mi}'\beta_m - \rho_{\varepsilon_m \varepsilon_f} x_{fi}'\beta_f}{1-\rho_{\varepsilon_m \varepsilon_f}^2}) \right\} + \sigma_{u_k \varepsilon_f} \left\{ \frac{1}{\sqrt{1-\rho_{\varepsilon_m \varepsilon_f}^2}} \lambda_{3i} (\lambda_{2i} + \frac{x_{fi}'\beta_f - \rho_{\varepsilon_m \varepsilon_f} x_{mi}'\beta_m}{1-\rho_{\varepsilon_m \varepsilon_f}^2}) \right\}.$$

When deriving the correction required to deal with the fact that we use estimated sample selectivity correction terms, we recognize that δ_{1i}^k and δ_{2i}^k (introduced when describing the conditional variances/covariances of u_k) are related to the derivatives of the conditional expectation of the error term, u_k, with respect to the male and female selection parameters β_m and β_f, specifically, $\frac{\partial \left(\sigma_{u_k \varepsilon_m} \lambda_{1i} + \sigma_{u_k \varepsilon_f} \lambda_{2i} \right)}{\partial \beta_m} = -\delta_{1i}^k x_{mi}$ and $\frac{\partial \left(\sigma_{u_k \varepsilon_m} \lambda_{1i} + \sigma_{u_k \varepsilon_f} \lambda_{2i} \right)}{\partial \beta_f} = -\delta_{2i}^k x_{fi}$ (where $\delta_{1i}^k \equiv -\frac{\partial \left(\sigma_{u_k \varepsilon_m} \lambda_{1i} + \sigma_{u_k \varepsilon_f} \lambda_{2i} \right)}{\partial (x_{mi}'\beta_m)}$ and $\delta_{2i}^k \equiv -\frac{\partial \left(\sigma_{u_k \varepsilon_m} \lambda_{1i} + \sigma_{u_k \varepsilon_f} \lambda_{2i} \right)}{\partial (x_{fi}'\beta_f)}$). Correspondingly, we defined $\delta_{3i}^k \equiv$

[3] Alternatively a GLS procedure could have been developed. Asymptotically efficient MLE estimates can be obtained using a one-step efficient Newton iteration from our OLS estimates.

$-\frac{\partial\left(\sigma_{u_k\varepsilon_m}\lambda_{1i}+\sigma_{u_k\varepsilon_f}\lambda_{2i}\right)}{\partial(\rho_{\varepsilon_m\varepsilon_f})}$ as it relates to the derivative of the conditional expectation of the error term with respect to the remaining parameter in our bivariate probit model.

For the computation of the covariance matrix using the two-step estimator, we require consistent estimates of the individual parameters $\sigma^2_{u_m}$, $\sigma^2_{u_f}$, $\rho_{u_m u_f}$, $\rho_{u_m\varepsilon_m}$, $\rho_{u_m\varepsilon_f}$, $\rho_{u_f\varepsilon_m}$, and $\rho_{u_f\varepsilon_f}$. Define $e_k = D_k(y_k - \hat{W}_k\hat{\pi}_k)$. A consistent estimate of the covariance matrix of the outcome equation errors (u_m, u_f) is given by

$$\begin{aligned} \widehat{\sigma}^2_{u_k} &= \frac{e_k' e_k}{\sum_{i=1}^n d_{ki}} + \overline{\hat{\delta}_k}, \text{ for } k = \text{male, female} \\ \widehat{\sigma}_{u_m u_f} &= \frac{e_m' e_f}{\sum_{i=1}^n d_{mi} d_{fi}} + \overline{\hat{\delta}_{mf}}, \end{aligned} \tag{7}$$

where $\overline{\hat{\delta}_k} = \sum_{i=1}^n d_{ki}(\hat{\sigma}_{u_k\varepsilon_m}\hat{\delta}^k_{1i} + \hat{\sigma}_{u_k\varepsilon_f}\hat{\delta}^k_{2i}) / \sum_{i=1}^n d_{ki}$ and $\overline{\hat{\delta}_{mf}} = \sum_{i=1}^n d_{mi}d_{fi}(\hat{\sigma}_{u_f\varepsilon_m}\hat{\delta}^m_{1i} + \hat{\sigma}_{u_f\varepsilon_f}\hat{\delta}^m_{2i}) / \sum_{i=1}^n d_{mi}d_{fi}$ (application of Slutsky's theorem and consistency result of the bivariate probit estimates). Consistent estimates of the correlations are given by

$$\begin{aligned} \hat{\rho}_{u_m u_f} &= \widehat{\sigma}_{u_m u_f} / (\widehat{\sigma}_{u_m}\widehat{\sigma}_{u_f}), \\ \hat{\rho}_{u_k\varepsilon_l} &= \widehat{\sigma}_{u_k\varepsilon_l} / \widehat{\sigma}_{u_k}, \text{ for } k, l = \text{male, female,} \end{aligned}$$

where it is noted, as in Greene (1981), that there are no restrictions to ensure that the estimates of the correlations are restricted to lie between $[-1, 1]$.

To analyze the gender wage differences using the Oaxaca-Ransom-Neumark wage-gap decomposition technique we need to formulate a pooled regression model, which appropriately accounts for sample selectivity bias. The aim is to obtain the nondiscriminatory wage structure as a suitably weighted average of the male and female wage structure (Neumark (1988)) net of selectivity. Our pooled regression model, can be written as

$$D\begin{pmatrix} y_m \\ y_f \end{pmatrix} = D\left[\begin{pmatrix} Z_m & \lambda_1|\lambda_2 & 0 \\ Z_f & 0 & \lambda_1|\lambda_2 \end{pmatrix}\begin{pmatrix} \theta_p \\ \sigma \end{pmatrix} + \begin{pmatrix} \xi_m \\ \xi_f \end{pmatrix}\right] \tag{8}$$

where θ_p reflects the nondiscriminatory wage structure, $\sigma = (\sigma_{u_m\varepsilon_m}, \sigma_{u_m\varepsilon_f}, \sigma_{u_f\varepsilon_m}, \sigma_{u_f\varepsilon_f})'$.[4] Our pooled regression model contains a composite error term, which in part reflects any differences between the true and nondiscriminatory wage structure. Specifically,

$$d_{ik}\xi_{ik} = d_{ik}\left(v_{ik} + z_{ki}'(\theta_k - \theta_p)\right) \;\; k = \text{ male, female,}$$

[4] The presence of gender dummies in the pooled regression model, as discussed by Fortin (2006), is equivalent to noting that the first two columns of Z_m and Z_f are given by the dummy variables male and female respectively, with θ_k suitably modified to incorporate this "additional" variable.

where v_{ik} (defined as before) satisfies $E(v_{ki}|d_{ki} = 1, x_{ki}, z_{ki}) = 0$. By construction $E(\xi_{ki}|d_{ki} = 1, x_{ki}, z_{ki}) = z'_{ki}(\theta_k - \theta_p)$, $k =$ male, female. The least squares estimate of the nondiscriminatory wage structure, $\hat{\theta}_p$ can be shown to equal

$$\hat{\theta}_p = \Omega_0\hat{\theta}_m + (I - \Omega_0)\hat{\theta}_f, \tag{9}$$

where

$$\Omega_0 = \left[(D_m Z_m)' M_{\lambda_m}(D_m Z_m) + (D_f Z_f)' M_{\lambda_f}(D_f Z_f)\right]^{-1} (D_m Z_m)' M_{\lambda_m}(D_m Z_m).$$

The weighting matrix, involves the idempotent matrices $M_{\lambda_k} = I - \lambda_k(\lambda'_k\lambda_k)^{-1}\lambda'_k$ with $\lambda_k = D_k [\lambda_1|\lambda_2]$ for $k =$ male, female. This weighting matrix, is comparable to that discussed in Oaxaca and Ransom (1994) and Neumark (1988)[5] with the notable difference that it controls for sample selectivity, by the inclusion of M_{λ_k}. This matrix appears in the partitioned regression formula of the parameter estimates $\hat{\theta}_k$ obtained from the sample selectivity corrected wage regressions $\hat{\theta}_k = [(D_k Z_k)' M_{\lambda_k}(D_k Z_k)]^{-1} (D_k Z_k)' M_{\lambda_k}(D_k y_k)$.

Empirical application of the gender gap decomposition using the Oaxaca-Ransom-Neumark procedure in the presence of sample selectivity to date (e.g., Mavromaras and Rudolph, 1997) has not accounted for the distinctive and jointly determined participation rules of men and women. Moreover, it has been the practice to impose in the pooled regression model identical parameter(s) on the selectivity correction. Since the objective is to obtain a non-discriminatory wage structure as a suitably weighted average between the male and female wage structure net of selection bias, our approach is needed. Not accounting for distinctive participation rules, just like ignoring a gender dummy in the pooled regression, could lead to the pooled coefficients capturing part of the "between" male and female effects (particularly with large gender differences in explanatory variables and/or participation rules) (see also Fortin, 2006). The asymptotic covariance matrix for the resulting two-step estimator of θ_p (and σ) can be derived along similar lines as before (see Appendix).

