Occupational Segregation, Gender Wage Differences and Trade Reforms

Empirical Applications for Urban Colombia

JAIRO GUILLERMO ISAZA CASTRO

In partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Economics

University of Sussex

September 2012

Statement

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature: …………………………………………………………..

Table of Contents

List of Tables

List of Figures

Summary

Occupational Segregation, Gender Wage Differences and Trade Reforms:

Empirical Applications for Urban Colombia

DPhil Thesis

Jairo Guillermo Isaza Castro

Email: jgisaza@gmail.com and jisaza@unisalle.edu.co

Department of Economics, University of Sussex

This DPhil thesis comprises three empirical essays that survey the evolution of gender differences in the labour market of urban Colombia since the 1980s. The first essay examines the evolution of gender segregation using occupational indices between 1986 and 2004, and presents a decomposition of their changes over time using a technique proposed by Deutsch *et al.* (2006). We find that a substantial proportion of the reduction in segregation indices is driven by changes in both the employment structure of occupations and the increasing participation of female labour observed over these years. The second essay assesses the effects of occupational segregation on the gender wage gap in urban Colombia between 1984 and 1999. The empirical strategy involves the estimation of a counterfactual distribution of female workers across occupations, as if they had been treated the same as their male counterparts. This provides a basis to formulate a decomposition of the gender wage gap in which the explained and unexplained portions of the gender distribution of jobs are explicitly incorporated. The results indicate that the unequal distribution of women and men across occupations actually helps, on average, to reduce gender pay differences in urban Colombia, particularly in the 'informal' segment where the labour income differential between women and men is the largest. The third and final essay examines the effects of trade liberalisation on the gender composition of employment across manufacturing industries in urban Colombia from 1981 to 2000. The empirical strategy involves a comparison of estimates drawn from different panel data techniques. As a main finding, we verify that increasing trade flows are associated with higher proportions of female employment.

Keywords: Occupational Segregation, Gender Wage Gap, Shapley Decomposition, Oaxaca-Blinder Decomposition, Fixed-Effects Instrumental Variables, Colombia

Introduction

This thesis comprises three empirical essays on the evolution of gender differences in the labour market using more than two decades of data from urban Colombia. In this sense, these data provide the basis to assess how long-term social and economic trends regarding the differentiated situation of women in the labour market have evolved in a semi-industrialised economy. The discussion revolves around three key gender aspects in the Colombian labour market, namely, (i) occupational segregation, (ii) gender wage differences, and (iii) female employment allocation across industries. These three issues constitute the backbone of the empirical chapters comprising this thesis.

In addition to a brief description of the chapters, we also provide in this introduction a succinct portrayal of the country and the main data sources used within the thesis. The main statistical source for the analyses presented in Chapters 1 and 2 are the microdata drawn from the National Household Survey, which was gathered on a quarterly basis between 1984 and 2000.[1] The household survey data have already been extensively used in empirical research on Colombia. The primary use of this information by the Colombian government is the measurement of employment levels in the main metropolitan areas of the country in order to provide statistically representative estimates for the urban population based on a stratified clustered multi-stage sampling design (DANE, 2004). These data are also used by many international agencies, including the International Labour Organisation, the Economic Commission for Latin America and the Caribbean, the United Nations and the World Bank, to provide estimates of labour market participation rates and welfare indicators for the urban population of this country. In order to facilitate comparisons over time, we

[1] After this year, household surveys are gathered permanently (not quarterly) using a different survey design.

restricted the sample to those cities regularly surveyed in all quarterly waves from the mid 1980s to 2000. These are represented by the seven metropolitan areas of Bogotá, Medellin, Cali, Barranquilla, Bucaramanga, Manizales and Pasto.

In Chapter 3 we use data from the Annual Manufacturing Survey collected by the Colombian Statistical Bureau (DANE, from its acronym in Spanish) as a census amongst all firms with more than ten employees, as well as firms with fewer than ten workers but with a production value above a given threshold. This survey is used for multiple purposes including national accounts, monitoring the performance of manufacturing industries in this country, and has also been widely used in applied econometric research (Cfr. Eslava et al., 2009, Roberts and Skoufias, 1997, Roberts and Tybout, 1997). For the purposes of this research, data from this survey can only be grouped and compared across the same ISIC Rev.2 codes between 1981 and 2000. Although it reports employment data disaggregated by gender and skill level, the survey does not provide information on labour costs (or wages) for men and women separately. In Chapter 3 we also use tariffs and trade data from the National Planning Department.

We now turn attention to some background information about urban Colombia. According to the *Human Development Report* for 2006, Colombia is a medium human development country with a life expectancy of about 73 years, an adult literacy rate close to 93 per cent (UNDP, 2006), and a GDP per capita of US$5,682 at purchasing power parity.[2] The country has experienced an improvement in most of its development indicators over the last decades. Its GDP per capita grew at an annual rate of 1.3 per cent between 1986 and 2004, although it is still below the average for Latin

[2] According to World Bank, World Development Indicators (WDI) October 2008, ESDS International (Mimas), University of Manchester (Last Access: 13 October 2008).

America and the Caribbean.[3] The human development index for urban areas in this country rose from 0.774 in 1991 to 0.794 in 2001 (PNUD, 2003) and by 2004 access to drinking water and sanitation was above 96 per cent of the urban population of the country.[4] Similarly, the percentage of the population with incomes below the national poverty line in the seven main cities of Colombia decreased from 59.9 per cent in 1990 to 48.0 per cent in 2004 with extreme poverty falling from 14.6 per cent to 12.0 per cent over the same period (Isaza et al., 2010).

The urban population in the seven main metropolitan areas of Colombia has exhibited a substantial demographic change over the last decades due to a strong decline in fertility rates and growth in life expectancy. Consequently, the composition of the labour market has also witnessed some marked changes in the seven main metropolitan areas of this country. While the proportion of those of working age by Colombian standards (12 years old and more) grew from 73.2 to 77.7 per cent between 1986 and 2004, such an increase in absolute numbers meant an addition of 4.7 million people, equivalent to an increment of 62.4 per cent in the number of potential workers.[5]

Colombia has undertaken an intensive process of market-oriented reforms since 1990, comprising a comprehensive package of trade liberalisation policies as well as a major restructuring of government functions, including privatisations and decentralisation of government functions and resources towards provinces (or *departamentos*) and municipalities (see Edwards, 2001 for a comprehensive review). The process of

[3] According to World Bank (*Ibid*), this is 70.1 per cent of the average for Latin America and the Caribbean in 2004.

[4] World Bank, (*Op. Cit.*).

[5] Population estimates based on household survey microdata for the seven largest metropolitan areas.

economic reform in Colombia has been accompanied by some progressive developments with an emphasis on the incorporation of gender and women's issues in Colombian legislation. The constitutional reform of 1991 established an inclusive policy of women in decision-making positions within public administration (Art.41), an explicit mandate to guarantee equal rights and opportunities for both gender groups, and the obligation of the State to assist and protect women in vulnerable situations (e.g., those in unemployment or acting as household heads (Article 43)). The 1991 constitution also endorsed the enforcement of international conventions on labour and their provisions in regard to gender equality (Article 53).

As indicated above, the main body of the research undertaken for this thesis is contained within three empirical chapters, which are now briefly outlined in turn.

In Chapter 1, we examine the evolution of gender segregation indices by occupation in the urban labour markets of Colombia between 1986 and 2004. For this purpose, we implement three different measures of occupational segregation for several sub-groups of the labour force in terms of age, schooling levels, sector of employment (government vs. private sector), and segment of employment (formal vs. informal). In addition to the conventional and widely used Duncan and Duncan (1955) dissimilarity index, we compute other measures of horizontal occupational segregation by gender, comprising the Gini coefficient based on the distribution of jobs by gender (see Deutsch et al., 1994) and the Karmel & MacLachlan (1988) index of labour market segregation. However, the analysis of segregation measures by occupation over time is subject to a number of methodological difficulties. On the one hand, segregation indices are sensitive to the number of occupations used in their computation, so the finer the classification of occupations, the higher the corresponding index value. On the other hand, absolute difference measures such as the Duncan and Duncan (1955) dissimilarity index are sensitive to changes in both female labour force participation and the structure of occupations. In order to address these issues, we implement a decomposition technique of segregation indices proposed by Deutsch *et al*. (2006) in which the effect of changes in 'net segregation', that is changes in the share of women within particular occupations, is separated from changes in 'gross segregation' in both, the gender composition of the overall labour force and the structure of occupations.

Although this study is focused on just one country, it exploits the advantage of having compatible data from household micro-level data surveys covering 19 years on 82 occupational groupings. We find that a substantial proportion of the reduction in segregation indices for this country is driven by changes in both the employment structure of occupations and the increasing female labour participation observed over these years, while changes in the gender composition of occupations have favoured mainly government employees and those with university education.

Chapter 2 is devoted to the analysis of the effects of occupational segregation on the gender wage gap in urban Colombia where improvements across both dimensions of gender inequality have been observed since the mid-1980s. In particular, we investigate whether female occupational intensity can be related to lower wages and whether the segregated nature of the distribution of jobs by gender explains some part of the gender pay gap in this country. On the one hand, our empirical strategy involves the estimation of a counterfactual distribution of female employment once the decision to participate in the labour market has been taken. We do this with a multinomial logit model in which the dependent variable is categorical in nature and comprises 23 occupation categories in the formal and 16 in the informal sector. Using the multinomial logit coefficients for the male subsample, we estimate a counterfactual distribution of female jobs (in the hypothetical scenario they were treated in the labour market in the same way as their male counterparts for the purposes of occupational allocation) to identify which portion of the wage gap can be attributed to both the explained and unexplained components of occupational segregation within an Oaxaca-Blinder decomposition framework. On the other hand, we also calculate log wage equations in which the percentage of female workers in a given occupation is included as an explanatory variable in addition to controls for occupation fixed effects. We find that the effects of occupational segregation on the magnitude of gender wage differences are modest, at the same time most of the unexplained portion of the gender wage gap is attributable to the presumable discriminatory treatment of human capital characteristics such as formal schooling and years of potential labour force experience. We find also that female occupation intensity is associated with lower wages for women in the formal sector across all years reviewed in this chapter, a result that is in line with the empirical findings from similar applications in other countries.

In Chapter 3 we exploit a natural experiment provided by the trade liberalisation that occurred in Colombia at the beginning of the 1990s to see its possible effects on the gender composition of the workforce across manufacturing industries. In order to account for the effects of changes in capital technology, our empirical strategy controls for different types of capital stock per worker (namely, machinery, office equipment and transport equipment) within a fixed-effects instrumental-variables framework in which estimates drawn from a variety of instruments are compared.[6] We also include a concentration index variable in order to account for changes in the degree of market power in order to assess Becker's hypothesis in relation to labour market discrimination, according to which increasing competition should erode monopolistic rents and reduce costly discrimination against women in the labour market. Our findings confirm that increasing levels of trade openness in the terms of both import penetration and export orientation tend to be associated with higher shares of female

[6] Estimates are also compared with dynamic panel coefficients based on the Generalised Method of Moments (GMM) proposed by Arellano and Bover (1995) and Blundell and Bond (1998).

employment although this effect appears to be differentiated in terms of skill level. Equally we find that manufacturing industries with higher levels of industry concentration tend to have lower female shares of jobs. Our variables for different types of the stock of capital per worker suggest that machinery and office equipment are associated with higher shares of female jobs, particularly in the white-collar workers category.

Acknowledgements

First and foremost, I am very grateful to my supervisor, Professor Barry Reilly for his valuable guidance and patience over the years of this process. His numerous reviews on previous drafts of the chapters included here are a real testimony of his teaching and research excellence. His teaching excellence in econometrics, well known in Sussex University and elsewhere, also inspired me to undertake the empirical approach that dominates most of this doctoral thesis. Other people have also contributed to this work. My special thanks to Alexander Cotte Poveda for his guidance on the implementation of GMM procedure in Stata for the econometric results presented in the Appendix 3.2 of Chapter 3. On the same chapter, I am also grateful for the comments received from Professor Richard Dickens. I want also to express my gratitude to the Department of Economics, in particular to Professor Andy McKay, Head of the Department, and Andrew Newell, former Head of the Department, for making every effort to improve the experience of DPhil students and providing us with plenty of opportunities to present our DPhil research work, particularly outside the UK. In this sense, I want to thank the comments to an earlier version of Chapter 1 by the participants at the Guanajuato Workshop for Young Economist, held at Universidad de Guanajuato, Mexico, on August 2010. My sincere thanks also to participants at the Third Alban

Conference, celebrated in Porto, Portugal on June, 2009, for their comments to one of the first versions of Chapter 2. This chapter also benefited from valuable suggestions that I received at the Economics DPhil Conference, organised by the Department of Economics, University of Sussex, on November 2009. I want to thank the comments received on an early draft of the Third Chapter presented at IV Symposium on Labour Studies, organised by the Department of Economics at Universidad de La Salle in Bogotá, Colombia, on April 2012.

