Occupational segregation, selection effects and gender wage differences: evidence from urban Colombia

Jairo Guillermo Isaza Castro*

_____

*   Associated professor, Universidad de La Salle. Economist, Universidad Católica de Colombia; MA in Development Studies, International Institute of Social Studies, The Netherlands; PhD in Economics, University of Sussex, United Kingdom. My sincere acknowledgements to Prof. Barry Reilly for his valuable suggestions on an earlier draft of this paper. The author is also thankful for the comments received from an anonymous referee and from the participants at the Sussex DPhil Conference organised by the Department of Economics, University of Sussex (Brighton, UK) on November 2009 and the Third Alban Conference celebrated in Porto (Portugal) on June 2009. Usual disclaimers apply. Contact: jgisaza@gmail.com and jisaza@lasalle.edu.co

73

Abstract

This paper assesses the effects of occupational segregation on the gender wage gap in urban Colombia between 1986 and 2000. The empirical methodology involves a two-step procedure whereby the occupational distributions of workers by gender are modelled using a multinomial logit model in the first stage. In the second stage, the multinomial logit estimates are used not only to derive a counterfactual occupational distribution of women in the absence of workplace discrimination but also to correct for selectivity bias in the wage equations for each occupational category using the procedure suggested by Lee (1983). Besides the explained and unexplained components in conventional decompositions of the gender wage gap, this methodology differentiates between the justified and unjustified effects of the gender allocation of workers across occupational categories. The results for urban Colombia indicate that controlling for selectivity bias at the occupational category level is found to be relevant in all years reviewed in this study. They also suggest that a changing composition of the female labour supply in terms of un observables (i.e., ability and motivation) is playing a role in the dramatic reduction of the observed wage gap.

Keywords: occupational segregation, gender wage gap, multinomial logit, selection bias, Colombia.

JEL: J16, J71, C24

Resumen

Este documento analiza los efectos de la segregación ocupacional en la brecha salarial por género en Colombia urbana entre 1986 y 2000. La metodología empírica involucra un procedimiento en dos etapas por el cual las distribuciones ocupacionales de trabajadores por género son modeladas a partir de un modelo logit multinomial en una primera etapa. En una segunda etapa, los estimativos del modelo logit multinomial se usan no solo para derivar una distribución ocupacional contra factual de las mujeres en ausencia de discriminación en el lugar de trabajo, sino también para corregir por sesgo de selección en las ecuaciones de ingresos laborales por género para cada categoría ocupacional, a través del procedimiento sugerido por Lee (1983). Además de los componentes explicados y no explicados en descomposiciones salariales convencionales, esta metodología diferencia entre los efectos justificados y no justificados de la distribución de trabajadores por género a lo largo de categorías ocupacionales. Los resultados para Colombia urbana indican que controlar por sesgo de selección en el nivel de categoría ocupacional es relevante en todos los años revisados en el este estudio. Ellos también sugieren que los cambios en la composición de la oferta de trabajo femenina en términos de variables no observables (por ejemplo, motivación y habilidad) están jugando un rol fundamental en la reducción de la brecha de ingresos por género registrada para este período.

Palabras clave: segregación ocupacional, brecha de ingresos por género, logit multinomial, sesgo de selección, Colombia.

INTRODUCTION

Most of the empirical literature on gender differences in the labour market has focused either exclusively on wage discrimination or occupational segregation. Empirical research linking both aspects of the differentiated situation of women in the labour market is still scarce. While much of the economic research has been motivated by the 'taste for discrimination' approach proposed by Becker (1971), the segregation dimension has merited less attention within this framework. The relative scarcity of applied economic research on the relationship between occupational segregation and gender/ethnic wage discrimination may be explained by the fact that Becker's original model of discrimination does not explicitly incorporate the segregation dimension. In this sense, Baldwin et al.

(2001) propose an extension to this approach by including a hierarchical dimension in which men dislike to work for women even in the case they do not object to work along side women. As a result, their model not only predicts that women's participation is exponentially decreasing with respect to job hierarchy but also that women's wage disadvantage in managerial positions is, at least, partially explained by a compensation mechanism of men's dislike for female supervision. Although we do not attempt an empirical test of all propositions derived from Baldwin et al.'s (2001) model, it is still relevant to highlight that the connection between gender wage discrimination and occupational segregation is theoretically grounded.

Since the seminal contributions of Oaxaca (1973) and Blinder (1973), a number of decomposition techniques of the gender

wage gap have been proposed to distinguish between the effects of explained differences in human capital and other characteristics, on the one hand, and the effects of unexplained differences in returns to those characteristics or discrimination, on the other. Blinder (1973), in particular, suggested the use of dummy variables in the gender wage equations to control for the effects of occupations. This dummy variable approach implies that the gender distribution of jobs is justified or randomly distributed. However, it has been argued that if the gender distribution of occupations is subject to some sort of discrimination, as implied by the occupational segregation literature, the dummy variable approach is inadequate (Meng and Miller, 1995, Miller, 1987, Reilly, 1991). Therefore, as occupational attainment is affected by discrimination, gender wage differences may be divided at least into two broad sources: first, within-occupations wage differences which are related to the explained and unexplained components mentioned above and, second, between-occupations wage differences due to job discrimination or occupational segregation. As in conventional wage decompositions, some part of the wage differential due to occupational segregation could be justified but some part may remain unexplained.

To the best of our knowledge, Brown et al. (1980) was the first to formulate a method to explicitly incorporate the effects of occupational segregation into the analysis of the gender wage gap. Using a multinomial logit, Brown et al. (1980) modelled the male occupational distribution to produce a counterfactual female occupational distribution based on the set of female average characteristics and the estimated coefficients from the male subsample. Thus, besides the explained and unexplained components, the gender wage gap is further disaggregated into a portion due to explained gender differences in the allocation of workers and a portion due to occupational segregation. Other studies (Liu et al., 2004, Neuman & Silber, 1996) have implemented a similar multinomial logit approach to decompose wage differences between ethnic groups while Miller (1987) relied on an ordered probit to model the gender distribution of occupations.

As suggested by Heckman (1979), the estimation of earning functions by gender posesa potential problemas they are drawn from a non-random sample of individuals who select themselves into the labour force. This creates well-known problems for gender wage gap decompositions given the fact that female labour force participation is markedly lower than male participation. But as suggested by the occupational segregation literature, access to some occupations may also be highly differentiated by gender (i.e., Anker, 1997, Borghans & Groot, 1999, Grazier & Sloane, 2007, Hakim, 1992). This suggests another important source of selectivity bias which has not been accounted for in most of the empirical

76

literature about the gender wage gaps and occupational segregation.

The aim of this paper is to contribute to this literature by implementing a methodology proposed by Reilly (1991) which simultaneously controls for the effects of occupational segregation and corrects for selection into different occupations using Lee's (1983) multinomial logit based procedure. With an empirical application to urban Colombia, we find that controlling for selectivity bias at the occupational category level is found to be relevant in all years reviewed in this study. The results also suggest that a changing composition of the female labour supply in terms of un observables (i.e., ability and motivation) is playing a significant role in the dramatic reduction of the wage gap in this country. The rest of this paper is organised as follows. The next section formulates the empirical model andthe third explains the data. The fourth provides some background of the country while the fifth presents the results and discusses the findings. Finally, the fifth section concludes.

METHODOLOGY

Brown et al. (1980) formulated an empirical framework in which human capital and productivity differences explain simultaneously the variation in wages and occupational attainment of men and women. Using a multinomial logit to produce a counterfactual of the female distribution of jobs using the coefficients for the male subsample, this approach provide a decomposition of the gender wage gap not only into differences in coefficients and average characteristics between men and women, but also into differences from explained and unexplained components of the gender occupational distribution. Formally, this approach may be formulated as follows. Let the male and female log wage equations be

$$W_j^m = X_j^m \, \beta_j^m + \varepsilon_j^m \qquad (1)$$

$$W_j^f = X_j^f \, \beta_j^f + \varepsilon_j^f \qquad (2)$$

where X is a vector of characteristics, β is a set of coefficients to be estimated across j occupations for the male ($= m$) and female ($= f$) subsamples and ε is an error term. Let P^m and P^f be the sample proportions for each gender. The gender wage gap for the overall labour force with k occupational categories may be expressed as

$$\Delta \overline{W} = \overline{W}^m - \overline{W}^f = \sum_{j=1}^{k} \left(P_j^m = \overline{W}_j^m - P_j^f + \overline{W}_j^f \right) \qquad (3)$$

After rearranging terms and having P^*_j as the counterfactual distribution of female employment in the absence of segregation, the gender wage gap may be decomposed into four terms:

$$\Delta W = \sum_{j=1}^{k} P_j^f \overline{X}_j^f \Delta \widehat{\beta}_j + \sum_{j=1}^{k} P_j^f \widehat{\beta}_j^m \Delta \overline{X}_j + \sum_{j-1}^{k} \overline{W}_j^m (\overline{P}_j^m - \overline{P}_j^*) + \sum_{j-1}^{k} \overline{W}_j^m (\overline{P}_j^* - \overline{P}_j^f) \quad (4)$$

where the fist term on the right-hand side represents the unexplained component of the gender wage gap due to differences in returns from observable characteristics, the second is the component due to the level differences in those characteristics, the third is the component due to the explained allocation of workers and the fourth represents the portion of the wage gap due to occupational segregation[1]. According to Liu et al. (2004), the first two components may be interpreted as the within occupation wage differential while the last two denote the between occupational wage differential.

It has been established that ignoring selection effects when present may yield a bias in the coefficient estimates. In the case of the labour market, this possibility is present in the endogeneity of a workers' occupational choice with respect to the income determination process within specific occupations (Gyourko and Tracy, 1988). While the Heckman (1979) procedure may be valid to correct for selection bias in terms of the binary case of participating or not participating in the labour market, this approach has evident limitations when dealing with segmented labour markets and/or where the decision to participate is taken across multiple choices such as economic activities and/or occupational categories. Using Lee's (1983) generalised method which allows a correction for selectivity bias across a set of mutually exclusive groups, Reilly (1991) refined the approach proposed by Brown et al. (1980) by using the same multinomial log it model results (also for women) in order to create a selection term as an additional regressor into each of the occupational wage equations[2]. Assuming that the probability P of attachment of the i individual to the j occupational choice is determined by the vector Z, the multinomial log it model of occupational choice may be represented as

$$P_i^j = \frac{\exp (Z_i \gamma_i)}{1 + \sum_{j=1}^{k-1} (Z_i \gamma_i)} \quad (5)$$

---

[1] This type of decomposition uses the male wage structure as the one prevailing in the absence of discrimination and we believe this is a reasonable approach in the context of this empirical application. However, the index number approach pursued here is subject to the conventional "index problem" as it is also possible to use the female wage structure or even the pooled sample wage structure. For a more extensive discussion of this topic, see: Appleton, Hoddinott & Krishnan (1999, p. 289-312).

[2] For an excellent review of alternative methods to correct for selectivity bias using multinomial logit models see: Bourguignon, Fournier & Gurgand (2007, p. 174-205).

Where the base outcome j is set equal to 0 as required by the Theil normalization. Estimates from (5) are used not only to construct the counterfactual occupational distribution $\bar{P}^*$ in (4) (with the coefficients drawn from the male subsample to simulate the occupational distribution for the female subsampleas if they were treated as their male counterparts) but also to correct the occupational wage equations for selection bias. Thus, the wage equations in (1) and (2) corrected for selectivity bias may be rewritten as

$$W_j^m = X_j^m \ \beta_j^m - \sigma_j^m \ \rho_j^m \ \frac{\varphi \ (J(z_j^m \ \gamma_j^m))}{F \ (z_j^m \ \gamma_j^m))} + \in_j^m \tag{6}$$

and

$$W_j^f = X_j^f \ \beta_j^f - \sigma_j^f \ \rho_j^f \ \frac{\varphi \ (J(z_j^f \ \gamma_j^f))}{F \ (z_j^f \ \gamma_j^f))} + \in_j^f \tag{7}$$

where φ and J denote the standard normal density function and the inverse of the cumulative density function, respectively, of predicted probabilities given by $Z_j^m \gamma_j^m$ and $Z_j^f \gamma_j^f$;s_j parameters represent the standard error in the jth wage equation for W and r_j parameters capture the strength of the correlation between the un observables determining W_j and those determining the jth occupational attachment modelled by the MNL. Lee's (1983) approach is analogous to Heckman's (1979) two-step procedure in the sense that the maximum likelihood estimates from (5) for each gender subsample are inserted in (6) and (7) in order to construct a selection term. Thus, (6) and (7) may be conveniently rewritten as

$$W_j^m = X_j^m \ \beta_j^m - \theta_j^m \ \widehat{\lambda}_j^m + \in_j^m \tag{8}$$

and

$$W_j^f = X_j^f \ \beta_j^f - \theta_j^f \widehat{\lambda}_j^f + \in_j^f \tag{9}$$

where, in the case of the male subsample, $\lambda = - \varphi \ (J(z_j^m \ \gamma_j^m)) \ F(z_j^f \gamma_j^f)$ and $\in_j^m$ is an error term[3]. With a large sample, heteroscedasticity consistent standard errors for (8) and (9) can be obtained using the White/Huber sandwich estimator (White, 1980) though, bootstrapping is also possible. We preferred White's (1980) procedure as this provides a reasonably accurate estimate of the variance/covariance matrix in this context.