The Oaxaca-Ransom-Neumark procedure entails decomposing the *observed male-female wage gap,* $\overline{y_m} - \overline{y_f}$ $(= E(y_m|d_m = 1) - E(y_f|d_f = 1))$ as

$$\overline{y_m} - \overline{y_f} = \left[(\overline{z}_m - \overline{z}_f)'\theta_p\right] + \left[\overline{z}'_m(\theta_m - \theta_p)\right] + \left[\overline{z}'_f(\theta_p - \theta_f)\right] + \left[E(u_m|d_m = 1) - E(u_f|d_f = 1)\right], \tag{10}$$

[5] Using our notation, their weighting matrix is given by

$$\left[(D_m Z_m)'(D_m Z_m) + (D_f Z_f)'(D_f Z_f)\right]^{-1} (D_m Z_m)'(D_m Z_m).$$

where $\overline{z}_m$ and $\overline{z}_f$ are the average endowments of the wage determining attributes for wage workers, i.e., $E(Z_m|d_m = 1)'$ and $E(Z_f|d_f = 1)'$ respectively. The first term gives an estimate of the productivity differences between the two groups in terms of θ_p, the nondiscriminatory wage structure; the second term is an estimate of the male advantage over and above the nondiscriminatory wage structure and the third term that of the female disadvantage. The remainder is due to the self-selection correction.[6] The observed male-female wage gap minus this self-selection correction gives the *offered male-female wage gap.*

In the Appendix a glossary can be found for the standard errors associated with the various components of the wage decomposition. This necessitates an evaluation of the covariance between the pooled parameter estimates, $\hat{\theta}_p$, and $(\hat{\theta}_m, \hat{\theta}_f)$, which can be derived easily using (9).

3 Data

The analysis uses Canadian data from the 1991 Census Public Use Microdata Family File (PUMF), which provides detailed information on socioeconomic characteristics on 345,351 families (3% of the population). While Baker et al., 1995, and Gunderson, 1998, amongst others, document a decline in the gender earnings differential in Canada[7], discrepancies between male and female wages still exist in Canada. This study restricts its attention to married couples, aged 17-65 years: a total of 161,928 couples. This subsample is chosen with a view to emphasize the impact of intra-household correlations on the estimation of gender specific earnings functions and the decomposition of the gender earning differences. Table A1 in the Appendix lists the variables used in this study and their descriptive statistics separated by gender and employment status ('paid worker' versus 'not paid worker').

The employment status 'paid worker' is defined as an individual who worked at least 1 week in 1990 as paid employee, is not attending school full time, and has reported positive annual earnings. The participation rates for men and women in the subsample equal 78 and 65 percent respectively, which are slightly below the overall participation rates for the age group (80 percent

[6] Neuman and Oaxaca (2004) discuss additional ways of decomposing the observed wage-gap as they consider various decompositions of the gender differences in selectivity effects in addition. This recognizes the fact that the gender differences in selection may represent discrimination as well.

[7] The decline has been accompanied by a sharp increase in the labour market participation of women (from 29% in 1961 to 60% in 1996), a stronger labour force attachment by women workers in the 1990s (fewer children and shorter work interruptions after birth), and a substantial gain in women's wage-determining or productivity-related characteristics (Statistics Canada, 1999).

for men and 68 percent for women). As dependent variable in the earnings equations we take the log of annual employment income received in 1990.[8] The observed log earnings differential is 0.751 (representing an observed earnings differential of $CA\$15399.75$). Given that annual earnings is the product of number of weeks worked, the hourly wage rate, and weekly hours of work, we will need to control for labor supply effects in our earnings equation. We do this by including the number of weeks worked in 1990 (men working on average 46.5 weeks versus 42.9 for women) and a dummy variable indicating whether the individual worked mainly full-time weeks in 1990 or not (96.8% of men working full time, versus 73.4% of women).

Most of the explanatory variables in the earnings equation are standard in the literature. We use the usual aspects of human capital: schooling, potential work experience (age-6-years of schooling), and experience squared. The major field of study, available for individuals with a postsecondary qualification, provides for a less traditional set of human capital controls we consider in addition for the earnings function yielding a more 'qualitative' measure to the human capital variable 'years of schooling'. The principal fields of study listed by men are engineering and commerce and administration, by women, secretarial and nursing.

This set of human capital regressors is augmented by other factors that may influence earnings. The ability to speak Canada's official languages might well be expected to lead to higher earnings (Shapiro and Stelcner, 1981, 1987, 1997). These effects are indicated by three dummy variables on language: French only, bilingual, and neither official language with the reference category given by English only. We also consider internal and international migration effects (see e.g., Baker and Benjamin, 1994, Bloom et al., 1995, and Shaprio and Stelcner, 1995) which we capture by two sets of dummy variables. To control for differences across labor markets in work opportunities, cost of living, and other pecuniary and nonpecuniary factors, a number of dummy variables for city size and geographic region are included. Including these geographic variables permits one to ascertain systematic regional or city-size differences in earnings.

Obviously the augmented set of human capital explanatory variables play a role in the participation decision as well. Additional variables we consider to affect the participation decision are family size, presence of children and non-labour income (total family income less earnings of the couple). These regressors are expected to influence the reservation wage of an individual and not the offered wage.

[8] A well-known problem with using Census data for earnings equations in that information on the "usual hours of work last year" are not available.

It is important to emphasize that the measure of wage discrimination implied by the decomposition of the earnings gap is limited is two important ways. First, differences in wage determining attributes are taken as given (exogenous) and are not measured explicitly as reflecting premarket factors, discriminatory or otherwise, that influence the acquisition of productive characteristics. Gender segregation in the educational system and training programs no doubt has a strong impact on labour market outcomes. Second, attempting to decompose the gender differences using the above set of regressors will enable us to obtain a measure of discrimination in the broad sense of 'unequal pay for equal productivity generating traits' (using Shapiro and Stelcner's, 1981, terminology), implicitly assuming that males and females are perfectly substitutable labour inputs. A narrower measure of discrimination in the sense of 'unequal pay for equal work' will require the introduction of variables which standardize for different working environments. Specifically, we consider the effect of introducing controls for occupation and industry. To the extent that women are crowded into (lowpaying) occupations and industries, the effect of this segregation will be reflected as a male attribute advantage and the measured degree of wage discrimination will be lower than that obtained without these variables. Managerial, natural science, machining, construction, and transport occupations are more prevalent among men, whereas teaching, medicine and health and services are more prevalent among women. Employed women are found principally in the health and social services industry, retail trade, manufacturing and educational services, employed men are found principally in the manufacturing industry, construction, retail and wholesale trade.

4 Empirical Results

Before turning our discussion to the gender decomposition of the wage differential, we first turn to the estimation of the selection and wage earning equations themselves. In our discussion of the results of the selection and wage earning regressions we focus in particular on the features that distinguish this study from others in this area rather than elaborating on all parameter estimates in detail.

Table A.2 in the Appendix provides the results for the selection of men and women into paid employment. The first two columns provide gender specific probit maximum likelihood estimates. A pooled probit regression, where we allow for gender specific heterogeneity, is reported in the third column. The results in this column provide clear evidence that male and female participation decisions are different (even when a gender dummy is included in the pooled probit model) and that gender specific heterogeneity is present. The LR test statistic of identical participation rules, which asymptotically is distributed as a $\chi^2(37)$, equals 6722.4 which is highly significant.[9] The final two columns provide the bivariate probit maximum likelihood estimates which accounts for correlations among spouses' unobserved characteristics. The highly significant parameter estimate of $\rho_{\varepsilon_m \varepsilon_f}$, 0.233, reveals that the unobservables in the participation decision of men and women are clearly correlated. The positive correlation found indicates that unobserved factors positively affecting the husbands' paid employment status affect his wife's paid employment status positively as well, pointing to similarity of partners' unobserved skills. Whether we account of intrahousehold correlations of the unobservables or not, the selection parameter estimates of (β_m, β_f) remain fairly stable.

The usual findings are obtained for selection of men and women into paid employment. Higher educational attainment increases participation in paid employment.[10] Variables considered to affect the participation decision without affecting the wage offer: family size, presence of children, age and education of spouse, and non-labour income are jointly highly significant and are of the expected sign. Having young children (younger than 6 years) in particular reduces the female participation in paid labour. While children of school age also significantly reduce her participation, their effect is reduced considerably as they become older. Finally, non-labour income has a significant negative impact on both men and women's participation. Noteworthy is that the use of an alternative definition of unearned income "household income minus the individual's earnings" yield significant differences between the individual and bivariate probit parameter estimates. Due to the correlatedness of the selection errors of men and women, the use of such a definition of unearned income should be avoided: the resulting endogeneity of such a variable (emanating from spousal income) would render both the individual and bivariate probit estimates inconsistent. This is a fact commonly ignored in other studies where the joint decision process faced by men and women is not taken into account.

In Table 1, three sets of estimation results are provided for the base human capital model. They represent the estimated earning functions for men and women under three alternative selec-

[9] The gender specific probit estimation results provide estimates of the male and female selection parameters up to scale, or $\beta_m/\sigma_{\varepsilon_m}$ and $\beta_f/\sigma_{\varepsilon_f}$. In order to test whether the male and female selection parameters are identical, $H_0 : \beta_m = \beta_f$, we do not want to impose identical variances. Consequently, the pooled probit allows for distinct variances for men and women where for identification purposes one of them is normalized to 1.

[10] This has been highlighted through the use of dummy variables reflecting the school attainment, where university education is considered to be above other non-university, and attainment of a degree, certificate or diploma is considered to be at a higher level than year completed or attended without an educational qualification.

tivity correction paradigms : In the columns labelled "OLS", no sample selectivity correction is introduced; in the columns labelled "Independent Selectivity" the standard Heckman selectivity correction terms are included; in the columns labelled "Dependent Selectivity" selectivity corrections are introduced which account for the spouse's correlated labour market decision. The table provides the parameter estimates for the human capital variables, the labour supply, and sample selectivity correction terms.