I acknowledge the financial support from Universidad de La Salle in Bogotá, Colombia for providing me with study leave to pursue doctoral studies in the United Kingdom. My special thanks to Hno. Carlos Gabriel Gómez Restrepo, Rector at Universidad de La Salle, for his unconditional support to my family and myself over my stay in the United Kingdom. His understanding of my personal circumstances deserves a special recognition which well exceeds what can be stated here. I am equally grateful for the sponsorship from the Programme Alβan, the European Union Programme of High Level Scholarships for Latin America, scholarship No.E06D100938CO.

Finally I want to dedicate this thesis to Nidia, my wife and the love of my life. She gave me not only the strength and encouragement to lift me up in the most difficult moments of my life. She also gave me two precious kids, Tomás and Samuel, to whom this thesis is also dedicated to. This thesis is also dedicated to my mother, Inés, who always believed on my and, to my father, Jairo Isaza for teaching me how to be an honest man.

Despite the valuable contributions of people mentioned above, I remain solely responsible for the contents of this doctoral thesis including its mistakes.

Chapter 1: Occupational Segregation by Gender –An Empirical Analysis for Urban Colombia (1986-2004)

1.1 Introduction

Gender discrimination in the labour market has several dimensions. The more widely studied is the gender wage gap itself while others, such as occupational gender segregation, have merited less attention in the empirical literature. This may be explained by methodological problems arising from the appropriate choice of the occupational aggregation level, as well as changes to the classifications of occupations over time. Despite these difficulties, differences in the pattern of jobs performed by men and women and their evolution over recent decades are still an important issue in the study of labour markets.

The existing literature suggests large and persistent gender differences in the distribution of jobs typically performed by men and women in all regions of the world although, the degree of occupational horizontal segregation by gender has exhibited a substantial decrease in recent decades (Deutsch et al., 2002, Tzannatos, 1999, Baunach, 2002, Anker et al., 2003, Semyonov and Jones, 1999). There is less agreement, however, on how to measure occupational segregation. It has been found that the Duncan and Duncan (1955) dissimilarity index and other absolute difference measures are sensitive to the number of occupations used in their computation (see Melkas and Anker (1997)). Another problem with conventional measures of segregation is that they are influenced by increases in the number of men and women entering the labour force and by the extent of female labour force participation. Blackburn and Harman (2005) found that in some developed countries, such as Sweden and Finland, high levels of occupational gender based segregation co-exist with high degrees of gender equality and low levels of the gender wage gaps. As argued by Semyonov and Jones (1999), in the gender analysis of occupations, nominal (or horizontal) segregation as measured by dissimilarity indexes is conceptually different from occupational inequality (or vertical segregation) and may be influenced in a different way by the labour market structure and the level of socio-economic development. From a statistical point of view, all of this suggests that measures of occupational segregation are sensitive not only to changes in female labour participation but also to changes in the structure of occupations. This is problematic from a policy analysis perspective, since changes in segregation indices over time may not be entirely explained by changes in the gender composition of particular occupations.

This chapter is devoted to enhancing understanding about the evolution of horizontal gender based occupational segregation over time through an empirical application using data from urban areas of Colombia over the period 1986 to 2004. In addition to the conventional Duncan and Duncan (1955) dissimilarity index, this chapter presents other measures of horizontal occupational segregation by gender comprising the Gini coefficient based on the distribution of jobs by gender (see Deutsch et al., 1994)) and the Karmel & MacLachlan (1988) index of labour market segregation. In order to address some of the biases mentioned above on segregation measures, we implement a decomposition technique proposed by Deutsch et al (2006) in which in the effect of changes in 'net segregation', this is changes in the share of women within particular occupations, is separated from changes in 'gross segregation' in both the gender composition of the overall labour force and the structure of occupations. Although this study is focused on only one country, it exploits the advantage of having compatible data from household micro-level data surveys covering 19 years on 82 occupational groupings. The remainder of the chapter is organized as follows. The next section presents a review of the existing literature on gender segregation in the labour market and its measurement. The third describes the data while the fourth provides some contextual background on Colombia. The fifth section reports the empirical results using three different measures of horizontal gender-based occupational segregation in urban Colombia and presents an analytical decomposition of their changes between 1986 and 2004. The final section offers some concluding remarks.

1.2. Literature Review

1.2.1 Gender-based occupational segregation: some basic concepts

A precise definition of occupational gender segregation should distinguish between three overlapping concepts: exposure, concentration and segregation (Blackburn and Jarman, 2005). Exposure is related to the degree of social contact and interaction that one gender group has with those from the other in the labour sphere. A high degree of occupational segregation by gender implies that male workers enjoy a low exposure to women. Concentration relates to the composition of the labour force by gender and is measured in one or more occupations. By definition, concentration can only be equal for men and women in the case that both gender groups are equally represented in absolute numbers. Segregation relates to the existence of a differentiated pattern of occupations predominantly performed by either women or men. Gender-based occupational segregation is clearly linked to gender inequality in the labour market. In this context, horizontal and vertical dimensions should be distinguished. Semyonov and Jones (1999) suggest that horizontal and vertical segregation should be interpreted as two different theoretical concepts. Based on data from a cross-sectional analysis of 56 countries, they conclude that the structural characteristics of the labour market affect both dimensions of gender segregation in different ways. For instance, while increasing female labour participation tends to be associated with lower levels of horizontal segregation, they find that in those countries where women comprise a large proportion of the labour force their access to 'high-status' occupations appears more restricted.

Blackburn and Jarman (2005) note the paradoxical case for some developed countries (e.g., Sweden and Finland) of high levels of horizontal segregation by gender co-existing with high degrees of gender equality and small gender pay gaps. In short, they explain that although women and men enjoy equal access to education and training opportunities, female career paths tend to specialize in female dominated jobs where their access to managerial positions is higher. In this way, high levels of horizontal segregation may be possible with high levels of gender equality in terms of gender pay gaps and female representation in managerial positions.[7]

1.2.2 Measuring occupational segregation: methodological issues

The dissimilarity index (hereafter, *DI*) is the most popular measure of horizontal occupational segregation in the literature (Anker et al., 2003, Anker and Melkas, 1997, Blackburn and Jarman, 2005, Mulekar et al., 2007, Silber, 1989, Karmel and MacLachlan, 1988). It was originally proposed by Duncan and Duncan (1955) to analyse the degree of geographical segregation of non-white communities in the United

[7] For an earlier discussion about gender occupational segregation in Nordic countries, see MELKAS, H. & ANKER, R. 1997. Occupational segregation by sex in Nordic countries: An emprical investigation. *International Labour Review,* 136.

States. The Duncan and Duncan or dissimilarity index, *DI*, is defined by the following formula:

$$DI = \frac{1}{2}\sum_{i=1}^{n}\left|\frac{F_i}{F} - \frac{M_i}{M}\right|, \qquad i = 1,2,\ldots,n \qquad (1.1)$$

where n is the number of occupations, F_i and M_i are the number of female and male workers in occupation i respectively, and F and M refer to the total number of female and male workers. This measure may be interpreted as the percentage of women and/or men who have to move to different occupations (activities) in order to generate a completely even distribution of jobs by gender group.

Despite its popularity, the *DI* has some methodological weaknesses. In particular, the index is sensitive to the number of categories used in its computation (Blackburn et al., 2001). For instance, the *DI* will increase, *ceteris paribus*, with the number of employment occupations. This entails obvious difficulties in trying to compare the degree of occupational segregation based on a crude measure of the *DI* across countries with different classification systems of occupations or in the case of time series analyses when a given classification system incorporates new occupations. One way in which this problem is addressed in the literature consists of limiting the computation to a small number of categories. For example, in a cross-sectional comparison of employment segregation by gender, Tzannatos (1999) uses six economic activities for the 61 countries included in the analysis, while Semyonov and Jones (1999) deploy seven major occupational categories to compare 56 countries. If the data are highly disaggregated more sophisticated procedures have been suggested by, among others, Blackburn et al. (2001), Blackburn and Jarman (2005) and Anker (2003).

Another caveat with the *DI* is that it equally weights each occupation regardless of its share in total employment (Silber, 1989: 239). Alternative measures have been suggested to incorporate in a more adequate way the heterogeneity of the occupations' relative weights by the use of concepts developed from the income inequality literature (see Silber (1989) for a detailed discussion). Specifically, these measures take advantage of the fact that the *DI* was developed originally from the concept of the segregation curve which, in the case of gender occupational segregation, is a graphical representation of the cumulative proportions of female and male workers in each occupation. The segregation curve is analogous to the Lorenz curve in the income distribution literature. A number of measures have been formulated but, in the current chapter, we use a Gini coefficient based on the distribution of jobs by gender. Formally, the Gini coefficient of the distribution of jobs by gender is given by the following expression:

$$GI = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{M_i}{M}\frac{M_j}{M}\left|\frac{F_i/M_i - F_j/M_j}{F/M}\right| \qquad (1.2)$$

where M_i and F_i are defined as in (1.1). It should be noted that because the weights used in the computation of *GI* are implicitly the shares of each occupation in total female employment, it represents a weighted relative mean of deviations of the male/female ratios from an average gender distribution of jobs within occupations. It follows that because the *DI* is a simple average of mean deviations from occupational gender ratios, *GI* and *DI* should yield similar results (Deutsch *et al*, 1994: 134). However, *GI* has the advantage of being less sensitive to changes in the weights of different occupations over time.

An additional problem with the *DI* relates to the practical feasibility of its interpretation. In the hypothetical scenario that the female (or male) labour force were re-distributed as suggested by the index, it would mean a change in the underlying structure of the labour force, either in terms of occupations or economic activities. In order to address this problem, Karmel and MacLachlan (1988) have formulated an index of the proportion of people required to change job in order to obtain the same distribution of jobs for men and women without altering the underlying occupational structure. This index may be expressed as:

$$KM = \sum_{i=1}^{n} \left| a\frac{M_i}{T} - (1-a)\frac{F_i}{T} \right| \tag{1.3}$$

where a $(=F/(M+F))$ represents the female participation in the labour force and $T = M + F$. This index dominates the traditional *DI* expressed in (1) because it takes into account that men and women have different participation rates in the labour force (Deutsch et al., 2002: 22). Thus, the *KM* index is less sensitive to changes in female labour force participation which is typically increasing over time and has the potential of biasing downwards the conventional *DI*. It is possible to derive an alternative measure of *KM* in which the female labour participation may be held constant over time so the changes in horizontal gender occupational segregation between different periods may be netted out from changes in the overall gender composition of the labour force. Assume two periods of time $t=1,2$ and their corresponding shares of female employment such that $a_1 < a_2$. Assume also that their corresponding indexes for the two periods are such as $KM_1 > KM_2$. Then, we have:

$$KM_1 = \sum_{i=1}^{n} \left| a_1\frac{M_i}{T} - (1-a_1)\frac{F_i}{T} \right| \tag{1.3a}$$

represents the index for t=1 and

$$KM_2 = \sum_{i=1}^{n} \left| a_2\frac{M_i}{T} - (1-a_2)\frac{F_i}{T} \right| \tag{1.3b}$$

is the corresponding index for t=2. We may also estimate an alternative segregation measure, *KM**, for t=2 in which the share of female employment is held constant at the level of period 1 such as

$$KM_2^* = \sum_{i=1}^{n} \left| a_1 \frac{M_i}{T} - (1-a_1)\frac{F_i}{T} \right|. \qquad (1.3c)$$

Thus, the total differential of employment segregation between t=1 and t=2 as measured by KM would be

$$\Delta = KM_2 - KM_1 \qquad (1.4)$$

and the differential net of changes in female labour force participation would be

$$\Delta^* = KM_2^* - KM_2 . \qquad (1.5)$$

Therefore, controlling for changes in female labour force participation makes the *KM* index amenable to inter-temporal decompositions in which changes in the level of female labour force participation may be an important factor in the evolution of occupational segregation. [8]

One of the methodological difficulties in the measurement of gender occupational segregation relates to the comparability of different classifications under which the data on occupations are collected over time. Even if an occupational classification remains unaltered over a long period of time, comparisons between different estimates of the same segregation measure for two or more periods are uncertain without reference to their variability. A similar concern applies when judging differences in dissimilarity indexes for different socio-demographic groups within the same population. Deutsch *et al.* (1994) suggest a bootstrap technique to compute standard errors and confidence intervals for the segregation measures (see also Deutsch et al., 2002). The technique consists of drawing a number of random samples (i.e., 500) with

[8] For computational convenience, it should be noted that $KM_2 = 2a_2(1-a_2)DI_2$ (see Karmel and MacLachlan, 1988: 189). Then it follows that $KM^*_2 = 2a_1(1-a_1)DI_2$ where a_1 is the female labour force participation rate for period 1.

replacement from the original sample for each year to compute for every sample a corresponding segregation measure. Subsequently, the distribution of bootstrapped segregation measures is used to compute relevant confidence intervals. In this empirical application, we implement this technique in order to assess differences between different groups of the labour force in terms of age, education and labour market segment (i.e., formal and informal workers). For this purpose, we draw 500 samples of size equal to the original sample for every one of the years included in this study to obtain standard errors and 99 per cent confidence intervals. This enables statistical inference about differences in segregation measures both over time and between the particular labour force groups outlined above.