---

[3]   It should be noted that our strategy to correct for selectivity bias based on a multinomial logit approach allows controlling for two sources of bias, one from selection into the labour force and another from the choice of the occupational category to which the worker belongs. This implies that all selection effects commented here in this paper refer to an average worker drawn at random from the sample with a given a set of average characteristics.

Data

The empirical application of the model described above relies on household survey microdata from the seven largest metropolitan areas of urban Colombia which represent around 36 per cent of the country's population and nearly one-half of its urban population. Household surveys in urban Colombia are gathered by the Government on a quarterly basis since the mid-1980s and our dataset comprises five repeated cross sections for all quarters of 1986, 1990, 1995, 2000 and 2004[4]. On average, each of these years provides microdata on 160,000 individuals between 18 and 65 years old and 66,000 workers in full time occupations (which, by Colombian standards is 40 or more hours per week). The microdata give information on labour earnings, number of weekly worked hours, industry, region, demographic variables (i.e., age, educational attainment, marital status),

and the socioeconomic strata of the incumbent household which is based on housing and neighbourhood characteristics. Furthermore, the surveys include information about occupations using a consistent classification of 82 categories over the entire period which, at the two-digit level, is identical to the International Standard Classification of Occupations ISCO-68. For the purposes of this research, we regrouped the 82 original categories into five broad occupations for computational convenience in order to guarantee a large enough sample size for each (see Table 1) plus one additional category which comprises both, individuals out of the labour force and part-time workers (less than 40 hours per week). We attempted also an alternative grouping using the ISCO-68 at one digit with nine categories but the small number of female observations in some cells made this approach unfeasible for accurate statistical inference.

---

[4] It should be observed that more recent microdata is available for this country. We restrict our study, however, this time frame as the sampling design of the household survey changed substantially in subsequent years, particularly since 2006 onwards when both, a different sampling methodology and population projections were implemented in this country. Similar exercises with more recent data are indeed possible and we are currently working on that on a new paper.

Table 1. Broad ad-hoc occupational categories and
average sample sizes per year and category

Categories	Description	ISCO 68 codes included (1 digit)	Average sample size per year: observations*	
			Males	**Females**
Professionals and managers	Professional, technical and related workers; military personnel, teachers, religion workers, writers, journalists, artists, musicians, athletes and sportsmen. Members of legislative bodies; directors and managerial workers	0-1, 2	4,006	2,301
Clerical and related workers	Clerks, sales personnel and trade related workers	3, 4	11,550	9,245
Service workers	Catering and lodging services, protective and security services; housekeepers, maids, hairdressers, barbers, beauticians, prostitutes, guides, undertakers, embalmers	5	4,699	7,963
Semi-skilled workers	Agricultural, Animal Husbandry and Forestry Workers, Fishermen and Hunters. Workers in mining and manufacturing industry; machinery operators. Shoemakers and leather goods makers; cabinet makers and related wood workers; electrical fitters and related electrical and electronics workers, jewellery and precious metal workers; glass formers, potters and related workers; broadcasting station and sound equipment operators	6, 7, 8, 9	10,066	3,361
Unskilled workers	Unskilled blue-collar workers in general: workers in rubber, plastic, paper, printing, construction, power generation and transport industries.	10	11,721	893

* Includes full-time workers (working 40 or more hours per week) aged 18 to 65 years. See (DANE, 2000) for a full description of the Colombian classification of occupations.
Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for labour force aged 18 and 65 years working full time (>200 hours per month).

EMPIRICAL RESULTS

Determinants of occupational attainment

As customary in similar empirical applications, the occupational attainment model described above is reduced form in nature and assumes that the occupation held by an individual is the result of supply and demand factors (Brown et al., 1980, Liu et al., 2004, Reilly, 1991). On the supply side, the decision of the individual to engage in a certain occupation is determined by a utility function in terms of expected income, personal tastes and family structure. Although most of those variables are not observed in the Colombian microdata, our model is rich invariables related to the household structure: marital status, a dummy for household heads, household unemployment rate, number of both infants and children in the household; and a dummy for households with a domestic servant[5]. Personal tastes and preferences are not directly observable but we attempt to proxy them with neighbourhood characteristics based on average log labour income and average education for the corresponding sampling segment[6]. We include also dummies for cities and neighbourhood socioeconomic strata which reveal housing conditions and

access to urban infrastructure. On the demand side, the willingness of an employer to hire a worker in a given position is assumed to be dictated by productivity signals as well as firm and job-specific tastes for discrimination. Our multinomial logit specification includes as productivity indicators a conventional proxy of labour experience (i.e., age – years of schooling – 6) with a quadratic term and, dummies for primary, secondary and university education. Firm and job-specific tastes for discrimination are, thus, relegated to the residual together with other un observables.

The maximum likelihood estimates of the multinomial logit described in (5) are presented in Tables 2a and 2b. In order to conserve space, we present only the results for 2004[7]. It is very well known that coefficient interpretation is not entirely transparent in the logit specification. For this reason, we prefer to present the marginal and impact effects computed at the average of sample characteristics rather than the crude multinomial logit coefficients. None the less, their interpretation should be taken with some caution asthe average characteristics may not be entirely appropriate at the extremes of the probability distribution of occupational

---

[5] The set of household characteristics included in the model plays an important role in our modelling strategy to control for selectivity bias as they represent the instruments not directly correlated with wages but highly correlated with the probability to participate in different occupational categories.

[6] Sampling segments, as the final sampling unit, are compact or continuous sets of ten houses which are surveyed in full. For details on the sample design of the Colombian household survey, see: DANE (2004).

[7] Marginal effects for other years are included in Appendix A.

outcomes. Beyond these caveats, some inferences can be safely made on the basis of these results. For instance, our proxy of labour market experience for both, men and women is statistically significant beyond conventional levels in all occupations. The significance of the quadratic term of labour market experience suggests an inverted U-shaped relationship for all categories and appears to be inversely correlated with the skill level. While the maximum probability of attachment to the professional & managerial category is located at 18 years of experience amongst men and 16 years amongst women, the peaks for those in semi-skilled and unskilled occupations are, respectively, at 26 years and 24 years for the male and female subsamples. On the other hand, the estimated effects for the dummies by educational level suggestin general that the higher the educational level of the individual, the more likely he or she is to be found in occupations where academic credentials are relevant. For example, possessing a university degree increases the probability of being in professional occupations by 34.7 percentage points amongst the male subsample compared to 38.7 percentage points amongst the female subsample.

Table 2a. Marginal/impact effects from multinomial log it coefficients
of occupational attainment, male subsample, 2004

	Professionals & managers	Clerks & Sales	Services	Semi-skilled	Unskilled
Experience	0.001	0.007	0.003	0.008	0.010
	(0.000)**	(0.000)**	(0.000)**	(0.008)**	(0.000)**
Experience^2	0.000	0.000	0.000	0.000	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Primary education	0.007	-0.020	-0.002	0.018	0.007
	(0.005)	(0.006)**	-0.004	(0.018)**	(0.004)
Secondary education	0.044	0.015	0.021	-0.002	-0.055
	(0.006)**	(0.006)*	(0.004)**	(-0.002)**	(0.005)**
University education	0.347	-0.056	-0.057	-0.107	-0.152
	(0.033)**	(0.008)**	(0.003)**	(-0.107)**	(0.003)**
Single	0.001	0.010	-0.002	-0.005	0.001
	(0.002)	(0.007)	-0.004	(-0.005)**	(0.005)
HH. head	0.015	0.051	0.033	0.035	0.032
	(0.001)**	(0.004)**	(0.002)**	(0.035)**	(0.003)**
HH. infants	0.000	0.006	0.004	0.006	0.009
	(0.001)	(0.002)**	(0.001)**	(0.006)**	(0.001)**
HH. children	-0.002	-0.003	0.000	0.002	0.001
	(0.001)**	(0.002)*	-0.001	(0.002)**	(0.001)
HH. Unemployment rate	-0.045	-0.320	-0.113	-0.186	-0.170
	(0.005)**	(0.014)**	(0.009)**	(-0.186)*	(0.011)**
Domestic servant	0.008	-0.026	0.003	-0.007	0.000
	(0.002)**	(0.005)**	-0.003	(-0.007)**	(0.004)
Log wage segment	-0.001	-0.003	0.001	-0.001	0.000
	(0.000)**	(0.002)*	-0.001	(-0.001)**	(0.001)
Education segment	0.003	0.001	-0.005	-0.009	-0.013
	(0.000)**	(0.001)	(0.000)**	(-0.009)**	(0.001)**
Barranquilla	0.001	0.045	0.007	0.020	0.011
	(0.002)	(0.006)**	(0.004)*	(0.02)**	(0.005)*
Bucaramanga	-0.003	0.000	-0.012	0.014	0.003
	-0.002	(0.006)	(0.003)**	(0.014)**	(0.005)
Manizales	-0.010	-0.024	-0.002	0.026	-0.017
	(0.002)**	(0.005)**	-0.004	(0.026)**	(0.005)**
Medellin	0.009	-0.007	-0.009	0.030	-0.003
	(0.002)**	(0.005)	(0.003)**	(0.03)**	(0.005)
Cali	0.002	0.003	-0.002	0.015	0.003
	(0.002)	(0.006)	-0.003	(0.015)**	(0.005)
Pasto	-0.013	-0.042	-0.018	-0.009	-0.005
	(0.001)**	(0.005)**	(0.003)**	(-0.009)**	(0.005)

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for people aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %. (i) denotes impact effects.

Table 2b. Marginal/impact effects from multinomial log it coefficients
of occupational attainment, female subsample, 2004

	Professionals & managers	Clerks & Sales	Services	Semi-skilled	Unskilled
Experience	0.000	0.004	0.004	0.003	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Experience^2	0.000	0.000	0.000	0.000	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Primary education	0.008	-0.004	-0.016	0.017	0.001
	-0.004	(0.005)	(0.002)**	(0.002)**	(0.001)
Secondary education	0.035	0.079	-0.050	0.011	0.001
	(0.007)**	(0.006)**	(0.003)**	(0.002)**	(0.001)
University education	0.387	0.006	-0.081	-0.019	-0.008
	(0.073)**	(0.013)	(0.001)**	(0.002)**	(-0.008)**
Single	0.000	0.011	0.030	0.007	0.003
	(0.001)	(0.004)**	(0.003)**	(0.002)**	(0.003)**
HH. head	0.003	0.020	0.031	0.013	0.004
	(0.001)**	(0.003)**	(0.003)**	(0.002)**	(0.004)**
HH. infants	-0.001	-0.008	-0.006	-0.002	0.000
	(0.000)**	(0.001)**	(0.001)**	(0.001)**	(0.000)**
HH. children	-0.001	-0.004	-0.005	-0.001	0.000
	(0.000)**	(0.001)**	(0.001)**	(0.001)	(0.000)**
HH. Unemployment rate	-0.023	-0.250	-0.107	-0.041	-0.017
	(0.002)**	(0.009)**	(0.007)**	(0.004)**	(-0.017)**
Domestic servant	-0.001	-0.037	0.220	-0.015	-0.004
	(0.000)**	(0.003)**	(0.005)**	(0.001)**	(-0.004)**
Log wage segment	0.000	0.001	0.001	0.002	0.000
	(0.000)**	(0.001)	(0.001)	(0.001)**	(0.000)**
Education segment	0.001	-0.003	0.001	-0.003	-0.001
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(-0.001)**
Barranquilla	0.006	-0.010	-0.005	-0.003	-0.004
	(0.001)**	(0.003)**	-0.003	-0.002	(-0.004)**
Bucaramanga	0.000	-0.005	-0.008	0.010	-0.002
	(0.001)	(0.004)	(0.003)**	(0.002)**	(-0.002)*
Manizales	-0.001	-0.034	-0.015	-0.004	-0.004
	(0.001)	(0.003)**	(0.002)**	(0.002)*	(-0.004)**
Medellin	0.006	-0.008	-0.007	0.020	-0.003
	(0.001)**	(0.003)**	(0.002)**	(0.002)**	(-0.003)**
Cali	0.000	-0.002	-0.006	-0.004	0.001
	(0.001)	(0.004)	(0.003)*	(0.002)*	(0.001)
Pasto	-0.002	-0.002	-0.012	-0.016	-0.006
	(0.001)**	(0.004)	(0.003)**	(0.001)**	(-0.006)**

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for people aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %. (i) denotes impact effects.

Variables pertaining to the family structure appear also to play an important role in the choice of occupation with a clear differentiation in terms of gender. Among such variables, both the number of infants and children are found to be negatively correlated with the probability of female participation in all categories. Conversely, the number of infantsis found to increase the probability of being in all occupations amongst menwhile the number of children reduces the probability of females being working in professionals, clerks and sales jobs. Marital status is also highly differentiated in terms of gender in as much as being single raises the probability of attachment of women to most occupations (except professionals & managers) while its effect is not statistically different from zero for men in all categories (except in semi-skilled occupations). In summary, the extent of differentiation in marginal/impact effects from variables related to the household structure somehow resembles asymmetric household responsibilities and gender roles which, in general, tend to reinforce patterns of occupational segregation[8].