The standard Heckman selection result reveal that sample selectivity corrections are called for, since the parameters on the inverse Mills' ratio are significantly different from zero. The negative coefficients indicate that unobservables captured by the error terms, which encourage participation in the wage sector, are associated with lower earnings. These negative correlations are found in the augmented (dependent) selectivity model as well, no significant differences are observed. The estimates for $(\rho_{u_m \varepsilon_m}, \rho_{u_f \varepsilon_f})$ equal $(-0.27, -0.35)$ when we allow for intrahousehold correlations versus $(-0.28, -0.34)$ when we do not. We find a positive correlation between the unobservables in the male and female wage equation, estimated at 0.270, revealing a similarity of spouse's unobservable productivity enhancing skills as with the participation decision. We also find significant, though much smaller, intrahousehold correlations $(\rho_{u_m \varepsilon_f}, \rho_{u_f \varepsilon_m})$, estimated at $(-0.09, -0.03)$.

Despite the fact that the hypothesis of no sample selection bias has to be rejected both for men and women, the returns to education (given by the estimate on the years of schooling) are fairly stable across the three selection correction paradigms, revealing a lower return to education for men, 3.7%, than women, 5.3%. Among the human capital variables, in particular the potential experience parameters are affected by the inclusion of selectivity corrections. The returns to education for women (men) show a marginally significant upward (downwards) bias when no selection correction is considered. The introduction of controlling variables on languages, region, city size, immigration and migration did not significantly affect the returns to education of men and women, the controls were jointly highly significant. As they may contribute towards explaining the observed wage differential we include them. Their estimates reveal that for men and women earnings are highest (ceterus paribus) in the provinces of Ontario and British Columbia, and in metropolitan areas. Immigration reduces the earnings as does (for men) the inability to speak Canada's official languages. While the negative effect of immigration on earnings is stronger for more recent immigrants, the effect is found for all immigrant cohorts, supporting the rather pessimistic picture of the immigrant experience in the Canadian labour market portrayed by Baker

	Human Capital Model					
	OLS		Independent Selectivity		Dependent Selectivity	
	Men	Women	Men	Women	Men	Women
Human Capital variables						
Potential Experience	0.035 (0.001)	0.022 (0.001)	0.032 (0.001)	0.018 (0.001)	0.032 (0.001)	0.018 (0.001)
Potential Experience^2/100	−0.051 (0.001)	−0.035 (0.002)	−0.043 (0.002)	−0.021 (0.002)	−0.043 (0.002)	−0.021 (0.002)
Years of Schooling	0.036 (0.001)	0.055 (0.001)	0.036 (0.001)	0.053 (0.001)	0.037 (0.001)	0.053 (0.001)
University Degree	0.232 (0.007)	0.245 (0.008)	0.232 (0.007)	0.239 (0.009)	0.230 (0.007)	0.239 (0.009)
Vocational Training	0.046 (0.005)	0.007 (0.007)	0.040 (0.005)	−0.006 (0.007)	0.039 (0.005)	−0.007 (0.007)
Labour Supply controls						
Weeks	0.029 (0.000)	0.035 (0.000)	0.029 (0.000)	0.035 (0.000)	0.029 (0.000)	0.035 (0.000)
Full-time	0.641 (0.017)	0.625 (0.006)	0.635 (0.011)	0.614 (0.006)	0.635 (0.011)	0.614 (0.006)
Sample selection controls						
λ_1 $[\sigma_{u_m\varepsilon_m}, \sigma_{u_f\varepsilon_m}]$			−0.198 (0.019)	0.000	−0.187 (0.019)	−0.024 (0.004)
λ_2 $[\sigma_{u_m\varepsilon_f}, \sigma_{u_f\varepsilon_f}]$			0.000	−0.270 (0.016)	−0.061 (0.003)	−0.272 (0.016)
$\sigma^2_{u_m}$	0.492		0.492		0.490	
$\sigma^2_{u_f}$	0.585		0.617		0.618	
ρ_{u_m,u_f}	0.000		0.000		0.270	
Other controls, constant*	...	...	...	...	...	...
$\bar{R}^2$	0.319	0.460	0.320	0.461	0.321	0.462

* Other variables included are: language, region of residence, citysize, mobility and immigration dummies

Table 1: Estimated earning functions by gender. Dependent variable is the log annual earnings in 1990. Numbers in parentheses are standard errors.

and Benjamin (1995). Bilingual men and women receive the highest earnings. Internal mobility is associated with higher earnings for men and lower earnings for women.

In the Appendix, Table A.3 provides two additional set of results for the male and female earnings functions with dependent selectivity in which further controls are introduced for field of study and occupational and industry. The introduction of these sets of controls has the obvious effect of reducing the "returns to education" relative to the base human capital model, in part by introducing a more qualitative measure of education (fields of study), in part by netting out occupational differences.[11] Men with postsecondary qualifications in the fields of health, engineering, commerce in particular fare better than their humanities, agriculture/biology, fine arts, and education counterparts ceteris paribus. Similar findings are observed for women with postsecondary qualifications where in addition the field of nursing holds positive returns ceteris paribus.

In Table 2, we present the decomposition results for the three earnings functions considered using the three selectivity paradigms (without selectivity correction (OLS), using the standard Heckman selectivity correction (independent selectivity), and using the augmented Heckman selectivity correction where intrahousehold correlations are considered (dependent selectivity)). The first three rows in the table present the observed wage gap, the selectivity correction gap, and the offered wage gap. Using the pooled wage parameter estimates as the non-discriminatory wage structure, we report the part of the wage gap that cannot be explained by differences in endowments, also labelled "discrimination". This part of the wage gap, that is due to differences in the returns to productivity enhancing skills (coefficients), is further disaggregated into the male advantage over and above the nondiscriminatory wage structure (preferential treatment) and the female disadvantage (discrimination). The decomposition results are provided for pooled wage structures which include and which do not include a dummy variable. For comparison, we also report the coefficient gap where either the male or female wage structure is used as the non-discriminatory wage structure as suggested by Oaxaca (1973). The table presents the coefficient gaps in levels (with associated standard errors) together with the percentage of the offered wage gap they represent.

The selectivity correction gap is significant both using the standard Heckman selectivity correction (independent selectivity) and using the selectivity correction we propose that takes account

[11] The returns to education are 3.2% and 4.4% for men and women respectively after controlling for field of study. The returns to education for men and women are 2.8% and 2.7% respectively (not statistically different) after controlling both for field of study and occupational differences.

	Human Capital Model			Controlling for Field of Study (FSTD)		
	No Selectivity	Independent Selectivity[1]	Dependent Selectivity[2]	No Selectivity	Independent Selectivity[1]	Dependent Selectivity[2]
Observed wage gap	0.751 (0.003)	0.751 (0.003)	0.751 (0.003)	0.751 (0.003)	0.751 (0.003)	0.751 (0.003)
Selectivity gap	0.000	0.067 (0.010)	0.071 (0.009)	0.000	0.052 (0.010)	0.056 (0.009)
Offered wage gap	0.751 (0.003)	0.685 (0.010)	0.680 (0.010)	0.751 (0.003)	0.700 (0.010)	0.695 (0.010)
Coefficient gap						
Weights:						
Pooled						
Incl dummy	0.453 60.3% (0.003)	0.386 56.3% (0.009)	0.382 56.2% (0.010)	0.453 60.3% (0.003)	0.401 57.3% (0.009)	0.394 56.7% (0.010)
Male Premium Female Penalty	30.1% 30.1%	32.3% 24.1%	25.6% 30.7%	30.2% 30.2%	33.6% 23.7%	25.8% 30.9%
Excl dummy	0.383 51.0% (0.003)	0.317 46.3% (0.009)	0.377 55.5% (0.010)	0.332 44.2% (0.003)	0.281 40.2% (0.003)	0.374 53.8% (0.010)
Male Premium Female Penalty	23.3% 27.7%	13.0% 33.4%	22.3% 33.2%	20.2% 24.0%	12.5% 27.7%	21.2% 32.6%
Male	0.460 61.2% (0.005)	0.393 57.4% (0.010)	0.388 57.1% (0.010)	0.445 59.2% (0.006)	0.392 56.1% (0.010)	0.388 55.9% (0.011)
Female	0.459 61.1% (0.003)	0.390 56.9% (0.010)	0.385 56.6% (0.010)	0.457 60.8% (0.005)	0.404 57.7% (0.011)	0.399 57.5% (0.011)

[1] Pooled regression: standard single selectivity correction
[2] Pooled regression: gender specific selectivity corrections

Table 2: Decomposition of the Gender Wage Gap. Numbers in parentheses are standard errors.