1.3. Data

The data used are derived from household surveys gathered in the seven main metropolitan areas of Colombia on a quarterly basis between 1986 and 2000 and on a monthly basis between 2001 and 2004. These cities represent around 36 per cent of the national population and almost one-half the country's urban inhabitants. These surveys provide micro-level data on more than 100,000 individuals within the labour force (aged between 15 and 65 years) per year and include information about occupations using a consistent classification of 82 categories over the entire period (see Appendix 1.1). At the two-digit level, it is identical to the International Standard Classification of Occupations ISCO-68. The Colombian classification of occupations was created by the National Learning Service –SENA and the International Labour Organisation in 1968 (DANE, 2000).

1.4. Background

As explained in the introductory section of this thesis, Colombia has undertaken an intensive process of market-oriented reforms since 1990. This process of economic reform in Colombia has also been accompanied by some progressive developments with an emphasis on the incorporation of gender and women's issues into Colombian legislation. The introduction of market-oriented reforms in Colombia initiated in 1990 ushered in a restructuring of the state through decentralization of state functions, privatisations and the introduction of private enterprises for the provision of social services. We do not aim formally to provide any conjectures about the effects of those reforms on the overall size of government employment. However, our data suggest that the number of people working for the government in urban Colombia has decreased either in absolute numbers or relative to total employment between 1986 and 2004. The number of government employees contracted in most of the years after 1991 when the reforms were initiated, while its share of total employment in urban Colombia fell from 11.7 per cent to 6.3 per cent over this period. By gender, the reductions in government employment affected mainly the male labour force while women increased participation in the public sector from 41 per cent of all government jobs in 1986 to around 50 per cent after 2000.[9] To some extent, these results suggest that the constitutional reforms implemented in Colombia after 1991 designed to enforce an inclusive policy for women at all levels of public administration of this country have led to a more egalitarian composition of government employment by gender. However, they also suggest that retrenchment in the public sector has hit hardest on male employment in urban Colombia, probably as a result of austere fiscal policies and/or institutional reforms.

---

[9] Ibid.

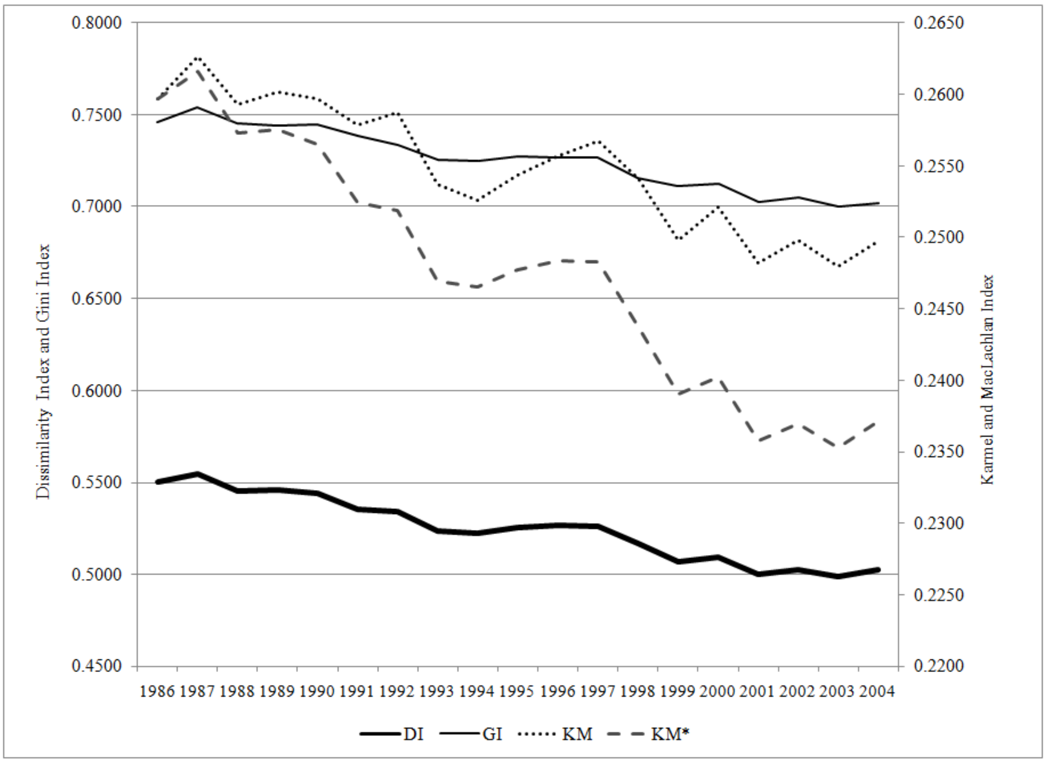
1.5. Empirical results

1.5.1 Occupational dissimilarity indices by gender in urban Colombia, 1986-2004

Horizontal gender-based occupational segregation has exhibited a marked decline in Colombia between 1986 and 2004. During this period, the *DI* for the entire labour market decreased 8.7 per cent during this period, while the *GI* and *KM* estimates contracted by 5.9 and 3.8 per cent, respectively. We also computed *KM** (see expression (1.3c)) in order to generate a counterfactual outcome for *KM* in which female labour force participation is held constant at the level of 1986 over the whole period. The results not only confirm a reduction in occupational segregation but also suggest that in holding female labour force participation constant at the 1986 level for all years, the extent of gender occupational segregation would be lower than that suggested by the original *KM* index (see Figure 1.1). This finding may be regarded as counterintuitive but it may simply reflect the fact that the increasing share of women into the labour force requires a larger proportion of people to move from jobs in order to have the same distribution of occupations across gender groups (see section 1.4, above).

The 99 per cent confidence intervals constructed through the bootstrap technique indicate that all segregation measures for 2004 are statistically different from those based on estimates for 1986 with negligible standard errors in all cases. For instance, the estimate for the *DI* for 2004 for all workers (.4999) lies outside the corresponding confidence interval for the same index in 1986 (0.5501 and 0.5506), which allows us to infer that the index in 2004 is significantly lower than that observed in 1986. The same consideration is valid for differences in segregation indicators estimated across different groups of the labour force (see Table 1.1).[10] It must be noted that the degree of association between the three dissimilarity measures is very high as is generally the case in the literature, and reveals similar patterns of change for most of the years.[11]

Figure 1.1: Indices of occupational segregation by gender in urban Colombia, total labour force, 1986-2004



Source: own calculations based on household surveys micro-data for seven main metropolitan areas.

In order to establish whether the pattern observed above is valid for all groups in the labour force, dissimilarity indicators were estimated separately for the formal and

[10] Estimates of bootstrapped standard errors and confidence intervals of segregation indices for all years are available from the authors.

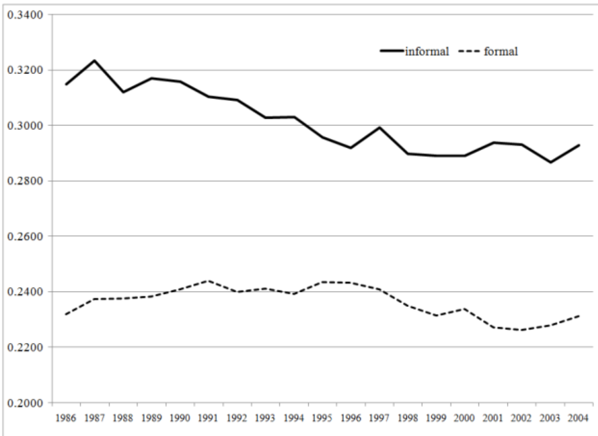
[11] The correlation coefficients between *DI* and *GI* and *KM* are 0.999 and 0.953, while the coefficient between *KM* and *GI* is .965.

informal segments of the labour market, by educational level, selected age groups and, government and private sector workers (see Table 1.1). In the first case, we defined the formal segment as comprising waged workers and skilled self-employed workers and consigned to the informal (or atypical employment) segment of the labour market all other workers (i.e., unskilled self-employed workers, family workers without remuneration and domestic servants). According to the *KM* index, horizontal occupational segregation by gender is highest in the informal sector in all observed years (see panel a in Figure 1.2). However, the same estimates also reveal that the extent of horizontal gender occupational segregation has decreased in both the formal and informal segments of the labour market in urban Colombia. According to the *KM* index, the reduction of the latter is 6.8 per cent compared to a contraction of about 4.1 per cent in the former between 1986 and 2004.

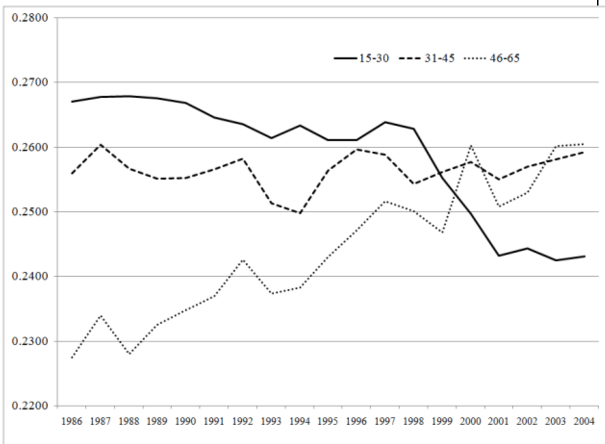
We also investigate the effects of demographic structure by dividing the labour force into three different age groups: 15 to 30 years old (the youngest group), 31 to 45 years old (the middle-age group) and, 46 to 65 years old (the oldest group). According to the *KM* index for 1986-2004, occupational segregation has decreased mainly amongst the youngest workers (15 to 30 years old) while exhibiting a substantial increase amongst the oldest. In the first case the reduction was about 10 per cent while in the second it grew 6.2 per cent. Those in the mid-range of age (31 to 45 years old) recorded a slight reduction (0.7 per cent) over this period. As in the cases documented above, the differences between age groups are statistically significant using the bootstrapped standard errors. Overall, trends by age groups indicate a reduction in the dispersion of segregation levels and a clear reduction amongst the youngest workers in particular (see Panel b in Figure 1.2).