Lastly, neighbourhood (or sampling segment) variables tend to reflect both, the tastes and the effects of the household social networks on the occupational outcome. Our results for 2004 suggests that women and men living in neighbourhoods where higher educational levels prevail are more likely to be attached to occupations where educational credentials are more relevant: professionals & managers and, clerks & sales personnel. Likewise, men and women living in neighbourhoods with higher educational levels are less likely to be found in semi-skilled and unskilled occupations.

Table 3 illustrates sample proportions of workers across the five occupational distributions defined for this study. It also displays the counterfactual distributions of female workers,, which arecomputed by applying the multinomial logit coefficients for the male subsample to the female subsample. Two categories, clerks & sales personnel and, service workers, represent more than 70 per cent of female employment in all years reviewed here. As the categorisation of occupations is quite broad, we observe also that the proportion of workers in professional & managerial positions in terms of gender is remarkably egalitarian and women have been increasing their share since 1986. Women appear underrepresented in the blue-collar semi-skilled and unskilled positions in all years. The counterfactual estimates indicate that women are over represented in service positions where

---

[8]   There is a major limitation in Colombian microdata as the family relationship of children is undetermined, except in the case of children with female household headship. This cast some doubt in the identification strategy of the selection model implemented in this paper given the scarcity of good instruments (i.e., highly correlated with the probability of participation but uncorrelated with wages) in the household surveys of this country. Unfortunately, this issue is present in most empirical works of the labour market for this country and it certainly deserves more research.

about two thirds of the female proportion of workers is 'unjustified'. Conversely, women are clearly underrepresented in unskilled jobs as the counterfactual

estimates indicate that the proportion of the female labour force allocated to this category should be, on average, more than 7.5 times higher.

Table 3. Male and female occupational distributions, urban Colombia (selected years)

Year	Occupation	P^m_i	P^f_i	P^*_i	$\lvert P^m_i - P^f_i \rvert$	$\lvert P^m_i - P^*_i \rvert$
1986	Professionals and managers	0.081	0.075	0.063	0.007	0.018
	Clerks and sales personnel	0.293	0.379	0.313	0.086	0.020
	Service workers	0.104	0.361	0.094	0.256	0.010
	Semi-skilled workers	0.252	0.148	0.251	0.104	0.001
	Unskilled workers	0.269	0.037	0.278	0.232	0.009
1990	Professionals and managers	0.092	0.085	0.085	0.008	0.007
	Clerks and sales personnel	0.272	0.386	0.297	0.114	0.025
	Service workers	0.105	0.334	0.096	0.229	0.009
	Semi-skilled workers	0.256	0.155	0.248	0.102	0.009
	Unskilled workers	0.274	0.040	0.274	0.234	0.000
1995	Professionals and managers	0.093	0.098	0.092	0.005	0.001
	Clerks and sales personnel	0.265	0.408	0.275	0.143	0.010
	Service workers	0.108	0.301	0.100	0.194	0.007
	Semi-skilled workers	0.235	0.151	0.233	0.084	0.002
	Unskilled workers	0.298	0.041	0.299	0.257	0.000
2000	Professionals and managers	0.100	0.105	0.107	0.005	0.007
	Clerks and sales personnel	0.274	0.391	0.281	0.117	0.007
	Service workers	0.124	0.347	0.119	0.223	0.005
	Semi-skilled workers	0.230	0.126	0.224	0.104	0.006
	Unskilled workers	0.272	0.031	0.269	0.241	0.003
2004	Professionals and managers	0.114	0.121	0.115	0.008	0.001
	Clerks and sales personnel	0.267	0.378	0.273	0.112	0.006
	Service workers	0.122	0.336	0.117	0.214	0.005
	Semi-skilled workers	0.219	0.127	0.217	0.093	0.002
	Unskilled workers	0.279	0.038	0.279	0.240	0.000

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for labour force aged 18 and 65 years working full time (>200 hours per month).

Our counterfactual modelling strategy assumes that the allocation of workers is primarily a function of personal, household and neighbourhood characteristics which are evaluated with the coefficients for the male subsample. Given that the distribution of those characteristics may differ in terms of gender, it is assumed

that some part of discrepancies in the gender distribution of workers across occupational categories is justified. For this reason, Table 4 presents factual and counterfactual estimates of dissimilarity indices to see more precisely the extent to which the gender distribution of jobs would be altered in the absence of

discrimination in the access to occupations. The first one is the widely known measure proposed by Duncan and Duncan (1955) which is interpreted as the percentage of workers of either sex who have to move from occupations in order to achieve an even distribution of jobs by gender. The second is a modified

version of the Duncan and Duncan index proposed by Karmel and Mac Lachlan (1988) which is interpreted as the proportion of female workers who would have to exchange jobs with male workers if an even gender distribution of jobs is to be achieved without altering the overall employment structure.

Table 4. Gender occupational segregation indices, urban Colombia

	1986	*1990*	*1995*	*2000*	*2004*
Duncan Index	0.342	0.344	0.341	0.345	0.333
Duncan Index*	0.029	0.025	0.010	0.014	0.008
Karmel and MacLachlan Index	0.153	0.155	0.157	0.163	0.158
Karmel and MacLachlan Index*	0.013	0.011	0.005	0.007	0.004

* counterfactual index.
Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for labour force aged 18 and 65 years working full time (>200 hours per month).

Estimates for the Duncan and Duncan index reveal little changes between 1986 and 2000, while those for the Karmel and Mac Lachlan suggest a slight increase in the extent of segregation. While the first one fell by just 2.8 per cent between 1986 and 2004 the second grew by 2.9 per cent over this same period. Furthermore, if women were treated as men in the occupational selection or sorting process, as assumed in the counterfactual scenario, both indices in 2004 would be reduced by more than 97 per cent. This gives an indication that most of the occupational segregation based on the five occupational categories implemented in our multinomial logit specification above is unjustified. It should be noted, however, that a set of different estimates using a classification of 82 occupations indicates that occupational segregation has experienced a substantial reduction during this period, particularly

amongst labour force with university education (Isaza & Reilly, 2009). The magnitude of standard errors computed using a bootstrap technique is remarkably small, with zeros in the first three decimal positions in most cases either in the case of the full 82 occupations or in the aggregated version of five categories presented here. For this reason, we judged useless to report them on the table.

Wage equations and selection effects

Now our attention turns to the interpretation of wage equations corrected for selectivity bias. Our log wage specification is quite austere in terms of human capital and productivity characteristics as the Colombian data do not include in all waves explicit information about labour market experience, and characteristics related to specific types of education and abilities

possessed by the individual. Hence, we have to rely only on those variables which are available in all survey waves between 1986 and 2000: age, its quadratic form, and the number of years of formal schooling in the form of a disjointed spline function with one knot at 11 years of formal schooling. The inclusion of education in the spline form aims at a consistent characterisation of the labour market across different occupational categories over the 19 year period analysed in this study. In the case of Colombia, the completion of 11 years of education constitutes a landmark in the educational system of this country as this enables access to professional and most vocational training programmes. Furthermore, access to compulsory military service for males is shortened for those with complete secondary education (i.e., 11 years of formal schooling). Complete secondary education, with a certificate of previous compliance of military service for males, are valuable credentials to access most of jobs in the formal sector. We include also seasonal dummies and controls for cities, as well as two dichotomous variables depicting some job characteristics: one for government employees, and another for own-account workers. As the literature on wage equations by gender is quite abundant, even in the Colombian case, our interpretation emphasises on the role of selection effects[9].

A distinctive feature of this empirical application is the inclusion of specific selection terms in the wage equations for each occupational category using the estimates from multinomial logit equations of occupational attainment by gender. Following the interpretation suggested in Gyourko and Tracy (1988) and Reilly (1991), the selection effect can be computed as the product of the mean value for the selection variable ($\varphi/F = \lambda$) multiplied by the coefficient value ($\sigma_j \rho_j = \theta$) (see expressions 6 and 7, above). According to this method, selectivity effects are found to be statistically significant for both, women and men in most years and across the majority of occupations (see Tables 5a and 5b). In the male subsample, the strongest effects are found amongst clerks & sales workers followed by those in the professional & managerial category. In the former, the results for 2004 indicate that men earn, on average and *ceteris paribus*, 95 per cent higher wages than an individual drawn at random from the male subsample while in the latter this differential is more than 89 per cent (see Table 5a). In the female subsample, selection effects are the strongest amongst professionals & managers whose hourly wages in 2004 are 59 per cent higher than those of an average woman randomly drawn from the female subsample (see Table 5b).

A possible rationale for the results outlined above is that individuals who are selected into occupations where academic credentials are highly relevant (i.e., professionals and clerks) possess some

---

9   A full set of estimates of the wage equations described in expressions (8) and (9) is presented in Appendix B.

productivity characteristics such as motivation and ability which are unobserved by the econometrician but highly rewarded in the labour market. If the set of unobservables is highly correlated with both, labour productivity and selection into some particular occupations, the magnitude of selection effects into these categories should be substantially larger compared to other occupations. This seems to be the case of men and women in professional & managerial occupations (and, to a lesser extent, men in clerks and sales personnel) whose selection coefficients are substantially larger than the rest of the labour force. To illustrate the opposite case, workers in the unskilled category report the lowest selection effects with -14.8 and -64.4 per cent in 2004 for the male and female subsamples respectively. Both selection effects are statistically different from zero at the one percent level. Selection effects in this category, as opposed to the case of the most qualified, imply that unskilled workers are endowed with a set of unobservables such as motivation and ability well below the average of their gender subsamples and consequently secure lower wages. This also suggests the possibility that differences in unobservables are more clearly differentiated when the labour force is divided into several occupational categories compared to the dichotomous approach of labour participation which tends to overshadow selection effects, particularly among male workers.

Overall the trends in the selection effects between 1986 and 2004 suggest opposite patterns across gender groups, as they are increasing for the male subsample and decreasing for the female subsample. A plausible explanation for these results is the diverging patterns of labour force participation in terms of gender over these years, with an increasing proportion of women enrolling in the labour market which makes selection effects less differentiated between those inside and outside the labour force. A decreasing trend of male labour force participation may be performing an opposite effect in the male subsample. Interestingly, selection effects appear to be larger for the male subsample and statistically significant in most cases. This finding contrasts with some empirical applications in urban Colombia where the conventional probit approach based on Heckman (1979) suggests that selection effects are mainly relevant only for the female subsamples (Nunez et al., 2005, Núñez & Sánchez, 2002, Ribero, 2003). On this it should be stressed that our multinomial logit approach is implicitly controlling for two different processes. The first one which relies on the individual decision to participate or not in the labour market and another which has to see with the sorting of individuals across different occupations once the decision to participate has been taken. While the former is a self-selection decision of the individual based on personal circumstances, the second is a selection process driven mainly by employers' preferences based on their valuations of different individual characteristics (i.e., educational level and previous labour market experience).

Table 5a. Selection effects from wage equations corrected
for selectivity bias, male subsample

Category		*1986*	*1990*	*1995*	*2000*	*2004*
Professionals						
& Managers	-θ	-0.377	-0.530	-0.536	-0.367	-0.631
		(0.041)**	(0.036)**	(0.042)**	(0.052)**	(0.04)**
	Λ	1.375	1.333	1.369	1.395	1.414
		(0.012)**	(0.011)**	(0.011)**	(0.012)**	(0.011)**
	Effect=-θ*Λ	0.518	0.706	0.734	0.512	0.892
		(0.043)**	(0.038)**	(0.043)**	(0.053)**	(0.042)**
Clerks						
& sales	-θ	-0.636	-0.584	-0.625	-0.509	-0.629
		(0.04)**	(0.04)**	(0.045)**	(0.05)**	(0.06)**
	λ	1.453	1.438	1.444	1.510	1.516
		(0.002)**	(0.002)**	(0.002)**	(0.002)**	(0.002)**
	Effect=-θ*Λ	0.924	0.840	0.903	0.769	0.953
		(0.04)**	(0.04)**	(0.045)**	(0.05)**	(0.06)**
Service						
workers	-θ	-0.026	0.209	-0.030	-0.359	-0.119
		(0.054)	(0.055)**	(0.064)	(0.1)**	(0.054)*
	λ	1.917	1.884	1.873	1.869	1.846
		(0.003)**	(0.003)**	(0.002)**	(0.003)**	(0.003)**
	Effect=-θ*Λ	0.050	-0.394	0.056	0.671	0.220
		(0.054)	(0.055)**	(0.064)	(0.1)**	(0.054)**
Semi-skilled						
workers	-θ	-0.007	-0.019	0.071	-0.041	-0.045
		(0.041)	(0.038)	(0.043)*	(0.068)	(0.056)
	λ	1.498	1.448	1.484	1.581	1.565
		(0.002)**	(0.002)**	(0.002)**	(0.003)**	(0.003)**
	Effect=-θ*Λ	0.010	0.028	-0.105	0.065	0.070
		(0.041)	(0.038)	(0.043)**	(0.068)	(0.056)
Unskilled						
workers	-θ	0.191	0.230	0.125	0.112	0.106
		(0.036)**	(0.031)**	(0.032)**	(0.046)**	(0.037)**
	λ	1.411	1.361	1.310	1.451	1.398
		(0.003)**	(0.003)**	(0.002)**	(0.003)**	(0.003)**
	Effect=-θ*Λ	-0.270	-0.313	-0.164	-0.163	-0.148
		(0.036)**	(0.031)**	(0.032)**	(0.046)**	(0.037)**

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for people aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %.