	Controlling for FSTD Occupation/Industy		
	No Selectivity	Independent Selectivity[1]	Dependent Selectivity[2]
Observed wage gap	0.751 (0.003)	0.751 (0.003)	0.751 (0.003)
Selectivity gap	0.000	0.029 (0.010)	0.032 (0.009)
Offered wage gap	0.751 (0.003)	0.722 (0.010)	0.719 (0.010)
Coefficient gap			
Weights:			
Pooled			
Incl dummy	0.405 (0.004) 53.9%	0.375 (0.009) 51.9%	0.371 (0.010) 51.6%
Male Premium	26.9%	31.0%	24.1%
Female Penalty	26.9%	21.0%	27.5%
Excl dummy	0.234 (0.002) 31.1%	0.206 (0.008) 28.6%	0.336 (0.010) 46.7%
Male Premium	14.2%	9.8%	18.6%
Female Penalty	16.9%	18.8%	28.1%
Male	0.363 (0.006) 48.3%	0.333 (0.011) 46.1%	0.331 (0.011) 46.0%
Females	0.453 (0.008) 60.3%	0.422 (0.012) 58.4%	0.419 (0.012) 58.2%

[1] Pooled regression: standard single selectivity correction
[2] Pooled regression: gender specific selectivity corrections

Table 2: Decomposition of the Gender Wage Gap (Cont'd).

of the joint decision process faced by men and women (dependent selectivity). The dependent selectivity correction appears to result in a larger selectivity gap (and thus smaller offered wage gap) compared to the independent selectivity correction, though the difference is not significant. When sequentially we add controls in the earning functions for field of study and occupation and industry, we see a reduction (significant) of the selectivity gap. The offered wage gap is significantly lower than the observed wage gap, when accounting for selectivity corrections, signalling that part of the observed wage differences is a direct result of their differential participation decisions.

The difference in returns to productivity enhancing skills ("discrimination") in significantly larger when we use as the nondiscriminatory wage structure the pooled wage structure which includes a gender dummy compared to when we use the pooled wage structure that does not include a gender dummy. For the human capital model, this coefficient gap amounts to 0.453 (or 60.3% of the offered wage gap) with inclusion of a gender dummy versus 0.383 (or 51.0%) without. Moreover, we observe that this observed significant disparity, is reduced when selectivity corrections are used that take account of the joint decision process faced by men and women.[12] In the human capital model with dependent selectivity, where the coefficient gap amounts to 0.382 (or 56.2%) with inclusion of a gender dummy versus 0.377 (or 55.5%) without, the disparity is even rendered insignificant. As conjectured, therefore, not accounting for distinctive participation rules, just like ignoring a gender dummy in the pooled regression, lead to the pooled coefficients capturing part of the "between" male and female effects (particularly with large gender differences in explanatory variables and/or participation rules), thereby reducing the evidence of discrimination. Moreover, taking account of the distinctive participation rules of men and women as opposed to introducing the standard pooled selectivity correction, reveals that a larger part of the level of "discrimination" should be attributed to a market undervaluation of women (female penalty) than a more favorable market evaluation of observed characteristics for men (male premium).

A further support for accounting for distinctive participation rules (and the inclusion of a gender dummy) in the pooled regression, is the comparability of the level of "discrimination" these models yield when compared to the level of "discrimination" when either the male or female wage structure (selectivity corrected) is used as the nondiscriminatory wage structure. When accounting for the distinctive participation rules in the pooled regression (dependent selectivity) the use of the male and female wage structure as the nondiscriminatory wage in the human

[12] The dependent selectivity results allow for intrahousehold correlations. Quantitatively (efficiency aside) the results are similar when intrahousehold correlations are ignored while accounting for gender specific selectivity rules in the pooled regressions.

capital model yield levels of "discrimination" of 57.1% and 56.6% respectively, comparable to the 56.2% [55.5%] using the pooled wage structure inclusive [exclusive] of a gender dummy. When no selectivity corrections are considered the use of the male and female wage structure as the nondiscriminatory wage in the human capital model yield levels of "discrimination" of 61.2% and 61.1% respectively which only compare well with the use of the pooled wage structure inclusive of a gender dummy, 60.3%. Not including a gender dummy in the pooled wage structure when no selectivity corrections are considered yield levels of "discrimination" significantly lower 51.0%. The same holds true for the independent selectivity model. Accounting for the distinctive participation rules in the pooled regression, in a way, controls for the omission of a gender dummy in the pooled wage regression.

The results, when controls are added for field of study, are fairly comparable to those obtained for the base human capital model, and indicate that around 43.5% of the offered wage gap can be explained by differences in endowments. When comparing these results to related Canadian gender wage decomposition studies, we note that our focus on married couples results in a larger part of the gender earnings gap being explained by wage-related characteristics.[13]

When additional controls for occupation and industry are added, we obtain significant differences in the level of "discrimination" depending on the reference wage structure used, providing evidence of the general concern with the Oaxaca decomposition approach that it is not invariant to the reference wage used. Without selectivity correction the coefficient gap equals 0.363 (or 48.3%) using male weights, 0.453 (or 60.3%) using female weight, and 0.405 (or 53.9%) using the pooled weight including a gender dummy variable. The use of pooled weight without a gender dummy variable seem to point to a much smaller level of "discrimination" equalling 0.234 (or 31.1%). With dependent selectivity correction the disparity is somewhat compressed, but still present. The coefficient gap equals 0.331 (or 46.0%) using male weights, 0.419 (or 58.2%) using female weight, 0.371 (or 51.6%) using the pooled weight including a dummy variable, and 0.336 (or 46.7%) using the pooled weight excluding a dummy variable. Adding additional controls for occupation and industry, though, does seem to reduce the evidence of wage discrimination. This evidence of reduction in "discrimination" (in the narrower sense of 'unequal pay for equal work'), in particular, is strongly significant when using either the male or the pooled wage structure

[13] For the period 1970-1990, studies like Baker et al. (1995), Gunderson (1998), Christofides and Swidinsky (1994), Coish and Haile (1995), Grenier and Joseph (1993), Miller (1987), and Shapiro and Stelcner (1987), using the standard decomposition technique, show that between 25 to 30 percent of the gender earnings gap in Canada can be explained by wage-related characteristics.

without a gender dummy as reference wage structure. The evidence of reduction is "discrimination" of adding controls for occupation and industry over and above controls for field of study is marginally significant when using the pooled wage structure with a gender dummy as reference wage structure.

5 Conclusions

In this paper we revisited the gender decomposition of wages in the presence of selection bias. We developed appropriate sample selection corrections when the labor market decisions of spouses are dependent, and discussed a two-step estimation strategy.

We follow Neumark (1988) and Oaxaca and Ransom (1994) in using the pooled nondiscriminatory wage structure in the gender decomposition analysis. Empirical application of the gender gap decomposition using the Oaxaca-Ransom-Neumark procedure in the presence of sample selectivity has not accounted for the distinctive and jointly determined participation rules of men and women. We showed that not accounting for distinctive participation rules (and their joint determinedness), just like ignoring a gender dummy in the pooled regression, reduced the evidence of discrimination, as the nondiscriminatory wage structure captures part of the "between" male and female effects (particularly with large gender differences in explanatory variables and/or participation rules). The influence that husbands' participation decision has on the female participation decision highlights the importance of using data on both spouses for the analysis of the gender wage gap. Adding additional controls for field of study does not significantly affect the decomposition analysis, but there does appear some reduction in "discrimination" (in the narrower sense of 'unequal pay for equal work') when additional controls for occupation and industry are added.

References

[1] Baker, M. and D. Benjamin (1994), "The Performance of Immigrants in the Canadian Labor Market," *Journal of Labor Economics*, 12, 369–405.

[2] Baker, M., D. Benjamin, A. Desaulniers, and M. Grant (1995), "The Distribution of the Male/Female Earnings Differential, *Canadian Journal of Economics*, 28, 479–500.

[3] Bloom, D.E., G. Grenier, and M. Gunderson (1995), "The Changing Labour Market Position of Canadian Immigrants," *Canadian Journal of Economics*, 28, 987–1005.

[4] Christofides, L.N. and R. Swidinsky (1994), "Wage Determination by Gender and Visible Minority Status: Evidence from the 1989 LMAS," *Canadian Public Policy*, 20, 34–51.

[5] Cotton, (1988): "On the Decomposition of Wage Differentials," *Review of Economics and Statistics*, 70, 236–243.

[6] Fortin, N.M. (2006): "Greed, Altruism, and the Gender Wage Gap," University of British Columbia, http://www.econ.usb.ca/nfortin/research.html.

[7] Greene, W.H. (1981): "Sample Selectivity Bias as a Specification Error: A Comment," *Econometrica*, 49, 795–798.

[8] Gunderson, M. (1998): *Women and the Canadian Labour market: Transition Towards the Future*, Statistics Canada, Census Monograph series, Catalogue Number 96-321-MPE No.2.

[9] Heckman, J.J. (1976): "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables, and a Simple Estimator for such Models," *Annals of Economic and Social Measurement*, 5, 475–492.

[10] _____ (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153–161.

[11] Mavromaras, K.G. and H. Rudolph (1997): "Wage Discrimination in the Reemployment Process," *The Journal of Human Resources*, 32, 812–860

[12] Miller, P.W. (1987), "Gender Differences in Observed and Offered Wages in Canada, 1980," *Canadian Journal of Economics*, 20, 225–244.

[13] Newman, S. and R.L. Oaxaca (2004): "Wage Decompositions with Selectivity Corrected Wage Equations: A Methodological Note," *Journal of Economic Inequality*, 2, 3–10.

[14] Neumark, D. (1988): "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination," *The Journal of Human Resources*, 23, 279–295.

[15] Oaxaca R.L. (1973), "Male-Female Wage Differentials in Urban Labor Markets, *International Economic Review*, 9, 693–709.

[16] Oaxaca, R.L. and M.R. Ransom (1994): "On Discrimination and the Decomposition of Wage Differentials," *Journal of Econometrics*, 61, 5–21.