Figure 1.2: Karmel and MacLachlan Index of occupational segregation by gender in urban Colombia, selected groups of the labour force, 1986-2004

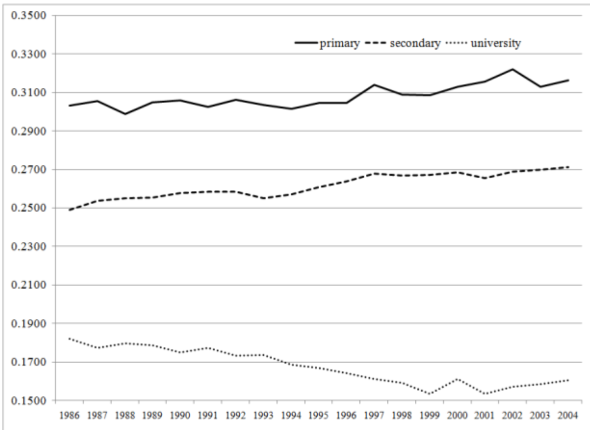
a) Informal and formal employment



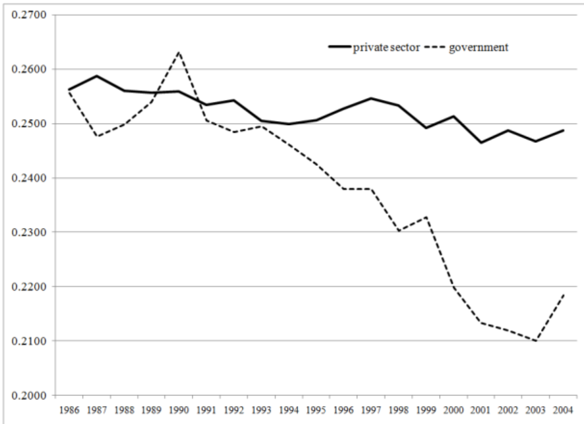
b) Age groups



c) Educational levels



d) Government and private sector



Source: own estimates based on household surveys micro-data for seven main metropolitan areas.

Table 1.1: Measures of dissimilarity in the distribution of occupations by gender and groups of the labour force in Urban Colombia: 1986 and 2004

Groups of the labour force	Year	Dissimilarity Index Mean	Standard Error	[99% Confidence Interval]		Gini Mean	Standard Error	[99% Confidence Interval]		Karmel and MacLachlan Mean	Standard Error	[99% Confidence Interval]	
All workers	1986	0.5504	0.0001	0.5501	0.5506	0.7458	0.0001	0.7456	0.7460	0.2597	0.0001	0.2595	0.2598
	2004	0.4994	0.0001	0.4991	0.4997	0.7011	0.0001	0.7009	0.7014	0.2448	0.0001	0.2446	0.2449
Informal workers	1986	0.6305	0.0002	0.6300	0.6310	0.8139	0.0001	0.8136	0.8143	0.3150	0.0001	0.3148	0.3152
	2004	0.5891	0.0002	0.5886	0.5896	0.7937	0.0003	0.7930	0.7943	0.2936	0.0001	0.2934	0.2939
Formal workers	1986	0.5215	0.0001	0.5212	0.5219	0.7031	0.0001	0.7028	0.7034	0.2320	0.0001	0.2318	0.2321
	2004	0.4690	0.0002	0.4686	0.4694	0.6366	0.0001	0.6362	0.6370	0.2225	0.0001	0.2223	0.2227
Private sector	1986	0.5466	0.0001	0.5463	0.5469	0.7456	0.0001	0.7453	0.7458	0.2563	0.0001	0.2562	0.2565
	2004	0.4973	0.0001	0.4970	0.4976	0.7011	0.0001	0.7008	0.7014	0.2428	0.0001	0.2426	0.2429
Government employees	1986	0.5305	0.0007	0.5286	0.5324	0.7129	0.0003	0.7121	0.7136	0.2557	0.0003	0.2548	0.2566
	2004	0.4362	0.0005	0.4348	0.4376	0.6097	0.0005	0.6084	0.6109	0.2157	0.0003	0.2150	0.2164
15 to 30 years old	1986	0.5477	0.0002	0.5473	0.5481	0.7572	0.0001	0.7568	0.7575	0.2670	0.0001	0.2669	0.2672
	2004	0.4860	0.0002	0.4855	0.4864	0.6702	0.0002	0.6697	0.6706	0.2403	0.0001	0.2400	0.2405
31 to 45 years old	1986	0.5446	0.0002	0.5441	0.5452	0.7384	0.0002	0.7380	0.7388	0.2559	0.0001	0.2557	0.2562
	2004	0.5169	0.0004	0.5159	0.5179	0.7187	0.0002	0.7182	0.7191	0.2540	0.0002	0.2536	0.2545
46 to 65 years old	1986	0.5630	0.0005	0.5615	0.5644	0.7609	0.0002	0.7603	0.7615	0.2275	0.0002	0.2268	0.2281
	2004	0.5154	0.0003	0.5147	0.5161	0.7414	0.0002	0.7409	0.7419	0.2416	0.0001	0.2412	0.2419
Primary education	1986	0.6426	0.0002	0.6422	0.6431	0.8375	0.0001	0.8372	0.8378	0.3032	0.0001	0.3030	0.3034
	2004	0.6385	0.0002	0.6379	0.6392	0.8285	0.0002	0.8281	0.8290	0.3119	0.0001	0.3115	0.3122
Secondary education	1986	0.5274	0.0002	0.5270	0.5278	0.7194	0.0001	0.7190	0.7197	0.2489	0.0001	0.2487	0.2491
	2004	0.5474	0.0002	0.5470	0.5478	0.7314	0.0001	0.7310	0.7317	0.2677	0.0001	0.2675	0.2679
University education	1986	0.3862	0.0004	0.3852	0.3871	0.5392	0.0003	0.5383	0.5401	0.1821	0.0002	0.1817	0.1826
	2004	0.3167	0.0003	0.3160	0.3173	0.4778	0.0003	0.4771	0.4786	0.1561	0.0001	0.1557	0.1564

Source: own calculations based on household survey data for labour force aged between 15 and 65 years in the seven main metropolitan areas. See text for definitions of different groups of the labour force.

The composition of the labour force in urban Colombia also recorded important transformations in terms of its educational structure. We thus calculated the same set of horizontal segregation measures for three schooling levels: workers with five or less years of schooling (i.e., primary education), workers with six to 11 years of schooling (i.e., secondary education), and workers with 12 or more years of schooling (i.e., university education). This particular disaggregation of the labour force provides the widest differences in the horizontal occupational segregation indices by gender and suggests that since the mid-1980s education has been a key factor in the evolution of gender occupational differences in the labour market of urban Colombia. On the one hand, estimates of the *KM* index for all years suggest an inverse relationship between educational levels and occupational segregation. As can be seen in Panel c of Figure 1.2, the *KM* index is the lowest for workers with university education and the highest among those with primary or less over all years (see also Table 1.1 for other indices). On the other, the reduction of segregation indicators alluded to above for the whole labour force is concentrated solely among those workers with university education. All of this suggests that increasing educational levels amongst female workers and, in particular, the rising proportion of these with university education appear as one of the main driving forces behind the reduction in gender based occupational segregation in urban Colombia.

As previously noted, our data suggest a re-structuring of government employment in urban Colombia in which women, after 2000, have steadily increased their share of public sector jobs to around 50 per cent (see Panel d on Figure 1.2). According to all the indices computed in this study, gender occupational segregation has exhibited a more marked decline amongst government workers compared to the rest of the labour force between 1986 and 2004. For instance, the *KM* index fell by 15.6 per cent in the former case compared to a reduction of 5.3 per cent for the latter case over this period. All other indices suggest a similar pattern of change (see Table 1.1 above). Interestingly, our measures of gender occupational segregation for government employees appear to be lower than those of the private sector after 1992, when most of the constitutional reforms towards a more egalitarian participation of women in government positions were put in place.

1.5.2 Decomposition of changes in segregation indices over time

As suggested above, the *DI* and other segregation measures may be sensitive to changes in both, the structure of occupations and the gender composition of the labour force. From an analytical point of view this represents a major problem since a reduction in occupational segregation indices may be possible without any changes in the gender (or ethnic) composition of particular occupations. In addressing this problem, Deutsch et al. (2006) proposed a generalisation of a decomposition technique originally introduced by Karmel and MacLachlan (1988) to identify what portion of a given change in a segregation index may be due to changes in 'net' segregation, this is changes in the gender/ethnic ratios of particular occupations, and what part of the change may be driven by 'gross' segregation which is due to changes in both the gender/ethnic composition of the overall labour force and the structure of occupations. According to Deutsch et al (2006), a change in a segregation measure over time may be defined as

$$\Delta I = I_v - I_p \tag{1.6}$$

where I_v and I_p represent, respectively, the indices for the final and initial periods of time. These two indices can be drawn from segregation matrices whose typical element in its internal structure, p_{ij}, represents the ratio T_{ij}/T where T_{ij} is the number of individuals in occupation *i* from gender *j* and T is the total number of workers. The margins of this matrix are defined by $p_i=T_i/T$ and $p_j=T_j/T$ which denote, respectively, the horizontal margins (or occupation shares) and the vertical margins (or gender shares).

The total variation ΔI may be expressed in terms of the variations in 'net' and 'gross' segregation: [12]

$$\Delta I = f(\Delta m, \Delta is) \tag{1.7}$$

where Δm and Δis represent, respectively, changes in the margins and in the internal structure of the segregation matrix. By applying the concept of the Shapley decomposition from the income distribution literature, Deutsch et al. (2006) propose that the change in a segregation index as in (1.6) and (1.7) may be expressed as

$$\Delta I = \Delta C_m + \Delta C_{is} \tag{1.8}$$

where ΔC_m represents the contribution of changes in the margins, or gross segregation, and ΔC_{is} represents the contribution changes in the internal structure, or net segregation. Deutsch et al. (2006) demonstrate that these two components can also be expressed as

$$\Delta C_m = \left(\frac{1}{2}\right) f(\Delta m) + \left(\frac{1}{2}\right) [f(\Delta m, \Delta is) - f(\Delta is)] \tag{1.9}$$

and

$$\Delta C_{is} = \left(\frac{1}{2}\right) f(\Delta is) + \left(\frac{1}{2}\right) [f(\Delta m, \Delta is) - f(\Delta m)] \tag{1.10}$$

where Δm denotes the change in the margins, Δis represents the change in the internal structure. According to Deutsch et al. (2006), the numeric solution for (1.8) to (1.10)

---

[12] As explained by Deutsch et al (2006), this technique could be applied to more than two groups of the labour force although our presentation here refers to the conventional gender dichotomous approach.

can be achieved through the derivation of a set of matrices which are obtained through the interaction of both the margins and the internal structure of P and V. In order to spell out this more clearly, let S be a matrix which has the internal structure of matrix P and the margins of matrix V. This matrix can be obtained by successive iterations (see Deming and Stephan, 1940) in which the first step is to multiply all elements p_{ij} by the ratios v_i/p_i to obtain a matrix X. Then, the elements of X are multiplied by the ratios $v_{.j}/x_{.j}$ where $v_{.j}$ and $x_{.j}$ are the vertical margins of the matrices V and X to obtain a matrix Y. After several iterations, the resultant matrix will converge to a matrix S with the internal structure of P and the margins of V. Similarly, a matrix W with the internal structure V and the margins of P may be obtained if we invert the process from V to P. Other necessary matrices are

- matrix L with the internal structure of P, the vertical margins of P and the horizontal margins of V;
- matrix K with the internal structure of P, the vertical margins of V and the horizontal margins of P;
- matrix C with the internal structure of V, the vertical margins of V and the horizontal margins of P and;
- matrix F with the internal structure of V, the vertical margins of P and the horizontal margins of V (see Deutsch et al. (2006) for details on the derivation).

Thus, the contribution of changes in internal structure, as in equation (1.8) above, can be conveniently re-expressed as

$$\Delta C_{is} = \left(\frac{1}{2}\right)\left(I_w - I_p\right) + \left(I_v - I_s\right) \qquad (1.11)$$

while in the case of the contribution to changes in the margins, this could be as

$$\Delta C_m = \left(\frac{1}{2}\right)\left[\left(I_s - I_p\right) + \left(I_v - I_w\right)\right]. \qquad (1.12)$$

However, from a policy perspective, it is interesting to differentiate between the specific contribution of changes in female labour force participation and those from changes in the structure of occupations. In other words, this is

$$\Delta C_m = C_h + C_t \tag{1.11}$$

where C_h and C_t represent the contributions from changes in the structure of occupations and gender totals, respectively. This could be expressed in terms of the index values I obtained from their corresponding matrix denoted by the subscript. Therefore, C_h and C_t can be estimated as

$$C_h = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left[(I_l - I_p) + (I_s - I_k)\right] + \left[(I_v - I_c) + (I_f - I_w)\right] \tag{1.12}$$

and

$$C_t = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left[(I_k - I_p) + (I_s - I_l)\right] + \left[(I_v - I_f) + (I_c - I_w)\right] \tag{1.13}$$

which together satisfy (1.11). To sum up, a change on a segregation index between two periods of time can be decomposed as

$$\Delta I = C_{is} + C_h + C_t, \tag{1.14}$$

which can be more explicitly divided into: (i) changes in the gender composition of occupations, C_{is}, (ii) changes in the labour market structure of occupations, C_h, and (iii) changes in female/male shares into the labour force, C_t.