Table 5b. Selection effects from wage equations corrected
for selectivity bias, female subsample

Category		1986	1990	1995	2000	2004
Professionals						
& Managers	$-\theta$	-0.105	-0.267	-0.283	-0.308	-0.359
		(0.047)*	(0.042)**	(0.044)**	(0.049)**	(0.044)**
	λ	1.774	1.675	1.626	1.616	1.643
		(0.016)**	(0.015)**	(0.013)**	(0.014)**	(0.012)**
	Effect=$-\theta*\Lambda$	0.186	0.447	0.460	0.498	0.590
		(0.05)**	(0.045)**	(0.046)**	(0.051)**	(0.046)**
Clerks						
& sales	$-\theta$	-0.096	-0.043	0.008	-0.032	-0.081
		(0.039)**	(0.032)	(0.032)	(0.037)	(0.043)*
	λ	1.686	1.613	1.531	1.600	1.608
		(0.003)**	(0.003)**	(0.003)**	(0.003)**	(0.003)**
	Effect=$-\theta*\Lambda$	0.162	0.069	-0.012	0.051	0.130
		(0.039)**	(0.032)*	(0.032)	(0.037)	(0.043)**
Service						
workers	$-\theta$	-0.001	0.085	0.075	0.058	0.003
		(0.013)	(0.012)**	(0.013)**	(0.018)**	(0.015)
	λ	1.271	1.329	1.421	1.360	1.441
		(0.008)**	(0.008)**	(0.007)**	(0.007)**	(0.006)**
	Effect=$-\theta*\Lambda$	0.001	-0.113	-0.107	-0.079	-0.004
		(0.015)	(0.014)**	(0.015)**	(0.019)**	(0.016)
Semi-skilled						
workers	$-\theta$	0.010	0.084	0.105	0.000	0.050
		(0.046)	(0.036)**	(0.04)**	(0.054)	(0.055)
	λ	2.049	1.990	1.968	2.038	2.008
		(0.005)**	(0.005)**	(0.005)**	(0.007)**	(0.006)**
	Effect=$-\theta*\Lambda$	-0.020	-0.167	-0.207	0.000	-0.100
		(0.046)	(0.036)**	(0.04)**	(0.054)	(0.055)*
Unskilled						
workers	$-\theta$	-0.029	-0.028	0.211	-0.010	0.256
		(0.101)	(0.143)	(0.101)*	(0.142)	(0.132)*
	λ	2.547	2.548	2.469	2.608	2.517
		(0.009)**	(0.007)**	(0.008)**	(0.01)**	(0.009)**
	Effect=$-\theta*\Lambda$	0.074	0.071	-0.521	0.026	-0.644
		(0.101)	(0.143)	(0.101)**	(0.142)	(0.132)**

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for people aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %.

Decomposition of gender wage gaps by occupational categories

In this section we present the results of the gender wage gap decomposition by occupational categories. As selection effects are statistically significant in most cases, we extend the decomposition by including the effects of the selection terms in the decomposition. Thus, the log wage gap for the *j* category may be expressed as

$$W_j^m = W_j^f = [\overline{X}_j^m - \overline{X}_j^f] \, \hat{\beta}_j^m + \overline{X}_j^{f\prime} [\hat{\beta}_j^m - \hat{\beta}_j^f] + [\hat{\theta}_j^m \, \overline{\lambda}_j^m - \hat{\theta}_j^f \, \overline{\lambda}_j^f] \qquad (10)$$

where the first term on the right hand side represents the portion of the wage gap due to differences in treatments, the second captures the portion due to differences in endowments and the third comprises the portion due to differences in selection effects. The results of expression (10) for all years are reported in Table 6.

Table 6. Decomposition of log wage gaps by occupational categories,
urban Colombia (selected years)

Categories	Components	1986	1990	1995	2000	2004
Professionals	*Treatments*	0.649	0.533	0.504	0.229	0.398
& Managers		(0.107)**	(0.088)**	(0.093)**	(0.098)*	(0.087)**
	Endowments	0.098	0.063	0.057	0.013	0.021
		(0.012)**	(0.007)**	(0.006)**	-0.008	(0.005)**
	Selection	-0.308	-0.235	-0.272	-0.064	-0.248
		(0.017)**	(0.025)**	(0.028)**	(0.023)**	(0.031)**
	Log wage gap	0.440	0.360	0.289	0.178	0.171
		(0.022)**	(0.02)**	(0.021)**	(0.026)**	(0.023)**
Clerks						
& Sales	*Treatments*	0.640	0.782	0.530	0.602	0.723
		(0.101)**	(0.086)**	(0.094)**	(0.106)**	(0.096)**
	Endowments	-0.003	0.011	0.003	-0.046	-0.046
		(0.004)	(0.003)**	(0.004)	(0.005)**	(0.005)**
	Selection	-0.461	-0.640	-0.386	-0.505	-0.630
		(0.014)**	(0.017)**	(0.021)**	(0.021)**	(0.024)**
	Log wage gap	0.176	0.153	0.147	0.050	0.046
		(0.011)**	(0.01)**	(0.01)**	(0.013)**	(0.013)**
Services	*Treatments*	-0.405	0.712	-0.055	0.023	-0.323
		(0.167)*	(0.128)**	(0.143)	-0.196	(0.133)*
	Endowments	0.119	0.066	0.076	0.096	0.094
		(0.01)**	(0.008)**	(0.008)**	(0.01)**	(0.007)**
	Selection	0.589	-0.447	0.250	0.119	0.484
		(0.021)**	(0.015)**	(0.02)**	(0.028)**	(0.022)**
	Log wage gap	0.303	0.331	0.271	0.238	0.255
		(0.012)**	(0.011)**	(0.011)**	(0.016)**	(0.012)**
Semi-skilled	*Treatments*	0.118	0.585	0.520	0.847	0.519
		(0.159)	(0.13)**	(0.137)**	(0.192)**	(0.167)**
	Endowments	0.012	0.003	-0.004	-0.029	-0.046
		(0.002)**	-0.002	-0.002	(0.004)**	(0.004)**
	Selection	0.114	-0.329	-0.258	-0.626	-0.336
		(0.022)**	(0.016)**	(0.02)**	(0.019)**	(0.022)**
	Log wage gap	0.243	0.259	0.258	0.193	0.138
		(0.012)**	(0.01)**	(0.01)**	(0.016)**	(0.014)**
Unskilled	*Treatments*	0.702	0.917	0.952	0.291	0.317
		(0.28)*	(0.236)**	(0.226)**	(0.304)	-0.294
	Endowments	-0.001	0.002	-0.001	-0.081	-0.068
		(0.004)	-0.004	-0.004	(0.007)**	(0.006)**
	Selection	-0.583	-0.791	-0.777	-0.264	-0.195
		(0.014)**	(0.015)**	(0.013)**	(0.034)**	(0.042)**
	Log wage gap	0.118	0.128	0.174	-0.054	0.054
		(0.019)**	(0.018)**	(0.016)**	(0.025)*	(0.025)*

The magnitude of the gender wage gap contracted in all occupational categories between 1986 and 2004. The largest reductionwas recorded amongst clerks and sales workers, where the differential fell by three quarters from 17.6 per cent in 1986 to 4.6 per cent in 2004, the lowest of all five groupings. In the case of professionals & managers, the log hourly wage gap fell by almost two thirds over the same period while it decreased by more than one half amongst unskilled workers. The smallest reduction is found for service workers, traditionally the occupational category with the largest proportion of women where, incidentally, the differential in 2004 was the highest of all occupational groups, with 25.5 per cent. It should be noted that those differentials are statistically significant in all cases, at least, at the five per cent level.

Regarding the decomposition methodology outlined in equation (10), differences in treatment (or what may be termed as differences in returns to observable characteristics which are presumably due to discrimination) explain most of gender wage gap for all occupational categories (except amongst service workers) over most of the five years included in this study. This component has also increased its magnitude between 1986 and 2004 in two occupational categories, clerks & sales and, semi-skilled workers while it got smaller for professionals & managers and, unskilled workers. It has been negative amongst service workers in most of the years which suggests that women's observable characteristics are better

rewarded compared to their male counterparts in this particular category. This may explain in part why this is the occupational category with the highest proportion of women in the labour market of urban Colombia.

Differences in treatments are partially offset by selection effects which are found to be negative, thus reducing the size of the wage gapin most years. In fact, differences in selection are found to reduce the wage gap in all occupational categories and years, except amongst service workers from 1995 onwards and unskilled workers in 2000. Implicitly, the inclusion of selection effects in the decomposition outlined in (10) assumes that the level of motivation, ability and other unobservables is homogeneous across the gender subsamples but their remuneration is gender differentiated (Gyourko& Tracy, 1988). This provides some evidence that unobserved variables consigned to the selection term might be playing an important role in reducing gender wage differences in urban Colombia.

The portion of the wage gap due to differences in endowments, which include human capital and other observable characteristics, experienced a net reduction in all occupational categories between 1986and 2004. This suggest that another force behind the reduction of the gender wage gap in urban Colombia over these years is the rapid educational improvement of women relative to men which has been documented in the literature for this

country (Isaza, 2002; Acosta & Isaza, 2007). This reduction is particularly important for workers in clerks & sales positions and, to a lesser extent, in semi-skilled jobs where the endowment component is negative for most years.

The effects of segregation on the gender wage gap

As previously noted, the gender hourly wage gap in urban Colombia presented a dramatic reduction across all occupational categories over the years examined in this study. Overall, it fell by more than two thirds from 27.1 per cent in 1986 to 9.8 per in 2004. This is a sizeable reduction by international standards if we consider, for

instance, that in the United States the gender annual wage gap contracted by just one third over the same years, from 35.0 per cent to 23.4 per cent (National Committee on Pay and Equity, 2008).

Table 7 presents the results of the decomposition outlined in equation (4) for all five cross sections analysed in this study. The first three components were already discussed in the preceding section, but we include two additional terms: one reflecting the explained distribution of jobs by gender and, another representing the unexplained distribution of jobs due to occupational segregation. Therefore, the inclusion of the selection component implies that expression (4) could be reformulated as

$$\Delta W = \sum_{j=1}^{k} P_j^f \, \bar{X}_j^f \, \Delta\hat{\beta}_j + \sum_{j=1}^{k} P_j^f \, \hat{\beta}_j^m \, \Delta\bar{X}_j + \sum_{j-1}^{k} P_j^f \, [\hat{\theta}_j^m \, \bar{\lambda}_j^m - \hat{\theta}_j^m \bar{\lambda}_j^m] + \sum_{j-1}^{k} \bar{W}_j^m (\bar{P}_j^m - \bar{P}_j^*) + \sum_{j-1}^{k} \bar{W}_j^m (\bar{P}_j^* - \bar{P}_j^f)$$

(4*)

where the third element on the right hand side depicts the portion of the

wage gap explained by differences in selection.

Table 7. Decomposition of log wage gaps, urban Colombia, selected years

Components	*1986*	*1990*	*1995*	*2000*	*2004*
Treatments	0.227	0.708	0.346	0.381	0.291
	(0.073)**	(0.058)**	(0.064)**	(0.083)**	(0.063)**
Endowments	0.052	0.034	0.033	0.010	0.008
	(0.004)**	(0.003)**	(0.003)**	(0.004)*	(0.003)**
Selection	-0.022	-0.493	-0.158	-0.248	-0.156
	(0.009)*	(0.009)**	(0.011)**	(0.013)**	(0.013)**
Explained allocation of workers	-0.002	-0.002	-0.002	-0.002	-0.002
	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)
Segregation	-0.016	-0.013	-0.005	-0.043	-0.048
	(0.004)**	(0.004)**	(0.004)	(0.005)**	(0.004)**
Log wage gap	0.240	0.235	0.213	0.098	0.094
	(0.006)**	(0.006)**	(0.006)**	(0.008)**	(0.007)**

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia amongst full-time workers (>200 hours per month) aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %.

As anticipated in the analysis by occupations presented above, the treatments component dominates most of the evolution of the gender wage gap over all cross sections examined here. The decomposition results for 2004 indicate that the gender wage gap solely explained by differences in treatment is roughly 29 per cent, which is higher than the magnitude of the same component in 1986, around 23 per cent. This indicates that most of the gender wage gap in urban Colombia concedes a large unexplained component, as far as it can be inferred from the variables included in the model, and probably suggests that the extent of discrimination is still quite pervasive.