[17] Shapiro, D.M. and M. Stelcner (1997), "Language and Earnings in Quebec: Trends over Twenty Years, 1970–1990," *Canadian Public Policy*, 23, 115–140.

[18] ____ (1987), "The Persistence of the Male-Female Earnings Gap in Canada, 1970–1980: The Impact of Equal Pay Laws and Language Policies," *Canadian Public Policy*, 13, 472–476.

[19] ____ (1981), "Male-Female Earning Differentials and the Role of Language in Canada, Ontario, and Quebec, 1970," *Canadian Journal of Economics*, 14, 341–348.

Appendix

In this appendix, we present various results discussed in this paper. First, we discuss the derivation of the asymptotic distribution of the extended sample selectivity correction model. Next, we provide a discussion of the pooled sample selectivity correction model, followed by a glossary for the standard errors of the decomposition terms. The tables A.1–A.3 can be found at the end of this Appendix.

A.1 Derivation of the standard errors of the two-step procedure

Two corrections need to be made to the standard errors obtained by a standard linear regression package, when applying OLS to our model

$$D\begin{pmatrix} y_m \\ y_f \end{pmatrix} = D\left[\begin{pmatrix} W_m & 0 \\ 0 & W_f \end{pmatrix}\begin{pmatrix} \pi_m \\ \pi_f \end{pmatrix} + \begin{pmatrix} v_m \\ v_f \end{pmatrix}\right], \tag{A.1}$$

where D is a diagonal matrix with elements $(d_{m1}, ..., d_{mN}, d_{f1}, ..., d_{fN})$ indicating selection into the group of wage earners, N equals the number of households, $W_k = [Z_k|\lambda_1|\lambda_2]$ and $\pi_k = \left[\theta_k', \sigma_{u_k\varepsilon_m}, \sigma_{u_k\varepsilon_f}\right]'$ with k = male, female. The first correction emanates from the fact that the errors in our wage regression are non-spherical. Specifically, our covariance matrix, exhibits heteroskedasticity and serial correlation of the form,

$$\begin{aligned} Var\left[D\begin{pmatrix} v_m \\ v_f \end{pmatrix}\right] &\equiv Var\left[\begin{pmatrix} D_m & 0 \\ 0 & D_f \end{pmatrix}\begin{pmatrix} v_m \\ v_f \end{pmatrix}\right] \\ &= \begin{bmatrix} D_m\left(\sigma^2_{u_m} I_N - \Delta_m\right) & D_m D_f(\sigma_{u_m u_f} I_N - \Delta_{mf}) \\ D_m D_f(\sigma_{u_m u_f} I_N - \Delta_{mf}) & D_f(\sigma^2_{u_f} I_N - \Delta_f) \end{bmatrix} \equiv V. \end{aligned} \tag{A.2}$$

The second correction emanates from the fact that we use the predicted sample selectivity correction terms $\lambda_1(X_m\hat{\beta}_m, X_f\hat{\beta}_f, \hat{\rho}_{\varepsilon_m\varepsilon_f})$ and $\lambda_2(X_m\hat{\beta}_m, X_f\hat{\beta}_f, \hat{\rho}_{\varepsilon_m\varepsilon_f})$ in place of the true selectivity correction terms. That is, our estimated regression model is given by

$$D\begin{pmatrix} y_m \\ y_f \end{pmatrix} = D\left[\begin{pmatrix} \hat{W}_m & 0 \\ 0 & \hat{W}_f \end{pmatrix}\begin{pmatrix} \pi_m \\ \pi_f \end{pmatrix} + \begin{pmatrix} \zeta_m \\ \zeta_f \end{pmatrix}\right] \tag{A.3}$$

where $\hat{W}_k = \left[Z_k|\hat{\lambda}_1|\hat{\lambda}_2\right]$, with the error vector ζ composed as

$$\begin{aligned} \zeta_k =& \nu_k + \eta_k \\ \eta_k =& \sigma_{u_k\varepsilon_m}\left[\lambda_1(X_m\beta_m, X_f\beta_f, \rho_{\varepsilon_m\varepsilon_f}) - \lambda_1(X_m\hat{\beta}_m, X_f\hat{\beta}_f, \hat{\rho}_{\varepsilon_m\varepsilon_f})\right] \\ &+ \sigma_{u_k\varepsilon_f}\left[\lambda_2(X_m\beta_m, X_f\beta_f, \rho_{\varepsilon_m\varepsilon_f}) - \lambda_2(X_m\hat{\beta}_m, X_f\hat{\beta}_f, \hat{\rho}_{\varepsilon_m\varepsilon_f})\right], \end{aligned} \tag{A.4}$$

$k =$ male,female.

Here, we derive the asymptotic distribution of the two step estimator. We consider

$$\begin{aligned} &\begin{pmatrix} \sqrt{N_m}\,(\widehat{\pi}_m - \pi_m) \\ \sqrt{N_f}\,(\widehat{\pi}_f - \pi_f) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\hat{W}_m' D_m \hat{W}_m}{N_m} & 0 \\ 0 & \frac{\hat{W}_f' D_f \hat{W}_f}{N_f} \end{pmatrix}^{-1}\left\{\begin{pmatrix} \frac{\hat{W}_m' D_m v_m}{\sqrt{N_m}} \\ \frac{\hat{W}_f' D_f v_f}{\sqrt{N_f}} \end{pmatrix} + \begin{pmatrix} \frac{\hat{W}_m' D_m \eta_m}{\sqrt{N_m}} \\ \frac{\hat{W}_f' D_f n_f}{\sqrt{N_f}} \end{pmatrix}\right\}, \end{aligned} \tag{A.5}$$

where $N_k = \sum_{i=1}^{N} d_{ki}$, $k =$ male, female as N, N_m and $N_f \to \infty$.

Obviously the asymptotic distribution of the first step estimator $(\hat{\beta}_m', \hat{\beta}_f', \hat{\rho}_{\varepsilon_m\varepsilon_f}) \equiv \hat{\varpi}$ plays a crucial role. Specifically, the multivariate probit maximum likelihood estimator $\widehat{\varpi}$ conditional on X and Z is asymptotic normally distributed, with mean $\varpi = (\beta_m', \beta_f', \rho_{\varepsilon_m\varepsilon_f})$ and variance given by the inverse of the information matrix, $I_{\varpi\varpi}^{-1}$. More concisely, using the simplifying notation

$$\begin{aligned} q_{ki} &= 2d_{ki} - 1 \\ \omega_{ki} &= q_{ki}\,(x_{ki}'\beta_k) \\ \rho^*_{\varepsilon_m\varepsilon_f} &= q_{mi}q_{fi}\rho_{\varepsilon_m\varepsilon_f}, \end{aligned} \tag{A.6}$$

and a standard asymptotic expansion for MLE estimators, we obtain

$$\sqrt{N}(\widehat{\varpi} - \varpi) = \left(\frac{I_{\varpi\varpi}}{N}\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{n} g_i(\varpi) + o_p(1), \text{ with}$$

$$g_i(\varpi) = \begin{pmatrix} q_{mi}\frac{\phi(\omega_{mi})\Phi\left(\frac{\omega_{fi}-\rho^*_{\varepsilon_m\varepsilon_f}\omega_{mi}}{\sqrt{1-\rho^2_{\varepsilon_m\varepsilon_f}}}\right)}{\Phi_2(\omega_{mi},\omega_{fi},\rho^*_{\varepsilon_m\varepsilon_f})}\cdot x_{mi} \\ q_{fi}\frac{\phi(\omega_{fi})\Phi\left(\frac{\omega_{mi}-\rho^*_{\varepsilon_m\varepsilon_f}\omega_{fi}}{\sqrt{1-\rho^2_{\varepsilon_m\varepsilon_f}}}\right)}{\Phi_2(\omega_{mi},\omega_{fi},\rho^*_{\varepsilon_m\varepsilon_f})}\cdot x_{fi} \\ q_{mi}q_{fi}\frac{\phi(\omega_{mi})\phi\left(\frac{\omega_{fi}-\rho^*_{\varepsilon_m\varepsilon_f}\omega_{mi}}{\sqrt{1-\rho^2_{\varepsilon_m\varepsilon_f}}}\right)}{\Phi_2(\omega_{mi},\omega_{fi},\rho^*_{\varepsilon_m\varepsilon_f})} \end{pmatrix} \equiv \begin{pmatrix} \lambda_{1i}x_{mi} \\ \lambda_{2i}x_{fi} \\ \lambda_{3i} \end{pmatrix} \tag{A.7}$$

where $\sum_{i=1}^{n} g_i(\varpi)$ denotes the score of the log-likelihood function.