Using the methods described above, we programmed the decomposition described in expression (1.14) in Mata, a matrix programming language in Stata, for the three segregation indices already used in this chapter between 1986 and 2004 but, for the sake of brevity we focus the analysis on the *KM*.[13]

---

[13] Decomposition results for Gini and KM indices are reported in Appendix 1.2.

For the labour force as a whole (see Table 1.2), we find that just 17.1 per cent of the variation in the index for all workers originated in changes of net segregation while the remaining 82.9 per cent comes from changes in gross segregation. The same decomposition results indicate that increasing female labour participation explains, by itself, 1.5 times the total variation in the *KM* index, a change which was just partially offset by changes in the structure of jobs. Decomposition results for the *DI* and Gini also indicate sizeable contributions of female labour participation to the total variation of these indices (see Appendix 1.2). These findings confirm that the increasing share of women workers in the labour force is actually driving most of the reduction in the segregation indices reported in this study for all workers in urban Colombia.

A broadly similar result is found between formal and informal workers where the variations in the *KI* are mainly driven by gross segregation. In the former, changes in female labour force participation and occupations' structure represent by themselves more than twice the reduction in the *KI* while changes in the net segregation operate in an opposite direction. In fact, decomposition results for the other two indices (see Appendix 1.2) reveal an increase in net segregation for formal workers. All of this suggests that, even though the reduction in segregation indices is the largest amongst formal workers, this result is driven by changes in the margins which mask an increase in net segregation for this segment of employment. On the other hand, the reduction in segregation measures for informal workers was modest compared to that of the formal sector. Our decomposition results indicate again that most of this reduction in the *KM* index is driven by changes in the margins with more than half coming from increases in female labour participation. To sum up, while all indices suggest a reduction in gender based occupational segregation for both segments of the labour force, a closer inspection of the decomposition analysis indicates that the gender composition of particular occupations is roughly the same over this 19 years period. In order to take this issue further, we have to determine whether this result holds for all groups of the labour force.

Table 1.2: Shapley decomposition of changes in Karmel and MacLachlan (1988) index between 1986 and 2004 in urban Colombia (seven largest metropolitan areas)

	(1)	(2)	(3)	(4)	(5)
Groups of the labour	***Female/male***	***Occupations'***	***(1 + 2)***	***Internal***	***Gross change***
All workers	-0.023	0.010	-0.012	-0.003	-0.015
	151.7%	-68.8%	82.9%	17.1%	100.0%
Formal workers	-0.006	-0.015	-0.020	0.011	-0.009
	58.1%	153.7%	211.8%	-111.8%	100.0%
Informal workers	-0.012	-0.008	-0.020	-0.001	-0.021
	57.1%	37.6%	94.7%	5.3%	100.0%
Primary education	-0.007	-0.013	-0.020	0.028	0.009
	-75.3%	-149.1%	-224.4%	324.4%	100.0%
Secondary education	0.006	0.002	0.009	0.010	0.019
	33.7%	11.8%	45.5%	54.5%	100.0%
University education	-0.007	-0.007	-0.014	-0.012	-0.026
	26.6%	28.0%	54.6%	45.4%	100.0%
Government workers	-0.004	-0.013	-0.017	-0.023	-0.040
	10.4%	32.5%	43.0%	57.0%	100.0%
Private sector workers	-0.004	-0.009	-0.013	-0.001	-0.014
	27.7%	66.4%	94.1%	5.9%	100.0%
Aged 15 to 30 years old	-0.015	-0.017	-0.032	0.005	-0.027
	54.7%	64.8%	119.5%	-19.5%	100.0%
Aged 31 to 45 years old	0.003	-0.003	0.000	-0.002	-0.002
	-154.3%	158.3%	4.0%	96.0%	100.0%
Aged 46 to 65 years old	0.015	-0.001	0.015	0.000	0.014
	107.1%	-4.1%	103.0%	-3.0%	100.0%

Source: own calculations based on household survey data for labour force aged between 15 and 65 years in the seven main metropolitan areas. *: as a percentage of the mean value of the indices.

By educational levels, it should be noted that the *KI* reported a significant reduction only amongst those workers with university education. In this group of workers, we observe a sizeable contribution of net segregation which by itself represents almost one half of the variation in the *KM* index between 1986 and 2004 (a result that is confirmed by the other two indices – see Appendix 1.2). For those with primary and secondary education, the decomposition of changes in the *KM* index suggests that, in both cases, the gender ratio of particular occupations became more segregated leading to an increase in net segregation between 1986 and 2004. In other words, decomposition results by educational levels indicate that only in the case of workers with university education there was a substantial change towards a more egalitarian gender composition of occupations over these years. As explained in the previous section, this was incidentally the group of the labour force with the lowest indices of occupational segregation by gender for all years reviewed in this study.

The division of the labour force between government and private sectors suggests that changes in net segregation explain 57.0 per cent of the variation of the *KM* index in the former compared to just 5.9 per cent in the latter. It should be remarked that the reduction in net segregation amongst government workers is the largest one of all subgroups of the labour force in urban Colombia between 1986 and 2004, not only for the *KM* but also for the other two indices. This is in line with the interpretation reported above regarding the effects of the gender legislation in Colombia which is, presumably, more enforceable in government institutions. In contrast, the reduction in the *KM* index amongst workers in the private sector is dominated by changes in gross segregation, with more than one quarter of it coming from increased female labour force participation and nearly two thirds from changes in the structure of occupations. In other words, just 5.9 per cent of the reduction in this index amongst workers outside government is explained by changes in the gender composition of particular occupations or net segregation.

Finally, the division of the labour force by age groups indicates that although the largest reduction in the *KM* index was reported for the youngest (15 to 30 years old), this variation is driven entirely by changes in gross segregation. This means that while the gender composition of occupations for this group became more segregated between

1986 and 2004, changes in the overall structure of occupations and the increasing labour force participation of women acted in an opposite direction to reduce the *KM* index across these two years. For those aged 46 and 65 years old, the slight increase in the *KM* index was driven mainly by changes in female labour participation which suggests that a substantial proportion of women entering the labour force within this age group did so in female dominated occupations. The only age group with a reduction in segregation is represented by those between 31 and 45 years old, in which most of the variation can be explained in terms of an improvement in the gender ratio of particular occupations. It is also interesting to observe in this age group that the effect of the increasing participation of women in the labour force contributed to raising the index, presumably as a result of more women joining female dominated occupations, an effect that was offset by changes in the structure of occupations in which less segregated occupations are increasing their share into the overall employment structure. To some extent, this finding is in line with our analysis from the alternative *KM* index in section 1.5.1, according to which the level of segregation would be higher for all years if female labour participation remained at the level of 1986. All of this suggests again that an increasing proportion of women in the labour market imply a proportionally larger reallocation of workers from both genders in order to keep the level of segregation measures at the same level.

1.6. Conclusions

According to the measures used in this study, gender occupational segregation has exhibited a statistically significant reduction in urban Colombia between 1986 and 2004. The use of datasets with a harmonised classification of occupations for the whole period provided an opportunity to implement a set of segregation measures that, in one way or another, overcame some of the more conventional difficulties in the measurement of occupational segregation by gender. In addition, the use of bootstrapped standard errors yielded a statistical basis to verify that most of the observed changes between 1986 and 2004, as well as the differences in point estimates between different groups of the labour force in terms of age, education and type of employment (formal and informal), are statistically significant.

From a methodological point of view, the implementation of different segregation measures such as the alternative version of *KM* in which changes in female labour force participation are held constant allows us to make some interesting qualifications about the observed trends in urban Colombia between 1986 and 2004. Even though conventional dissimilarity indices suggest a reduction in occupational segregation by gender for all age groups, once the effects from the rising share of women in the labour force are controlled for it becomes fairly evident that an important proportion of those women entering the labour force are doing so into highly segregated occupations. Results disaggregated by education also reveal that only in the case of workers with a university education is there an unambiguous reduction in the extent of horizontal gender based occupational segregation as measured by the indices used in this study. But clearly, from all subdivisions of the labour force presented here, the largest reduction in the *KM* and *DI* was found amongst government workers.

The decomposition of the changes in occupational segregation measures between 1986 and 2004 indicates that the main underlying force in the reduction of gender occupational segregation indices for all workers during this period was the increasing female labour participation. The same decomposition results (and the level of indices by themselves) suggest that the majority of women in urban Colombia are still employed in female-dominated occupations and that a substantial proportion of those entering the labour force are doing so into these type of jobs. This explains why horizontal segregation measures by gender remain so persistently high in the urban areas of Colombia. We have found convincing evidence that the lowest levels of occupational segregation are found amongst workers with university education and those employed by the government. The decomposition results indicate that those are the groups in which a less segregated gender composition of individual occupations (net segregation) played a major role in the reduction of segregation indices. In the case of government workers, we find suggestive evidence that the introduction of gender equality legislation at the beginning of the 1990s and its interaction with a more regulated institutional environment to enforce these provisions in the public sector are fundamental forces behind this result. In the case of workers with university education the reduction in the indices is less pronounced than in the case of government workers but they remain as the least segregated in terms of gender. The increased access of women to university education has favoured their access to a wider variety of occupations in which academic credentials, rather than gender roles, are more relevant. It is also true that more educated workers in general are more likely to be aware of, and eventually demand, their gender rights.

All of this suggests that institutions play a differentiated role in the level of horizontal gender segregation amongst some groups within the labour force. All horizontal segregation indices are consistently lower amongst those with university education and those in government jobs. Interestingly, the differences in point estimates between government employees and the rest of the labour force are rather small before 1991 but subsequently the level of segregation in the former exhibited a substantial reduction only equalled by workers with university education. To some extent, this evidence is consistent with the fact that the new Colombian constitution enacted in 1991 mandated an inclusive employment policy for women in all levels of public administration. In the same vein of analysis, it becomes clear why all measures of gender occupational segregation are the highest amongst informal workers, given the unregulated nature of this segment of the labour market.

We were able to provide an optimistic story about the evolution of horizontal occupational segregation in urban Colombia, as far as the formal and the informal or atypical employment segments of the labour market are concerned. However, we do not provide any conjectures about the extent of vertical segregation and the access of women to managerial and decision-making positions within occupation. In this respect, the story may be somehow less positive, in particular among vulnerable groups such as unskilled older workers outside the formal sector.