The effects of the unequal returns to human capital and other observable characteristics is counterbalanced by the effects of the selection component, which comes up negative in magnitude for all five cross sections and suggest that the rewards to unobservable characteristics are actually helping to reduce wage differences between men and women in urban Colombia. In fact, it could be argued that most of the reduction in the gender wage gap in this country between 1986 and 2004 is driven by selection effects, reaffirming our conclusion from the previous section that gender differences in returns to unobservable characteristics such as ability and motivation are actually playing a central role in the reduction of wage differences by gender in this country. If we set the selection effects to zero (see Gyourko and Tracy, 1988), the story would be less

promising: the gender wage offer gap would have changed from 29.9 per cent in 1986 to 28.4 per cent in 2004, a rather marginal reduction over 19 years. Ignoring selection effects implies that the comparison is performed using wage offers and implies that both, the returns to, and the level of unobservables are the same across gender groups so $\hat{\theta}_j^m \bar{\lambda}_j^m$ and $\hat{\theta}_j^f \bar{\lambda}_j^f$ in (4*) cancel each other. To reconcile this line of analysis with what the statistician gets from the actual labour market data as the mean wage gender gap, we need to emphasize that the wage offer estimates control for selection effects which are then set to zero. This differential provides an estimate of the expected wage gap in the labour market for a couple of men and women with average characteristics randomly drawn from the sample. This is a less intuitive approach in as much as the traditional comparison of the conventional wage offer gap is replaced by a less appealing measure of wage offers which are not readily observable in the data.

However, the comparison of both approaches provides some light in the understanding of the evolution of gender wage differences in urban Colombia and the paramount role of selection effects. If the evolution of hourly wage offers indicates almost no change in the magnitude of the wage offer gap but the dynamic of observed mean hourly wages points to a dramatic reduction, selection effects are actually playing a key role in this story, perhaps, by selecting the best

individuals in terms of unobservables. The female labour supply in urban Colombia is improving its average productivity over time through the accumulation of unobservable characteristics that enhances productivity and so, it manages to reduce the magnitude of the observed hourly wage gap by nearly two thirds over 19 years, from 1986 and 2004. Although conditions in the labour market have not improved as far as differences in treatment are concerned, the supply response has been sufficiently strong to compensate and actually reduce the extent of wage inequality by gender.

As in expression (4*), wage differences can also be due to differentials in endowments which reflect the effects of differences in human capital and other observable characteristics. This component has decreased its importance on the magnitude of the wage gap over the years covered in this study. This is, at least in part, the result of a better human capital improvement of the female labour force with respect to men commented on the literature of this country (Isaza, 2002, 2013; Isaza & Reilly, 2009). Therefore, the contribution of the endowment component to the overall gender wage gap fell from just above one fifth to less than one tenth between 1986 and 2004.

Regarding the unequal distribution of jobs by gender, our decomposition results indicate that the explained allocation of workers component is not statistically different from zero along all years

examined here. In contrast, the unexplained component originated as a result of occupational segregation is increasingly contributing to the reduction of the overall gender wage gap in urban Colombia. In other words, the unjustified portion of the dissimilar allocation of workers across the five occupational categories defined in this study (this is, the pure segregation component) has actually helped to reduce the wage disadvantage against women in the labour market of urban Colombia, in particular, in 2000 and 2004 when this helped by itself to reduce the hourly log wage gap by one half. Perhaps, these results are driven by a highly aggregated classification of occupations based on only five broad groupings in which the effect of a changing balance in the gender composition of more narrowly defined occupations is obscured. It must be noted that our segregation indicators presented above reveal little change between 1986 and 2004, while the set of estimates (not presented here) based on 82 occupations defined by the Colombian classification reveal a substantial decrease in occupational segregation over the same years.

FINAL REMARKS

This paper has found that controlling for selectivity bias is justified for both genders as opposed to the conventional probit approach using the Heckman procedure in which selection effects are frequently found unwarranted for the male subsample. Our findings suggests that a

finer characterization of the selection process into the labour market, either in terms of broad occupational categories or segments (i.e., formal/informal) may prove fruitful to better inform the nature of gender wage differences. We have adopted a multinomial logit approach as a departure point but other alternatives such as the ordered or multinomial probit may be more appropriate in cases where the categorization of different occupations is more clearly informed by hierarchical variables of productivity, prestige or social perceived status.

Our modelling strategy has provided interesting insights into the underlying forces behind a dramatic reduction of the overall wage gap in urban Colombia which, by international standards, is by itself a remarkable case. The decomposition results indicate that both, at the level of individual occupational categories and for the overall average for all categories, the unexplained component of the gender wage gap is the main force behind persistent levels of wage disadvantage against women in urban Colombia. This component is by itself in 2004 larger in absolute and relative terms compared to 1986. A decreasing portion of the wage gap is also explained by differences in endowments in favour of men, which are being reduced over time thanks to, most likely, better progress of women in terms of education compared to men.

Our decomposition results also suggest that selection effects are a key element in the history of the evolution of gender wage gaps in urban Colombia asthey appear to reduce the magnitude of labour income differences between women and men over all years. Although our modelling strategy renders direct comparison to other papers on this topic in Colombia, some authors have found selection effects as a key determinant of gender wage gaps in this country[10]. To some extent, this finding suggests that the composition of the female labour supply is changing in terms of unobservables in order to yield a net reduction in the overall wage gap. Probably, the pool of female workers entering the labour force is changing its composition as employers may be selecting the most endowed in terms of motivation, innate ability and other correlates of productivity not observed in our data. The effect of this process is not only to improve the quality of the female labour force as a whole but also to reduce the magnitude of the gender wage inequality. Probably as a matter for further research, it would be interesting to see with more complete datasets how the evolution of variables related to the level of motivation and intrinsic ability (such as scores of standardised tests) interact with labour market earnings and their influence on the magnitude of the gender wage gaps over time.

---

[10] For instance, Badel & Peña (2010), using a quintile regression approach arrive at a similar conclusion based on household survey data for 2006.

Quite surprisingly, our results indicate that the dissimilar allocation of workers actually helps to reduce the magnitude of the gender wage gap. We learned from our multinomial logit estimates that all segregation indices would be reduced by more than 90 per cent if women were treated as men in the selection process into the five occupational categories defined in this study. However, the decomposition results indicate that the wage gap would be higher in all years if women were distributed across the five occupational categories as implied by our counterfactual estimates. In other words, the way women and men are distributed across all five jobs groups actually helps to reduce, rather than increase the size of the wage gap over all years. This result is potentially affected by the highly aggregated nature of our occupational categorisation. We were not able to properly account for the dynamics of segregation levels in urban Colombia which are visible using a finer categorisation of occupations. Also, the effects of the limited access of women to high salaries positions largely dominated by men are also obscured by our gross categorisation of occupations. However, we believe that our application has its own merits in providing a better framework to correct for selectivity bias and simultaneously incorporate the effects of dissimilar job distributions by gender. There is clearly an opportunity for further research in the development of more efficient algorithms in the estimation of multiple choice models with a better capacity to model a large number of occupational outcomes.

Referencias

Acosta, C. E. & Isaza, J. G. (2006). Gender, poverty and labour market dynamics: a microsimulations analysis for urban Colombia between 1990 and 2004. *Social Science Research Network -SSRN*. Retrieved fromhttp://ssrn.com/abstract=1281545.

Anker, R. (1997). Theories of occupational segregation by sex: an overview. *International Labour Review, 136*, 315.

Appleton, S., Hoddinott, J. & Krishnan, P. (1999). The Gender Wage Gap in Three African Countries. *Economic Development and Cultural Change, 47*, 289-312.

Becker, G. (1971). *The economics of discrimination*. Chicago: University of Chicago Press.

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

Borghans, L. & Groot, L. (1999). Educational presorting and occupational segregation. *Labour Economics, 6*, 375-395.

Bourguignon, F., Fournier, M. & Gurgand, M. (2007). Selection bias corrections based on the multinomial logit model: Monte Carlo comparisons. *Journal of Economic Surveys, 21,* 174-205.

Brown, R. S., Moon, M. & Zoloth, B. S. (1980). Incorporating Occupational Attainment in Studies of Male-Female Earnings Differentials. *The Journal of Human Resources, 15,* 3-28.

DANE (2000). *Encuesta nacional de hogares, ciudades y areas metropolitanas, Manual de Crítica y Codificación, Etapa 108.* Bogotá: Dirección Técnica de Estadísticas Básicas, DANE.

DANE (2004). *Metodología de la encuesta de hogares.* Bogotá: Dirección Técnica de Estadísticas Básicas, DANE. Recuperado de http://suamox03.dane.gov.co:7778/imgcurso/mettran/indice.html

Duncan, O. D. & Duncan, B. (1955). A Methodological analysis of segregation indexes. *American Sociological Review, 20*, 210-217.

Edwards, S. (2001). *The economics and politics of transition to an open market economy: Colombia Paris.* Organisation for Economic Co-Operation and Development OECD.

Grazier, S. & Sloane, P. J. (2007). *Accident risk, gender, family status and occupational choice in the UK.* Labour Economics, In Press, Corrected Proof.

Gyourko, J. & Tracy, J. (1988). An analysis of public- and private-sector wages allowing for endogenous choices of both government and union status. *Journal of Labor Economics, 6,* 229-253.

Hakim, C. (1992). Explaining trends in occupational segregation: the measurement, causes, and consequences of the sexual division of labour. *European Sociological Review, 8,* 127-152.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47,* 153-161.

Isaza, J. G. (2002) Women workers in Bogota's informal sector: gendered impact of structural adjustment policies in the 1990s. *Planeación y Desarrollo 33*, 411-468.

Isaza, J. G. & Reilly, B. (2009). *Occupational segregation by gender: an empirical analysis for urban Colombia (1986-2004)*: Brighton, Department of Economics, University of Sussex.

Isaza, J. G. (2013). *Occupational segregation, gender wage differences and trade reforms: empirical applications for urban Columbia.* Doctoral thesis, University of Sussex. Retrieved from: http://sro.sussex.ac.uk/44798/

Karmel, T. & Maclachlan, M. (1988). Occupational sex segregation-increasing or decreasing? *Economic Record, 64*, 187.

Lee, L.-F. (1983). *Generalized econometric models with selectivity*. *Econometrica, 51,* 507-512.

Liu, P.-W., Zhang, J. & Chong, S.-C. (2004). Occupational segregation and wage differentials between natives and immigrants: evidence from Hong Kong. *Journal of Development Economics, 73,* 395-413.

Meng, X. & Miller, P. (1995). Occupational segregation and its impact on gender wage discrimination in China's rural industrial sector. *Oxford Economic Papers, 47.*

Miller, P. W. (1987). The wage effect of the occupational segregation of women in Britain. *The Economic Journal, 97*, 885-896.

Neuman, S. & Silber, J. G. (1996). Wage discrimination across ethnic groups: Evidence from Israel. *Economic Inquiry, 34*, 648.

Nunez, J., Ramírez, J. C. & Cuesta, L. (2005). Determinantes de la Pobreza en Colombia, 1996-2004. *Documento CEDE*, 1-45.

Núñez, J. & Sánchez, F. (2002). A dynamic analysis of households decision making in Colombia, 1976-1998: changes in household structure, human capital and its returns, and female labor force participation. *Documento CEDE*, 15, 1-50.

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

PNUD (2003). *El conflicto, callejón con salida (Informe Nacional de Desarrollo Humano Colombia - 2003).* Bogotá: Programa de Naciones Unidas para el Desarrollo.

Reilly, B. (1991). Occupational segregation and selectivity bias in occupational wage equations: an empirical analysis using Irish data. *Applied Economics, 23,* 1.

Ribero, R. (2003). Gender dimensions of non-formal employment in Colombia. *Documento CEDE,* 1-84.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48,* 817-838.

102

APPENDIX A

Table A1. Marginal/impact effects from multinomial logit
coefficients of occupational attainment, male subsample, 1986

	Professionals & managers	*Clerks & Sales*	*Services*	*Semi-skilled*	*Unskilled*
experience	0.001	0.007	0.003	0.007	0.009
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
experience2	0.000	0.000	0.000	0.000	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
primary education	0.007	-0.002	0.005	0.018	-0.011
	(0.002)**	(0.004)	(0.002)*	(0.004)**	(0.003)**
secondary education	0.070	0.059	-0.013	-0.039	-0.091
	(0.008)**	(0.006)**	(0.003)**	(0.004)**	(0.003)**
university education	0.493	-0.105	-0.050	-0.127	-0.144
	(0.033)**	(0.006)**	(0.002)**	(0.003)**	(0.002)**
single	0.000	0.007	-0.003	0.015	-0.011
	(0.002)	(0.009)	(0.005)	(0.008)	(0.006)
HH. head	0.005	0.064	0.032	0.053	0.029
	(0.001)**	(0.004)**	(0.002)**	(0.003)**	(0.003)**
HH. infants	0.000	0.004	0.003	0.005	0.007
	(0.000)**	(0.002)*	(0.001)**	(0.001)**	(0.001)**
HH. children	0.000	-0.007	-0.001	-0.001	0.002
	(0.000)**	(0.002)**	(0.001)	(0.001)	(0.001)*
HH. Unemployment rate	-0.024	-0.366	-0.089	-0.211	-0.173
	(0.003)**	(0.015)**	(0.008)**	(0.012)**	(0.011)**
domestic servant	0.002	-0.003	-0.003	-0.027	-0.008
	(0.001)*	(0.005)	(0.003)	(0.004)**	(0.004)
log wage segment	-0.001	-0.001	-0.001	-0.011	-0.005
	(0.000)**	(0.002)	(0.001)	(0.002)**	(0.002)*
education segment	0.001	0.006	-0.005	-0.009	-0.013
	(0.000)**	(0.001)**	(0.000)**	(0.001)**	(0.001)**
Barranquilla	-0.002	0.058	-0.013	-0.017	-0.011
	(0.001)**	(0.005)**	(0.002)**	(0.004)**	(0.004)**
Bucaramanga	0.004	0.030	-0.018	-0.009	-0.003
	(0.001)**	(0.006)**	(0.003)**	(0.005)	(0.004)
Manizales	-0.002	-0.038	-0.013	-0.024	-0.044
	(0.001)	(0.006)**	(0.003)**	(0.006)**	(0.004)**
Medellin	0.003	0.002	-0.008	0.032	0.000
	(0.001)**	(0.004)	(0.002)**	(0.004)**	(0.003)
Cali	0.000	0.011	0.004	0.014	0.008
	(0.001)	(0.005)*	(0.003)	(0.004)**	(0.004)*
Pasto	-0.008	-0.073	-0.026	-0.042	-0.014
	(0.001)**	(0.005)**	(0.003)**	(0.005)**	(0.005)**
Observations	2,959	10,566	4,188	9,905	10,548

Total observations: 89.433. Individuals out of the waged employment represent the base category of the multinomial logit.
Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for people aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %. (i) denotes impact effects.