The consistency of our first step estimates $\widehat{\varpi} = (\widehat{\beta}'_m, \widehat{\beta}'_f, \widehat{\rho}_{\varepsilon_m \varepsilon_f})'$ ensure that

$$\underset{\substack{N_k \to \infty \\ N \to \infty}}{\text{plim}} \frac{1}{N_k} \widehat{W}'_k D_k \widehat{W}_k = \lim_{N_k \to \infty} \frac{1}{N_k} W'_k D_k W_k, \tag{A.8}$$

where $W_k = \begin{bmatrix} Z_k & \lambda_1 & \lambda_2 \end{bmatrix}$ by Slutsky theorem. We assume this to be a finite positive definite matrix (imposes a regularity condition on the regressors). Moreover, these conditions ensure that

$$\left.\begin{pmatrix} \frac{\hat{W}'_m D_m v_m}{\sqrt{N_m}} \\ \frac{\hat{W}'_f D_f v_f}{\sqrt{N_f}} \end{pmatrix}\right|_{X,Z} \xrightarrow{d} N\left(0, \begin{pmatrix} \underset{\substack{N \to \infty \\ N_m \to \infty}}{\text{plim}} \frac{W'_m D_m V_{mm} D_m W_m}{N_m} & c \underset{\substack{N \to \infty \\ N_{mf} \to \infty}}{\text{plim}} \frac{W'_m D_m V_{mf} D_f W_f}{N_{mf}} \\ c \underset{\substack{N \to \infty \\ N_{mf} \to \infty}}{\text{plim}} \frac{W'_f D_f V_{fm} D_m W_m}{N_{mf}} & \underset{\substack{N \to \infty \\ N_f \to \infty}}{\text{plim}} \frac{W'_f D_f V_{ff} D_f W_f}{N_f} \end{pmatrix}\right), \tag{A.9}$$

where $N_{mf} = \sum_{i=1}^{N} d_{mi} d_{fi}$, $c = \lim_{N \to \infty} \frac{N_{mf}}{\sqrt{N_m N_f}}$ $0 < c \leq 1$, and V, the covariance matrix of $E(Dv)$, is partitioned as $\begin{bmatrix} V_{mm} & V_{mf} \\ V_{fm} & V_{ff} \end{bmatrix}$.

To derive the limiting distribution of $\begin{pmatrix} \frac{\hat{W}'_m D_m \eta_m}{\sqrt{N_m}} \\ \frac{\hat{W}'_f D_f n_f}{\sqrt{N_f}} \end{pmatrix}$, we note that a one term Taylor expansion of $\sigma_{u_k \varepsilon_m} \hat{\lambda}_1 + \sigma_{u_k \varepsilon f} \hat{\lambda}_2$ around $\sigma_{u_k \varepsilon_m} \lambda_1 + \sigma_{u_k \varepsilon f} \lambda_2$, yields

$$\eta_k = -\frac{\partial(\sigma_{u_k \varepsilon_m} \lambda_1 + \sigma_{u_k \varepsilon f} \lambda_2)}{\partial \varpi'} (\widehat{\varpi} - \varpi) + O_p(N^{-1}). \tag{A.10}$$

Given the asymptotic distribution of the multivariate probit maximum likelihood estimator,

$$\left.\frac{\hat{W}'_k D_k \eta_k}{\sqrt{N_k}}\right|_{X,Z} \xrightarrow{d} \tag{A.11}$$

$$N\left(0, c_k \cdot \lim_{\substack{N \to \infty \\ N_k \to \infty}} \frac{W'_k D_k \frac{\partial(\sigma_{u_k \varepsilon_m} \lambda_1 + \sigma_{u_k \varepsilon f} \lambda_2)}{\partial \varpi'}}{N_k} \left(\lim_{N \to \infty} \frac{I_{\varpi\varpi}}{N}\right)^{-1} \lim_{\substack{N \to \infty \\ N_k \to \infty}} \frac{\frac{\partial(\sigma_{u_k \varepsilon_m} \lambda_1 + \sigma_{u_k \varepsilon f} \lambda_2)'}{\partial \varpi'} D_k W_k}{N_k}\right),$$

where $c_k = \lim_{N \to \infty} \frac{N_m}{N}$, $0 < c_k \leq 1$ $k =$ male and female.[14] Since η_m and η_f are correlated for individuals coming from the same household, $\frac{\hat{W}'_m D_m \eta_m}{\sqrt{N_m}}$ and $\frac{\hat{W}'_f D_f \eta_f}{\sqrt{N_f}}$ are asymptotically correlated

[14]Using, (A.10), we note for instance $\frac{\hat{W}'_m D_m \eta_m}{\sqrt{N_m}} \overset{a}{=} \sqrt{\frac{N_m}{N}} \frac{W'_m D_m \frac{\partial(\sigma_{u_k \varepsilon_m} \lambda_{1i} + \sigma_{u_k \varepsilon f} \lambda_{2i})}{\partial \varpi'}}{N_m} \sqrt{N} (\hat{\varpi} - \varpi)$. The continuous mapping theorem yields the result provided, where the usual regularity conditions ensure that $\lim_{N, N_k \to \infty} \frac{W'_k D_k \frac{\partial(\sigma_{u_k \varepsilon_m} \lambda_{1i} + \sigma_{u_k \varepsilon f} \lambda_{2i})}{\partial \varpi'}}{N_k}$ exists.

as well. In particular,

$$ACov(\frac{\hat{W}_m' D_m \eta_m}{\sqrt{N_m}}, \frac{\hat{W}_f' D_f \eta_f}{\sqrt{N_f}})\Big]_{X,Z}$$

$$= c_{mf} \cdot \lim_{\substack{N\to\infty \\ N_m\to\infty}} \frac{W_m' D_m \frac{\partial(\sigma_{u_m\varepsilon_m}\lambda_1+\sigma_{u_m\varepsilon f}\lambda_2)}{\partial\varpi'}}{N_m} \left(\lim_{N\to\infty} \frac{I_{\varpi\varpi}}{N}\right)^{-1} \lim_{\substack{N\to\infty \\ N_f\to\infty}} \frac{\frac{\partial(\sigma_{u_f\varepsilon_m}\lambda_1+\sigma_{u_f\varepsilon f}\lambda_2)}{\partial\varpi'}' D_f W_f}{N_f}, \tag{A.12}$$

where $c_{mf} = \lim_{N\to\infty} \frac{\sqrt{N_m N_f}}{N}$, $0 < c_{mf} \le 1$.

The sample selectivity correction we proposed in this study allows us to ascertain that η and ν are uncorrelated. Consequently, we can conclude that conditional on the data, $\widehat{\pi}$ is asymptotically normal with mean π and asymptotic variance given by

$$\text{Avar}(\widehat{\pi}) = \begin{pmatrix} (W_m' D_m W_m)^{-1} W_m' D_m \\ \left(W_f' D_f W_f\right)^{-1} W_f' D_f \end{pmatrix} [V+R] \begin{pmatrix} (W_m' D_m W_m)^{-1} W_m' D_m \\ \left(W_f' D_f W_f\right)^{-1} W_f' D_f \end{pmatrix}', \tag{A.13}$$

with

$$R = \begin{pmatrix} \frac{\partial(\sigma_{u_m\varepsilon_m}\lambda_1+\sigma_{u_m\varepsilon f}\lambda_2)}{\partial\varpi'} \\ \frac{\partial(\sigma_{u_f\varepsilon_m}\lambda_1+\sigma_{u_f\varepsilon f}\lambda_2)}{\partial\varpi'} \end{pmatrix} I_{\varpi\varpi}^{-1} \begin{pmatrix} \frac{\partial(\sigma_{u_m\varepsilon_m}\lambda_1+\sigma_{u_m\varepsilon f}\lambda_2)}{\partial\varpi'} \\ \frac{\partial(\sigma_{u_f\varepsilon_m}\lambda_1+\sigma_{u_f\varepsilon f}\lambda_2)}{\partial\varpi'} \end{pmatrix}'$$

$$\equiv \begin{pmatrix} \Delta_1^m X_m & \Delta_2^m X_f & \Delta_3^m \\ \Delta_1^f X_m & \Delta_2^f X_f & \Delta_3^f \end{pmatrix} I(\beta_m, \beta_f, \rho_{\varepsilon_m\varepsilon_f})^{-1} \begin{pmatrix} \Delta_1^m X_m & \Delta_2^m X_f & \Delta_3^m \\ \Delta_1^f X_m & \Delta_2^f X_f & \Delta_3^f \end{pmatrix}'. \tag{A.14}$$

Let us turn to the important result implied by our sample selectivity correction, the uncorrelatedness of η and ν. First, we note that η is asymptotically a linear function of (d_f, d_m) on account of (A.7) and (A.10). Next, we realize that our sample selectivity correction defined $v = (v_m', v_f')$, with

$$v_{mi} = u_{mi} - d_{fi}E(u_{mi}|d_{mi}=1, d_{fi}=1, x_i, z_i) - (1-d_{fi})E(u_{mi}|d_{mi}=1, d_{fi}=0, x_i, z_i) \tag{A.15}$$
$$v_{fi} = u_{fi} - d_{mi}E(u_{fi}|d_{mi}=1, d_{fi}=1, x_i, z_i) - (1-d_{mi})E(u_{fi}|d_{mi}=1, d_{fi}=0, x_i, z_i),$$

for $i = 1, ..., N$. Clearly, these formulations show that ν is uncorrelated with (d_f, d_m). This convenient result is absent when the standard selectivity correction is applied, since $u_{mi} - E(u_{mi}|d_{mi} = 1, x_i, z_i)$ in general is still correlated with d_{fi} while being uncorrelated with d_{mi}.