Appendix 1.1

Table A1.1: Colombian Classification of Occupations

1 – Physical Scientists and Related Technicians	54 - Maids and Related Housekeeping Service
2 - Architects, Engineers and Related Technicians	55 - Building Caretakers, Charworkers, Cleaners and
3 - Engineering technicians, Surveyors,	56 - Launderers, Dry-Cleaners and Pressers
4 - Aircraft and Ships' Officers	57 - Hairdressers, Barbers, Beauticians and Related
5 - Life Scientists and Related Technicians	58 - Protective Service Workers
6 - Medical, Dental, Veterinary and Related	59 - Service Workers n.e.c.
7 - Professional nurses, optometrists,	60 - Farm Managers and Supervisors
8 - Statisticians, Mathematicians, Systems	61 - Farmers
9 – Economists	62 - Agricultural and Animal Husbandry Workers
11 – Accountants	63 - Forestry Workers
12 - Jurists, lawyers and judges	64 - Fishermen, Hunters and Related Workers
13 – Teachers	70 - Production Supervisors and General Foremen
14 - Workers in Religion	71 - Miners, Quarrymen, Well Drillers and Related
15 - Authors, Journalists and Related Writers	72 - Metal Processers
16 - Sculptors, Painters, Photographers and	73 - Wood Preparation Workers and Paper Makers
17 - Composers and Performing Artists	74 - Chemical Processers and Related Workers
18 - Athletes, Sportsmen and Related Workers	75 - Spinners, Weavers, Knitters, Dyers and Related
19 - Professional, Technical and Related Workers	76 - Tanners, Fellmongers and Pelt Dressers
20 - Legislative Officials and Government	77 - Food and Beverage Processers
21 - General managers	78 - Tobacco Preparers and Tobacco Product Makers
30 - Production managers (except farm)	79 - Tailors, Dressmakers, Sewers, Upholsterers and
31 - Government Executive Officials	80 - Shoemakers and Leather Goods Makers
32 - Stenographers, Typists and Card- and Tape-	81 - Cabinetmakers and Related Woodworkers
33 - Bookkeepers, Cashiers and Related Workers	82 - Stone Cutters and Carvers
34 - Computing Machine Operators	83 - Blacksmiths, Toolmakers and Machine-Tool
35 - Transport and Communications Supervisors	84 - Machinery Fitters, Machine Assemblers and
36 - Transport Conductors	85 - Electrical Fitters and Related Electrical and
37 - Mail Distribution Clerks	86 - Broadcasting Station and Sound Equipment
38 - Telephone and Telegraph Operators	87 - Plumbers, Welders, Sheet Metal and Structural
39 - Clerical and Related Workers n.e.c.	88 - Jewellery and Precious Metal Workers
40 - Managers (Wholesale and Retail Trade)	89 - Glass Formers, Potters and Related Workers
41 - Working Proprietors (Wholesale and Retail	90 - Rubber and Plastics Product Makers
42 - Sales Supervisors and Buyers	91 - Paper and Paperboard Products Makers
43 - Technical Salesmen, Commercial Travellers	92 - Printers and Related Workers
44 - Insurance, Real Estate, Securities and	93 - Painters (buildings, construction, etc)
45 - Salesmen, Shop Assistants and Related	94 - Production and Related Workers n.e.c.
49 - Sales Workers n.e.c.	95 - Bricklayers, Carpenters and Other Construction
50 - Managers (Catering and Lodging Services)	96 - Stationary Engine and Related Equipment
51 - Working Proprietors (Catering and Lodging	97 - Material-Handling and Related Equipment
52 - Housekeeping and Related Service	98 - Transport Equipment Operators
53 - Cooks, Waiters, Bartenders and Relaters	99 - Labourers and workers n.e.c.

Source: adapted from DANE (2000).

Appendix 1.2

Table A1.2.1: Shapley decomposition of changes in Gini segregation index (Deutch et al., 1994) between 1986 and 2004 in urban Colombia (seven largest metropolitan areas)

Groups of the labour	(1) *Female/male*	(2) *Occupations'*	(3) *(1 + 2)*	(4) *Internal*	(5) *Gross change*
All workers	-0.043	-0.002	-0.045	0.001	-0.045
	97.3%	4.2%	101.5%	-1.5%	100.0%
Formal workers	-0.038	-0.043	-0.081	0.015	-0.066
	57.2%	65.1%	122.3%	-22.3%	100.0%
Informal workers	-0.007	-0.005	-0.013	-0.008	-0.020
	35.4%	26.8%	62.2%	37.8%	100.0%
Primary education	-0.032	-0.033	-0.065	0.056	-0.009
	360.6%	367.5%	728.1%	-628.1%	100.0%
Secondary education	-0.005	-0.006	-0.011	0.023	0.012
	-42.0%	-51.7%	-93.7%	193.7%	100.0%
University education	-0.018	-0.015	-0.033	-0.028	-0.061
	29.7%	24.7%	54.4%	45.6%	100.0%
Government workers	-0.039	-0.048	-0.087	-0.016	-0.103
	38.0%	46.5%	84.6%	15.4%	100.0%
Private sector workers	-0.025	-0.025	-0.050	0.006	-0.044
	55.9%	56.6%	112.5%	-12.5%	100.0%
Aged 15 to 30 years old	-0.040	-0.041	-0.081	-0.006	-0.087
	46.4%	47.3%	93.7%	6.3%	100.0%
Aged 31 to 45 years old	-0.006	-0.008	-0.013	-0.007	-0.020
	27.9%	39.0%	67.0%	33.0%	100.0%
Aged 46 to 65 years old	-0.012	-0.012	-0.024	0.005	-0.019
	61.2%	63.3%	124.5%	-24.5%	100.0%

Source: own calculations based on household survey data for labour force aged between 15 and 65 years in the seven main metropolitan areas.

Table A1.2.2: Shapley decomposition of changes in Duncan & Duncan (1955) segregation index between 1986 and 2004 in urban Colombia (seven largest metropolitan areas)

Groups of the labour	(1) *Female/male*	(2) *Occupations'*	(3) *(1 + 2)*	(4) *Internal*	(5) *Gross*
All workers	-0.047	0.001	-0.046	-0.005	-0.051
	91.4%	-2.0%	89.4%	10.6%	100.0%
Formal workers	-0.034	-0.041	-0.075	0.022	-0.053
	64.3%	78.2%	142.5%	-42.5%	100.0%
Informal workers	-0.024	-0.016	-0.039	-0.002	-0.041
	56.8%	37.8%	94.5%	5.5%	100.0%
Primary education	-0.028	-0.034	-0.062	0.058	-0.004
	686.2%	818.8%	1505.1%	-1405.1%	100.0%
Secondary education	0.000	-0.002	-0.001	0.021	0.020
	2.0%	-7.9%	-6.0%	106.0%	100.0%
University education	-0.025	-0.020	-0.045	-0.025	-0.069
	35.3%	29.1%	64.3%	35.7%	100.0%
Government workers	-0.018	-0.031	-0.049	-0.046	-0.094
	18.8%	32.6%	51.4%	48.6%	100.0%
Private sector workers	-0.022	-0.025	-0.047	-0.002	-0.049
	44.6%	51.5%	96.1%	3.9%	100.0%
Aged 15 to 30 years old	-0.035	-0.038	-0.072	0.011	-0.062
	56.2%	60.8%	117.0%	-17.0%	100.0%
Aged 31 to 45 years old	-0.010	-0.014	-0.024	-0.004	-0.028
	35.1%	50.0%	85.1%	14.9%	100.0%
Aged 46 to 65 years old	-0.020	-0.027	-0.047	-0.001	-0.048
	42.4%	56.0%	98.4%	1.6%	100.0%

Source: own calculations based on household survey data for labour force aged between 15 and 65 years in the seven main metropolitan areas.

Chapter 2: Occupational segregation and gender wage differences: evidence from Urban Colombia

2.1 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 was relatively scarce until recently. 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.

While the occupational segregation literature in general suggests that access to occupations tends to be highly differentiated by gender (i.e., Anker, 1997, Borghans and Groot, 1999, Grazier and Sloane, 2007, Hakim, 1992), a number of advances in the literature suggest this is also a key element in understanding gender wage differences. In this sense, Baldwin et al (2001) proposed an extension to the conventional approach by including a hierarchical dimension in which men dislike to be supervised by women even in the case where they do not object to working alongside women. As a result, their model not only predicts that women's participation is decreasing with respect to job hierarchy but also that female wage disadvantage in managerial positions is, at least, partially explained by a compensation mechanism of men's dislike for female supervisors.[14]

Recognition that the gender wage gap is not homogeneous across the entire wage distribution has lead to an investigation of the existence of 'glass ceilings' (Arulampalam et al., 2007)[15] and 'sticky floors' (i.e., Booth et al., 2003) where the former pertains to the barriers women face in access to jobs at the highest occupational and pay levels while the latter refers to the concentration of women at the bottom of the occupational wage structure even in cases when they get promoted. On the other hand, it has been found that occupations with a high concentration of women tend to offer lower wages for both genders compared to male dominated occupations (i.e., Bayard et al., 2003, Jurajda, 2003, Lucifora and Reilly, 1990, MacPherson and Hirsch, 1995, Baker and Fortin, 2001). The empirical strategy in most of those studies relies on conventional wage equations in which the share of female workers in the incumbent worker's occupation is included as an additional regressor. This literature suggests, in general, that the wage penalty arising from the female occupational intensity tends to be lower when controls for occupation characteristics, industry affiliation and firm characteristics are introduced in the econometric specification. According to Jurada (2005) and Jurajda and Harmgart (2007), the wage penalty on female jobs can be explained by (i) discrimination practices which restrict the access of women to high-wage positions, (ii) lower productivity levels of workers who typically engage in these jobs and, (iii) nonwage characteristics (i.e., flexible work times) which are functional to

[14] Although we do not attempt an empirical test of all propositions derived from Baldwin et al.'s (2001) model, it is appropriate to highlight that the connection between gender wage discrimination and occupational segregation is theoretically well grounded.

[1515] For instance, Baron and Cobb-Clark (2008) found in Australia that the gender wage gap for those at the top of the conditional wage distribution remains "largely unexplained" after controlling for differences in human capital and other relevant characteristics.

female roles in society but impose a wage penalty. Some of the studies outlined above (i.e., Baker and Fortin, 2001) show indeed that controlling for characteristics outlined in (ii) and (iii) reduce the coefficient effect from female occupational intensity. Nevertheless, female dominated occupations do not always offer lower wages. For instance, Shauman (2006) finds in the case of US college graduates that some female dominated jobs where verbal abilities are important tend to have higher pay levels although, this effect is counterbalanced by characteristics such as the availability to work part-time or people-oriented skills which tend to be associated with lower wages. Jurada and Harmgart (2007) find also that female dominated occupations in West Germany do not tend to pay lower wages while many of them actually pay higher wages in East Germany, which was subject to a centrally planned socialist system. The aggregate evidence from this literature suggests that women tend to cluster in occupations with low pay levels and/or costly characteristics and amenities that make them less likely to be positioned at the upper end of the wage distribution.[16]

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. Although this dummy variable approach implies that the gender distribution of jobs is justified or randomly distributed (or allocated), a number of econometric methods allow us to differentiate the portion of the unequal distribution of jobs by gender and determine if it is justified based on the set of observed

[16] There is also a vast range of literature using quantile regression approaches showing precisely this point but that discussion lies outside the focus of the current chapter.

characteristics and what part of it may be deemed as a result of segregation.[17] Therefore, as occupational attainment is affected by segregation, gender wage differences may be divided at least into two broad sources: first, within-occupation wage differences which are related to the explained and unexplained components mentioned above and, second, between-occupation 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) were 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 occupational attainment model, Brown *et al.* (1980) modelled the male occupational distribution to produce a counterfactual female 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 (see also Meng and Miller, 1995, Miller, 1987, Reilly, 1991, Akter, 2005). Other studies (Liu et al., 2004, Neuman and 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.

The aim of this chapter is to contribute to this literature with an empirical application for urban Colombia, where some improvements in both gender wage gaps and gender-

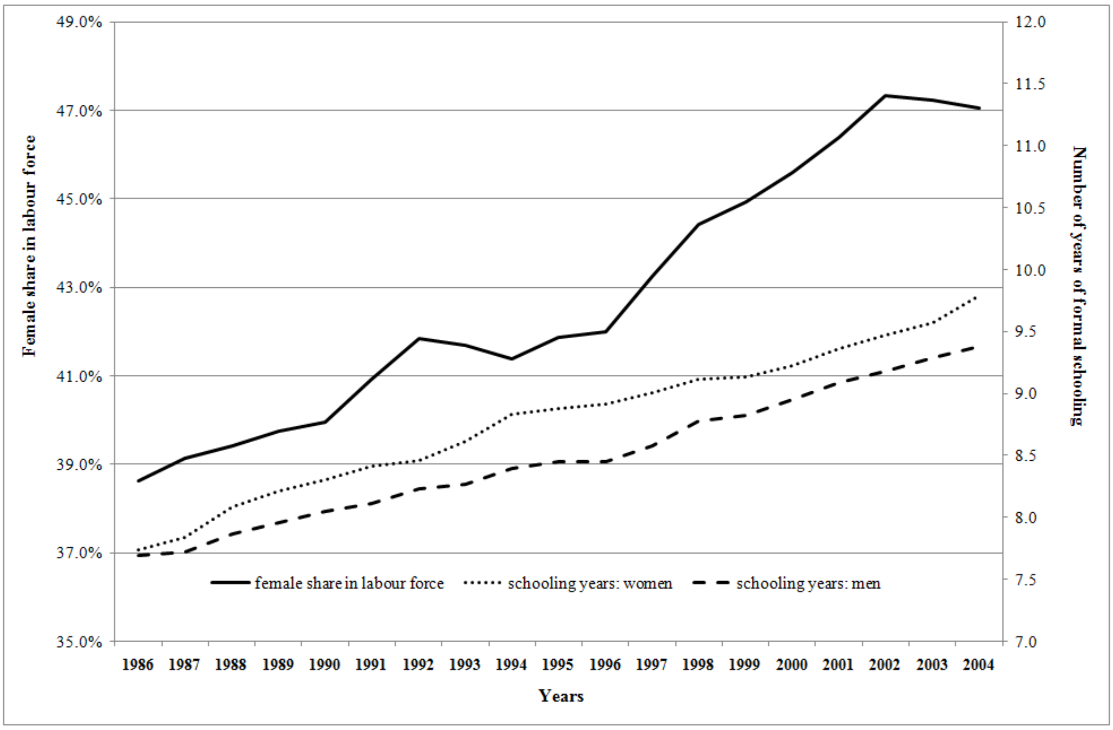
[17] However, it has been argued that if the gender distribution of occupations is subject to some sort of discrimination, as implied in the occupational segregation literature, the dummy variable approach might be inadequate (Meng and Miller, 1995, Miller, 1987, Reilly, 1991).

based occupational segregation indices have been observed since the mid-1980s. In particular, we investigate whether lower wages can be related to female occupational intensity and whether the segregated nature of the distribution of jobs by gender explains some part of the gender wage gap in this country. On the one hand, our empirical strategy involves the estimation of a counterfactual distribution of female employment once the decision to participate in the labour market has been taken. Then, we use these results to identify which portion of the wage gap can be attributed to both, the explained and unexplained components of the gender distribution of jobs within an Oaxaca-Blinder decomposition framework. We also estimate log wage equations in which the percentage of female workers in a given occupation is included instead of controls for occupation effects. The rest of this chapter is organised as follows. The second section presents some labour market background of the country. The third explains the empirical model. The fourth describes the microdata used for this study. The fifth presents and discusses the results and, finally, a sixth section summarises the main findings and identifies some issues for further research.