Table A2. Marginal/impact effects from multinomial logit coefficients
of occupational attainment, female subsample, 1986

	Professionals & managers	*Clerks & Sales*	*Services*	*Semi-skilled*	*Unskilled*
experience	0.000	0.004	0.001	0.003	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
experience2	0.000	0.000	0.000	0.000	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
primary education	0.006	0.021	-0.032	0.015	0.003
	(0.001)**	(0.003)**	(0.001)**	(0.001)**	(0.001)**
secondary education	0.066	0.101	-0.074	-0.002	-0.001
	(0.016)**	(0.006)**	(0.001)**	(0.002)	(0.001)
university education	0.589	-0.029	-0.048	-0.022	-0.006
	(0.085)**	(0.011)**	(0.001)**	(0.002)**	(0.000)**
single	0.001	0.030	0.031	0.015	0.003
	(0.000)**	(0.004)**	(0.003)**	(0.002)**	(0.001)**
HH. head	0.001	0.049	0.035	0.023	0.006
	(0.000)**	(0.004)**	(0.003)**	(0.002)**	(0.001)**
HH. infants	-0.001	-0.013	-0.004	-0.003	-0.001
	(0.000)**	(0.001)**	(0.001)**	(0.001)**	(0.000)**
HH. children	0.000	-0.007	-0.002	-0.002	-0.001
	(0.000)**	(0.001)**	(0.001)**	(0.001)**	(0.000)**
HH. Unemployment rate	-0.008	-0.156	-0.070	-0.034	-0.010
	(0.001)**	(0.009)**	(0.005)**	(0.004)**	(0.002)**
domestic servant	0.000	-0.016	0.249	-0.009	-0.002
	(0.000)**	(0.002)**	(0.005)**	(0.001)**	(0.001)**
log wage segment	0.000	0.002	0.004	0.003	0.000
	(0.000)**	(0.001)	(0.001)**	(0.001)**	(0.000)**
education segment	0.000	-0.003	0.004	-0.003	-0.001
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Barranquilla	0.001	-0.003	-0.008	-0.010	-0.002
	(0.000)**	(0.003)	(0.002)**	(0.001)**	(0.001)**
Bucaramanga	0.001	0.010	-0.006	0.010	-0.004
	(0.000)**	(0.004)**	(0.002)**	(0.002)**	(0.000)**
Manizales	0.000	-0.037	-0.007	-0.016	-0.002
	(0.000)**	(0.003)**	(0.002)**	(0.002)**	(0.001)**
Medellin	0.002	-0.013	-0.007	0.010	-0.002
	(0.000)**	(0.002)**	(0.001)**	(0.002)**	(0.000)**
Cali	0.000	-0.004	-0.006	-0.002	-0.001
	(0.000)**	(0.003)	(0.002)**	(0.002)	(0.001)**
Pasto	-0.002	-0.001	-0.013	-0.015	-0.006
	(0.000)**	(0.004)	(0.002)**	(0.002)**	(0.000)**
Observations	1,439	7,155	7,994	3,066	802

Total observations: 107.102. Individuals out of waged employment represent the base category of the multinomial logit.
Sorce: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for people aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %. (i) denotes impact effects.

104

Table A3. Marginal/impact effects from multinomial logit coefficients
of occupational attainment, male subsample, 2004

	Professionals & managers	*Clerks & Sales*	*Services*	*Semi-skilled*	*Unskilled*
experience	0,001	0,007	0,003	0,008	0,010
	(0.000)***	(0.000)***	(0.000)***	(0.008)***	(0.000)***
experience2	0,000	0,000	0,000	0,000	0,000
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***
primaryeducation	0,007	-0,020	-0,002	0,018	0,007
	(0.005)	(0.006)***	(0.004)*	(0.018)***	(0.004)
secondaryeducation	0,044	0,015	0,021	-0,002	-0,055
	(0.006)***	(0.006)**	(0.004)***	(-0.002)***	(0.005)***
universityeducation	0,347	-0,056	-0,057	-0,107	-0,152
	(0.033)***	(0.008)***	(0.003)***	(-0.107)***	(0.003)***
single	0,001	0,010	-0,002	-0,005	0,001
	(0.002)	(0.007)	(0.004)*	(-0.005)***	(0.005)
HH. head	0,015	0,051	0,033	0,035	0,032
	(0.001)***	(0.004)***	(0.002)***	(0.035)***	(0.003)***
HH. infants	0,000	0,006	0,004	0,006	0,009
	(0.001)	(0.002)***	(0.001)***	(0.006)***	(0.001)***
HH. children	-0,002	-0,003	0,000	0,002	0,001
	(0.001)***	(0.002)**	(0.001)*	(0.002)***	(0.001)
HH. Unemploymentrate	-0,045	-0,320	-0,113	-0,186	-0,170
	(0.005)***	(0.014)***	(0.009)***	(-0.186)**	(0.011)***
domesticservant	0,008	-0,026	0,003	-0,007	0,000
	(0.002)***	(0.005)***	(0.003)*	(-0.007)***	(0.004)
log wagesegment	-0,001	-0,003	0,001	-0,001	0,000
	(0.000)***	(0.002)**	(0.001)*	(-0.001)***	(0.001)
educationsegment	0,003	0,001	-0,005	-0,009	-0,013
	(0.000)***	(0.001)	(0.000)***	(-0.009)***	(0.001)***
Barranquilla	0,001	0,045	0,007	0,020	0,011
	(0.002)	(0.006)***	(0.004)**	(0.02)***	(0.005)**
Bucaramanga	-0,003	0,000	-0,012	0,014	0,003
	(0.002)*	(0.006)	(0.003)***	(0.014)***	(0.005)
Manizales	-0,010	-0,024	-0,002	0,026	-0,017
	(0.002)***	(0.005)***	(0.004)*	(0.026)***	(0.005)***
Medellin	0,009	-0,007	-0,009	0,030	-0,003
	(0.002)***	(0.005)	(0.003)***	(0.03)***	(0.005)
Cali	0,002	0,003	-0,002	0,015	0,003
	(0.002)	(0.006)	(0.003)*	(0.015)***	(0.005)
Pasto	-0,013	-0,042	-0,018	-0,009	-0,005
	(0.001)***	(0.005)***	(0.003)***	(-0.009)***	(0.005)
Observations	2,905	7,615	3,856	6,550	8,285

Total observations: 68.833. Individuals out of waged employment represent the base category of the multinomial logit.
Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for people aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %. (i) denotes impact effects.

Table A4. Marginal/impact effects from multinomial logit coefficients
of occupational attainment, female subsample, 2004

	Professionals & managers	Clerks & Sales	Services	Semi-skilled	Unskilled
experience	0,000	0,004	0,004	0,003	0,000
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***
experience2	0,000	0,000	0,000	0,000	0,000
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***
primaryeducation	0,008	-0,004	-0,016	0,017	0,001
	(0.004)*	(0.005)	(0.002)***	(0.002)***	(0.001)
secondaryeducation	0,035	0,079	-0,050	0,011	0,001
	(0.007)***	(0.006)***	(0.003)***	(0.002)***	(0.001)
universityeducation	0,387	0,006	-0,081	-0,019	-0,008
	(0.073)***	(0.013)	(0.001)***	(0.002)***	(-0.008)***
single	0,000	0,011	0,030	0,007	0,003
	(0.001)	(0.004)***	(0.003)***	(0.002)***	(0.003)***
HH. head	0,003	0,020	0,031	0,013	0,004
	(0.001)***	(0.003)***	(0.003)***	(0.002)***	(0.004)***
HH. infants	-0,001	-0,008	-0,006	-0,002	0,000
	(0.000)***	(0.001)***	(0.001)***	(0.001)***	(0.000)***
HH. children	-0,001	-0,004	-0,005	-0,001	0,000
	(0.000)***	(0.001)***	(0.001)***	(0.001)	(0.000)***
HH. Unemploymentrate	-0,023	-0,250	-0,107	-0,041	-0,017
	(0.002)***	(0.009)***	(0.007)***	(0.004)***	(-0.017)***
domesticservant	-0,001	-0,037	0,220	-0,015	-0,004
	(0.000)***	(0.003)***	(0.005)***	(0.001)***	(-0.004)***
log wagesegment	0,000	0,001	0,001	0,002	0,000
	(0.000)***	(0.001)	(0.001)	(0.001)***	(0.000)***
educationsegment	0,001	-0,003	0,001	-0,003	-0,001
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(-0.001)***
Barranquilla	0,006	-0,010	-0,005	-0,003	-0,004
	(0.001)***	(0.003)***	(0.003)*	(0.002)*	(-0.004)***
Bucaramanga	0,000	-0,005	-0,008	0,010	-0,002
	(0.001)	(0.004)	(0.003)***	(0.002)***	(-0.002)**
Manizales	-0,001	-0,034	-0,015	-0,004	-0,004
	(0.001)	(0.003)***	(0.002)***	(0.002)**	(-0.004)***
Medellin	0,006	-0,008	-0,007	0,020	-0,003
	(0.001)***	(0.003)***	(0.002)***	(0.002)***	(-0.003)***
Cali	0,000	-0,002	-0,006	-0,004	0,001
	(0.001)	(0.004)	(0.003)**	(0.002)**	(0.001)
Pasto	-0,002	-0,002	-0,012	-0,016	-0,006
	(0.001)***	(0.004)	(0.003)***	(0.001)***	(-0.006)***
Observations	2,102	6,895	7,046	2,564	773

Total observations: 84.246. Individuals out of waged employment represent the base category of the multinomial logit.
Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for people aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %. (i) denotes impact effects.

106

APPENDIX B

Table B1. Wage equations corrected for selectivity bias, male subsample, 1986

	Professionals & managers	*Clerks & Sales*	*Services*	*Semi-skilled*	*Unskilled workers*
age	0.067	0.050	0.040	0.065	0.060
	(0.010)**	(0.004)**	(0.006)**	(0.004)**	(0.004)**
age squared	-0.001	0.000	0.000	-0.001	-0.001
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
<=11 years education	0.018	0.060	0.049	0.055	0.039
	(0.012)	(0.003)**	(0.003)**	(0.002)**	(0.003)**
>11 years education	0.022	0.141	0.112	0.103	0.030
	(0.011)	(0.007)**	(0.016)**	(0.012)**	(0.022)
Barranquilla	0.015	-0.132	0.056	-0.031	-0.051
	(0.039)	(0.021)**	(0.03)	(0.019)	(0.018)**
Bucaramanga	0.009	-0.060	0.037	-0.040	0.010
	(0.040)	(0.027)*	(0.034)	(0.025)	(0.021)
Manizales	-0.064	-0.160	-0.237	-0.242	-0.237
	(0.052)	(0.037)**	(0.040)**	(0.033)**	(0.033)**
Medellin	-0.091	-0.022	0.064	-0.009	0.041
	(0.038)*	(0.019)	(0.022)**	(0.016)	(0.015)**
Cali	-0.053	-0.051	0.016	-0.013	-0.011
	(0.036)	(0.019)**	(0.022)	(0.018)	(0.016)
Pasto	-0.276	-0.246	-0.070	-0.461	-0.417
	(0.070)**	(0.039)**	(0.041)	(0.034)**	(0.029)**
Government	-0.099	0.029	0.121	0.221	0.163
	(0.028)**	(0.021)	(0.018)**	(0.025)**	(0.021)**
Self employed	-0.166	-0.248	-0.227	-0.191	-0.078
	(0.035)**	(0.017)**	(0.044)**	(0.018)**	(0.015)**
Second quarter	0.096	0.054	0.020	0.040	0.076
	(0.034)**	(0.018)**	(0.022)	(0.016)*	(0.015)**
Third quarter	0.132	0.116	0.063	0.057	0.092
	(0.035)**	(0.019)**	(0.023)**	(0.017)**	(0.015)**
Fourth quarter	0.163	0.120	0.084	0.075	0.103
	(0.034)**	(0.018)**	(0.022)**	(0.016)**	(0.015)**
φ/F	-0.377	-0.636	-0.026	-0.007	0.191
	(0.041)**	(0.040)**	-0.054	-0.041	(0.036)**
Constant	4.334	4.072	3.379	2.955	2.804
	(0.297)**	(0.130)**	(0.171)**	(0.115)**	(0.100)**
Observations	2,959	10,566	4,188	9,905	10,548
R-squared	0.264	0.296	0.172	0.192	0.143

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for full time workers (200 or more hours per month) aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %.