A.2 The pooled earnings regression

Our pooled regression model, is given by

$$D\binom{y_m}{y_f} = D\left[\begin{pmatrix} Z_m & \lambda_1|\lambda_2 & 0 \\ Z_f & 0 & \lambda_1|\lambda_2 \end{pmatrix}\binom{\theta_p}{\sigma} + \binom{\xi_m}{\xi_f}\right], \tag{A.16}$$

where θ_p reflects the nondiscriminatory wage structure and $\sigma = (\sigma_{u_m\varepsilon_m}, \sigma_{u_m\varepsilon_f}, \sigma_{u_f\varepsilon_m}, \sigma_{u_f\varepsilon_f})'$. Using the partitioned regression result,

$$\hat{\theta}_p = \left[\begin{pmatrix} D_m Z_m \\ D_f Z_f \end{pmatrix}' M_{\lambda_{mf}} \begin{pmatrix} D_m Z_m \\ D_f Z_f \end{pmatrix}\right]^{-1} \begin{pmatrix} D_m Z_m \\ D_f Z_f \end{pmatrix}' M_{\lambda_{mf}} \begin{pmatrix} D_m y_m \\ D_f y_f \end{pmatrix}, \tag{A.17}$$

where $M_{\lambda_{mf}} = I - \lambda_{mf}\left[\lambda_{mf}'\lambda_{mf}\right]^{-1}\lambda_{mf}'$ with $\lambda_{mf} = \begin{pmatrix} D_m\lambda & 0 \\ 0 & D_f\lambda \end{pmatrix}$. As $M_{\lambda_{mf}}$ can be written as $\begin{pmatrix} M_{\lambda_m} & 0 \\ 0 & M_{\lambda_f} \end{pmatrix}$, where $M_{\lambda_k} = I - \lambda_k\left[\lambda_k'\lambda_k\right]^{-1}\lambda_k'$ with $\lambda_k = (D_k\lambda)$, we can simplify $\hat{\theta}_p$ as being given by

$$\begin{aligned}\hat{\theta}_p = &\left[(D_m Z_m)' M_{\lambda_m}(D_m Z_m) + (D_f Z_f)' M_{\lambda_f}(D_f Z_f)\right]^{-1} \\ &\left[(D_m Z_m)' M_{\lambda_m}(D_m y_m) + (D_f Z_f)' M_{\lambda_f}(D_f y_f)\right]\end{aligned} \tag{A.18}$$

Since the partitioned regression result of the male and female wage regression, similarly allowed us to denote

$$\hat{\theta}_k = \left[(D_k Z_k)' M_{\lambda_k}(D_k Z_k)\right]^{-1}(D_k Z_k)' M_{\lambda_k}(D_k y_k), \ k = \text{ male, female}, \tag{A.19}$$

we realize that the nondiscriminatory wage structure indeed is a weighted average of the male and female wage structure, $\hat{\theta}_p = \Omega_0\hat{\theta}_m + (I - \Omega_0)\hat{\theta}_f$, where

$$\Omega_0 = \left[(D_m Z_m)' M_{\lambda_m}(D_m Z_m) + (D_f Z_f)' M_{\lambda_f}(D_f Z_f)\right]^{-1}(D_m Z_m)' M_{\lambda_m}(D_m Z_m). \tag{A.20}$$

Given the limiting distribution of $\pi = (\theta_m', \theta_m', \sigma')'$ obtained in the previous section, the limiting distribution of $\hat{\theta}_p$ readily follows, as $\hat{\theta}_p = [\Omega_0|I - \Omega_0|0]\,\hat{\pi}$.

A.3 Glossary, standard errors of the decomposition terms

We denote $\text{AVar}(\widehat{\theta}_k) = V_k$, $\text{ACov}(\widehat{\theta}_k, \widehat{\theta}_l) = V_{kl}$, $k, l =$ male, female. Recognizing, that the nondiscriminatory wage structure $\widehat{\theta}_p$ is given by a weighted average of the male and female wage structure, $\hat{\theta}_p = \Omega_0\hat{\theta}_m + (I - \Omega_0)\hat{\theta}_f$, we obtain the following expressions for the variance associated with the distinct components of the Oaxaca-Ransom decomposition discussed in the text:

(1) Endowment gap := $(\overline{z}_m - \overline{z}_f)'\theta_p$

$$
\begin{aligned}
&\text{Var}(\widehat{\text{Endowment gap}}) \\
&= \left[(\overline{z}_m - \overline{z}_f)'\text{AVar}(\widehat{\theta}_p)(\overline{z}_m - \overline{z}_f)\right] \qquad \text{(A.21)} \\
&= \left[(\overline{z}_m - \overline{z}_f)'\left\{\Omega_0 V_m \Omega_0 + (I-\Omega_0)V_f(I-\Omega_0)' + \Omega_0 V_{mf}(I-\Omega_0)' + (I-\Omega_0)V_{fm}\Omega_0{}'\right)\right\}(\overline{z}_m - \overline{z}_f)\right]
\end{aligned}
$$

(2) Coefficients (male advantage) := $\overline{z}_m'(\theta_m - \theta_p)$

$$
\begin{aligned}
\text{Var}(\text{Male }\widehat{\text{advantage}}) &= \left[\overline{z}_m'\left\{\text{AVar}(\widehat{\theta}_m - \widehat{\theta}_p)\right\}\overline{z}_m\right] \qquad \text{(A.22)} \\
&= \left[\overline{z}_m'\left\{(I-\Omega_0)\left\{V_m + V_f - V_{mf} - V_{fm}\right\}(I-\Omega_0)'\right\}\overline{z}_m\right]
\end{aligned}
$$

(3) Coefficients (female disadvantage) := $\overline{z}_f'(\theta_p - \theta_f)$

$$
\begin{aligned}
\text{Var}(\text{Female }\widehat{\text{disadvantage}}) &= \left[\overline{z}_f'\left\{\text{AVar}(\widehat{\theta}_p - \widehat{\theta}_f)\right\}\overline{z}_f\right] \qquad \text{(A.23)} \\
&= \left[\overline{z}_f'\left\{(\Omega_0\left\{V_m + V_f - V_{mf} - V_{fm}\right\}\Omega_0{}'\right\}\overline{z}_f\right]
\end{aligned}
$$

(4) Coefficients (2)+(3) := $\overline{z}_m'(\theta_m - \theta_p) + \overline{z}_f'(\theta_p - \theta_f)$

$$
\begin{aligned}
\text{Var}(\widehat{\text{Coefficients}}) &= (2) + (3) + 2 * \left[\overline{z}_m'\left\{\text{ACov}\left(\widehat{\theta}_m - \widehat{\theta}_p, \widehat{\theta}_p - \widehat{\theta}_f\right)\right\}\overline{z}_f'\right] \qquad \text{(A.24)} \\
&= (2) + (3) + 2 * \left[\overline{z}_m'\left\{(I-\Omega_0)\left\{V_m + V_f - V_{mf} - V_{fm}\right\}\Omega_0\right\}\overline{z}_f'\right]
\end{aligned}
$$

A.4 Tables

Endogenous Variables:	Not Paid Men N= 31635		Paid Men N=130293		Not Paid Wom N= 52233		Paid Wom N=109695	
Paid worker	0.000		1.000	80.5%	0.000		1.000	67.7%
Log yearly earnings			10.280	(0.838)			9.529	(1.040)
Exogenous Variables:								
▷ **Selection eqn only**								
Family Structure:								
HH size	3.163	(1.222)	3.342	(1.141)	3.373	(1.235)	3.275	(1.120)
No. children <6	0.271	(0.582)	0.357	(0.640)	0.397	(0.687)	0.314	(0.600)
No. children 6-14	0.426	(0.782)	0.532	(0.826)	0.510	(0.832)	0.513	(0.813)
No. children 15-17	0.140	(0.391)	0.161	(0.412)	0.140	(0.389)	0.165	(0.416)
Non-labour income	0.16	(0.21)	0.08	(0.13)	0.12	(0.18)	0.08	(0.14)
▷ **Selection/earnings eqn**								
Human Capital:								
Age	46.62	(12.77)	41.26	(10.56)	42.40	(12.28)	38.47	(10.00)
Potential Experience	29.84	(14.72)	23.33	(11.90)	25.89	(13.80)	20.50	(11.09)
Years of Schooling	11.774	(4.587)	12.928	(4.079)	11.518	(3.741)	12.970	(3.396)
University Degree	0.132	(0.339)	0.160	(0.366)	0.079	(0.270)	0.133	(0.340)
Vocational Training	0.201	(0.401)	0.239	(0.427)	0.105	(0.307)	0.134	(0.341)
Linguistics: (ref English)								
French only	0.257	(0.437)	0.257	(0.437)	0.295	(0.456)	0.240	(0.427)
Bilingual	0.345	(0.475)	0.376	(0.484)	0.342	(0.474)	0.383	(0.486)
Neither domestic lang	0.190	(0.392)	0.162	(0.368)	0.153	(0.360)	0.174	(0.379)
Mobility: (ref CD same)								
Prov. same, CD not	0.119	(0.324)	0.119	(0.324)	0.119	(0.324)	0.119	(0.324)
Prov. different	0.514	(0.500)	0.602	(0.489)	0.546	(0.498)	0.603	(0.489)

Table A.1: Descriptive Statistics by gender and employment status. Means and standard deviations in parentheses.