2.2 Background of the country

As indicated at the introductory section of this thesis, Colombia experienced a dramatic process of demographic change. This process was accompanied by a massive absorption of women into the labour market of urban Colombia as well as increasing educational levels for both gender groups. Women increased their share of the total labour force from 38.1 per cent in 1986 to 46.1 per cent in 2004. At the same time, the percentage of female workers with university education doubled from 15.7 per cent in 1986 to 32.1 per cent in 2004, while their average number of years of formal schooling rose from 7.7 to 9.8 years over this period (see Figure 2.1). It should be noted that the proportion of female workers with university education in urban Colombia has been increasing at a faster rate than that among their male counterparts over this period. The labour market absorption of women has been in conjunction with a structural transformation of the nature of employment in urban Colombia.

Figure 2.1: Female share in labour force and number years of formal schooling years by gender in urban Colombia, seven metropolitan areas



Own estimates based on household survey microdata from seven main metropolitan areas in urban Colombia for people aged 18 and 65 years. Population projections are not strictly comparable before and after 2000 due to changes in sampling design.

As a result of a massive migration process from rural areas during the last century, Colombia has one of the highest rates of urbanisation in Latin America (Hanratty and Meditz, 1988). Urban labour supply in this country has increased over the last decades and, according to Florez (2003), Gilbert (1997) and Isaza (2002), most of it has been absorbed by the so-called 'informal sector'. Government estimates for the first quarter of 2010 indicate that the informal sector represents 51.6 per cent of the labour force in the 13 largest cities of this country (DANE, 2010).[18] All of this suggests that the informal sector is a major feature of the urban employment structure of the country which comprises different dimensions of inequality within the labour force. For instance, 89.1 per cent of those in the informal sector in 2010 had no access to social security, compared to 23.9 per cent in the formal sector (DANE, 2010).

2.3 Methodology

There are two fundamental questions we want to address in this empirical application. On the one hand, we want to know how much of the overall gender hourly wage gap can be attributed to a differentiated pattern in the distribution of occupations between women and men once the effects of relevant characteristics are controlled. On the other, we want to verify whether, *ceteris paribus*, the proportion of women in a given occupation is to some extent associated with lower pay levels for both men and women.

In the first question, our strategy draws on developments in the literature in which the effects of occupational segregation are incorporated into the analysis of the wage gaps by gender and ethnic groups (Akter, 2005, Brown et al., 1980, Meng and Miller, 1995, Miller, 1987, Neuman and Silber, 1996, Reilly, 1991, Silber, 1989). In general, this

[18] The definition of the informal sector according to the Colombian government follows the guidelines from the International Labour Organisation according to which informal workers are those who (i) work in enterprises with less than five workers, (ii) those in unpaid work and, (iii) domestic servants. This definition does not take into consideration contractual issues and membership to social security, which is mandatory for all workers according to the Colombian labour legislation.

literature suggests that some part of the differentiated distribution of jobs by gender (or ethnic groups) is justified on the basis of differences in observable characteristics between groups of the labour force. However, another part of the distribution of jobs across gender (ethnic) groups remains unexplained due to the differentiated treatment that their corresponding characteristics receive in the labour market for the purposes of occupational allocation. The empirical challenge is, therefore, to find a counterfactual for the distribution of jobs in the hypothetical case that female (or minority group) characteristics were treated as the male (or majority group) characteristics. This can be achieved by estimating a multiple choice outcome model in which the dependent variable is categorical in nature and depicts j occupation categories based on a set of observable characteristics for the male (or majority) subsample. Then, the coefficients from the male (majority) subsample can be applied to the female (minority) subsample in order to obtain a counterfactual distribution of female (minority) workers, $\hat{P}^{*}$, across the j occupational categories in the hypothetical case that they were equally treated as their male (or majority) counterparts. Thus, using the observed occupational distributions of male, $\bar{P}^{m}$, and female, $\bar{P}^{f}$, jobs plus the counterfactual distribution female jobs in the absence of a differentiated treatment of their observed characteristics, $\hat{P}^{*}$, the Oaxaca-Blinder decomposition of the gender wage gap can be extended to isolate the portion of the observed log hourly wage gap that can be attributed to differences in the explained $\left(\bar{P}^{m} - \hat{P}^{*}\right)$ and unexplained $\left(\hat{P}^{*} - \bar{P}^{f}\right)$ allocation of workers. In addition, we include in this empirical application an additional term to depict differences in the returns that women and men are expected to receive from specific occupations in the labour market once the effects of observable characteristics included in the model are controlled for.

Thus, our empirical strategy involves the estimation of wage equations for each gender group:

$$W_i^m = X_i^m \beta^m + P_i^m \gamma^m + \varepsilon_i^m \tag{2.1}$$

$$W_i^f = X_i^f \beta^f + P_i^f \gamma^f + \varepsilon_i^f \tag{2.2}$$

where X is a set of observable characteristics, β is its corresponding coefficients, P is a vector of controls of occupational categories with their γ intercept terms to be estimated for the male(= m) and female(= f) subsamples and, ε is an error term. It should be noted that because the full set of occupation dummies is included in the equation, the constant term must be dropped. Thus, the gender wage gap for the overall labour force may be expressed as

$$\Delta\bar{W} = \bar{W}^m - \bar{W}^f = (\bar{X}^m \hat{\beta}^m - \bar{X}^f \hat{\beta}^f) + (\bar{P}^m \hat{\gamma}^m - \bar{P}^f \hat{\gamma}^f) \tag{2.3}$$

After rearranging terms and having P^* as the counterfactual vector of shares on female employment in the absence of segregation, the gender wage gap may be conveniently decomposed into five terms:

$$\Delta\bar{W} = \bar{X}^f(\hat{\beta}^m - \hat{\beta}^f) + (\bar{X}^m - \bar{X}^f)\hat{\beta}^m + \bar{P}^f(\hat{\gamma}^m - \hat{\gamma}^f) + (\bar{P}^m - \hat{P}^*)\hat{\gamma}^m + (\hat{P}^* - \bar{P}^f)\hat{\gamma}^m \tag{2.4}$$

where the first two terms on the right hand side represent the conventional unexplained (*treatments*) and explained (*endowments*) portions of the gender wage gap due to differences in, respectively, returns and levels of observed characteristics. The other three terms show a breakdown of the segregation dimension in the gender wage gap. $\bar{P}^f(\hat{\gamma}^m - \hat{\gamma}^f)$ is the portion of the wage gap that can be attributed to differentiated returns/intercepts for men and women within particular occupations, $(\bar{P}^m - \hat{P}^*)\hat{\gamma}^m$ represents the part of the gender wage gap that can be explained by differences in the characteristics of the gender subsamples and render them to be distributed differently across the distribution of jobs and, finally, $(\hat{P}^* - \bar{P}^f)\hat{\gamma}^m$ represents that part of the gender wage gap attributable to the unequal treatment that those characteristics encounter in the labour market in the distribution of jobs by gender. The latter term could be framed as the ‘pure’ segregation effect on the gender wage gap. In the same vein, the first three components may be interpreted as the overall within occupational wage differential and, the remaining two denote the between occupational wage differential. 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 the urban Colombian labour market. However, the index number approach pursued here is subject to the conventional “index number problem” as it is also possible to use the female wage structure or even the pooled sample wage structure (see Appleton et al., 1999 for a detailed discussion on this).

As indicated above, the decomposition outlined in expression (2.4) requires the estimation of $\hat{P}^{*}$, a counterfactual distribution of female employment in the absence of unequal treatment. For this purpose, we estimate a multinomial logit model in which the dependent variable features *j* occupational outcomes (see Table 2.1) for the male subsample,

$$P_i^j = \frac{exp(Z_i\gamma_i)}{1+\sum_{j=1}^{k-1}(Z_i\gamma_i)} \qquad (2.5)$$

where outcome *j* is set equal to 0 as required by the Theil normalization, *Z* is a vector of characteristics for the male subsample and γ is the estimated coefficients. Subsequently, we use (2.5) to estimate the predicted probabilities for the female subsample to obtain $\hat{P}^{*}$.

Similar empirical applications in which the effects of occupational segregation on the gender (ethnic) wage gap are analysed (i.e., Brown et al., 1980, Liu et al., 2004, Reilly, 1991) rely on a limited number of occupational categories (ranging from five to eight broad groups). This enables the estimation of specific log wage equations for each one of the categories and gender groups included in the analysis. One of the advantages from this empirical approach is that the endogenous nature of gender or ethnic distribution of jobs can be controlled for selection bias into broad occupation categories with econometric procedures that are analogous to the Heckman correction method. A major drawback with this approach is that the extent of occupational segregation is clearly underestimated, as conventional dissimilarity indices such as the index proposed by Duncan and Duncan (1955) as well as other absolute difference measures are sensitive to the number of occupations used in their computation (Melkas and Anker, 1997).

In order to avoid this type of aggregation bias to the greatest extent possible we opted to include a larger number of occupational categories in our analysis (23 in the formal sector and 16 in the informal one – see the next section on data for details). Within this framework, our modelling strategy requires only one log wage equation per gender group with dummies for the full set of occupational categories and no constant term. This allows a clear identification of the occupational wage effects. It should be noted that the use of a multinomial logit model within the empirical approach developed here is just a mechanical device which allows us to obtain a counterfactual distribution of female workers in the absence of unequal treatment on observable characteristics. There are also obvious limitations on the number of occupational outcomes being modelled to obtain $\hat{P}^*$ in terms of both, sample cell sizes and computational time. In our case, the estimation of a model with 23 occupation categories for males in the formal sector (and 16 categories in the informal sector) based on samples with more than 100,000 observations took around 28 hours (and 22 hours, respectively) for each of the two datasets used along this chapter (see Appendix 2.1). All in all, we believe that this strategy is by itself a contribution to the existing empirical literature as it provides more precise measurements of the effects of segregation on gender wage differences based on a more refined classification of occupations with respect to previous studies in this field.