Table B2. Wage equations corrected for selectivity bias, female subsample, 1986

	Professionals & managers	Clerks & Sales	Services	Semi-skilled	Unskilled workers
age	0.068	0.062	0.025	0.029	0.049
	(0.012)**	(0.006)**	(0.004)**	(0.009)**	(0.015)**
age squared	-0.001	-0.001	0.000	0.000	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)*	(0.000)*
<=11 years education	0.017	0.072	0.037	0.038	0.040
	(0.018)	(0.004)**	(0.003)**	(0.004)**	(0.008)**
>11 years education	0.091	0.081	0.093	0.142	0.123
	(0.011)**	(0.006)**	(0.027)**	(0.026)**	(0.059)*
Barranquilla	-0.138	0.038	-0.012	-0.063	-0.028
	(0.045)**	(0.023)	(0.026)	(0.041)	(0.054)
Bucaramanga	-0.110	-0.086	-0.157	-0.134	-0.276
	(0.051)*	(0.025)**	(0.026)**	(0.038)**	(0.125)*
Manizales	-0.143	-0.152	-0.281	-0.466	-0.184
	(0.066)*	(0.039)**	(0.040)**	(0.086)**	(0.077)*
Medellin	-0.023	0.014	-0.005	-0.023	-0.023
	(0.042)	(0.019)	(0.021)	(0.024)	(0.045)
Cali	0.030	-0.036	-0.043	-0.030	-0.036
	(0.048)	(0.021)	(0.021)*	(0.032)	(0.057)
Pasto	-0.465	-0.508	-0.411	-0.536	-0.999
	(0.181)*	(0.042)**	(0.029)**	(0.096)**	(0.346)**
Government	0.024	0.204	0.324	0.481	0.153
	(0.033)	(0.018)**	(0.024)**	(0.086)**	(0.096)
Self employed	-0.055	-0.451	-0.363	-0.348	-0.152
	(0.068)	(0.029)**	(0.034)**	(0.042)**	(0.134)
Second quarter	0.090	0.053	0.112	0.009	-0.004
	(0.043)*	(0.020)**	(0.022)**	(0.028)	(0.047)
Third quarter	0.056	0.082	0.094	0.017	0.002
	(0.040)	(0.020)**	(0.021)**	(0.028)	(0.045)
Fourth quarter	0.166	0.098	0.057	0.032	0.061
	(0.041)**	(0.019)**	(0.020)**	(0.027)	(0.050)
φ/F	-0.105	-0.096	-0.001	0.010	-0.029
	(0.047)*	(0.039)*	-0.013	-0.046	-0.101
Constant	3.457	2.860	3.603	3.611	3.308
	(0.362)**	(0.146)**	(0.072)**	(0.208)**	(0.366)**
Observations	1,439	7,155	7,994	3,066	802
R-squared	0.264	0.344	0.110	0.154	0.148

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for full time workers (200 or more hours per month) aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %.

Table B3. Wage equations corrected for selectivity bias, male subsample, 1990

	Professionals & managers	*Clerks & Sales*	*Services*	*Semi-skilled*	*Unskilled workers*
age	0.044	0.048	0.039	0.050	0.055
	(0.009)**	(0.004)**	(0.005)**	(0.004)**	(0.003)**
age squared	0.000	0.000	0.000	0.000	-0.001
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
<=11 years education	0.007	0.060	0.038	0.050	0.030
	(0.011)	(0.003)**	(0.003)**	(0.002)**	(0.002)**
>11 years education	-0.013	0.134	0.077	0.087	0.043
	(0.010)	(0.006)**	(0.020)**	(0.010)**	(0.015)**
Barranquilla	-0.136	-0.107	-0.007	-0.038	-0.076
	(0.032)**	(0.019)**	(0.022)	(0.016)*	(0.016)**
Bucaramanga	-0.043	0.035	0.008	0.023	0.007
	(0.042)	(0.023)	(0.030)	(0.020)	(0.018)
Manizales	0.042	-0.025	-0.052	-0.086	-0.101
	(0.058)	(0.033)	(0.029)	(0.024)**	(0.022)**
Medellin	-0.031	0.034	0.131	0.079	0.056
	(0.033)	(0.019)	(0.022)**	(0.016)**	(0.015)**
Cali	-0.086	0.060	0.040	0.023	0.041
	(0.037)*	(0.022)**	(0.024)	(0.019)	(0.017)*
Pasto	-0.145	-0.245	-0.282	-0.510	-0.532
	(0.061)*	(0.036)**	(0.039)**	(0.036)**	(0.029)**
Government	-0.112	-0.008	0.173	0.181	0.172
	(0.024)**	(0.019)	(0.016)**	(0.024)**	(0.018)**
Self employed	-0.173	-0.176	-0.217	-0.169	-0.069
	(0.035)**	(0.016)**	(0.040)**	(0.016)**	(0.014)**
Second quarter	0.039	0.043	0.085	0.057	0.054
	(0.031)	(0.016)**	(0.020)**	(0.015)**	(0.014)**
Third quarter	0.119	0.076	0.076	0.097	0.081
	(0.032)**	(0.017)**	(0.021)**	(0.015)**	(0.014)**
Fourth quarter	0.108	0.129	0.084	0.102	0.103
	(0.031)**	(0.017)**	(0.019)**	(0.015)**	(0.014)**
φ/F	-0.530	-0.584	0.209	-0.019	0.230
	(0.036)**	(0.040)**	(0.055)**	-0.038	(0.031)**
Constant	6.258	4.824	3.892	4.093	3.824
	(0.258)**	(0.124)**	(0.161)**	(0.107)**	(0.080)**
Observations	3,365	10,294	4,183	9,976	10,758
R-squared	0.269	0.301	0.180	0.203	0.152

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for full time workers (200 or more hours per month) aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %.

Table B4. Wage equations corrected for selectivity bias, female subsample, 1990

	Professionals & managers	Clerks & Sales	Services	Semi-skilled	Unskilled workers
age	0.023	0.047	0.035	0.028	0.017
	(0.012	(0.005)**	(0.004)**	(0.007)**	(0.017)
age squared	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)**	(0.000)**	(0.000)**	(0.000)
<=11 years education	0.010	0.073	0.031	0.033	0.028
	(0.018)	(0.004)**	(0.003)**	(0.003)**	(0.006)**
>11 years education	0.057	0.067	0.060	0.063	0.123
	(0.012)**	(0.005)**	(0.024)*	(0.022)**	(0.042)**
Barranquilla	-0.146	0.016	-0.377	-0.021	0.018
	(0.038)**	(0.019)	(0.020)**	(0.026)	(0.063)
Bucaramanga	-0.050	-0.001	-0.187	-0.059	-0.067
	(0.046)	(0.022)	(0.022)**	(0.032)	(0.068)
Manizales	0.039	-0.064	-0.223	-0.053	-0.044
	(0.069)	(0.032)*	(0.030)**	(0.037)	(0.057)
Medellin	0.075	0.026	0.062	0.087	0.107
	(0.040)	(0.017)	(0.018)**	(0.022)**	(0.046)*
Cali	0.034	0.012	-0.020	0.002	0.050
	(0.058)	(0.021)	(0.022)	(0.029)	(0.058)
Pasto	-0.119	-0.539	-0.471	-0.641	-0.647
	(0.073)	(0.036)**	(0.028)**	(0.095)**	(0.182)**
Government	0.016	0.204	0.243	0.286	0.320
	(0.031)	(0.017)**	(0.022)**	(0.094)**	(0.107)**
Self employed	0.095	-0.295	-0.416	-0.301	-0.195
	(0.049)*	(0.025)**	(0.028)**	(0.031)**	(0.103)
Second quarter	0.095	-0.001	0.028	0.039	0.103
	(0.037)*	(0.017)	(0.017)	(0.022)	(0.035)**
Third quarter	0.056	0.030	0.030	0.074	0.035
	(0.041)	(0.017)	(0.018)	(0.023)**	(0.052)
Fourth quarter	0.190	0.077	0.095	0.065	0.108
	(0.038)**	(0.017)**	(0.017)**	(0.023)**	(0.045)*
φ/F	-0.267	-0.043	0.085	0.084	-0.028
	(0.042)**	-0.032	(0.012)**	(0.036)*	-0.143
Constant	5.623	3.913	4.286	4.308	4.704
	(0.350)**	(0.122)**	(0.064)**	(0.160)**	(0.532)**
Observations	1,680	7,603	7,280	3,229	852
R-squared	0.241	0.328	0.229	0.166	0.142

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for full time workers (200 or more hours per month) aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %.

Table B5. Wage equations corrected for selectivity bias, male subsample, 1995

	Professionals & managers	*Clerks & Sales*	*Services*	*Semi-skilled*	*Unskilled workers*
age	0.050	0.032	0.060	0.053	0.042
	(0.010)**	(0.004)**	(0.005)**	(0.004)**	(0.003)**
age squared	0.000	0.000	-0.001	-0.001	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
<=11 years education	-0.010	0.052	0.052	0.049	0.036
	(0.011)	(0.003)**	(0.003)**	(0.002)**	(0.002)**
>11 years education	0.020	0.166	0.126	0.099	0.050
	(0.012)	(0.006)**	(0.017)**	(0.011)**	(0.014)**
Barranquilla	-0.159	-0.232	-0.166	-0.189	-0.252
	(0.035)**	(0.020)**	(0.024)**	(0.018)**	(0.016)**
Bucaramanga	-0.057	0.089	-0.061	-0.018	0.003
	(0.044)	(0.023)**	(0.030)*	(0.020)	(0.016)
Manizales	-0.009	-0.217	-0.272	-0.207	-0.272
	(0.081)	(0.033)**	(0.034)**	(0.025)**	(0.024)**
Medellin	-0.061	-0.018	-0.031	-0.058	-0.055
	(0.038)	(0.020)	(0.023)	(0.017)**	(0.015)**
Cali	0.014	0.069	0.001	0.014	0.042
	(0.042)	(0.023)**	(0.026)	(0.020)	(0.017)*
Pasto	-0.118	-0.206	-0.290	-0.472	-0.502
	(0.048)*	(0.033)**	(0.045)**	(0.029)**	(0.023)**
Government	-0.163	0.047	0.166	0.214	0.165
	(0.029)**	(0.024)*	(0.022)**	(0.030)**	(0.025)**
Self employed	-0.149	-0.137	-0.120	-0.094	0.018
	(0.036)**	(0.017)**	(0.036)**	(0.016)**	(0.012)
Second quarter	0.035	0.073	0.027	0.035	0.033
	(0.035)	(0.018)**	(0.021)	(0.015)*	(0.013)*
Third quarter	0.140	0.048	0.039	0.029	0.045
	(0.036)**	(0.018)**	(0.021)	(0.016)	(0.013)**
Fourth quarter	0.085	0.067	0.039	0.019	0.048
	(0.033)*	(0.018)**	(0.021)	(0.015)	(0.013)**
φ/F	-0.536	-0.625	-0.030	0.071	0.125
	(0.042)**	(0.045)**	-0.064	-0.043	(0.032)**
$\bar{\text{Constant}}$	7.314	6.372	5.045	5.140	5.378
	(0.279)**	(0.130)**	(0.167)**	(0.123)**	(0.079)**
Observations	3,477	10,560	4,568	9,620	12,273
R-squared	0.283	0.285	0.201	0.188	0.140

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for full time workers (200 or more hours per month) aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %.