	Not Paid Men		Paid Men		Not Paid Wom		Paid Wom	
Immigration: (ref born)								
Immigrated <1961	0.138	(0.345)	0.124	(0.330)	0.199	(0.400)	0.134	(0.341)
Immig. 1961-1970	0.176	(0.381)	0.198	(0.398)	0.144	(0.351)	0.177	(0.382)
Immig. 1971-1980	0.013	(0.115)	0.007	(0.084)	0.019	(0.135)	0.008	(0.091)
Immig. 1981-1985	0.075	(0.264)	0.053	(0.224)	0.054	(0.227)	0.038	(0.192)
Immig. 1986-1990	0.049	(0.216)	0.054	(0.227)	0.047	(0.212)	0.052	(0.222)
Province: (ref Atlantic)								
Quebec	0.049	(0.215)	0.056	(0.231)	0.048	(0.213)	0.058	(0.234)
Ontario	0.018	(0.134)	0.018	(0.133)	0.020	(0.139)	0.019	(0.137)
Prairies	0.036	(0.187)	0.025	(0.157)	0.039	(0.193)	0.026	(0.160)
British Columbia	0.043	(0.202)	0.049	(0.215)	0.049	(0.216)	0.047	(0.211)
Census Metropolitan	0.100	(0.301)	0.118	(0.323)	0.111	(0.315)	0.124	(0.329)
▷ **Earnings eqn only**								
Labour Supply:								
Weeks			46.525	(10.904)			42.895	(14.031)
Full-time			0.968	(0.176)			0.734	(0.442)
Field of study:								
Fine Arts			0.014	(0.116)			0.030	(0.170)
Humanities			0.021	(0.145)			0.026	(0.160)
Social Science			0.037	(0.189)			0.036	(0.186)
Commerce-Admin			0.074	(0.262)			0.060	(0.237)
Secretarial			0.004	(0.063)			0.080	(0.271)
Agriculture/Biology			0.021	(0.144)			0.018	(0.132)
Engineer			0.254	(0.435)			0.019	(0.137)
Nursing			0.003	(0.052)			0.070	(0.255)
Other Health			0.013	(0.112)			0.026	(0.158)
Maths/Physics			0.020	(0.140			0.009	(0.096

Table A.1: Descriptive Statistics by gender and employment status (Cont'd).

	Paid Men		Paid Wom	
Occupation (ref Clerical)				
Managerial, administrative and related	0.178	(0.382)	0.110	(0.313)
Natural sciences, engineering and mathematics	0.066	(0.249)	0.017	(0.131)
Social sciences and related	0.014	(0.117)	0.027	(0.163)
Teaching and related	0.035	(0.184)	0.073	(0.260)
Medicine and health	0.016	(0.124)	0.098	(0.297)
Artistic, literary, recreational and related	0.012	(0.108)	0.012	(0.109)
Sales	0.083	(0.276)	0.085	(0.279)
Services	0.080	(0.272)	0.131	(0.338)
Farming and Horticulture	0.013	(0.115)	0.013	(0.115)
Other Primary	0.024	(0.153)	0.002	(0.048)
Processing	0.044	(0.206)	0.020	(0.138)
Machining, product fabricating, and assembling	0.135	(0.342)	0.038	(0.191)
Construction	0.109	(0.311)	0.003	(0.053)
Transport equipment operating	0.065	(0.247)	0.008	(0.090)
Other Occupations	0.066	(0.249)	0.024	(0.153)

Table A.1: Descriptive Statistics by gender and employment status (Cont'd).

	Paid Men		Paid Wom	
Industry (ref Govt Services)				
Agriculture	0.018	(0.134)	0.022	(0.147)
Other Primary	0.043	(0.202)	0.009	(0.093)
Manufacturing	0.223	(0.417)	0.111	(0.314)
Construction	0.104	(0.305)	0.020	(0.140)
Transportation and storage	0.069	(0.253)	0.020	(0.140)
Communications and other utilities	0.049	(0.217)	0.028	(0.166)
Wholesale trade	0.061	(0.240)	0.031	(0.174)
Retail trade	0.092	(0.289)	0.136	(0.343)
Finance, insurance and real estate	0.045	(0.208)	0.089	(0.284)
Business services	0.049	(0.215)	0.054	(0.226)
Educational services	0.057	(0.233)	0.109	(0.311)
Health and social services	0.029	(0.167)	0.171	(0.376)
Accommodation, food and beverage services	0.025	(0.155)	0.062	(0.242)
Other services	0.039	(0.193)	0.065	(0.246)

Table A.1: Descriptive Statistics by gender and employment status (Cont'd).

	Probit		Probit	Bivariate Probit	
	Men	Women	Pooled	Men	Women
$\sigma^2_{\varepsilon_m}$ ($\sigma^2_{\varepsilon_f} = 1$)	1.000		0.770 (0.009)	1.000	
$\rho_{\varepsilon_m,\varepsilon_f}$	0.000		0.000	0.233 (0.005)	
Human capital					
Secondary School, with Grad Cert	0.232 (0.013)	0.313 (0.010)	0.293 (0.009)	0.225 (0.013)	0.306 (0.010)
Secondary School, with Trade Cert	0.233 (0.016)	0.389 (0.020)	0.356 (0.014)	0.233 (0.016)	0.377 (0.020)
Other Non-Univ, no Certificate	0.065 (0.017)	0.290 (0.014)	0.227 (0.012)	0.053 (0.017)	0.282 (0.014)
Other Non-Univ, with Trade Cert	0.159 (0.013)	0.305 (0.016)	0.268 (0.012)	0.154 (0.013)	0.300 (0.016)
Other Non-Univ, with Other Cert	0.280 (0.015)	0.495 (0.012)	0.451 (0.010)	0.280 (0.015)	0.485 (0.011)
Univ, no Cert	0.084 (0.022)	0.249 (0.020)	0.201 (0.016)	0.079 (0.022)	0.248 (0.020)
Univ, with below BA Cert	0.171 (0.019)	0.483 (0.016)	0.402 (0.013)	0.168 (0.018)	0.481 (0.016)
Univ, with BA Cert	0.106 (0.014)	0.461 (0.014)	0.342 (0.011)	0.109 (0.014)	0.462 (0.014)
Univ, with above BA Cert	0.181 (0.031)	0.593 (0.032)	0.458 (0.025)	0.181 (0.031)	0.600 (0.032)
Univ, with MA Cert	0.202 (0.024)	0.476 (0.029)	0.391 (0.021)	0.212 (0.023)	0.489 (0.029)
Univ, with PhD Cert	0.543 (0.048)	0.670 (0.094)	0.739 (0.051)	0.563 (0.047)	0.667 (0.093)
Family Characteristics					
HH size	0.163 (0.007)	0.031 (0.006)	0.092 (0.005)	0.162 (0.007)	0.032 (0.006)
No. children <6	−0.205 (0.009)	−0.405 (0.008)	−0.363 (0.007)	−0.205 (0.009)	−0.407 (0.008)
No. children 6-14	−0.213 (0.008)	−0.208 (0.007)	−0.234 (0.006)	−0.211 (0.008)	−0.209 (0.007)
No. children 15-17	−0.186 (0.012)	−0.095 (0.011)	−0.146 (0.009)	−0.184 (0.012)	−0.094 (0.011)
Non-labour income	−1.321 (0.025)	−0.654 (0.025)	−1.038 (0.020)	−1.325 (0.025)	−0.657 (0.025)
*Other controls**, *constant*	...	...	...	...	...

* Other variables included are: age, agesq, language, region of residence, citysize mobility, immigration dummies, age and education of spouse, and gender dummy (pooled probit)

Table A.2: Selection of into Paid Employment by gender. Numbers in parentheses are standard errors.

	Human Capital Model		Controlling for Field of Study		Controlling for Field of Study Occ & Indust	
	Men	Women	Men	Women	Men	Women
Human Capital variables						
Potential Experience	0.032 (0.001)	0.018 (0.001)	0.034 (0.001)	0.019 (0.001)	0.029 (0.001)	0.014 (0.001)
Potential Experience^2/100	−0.043 (0.002)	−0.021 (0.002)	−0.046 (0.002)	−0.024 (0.002)	−0.038 (0.002)	−0.016 (0.002)
Years of Schooling	0.037 (0.001)	0.053 (0.001)	0.032 (0.001)	0.044 (0.001)	0.028 (0.001)	0.027 (0.001)
University Degree	0.230 (0.007)	0.239 (0.009)	0.208 (0.008)	0.284 (0.009)	0.161 (0.008)	0.217 (0.009)
Vocational Training	0.039 (0.005)	−0.007 (0.007)	−0.067 (0.007)	−0.041 (0.008)	−0.033 (0.007)	−0.031 (0.008)
Labour Supply controls						
Weeks	0.029 (0.000)	0.035 (0.000)	0.029 (0.000)	0.034 (0.000)	0.029 (0.000)	0.033 (0.000)
Full-time	0.635 (0.011)	0.614 (0.006)	0.627 (0.011)	0.630 (0.006)	0.579 (0.011)	0.585 (0.006)
Sample selection controls						
λ_1 $[\sigma_{u_m\varepsilon_m}, \sigma_{u_f\varepsilon_m}]$	−0.187 (0.019)	−0.024 (0.004)	−0.152 (0.019)	−0.016 (0.004)	−0.155 (0.019)	−0.017 (0.004)
λ_2 $[\sigma_{u_m\varepsilon_f}, \sigma_{u_f\varepsilon_f}]$	−0.061 (0.003)	−0.272 (0.016)	−0.057 (0.003)	−0.218 (0.017)	−0.049 (0.003)	−0.170 (0.016)
$\sigma^2_{u_m}$	0.490		0.482		0.461	
$\sigma^2_{u_f}$	0.618		0.600		0.557	
ρ_{u_m,u_f}	0.270		0.261		0.268	
Other controls, constant*	...	...	...	...	...	...
$\bar{R}^2$	0.321	0.462	0.327	0.468	0.357	0.498

* Other variables included are: language, region of residence, citysize, mobility and immigration dummies

Table A.3: Estimated earning functions by gender. Dependent variable log annual earnings in 1990. Numbers in parentheses are standard errors.