Table B6. Wage equations corrected for selectivity bias, female subsample, 1995

	Professionals & managers	Clerks & Sales	Services	Semi-skilled	Unskilled workers
age	0.039	0.045	0.028	0.021	0.055
	(0.011)**	(0.004)**	(0.004)**	(0.006)**	(0.019)**
age squared	0.000	0.000	0.000	0.000	-0.001
	(0.000)*	(0.000)**	(0.000)**	(0.000)*	(0.000)*
<=11 years education	-0.012	0.076	0.036	0.030	0.021
	(0.019)	(0.004)**	(0.003)**	(0.003)**	(0.007)**
>11 years education	0.073	0.107	0.132	0.118	0.093
	(0.012)**	(0.005)**	(0.016)**	(0.023)**	(0.039)*
Barranquilla	-0.162	-0.063	-0.290	-0.120	-0.210
	(0.041)**	(0.020)**	(0.018)**	(0.032)**	(0.052)**
Bucaramanga	0.019	-0.051	-0.067	-0.100	-0.212
	(0.052)	(0.022)*	(0.022)**	(0.027)**	(0.061)**
Manizales	-0.064	-0.228	-0.286	-0.143	-0.210
	(0.072)	(0.031)**	(0.027)**	(0.035)**	(0.083)*
Medellin	-0.045	0.005	-0.057	-0.009	-0.096
	(0.039)	(0.017)	(0.018)**	(0.020)	(0.042)*
Cali	-0.041	0.038	-0.050	0.000	-0.023
	(0.044)	(0.019)	(0.021)*	(0.028)	(0.042)
Pasto	-0.181	-0.542	-0.482	-0.666	-0.742
	(0.080)*	(0.030)**	(0.028)**	(0.081)**	(0.158)**
Government	0.018	0.215	0.299	0.274	0.320
	(0.030)	(0.019)**	(0.023)**	(0.172)	(0.106)**
Self employed	-0.015	-0.240	-0.313	-0.106	-0.088
	(0.049)	(0.023)**	(0.022)**	(0.028)**	(0.081)
Second quarter	0.023	0.005	0.045	0.042	-0.006
	(0.039)	(0.017)	(0.017)**	(0.022)	(0.040)
Third quarter	0.090	0.050	0.031	0.038	-0.034
	(0.039)*	(0.017)**	(0.017)	(0.021)	(0.040)
Fourth quarter	0.050	0.054	0.057	0.036	-0.049
	(0.037)	(0.016)**	(0.017)**	(0.021)	(0.039)
φ/F	-0.283	0.008	0.075	0.105	0.211
	(0.044)**	-0.032	(0.013)**	(0.040)**	(0.101)*
Constant	6.824	4.939	5.536	5.588	4.907
	(0.345)**	(0.114)**	(0.066)**	(0.157)**	(0.417)**
Observations	2,126	9,147	7,344	3,575	961
R-squared	0.230	0.293	0.237	0.139	0.132

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for full time workers (200 or more hours per month) aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %.

Table B7. Wage equations corrected for selectivity bias, male subsample, 2000

	Professionals & managers	Clerks & Sales	Services	Semi-skilled	Unskilled workers
age	0.063	0.041	0.059	0.053	0.056
	(0.019)**	(0.006)**	(0.009)**	(0.006)**	(0.005)**
age squared	-0.001	0.000	-0.001	-0.001	-0.001
	(0.000)*	(0.000)**	(0.000)**	(0.000)**	(0.000)**
<=11 years education	0.006	0.063	0.046	0.059	0.045
	(0.016)	(0.003)**	(0.005)**	(0.004)**	(0.003)**
>11 years education	0.057	0.127	0.113	0.099	0.043
	(0.011)**	(0.007)**	(0.019)**	(0.012)**	(0.013)**
Barranquilla	-0.130	-0.031	-0.006	-0.041	-0.076
	(0.063)*	(0.026)	(0.035)	(0.028)	(0.025)**
Bucaramanga	-0.088	0.059	0.106	-0.067	0.007
	(0.062)	(0.029)*	(0.043)*	(0.032)*	(0.026)
Manizales	-0.106	-0.071	-0.125	-0.187	-0.167
	(0.068)	(0.037)	(0.049)*	(0.038)**	(0.029)**
Medellin	-0.043	0.003	0.016	0.005	-0.054
	(0.065)	(0.030)	(0.039)	(0.028)	(0.026)*
Cali	-0.012	0.019	-0.062	-0.058	-0.102
	(0.064)	(0.029)	(0.043)	(0.031)	(0.027)**
Pasto	-0.202	-0.139	-0.208	-0.354	-0.333
	(0.063)**	(0.031)**	(0.040)**	(0.030)**	(0.028)**
Government	0.066	0.165	0.322	0.157	0.373
	(0.037)	(0.051)**	(0.034)**	(0.111)	(0.044)**
Self employed	-0.301	-0.357	-0.290	-0.288	-0.254
	(0.045)**	(0.020)**	(0.034)**	(0.021)**	(0.015)**
Second quarter	-0.033	-0.054	-0.030	-0.003	-0.027
	(0.044)	(0.022)*	(0.032)	(0.024)	(0.020)
Third quarter	-0.006	-0.020	0.001	-0.023	-0.006
	(0.049)	(0.022)	(0.031)	(0.024)	(0.019)
Fourth quarter	-0.042	-0.034	-0.001	-0.034	-0.020
	(0.045)	(0.022)	(0.029)	(0.026)	(0.019)
φ/F	-0.367	-0.509	-0.359	-0.041	0.112
	(0.052)**	(0.050)**	(0.100)**	-0.068	(0.046)*
Constant	7.143	6.592	6.224	5.786	5.562
	(0.450)**	(0.167)**	(0.305)**	(0.198)**	(0.140)**
Observations	2,619	7,962	3,915	6,898	8,255
R-squared	0.229	0.306	0.188	0.163	0.159

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for full time workers (200 or more hours per month) aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %.

Table B8. Wage equations corrected for selectivity bias, female subsample, 2000

	Professionals & managers	Clerks & Sales	Services	Semi-skilled	Unskilled workers
age	0.043	0.046	0.027	0.049	0.062
	(0.014)**	(0.005)**	(0.006)**	(0.009)**	(0.023)**
age squared	0.000	0.000	0.000	-0.001	-0.001
	(0.000)	(0.000)**	(0.000)**	(0.000)**	(0.000)*
<=11 years education	0.039	0.082	0.035	0.030	0.051
	(0.023)	(0.005)**	(0.003)**	(0.005)**	(0.011)**
>11 years education	0.061	0.089	0.070	0.082	0.110
	(0.013)**	(0.005)**	(0.016)**	(0.024)**	(0.032)**
Barranquilla	-0.082	0.015	-0.018	0.023	0.024
	(0.056)	(0.025)	(0.034)	(0.046)	(0.074)
Bucaramanga	-0.184	-0.052	0.030	-0.164	-0.102
	(0.056)**	(0.025)*	(0.037)	(0.042)**	(0.088)
Manizales	-0.163	-0.129	-0.120	-0.157	0.062
	(0.059)**	(0.029)**	(0.041)**	(0.048)**	(0.076)
Medellin	-0.058	-0.013	0.019	-0.003	0.010
	(0.061)	(0.026)	(0.041)	(0.035)	(0.070)
Cali	-0.127	-0.071	0.064	-0.062	-0.034
	(0.061)*	(0.027)**	(0.038)	(0.045)	(0.074)
Pasto	-0.264	-0.301	-0.270	-0.385	-0.538
	(0.066)**	(0.028)**	(0.038)**	(0.066)**	(0.167)**
Government	0.059	0.300	0.486	0.277	0.545
	(0.038)	(0.030)**	(0.033)**	(0.118)*	(0.149)**
Self employed	-0.155	-0.545	-0.535	-0.462	-0.217
	(0.060)*	(0.027)**	(0.026)**	(0.039)**	(0.089)*
Second quarter	0.014	0.006	-0.048	-0.048	-0.178
	(0.047)	(0.020)	(0.024)*	(0.032)	(0.060)**
Third quarter	0.002	-0.002	-0.007	0.005	-0.007
	(0.049)	(0.021)	(0.023)	(0.032)	(0.065)
Fourth quarter	-0.001	0.020	-0.059	-0.051	-0.133
	(0.051)	(0.020)	(0.025)*	(0.033)	(0.057)*
φ/F	-0.308	-0.032	0.058	0.000	-0.010
	(0.049)**	-0.037	(0.018)**	-0.054	-0.142
Constant	6.926	5.534	6.154	5.956	5.573
	(0.419)**	(0.142)**	(0.114)**	(0.246)**	(0.640)**
Observations	1,850	7,234	7,273	2,469	618
R-squared	0.234	0.377	0.135	0.208	0.228

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for full time workers (200 or more hours per month) aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %.

Table B9. Wage equations corrected for selectivity bias, male subsample, 2004

	Professionals & managers	*Clerks & Sales*	*Services*	*Semi- skilled*	*Unskilled workers*
age	0.008	0.034	0.057	0.044	0.052
	(0.01	(0.005)**	(0.006)**	(0.006)**	(0.005)**
age squared	0.000	0.000	-0.001	0.000	-0.001
	(0.000)	(0.000)**	(0.000)**	(0.000)**	(0.000)**
<=11 years education	-0.013	0.056	0.040	0.049	0.041
	(0.013	(0.003)**	(0.003)**	(0.003)**	(0.003)**
>11 years education	0.044	0.124	0.090	0.119	0.047
	(0.008)**	(0.006)**	(0.012)**	(0.011)**	(0.011)**
Barranquilla	-0.240	-0.079	0.136	0.004	-0.053
	(0.042)**	(0.028)**	(0.029)**	(0.030)	(0.025)*
Bucaramanga	0.002	0.051	0.102	0.052	0.047
	(0.052)	(0.029)	(0.032)**	(0.028)	(0.023)*
Manizales	-0.026	-0.112	0.006	-0.098	-0.134
	(0.050)	(0.031)**	(0.030)	(0.028)**	(0.025)**
Medellin	-0.049	0.024	0.125	0.044	0.081
	(0.045)	(0.029)	(0.031)**	(0.028)	(0.023)**
Cali	-0.094	-0.005	0.080	0.014	0.032
	(0.051)	(0.030)	(0.031)**	(0.032)	(0.026)
Pasto	-0.099	0.003	-0.034	-0.190	-0.182
	(0.048)*	(0.033)	(0.036)	(0.030)**	(0.024)**
Government	-0.014	0.322	0.426	0.353	0.325
	(0.036)	(0.044)**	(0.023)**	(0.057)**	(0.051)**
Self employed	-0.291	-0.344	-0.414	-0.368	-0.245
	(0.034)**	(0.019)**	(0.032)**	(0.020)**	(0.014)**
Second quarter	0.036	0.039	0.067	0.020	0.006
	(0.037)	(0.021)	(0.022)**	(0.021)	(0.018)
Third quarter	0.023	0.061	0.104	0.029	0.065
	(0.038)	(0.021)**	(0.024)**	(0.021)	(0.017)**
Fourth quarter	0.045	0.037	0.072	0.044	0.061
	(0.036)	(0.021)	(0.023)**	(0.020)*	(0.017)**
φ/F	-0.631	-0.629	-0.119	-0.045	0.106
	(0.040)**	(0.060)**	(0.054)*	-0.056	(0.037)**
Constant	9.017	7.123	6.004	6.140	5.781
	(0.307)**	(0.167)**	(0.189)**	(0.177)**	(0.117)**
Observations	2,905	7,615	3,856	6,550	8,285
R-squared	0.377	0.316	0.314	0.211	0.161

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for full time workers (200 or more hours per month) aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %.

Table B10. Wage equations corrected for selectivity bias, female subsample, 2004

	Professionals & managers	Clerks & Sales	Services	Semi-skilled	Unskilled workers
age	0.040	0.038	0.029	0.039	0.071
	(0.012)**	(0.005)**	(0.004)**	(0.010)**	(0.019)**
age squared	0.000	0.000	0.000	0.000	-0.001
	(0.000)*	(0.000)**	(0.000)**	(0.000)**	(0.000)**
<=11 years education	0.006	0.073	0.028	0.031	0.045
	(0.022)	(0.006)**	(0.003)**	(0.004)**	(0.009)**
>11 years education	0.075	0.096	0.100	0.080	0.056
	(0.011)**	(0.005)**	(0.012)**	(0.017)**	(0.030)
Barranquilla	-0.373	0.080	-0.039	0.023	0.047
	(0.049)**	(0.028)**	(0.025)	(0.048)	(0.071)
Bucaramanga	-0.208	0.013	-0.062	-0.170	-0.156
	(0.055)**	(0.027)	(0.026)*	(0.044)**	(0.073)*
Manizales	-0.144	-0.072	-0.091	-0.016	-0.113
	(0.052)**	(0.028)**	(0.026)**	(0.041)	(0.078)
Medellin	-0.112	0.051	0.057	0.109	-0.040
	(0.050)*	(0.024)*	(0.023)*	(0.039)**	(0.073)
Cali	-0.135	-0.041	-0.053	0.026	-0.027
	(0.070)	(0.029)	(0.026)*	(0.046)	(0.072)
Pasto	-0.200	-0.183	-0.387	-0.165	-0.348
	(0.060)**	(0.029)**	(0.026)**	(0.070)*	(0.106)**
Government	0.024	0.341	0.491	0.102	0.363
	(0.038	(0.036)**	(0.035)**	(0.085)	(0.180)*
Self employed	-0.276	-0.562	-0.583	-0.546	-0.644
	(0.048)**	(0.029)**	(0.025)**	(0.040)**	(0.089)**
Second quarter	-0.024	0.047	0.060	-0.001	0.045
	(0.038)	(0.021)*	(0.019)**	(0.032)	(0.066)
Third quarter	-0.026	0.049	0.034	-0.013	0.051
	(0.041)	(0.021)*	(0.019)	(0.031)	(0.063)
Fourth quarter	0.004	0.022	0.066	-0.012	-0.003
	(0.040)	(0.021)	(0.019)**	(0.032)	(0.062)
φ/F	-0.359	-0.081	0.003	0.050	0.256
	(0.044)**	-0.043	-0.015	-0.055	-0.132
Constant	7.729	5.996	6.471	6.158	4.919
	(0.368)**	(0.159)**	(0.084)**	(0.259)**	(0.503)**
Observations	2,102	6,895	7,046	2,564	773
R-squared	0.330	0.366	0.231	0.205	0.223

Source: Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for full time workers (200 or more hours per month) aged 18 and 65 years. Robust standard errors in parentheses. * significant at 5 %; ** significant at 1 %.