There are, however, at least two obvious limitations in our empirical strategy. First, as indicated above, the dummy variable approach to capture the wage effects of particular occupations assumes the distribution of workers across occupations as randomly distributed. This may not be an entirely realistic assumption given the segregated nature of the gender distribution jobs. Second, the dummy variable approach also assumes that the returns from observable characteristics are the same for all occupation categories. This assumption ignores a potential source of heterogeneity in the relationship between wage and productivity covariates for specific occupations (i.e., education and potential labour experience) included in our specification of wage equations. An alternative way to address these problems could be the implementation of a system of wage equations by gender across a given number of occupation categories in which occupational attachment to specific occupations is endogenously determined by a set of observable characteristics. This involves the estimation of selection equations which are either estimated simultaneously with the log wage equations through maximum likelihood methods or obtained through two-step procedures. All of this suggests that the wage equations described in (2.1) and (2.2) should be corrected for selectivity bias in the first place. We implemented several procedures to do so, including the conventional Heckman (1979) univariate probit approach based on the full sample of participants and non-participants. We also attempted to use a bivariate probit strategy which allows for the simultaneous determination of both, selection into (i) labour force and (ii) sector allocation between the formal and informal segments (see Main and Reilly, 1992 for an empirical application to Britain). In addition, we also tried different specifications using the multinomial logit approach proposed by Lee (1983) to control for both selection into labour force and then, selection into specific occupations (see, for instance, Liu et al., 2004, Reilly, 1991). Unfortunately, estimates (not reported here) from these selection models provide implausible results that are hard to rationalise in terms of both economic theory and the specific characteristics of the labour market in urban Colombia.

At the heart of this, we believe that a lack of adequate identifying instruments for the task at hand made unfeasible the implementation of a more comprehensive model suited to address the limitations mentioned above.[19] We should also stress that our formal/informal separation is somehow problematic as any correction for selection bias should ideally rely on good instruments that are able to shift the probability of engaging into the formal or informal segments but uncorrelated to both the probability of labour participation and the wage determination process.[20] In the light of these insurmountable problems, we have opted to report uncorrected OLS estimates for (2.1) and (2.2) with robust standard errors using White's (1980) procedure.

[19] To illustrate this point, the identification strategy involved in such selection models relied on conventional instruments such as the existence of children and infants in the household which are presumably highly correlated with labour participation decisions by gender. Unfortunately, the Colombian household survey microdata only allow identifying the relationship of children with respect to the household head but not with other members. We believe that the inaccuracy of such an important instrument might, to some extent, explain some of the problems encountered in our selection models. Other identifiers were based on overall households' characteristics such as the existence of unemployed members and the income from the rest of the household.

[20] In this sense, our strategy focused on the incorporation of instruments based on neighbourhood informality and unemployment rates, based on clusters and sectors drawn from the sampling design of household surveys.

2.4 Data and description of the sample

The empirical application in this chapter relies on the use of 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 inhabitants. Household surveys in urban Colombia are conducted by the Government on a quarterly basis since the mid-1980s. Our estimates come from two datasets with pooled observations from all waves between 1986 and 1989 and another for all waves between 1996 and 1999.[21] On average, each year comprises 88,000 individuals in the labour force aged between 18 and 65 years. The microdata provide information on labour earnings, number of weekly worked hours, industry, region and demographic variables such as age, educational attainment and marital status. Given the highly segmented nature of the labour market in urban Colombia (see below), we opted to divide our estimates between two sectors of employment, one for waged workers, their employers and all those in professional occupations which we label "formal workers" and another for those in 'atypical' employment conditions, this is own-account workers (except professionals) and domestic servants which we call "informal workers".[22] Furthermore, the surveys include information about occupations using a consistent classification of 82 categories over the entire period which, at the

---

21 This data is not a panel but we eliminated a small number of repeated households which may appear in more than one wave.

[22] Other divisions of the labour force from the point of view of informality are possible with the Colombian data including the ILO definition commented above. As the full set of questions for the informal sector, including the number of workers in the enterprise, is only gathered every two years in urban Colombia, we have to rely on a simpler classification criterion. For a detailed discussion on definitions and measurement of the informal sector in Colombia, see Florez (2002) and Ribero (2003).

two-digit level, is identical to the International Standard Classification of Occupations ISCO-68. For the reasons explained above, we regrouped the 82 original categories into 23 occupations for computational convenience in order to guarantee a large enough sample size for each occupation group by gender (see Table 2.1).[23]

Table 2.1: Broad ad-hoc occupational categories and their equivalents in the ISCO-68 and sub-sample sizes

Occupation categories	*ISCO-68 codes**	*Number of observations*	
		1986-1989	*1996-1999*
1. Engineers, technicians and physical scientist	*1 to 4*	5,630	4,627
2. Medical workers and life scientist	*5 to 7*	4,518	4,471
3. Social sciences and humanities	*8, 9, 14 to 19*	7,614	7,177
4. Accountants	*11*	3,230	2,696
5. Jurists	*12*	2,968	2,538
6. Teachers	*13*	14,392	13,557
7. Managers and directors	*20, 21, 30, 40, 50, 60, 70*	14,920	11,840
8. Bookkeepers, cashiers and computing	*33, 34*	13,045	11,426
9. Clerical workers	*31, 32, 39*	25,801	18,938
10. Transport and communication workers	*35 to 38*	7,278	5,971
11. Wholesale and Retail Trade workers	*41*	18,664	15,735
12. Sales workers	*42 to 44, 49*	5,415	5,656
13. Shop assistants	*45*	48,330	37,150
14. Catering and lodging workers	*51*	2,646	1,957
15. Housekeeping workers	*52, 54 to 56*	38,828	32,631
16. Other personal services workers	*53, 57 to 59*	31,674	27,105
17. Farm and related workers	*61 to 64*	5,145	3,400
18. Spinners, Weavers, Knitters, Dyers	*75*	4,874	1,699
19. Food and Beverage Processers	*77*	5,086	4,246
20. Tailors, Dressmakers, Sewers	*79*	17,925	13,096
21. Shoemakers and Leather Goods Makers	*80*	8,924	4,937
22. Material-Handling Equipment Operators	*97*	10,186	7,733
23. Other blue collar workers	*71 to 74, 76, 78, 81 to 96, 98*	97,164	72,494
Total		394,257	311,080

*For a description of ISCO-68 codes, see Table A1.1 in Appendix 1.1, Chapter One. See also ILO (1969) *International Standard Classification of Occupations* (Revised Edition 1968). Geneva, International Labour Office (available from: http://www.ilo.org/public/libdoc/ilo/1969/69B09_35_engl.pdf -last access: 21 May 2010).

[23] It is worth to mention that this conflation of occupations was also based on similarities of jobs in terms of formal training, areas of knowledge and, industry.

Besides presenting the observed sample proportions of women and men across occupational categories, we also provide estimates of $\hat{P}^*$, the counterfactual distribution of female workers in the absence of a differentiated treatment of their observed characteristics as explained in the previous section. Estimates of $\hat{P}^*$ were obtained using the multinomial logit coefficients from the male subsample to obtain predicted probabilities for the female subsample based on a set of observable characteristics (see Appendix 2.1). The variables included in the model are potential experience (age – years of formal schooling – 6) and its quadratic term, splines for education (0-11 years of formal schooling and 12 or more years), household head, number of infants (less than 2 years old) and children (2 to 5 years old) in the household, average schooling years amongst adult members from the rest of the household and geographical (cities) controls. It has to be said that the inclusion of all covariates is justified not only on the basis of previous labour market research in Colombia but also the statistical significance of the estimated coefficients which appears well determined in most cases. Our estimates in general reveal little in terms of surprise (see Tables 2a and 2b, above) in the sense that they predict substantially lower (higher) female sample proportions in highly female (male) dominated occupations. In the case of the informal sector, the predicted proportion of women as *housekeeping workers* using the multinomial logit coefficients from the informal male subsample is substantially lower than the observed one. Similarly, the predicted proportion of women described as *Clerical workers* in the formal sector is significantly lower than the actual sample proportion. In both segments of employment, the counterfactual estimates indicate that women are clearly overrepresented in the *Tailors, dressmakers and sewers* category and underrepresented in the *Other blue collar workers* category.

Table 2.2a: Employment proportions by gender across 16 ad-hoc occupational categories, informal workers (own-account workers -except professionals- and domestic servants), urban Colombia: 1986-1989 and 1996-1999

Occupation categories	*1986-1989*			*1996-1999*		
	Men	***Women***	$\hat{P}^{*}$	***Men***	***Women***	$\hat{P}^{*}$
Bookkeepers, cashiers and computing	0.003	0.001	0.002	0.005	0.007	0.005
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Clerical workers	0.005	0.003	0.005	0.006	0.004	0.006
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Transport and communication workers	0.001	0.000	0.002	0.004	0.001	0.007
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wholesale and Retail Trade workers	0.154	0.086	0.115	0.114	0.12	0.098
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Sales workers	0.021	0.004	0.017	0.014	0.006	0.013
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Shop assistants	0.246	0.144	0.239	0.202	0.144	0.178
	(0.002)	(0.001)	(0.000)	(0.002)	(0.002)	(0.000)
Catering and lodging workers	0.011	0.011	0.007	0.007	0.011	0.007
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Housekeeping workers	0.005	0.531	0.028	0.008	0.476	0.019
	(0.000)	(0.002)	(0.000)	(0.000)	(0.002)	(0.000)
Other personal services workers	0.026	0.053	0.028	0.035	0.072	0.04
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Farm and related workers	0.022	0.002	0.026	0.019	0.003	0.023
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Spinners, Weavers, Knitters, Dyers	0.001	0.015	0.002	0.001	0.008	0.001
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Food and Beverage Processers	0.011	0.006	0.01	0.011	0.014	0.01
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Tailors, Dressmakers, Sewers	0.017	0.117	0.016	0.013	0.099	0.014
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Shoemakers and Leather Goods Makers	0.029	0.004	0.032	0.021	0.006	0.022
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Material-Handling Equipment Operators	0.022	0	0.029	0.032	0.001	0.036
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Other blue collar workers	0.426	0.023	0.443	0.508	0.03	0.522
	(0.002)	(0.001)	(0.000)	(0.002)	(0.001)	(0.000)

Standard errors in parentheses. $\hat{P}^{*}$ is a counterfactual distribution of female workers based on multinomial logit coefficients from male subsample (see tables A2.1.1b and A2.1.1d in Appendix 2.1).

Table 2.2b: Employment proportions by gender across 23 ad-hoc occupational categories, formal workers (professionals, managers, employers and waged workers), urban Colombia: 1986-1989 and 1996-1999

Occupation categories	*1986-1989*			*1996-1999*		
	Men	***Women***	***P****	***Men***	***Women***	***P****
Engineers, technicians	0.026	0.008	0.034	0.031	0.010	0.042
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Medical workers and life scientist	0.013	0.022	0.015	0.017	0.029	0.021
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Social sciences and humanities	0.027	0.028	0.036	0.035	0.035	0.048
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Accountants	0.012	0.011	0.013	0.012	0.015	0.016
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Jurists	0.011	0.010	0.011	0.013	0.012	0.014
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Teachers	0.027	0.096	0.037	0.04	0.103	0.054
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Managers and directors	0.06	0.04	0.06	0.061	0.052	0.062
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Bookkeepers, cashiers and computing	0.036	0.063	0.051	0.037	0.075	0.05
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Clerical workers	0.039	0.185	0.05	0.04	0.16	0.049
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Transport and communication workers	0.035	0.007	0.047	0.039	0.011	0.046
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Wholesale and Retail Trade workers	0.021	0.011	0.013	0.02	0.013	0.013
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Sales workers	0.016	0.012	0.018	0.02	0.025	0.024
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Shop assistants	0.075	0.126	0.088	0.072	0.122	0.079
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Catering and lodging workers	0.005	0.005	0.003	0.005	0.005	0.003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Housekeeping workers	0.011	0.071	0.01	0.015	0.069	0.013
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Other personal services workers	0.093	0.105	0.074	0.103	0.106	0.088
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Farm and related workers	0.018	0.005	0.014	0.016	0.004	0.013
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Spinners, Weavers, Knitters, Dyers	0.014	0.015	0.012	0.007	0.005	0.006
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Food and Beverage Processers	0.018	0.008	0.014	0.018	0.009	0.014
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tailors, Dressmakers, Sewers	0.011	0.086	0.011	0.011	0.072	0.01
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Shoemakers and Leather Goods Makers	0.027	0.022	0.026	0.018	0.016	0.016
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Material-Handling Equipment Operators	0.037	0.022	0.032	0.036	0.018	0.033
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Other blue collar workers	0.369	0.041	0.329	0.337	0.034	0.286
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Standard errors in parentheses. P* is a counterfactual distribution of female workers based on multinomial logit coefficients from male subsample (see tables A2.1.1a and A2.1.1c in Appendix 2.1).

Table 2.3: Indices of occupational segregation by gender, urban Colombia: 1986-1989 and 1996-1999

Index	1986-1989		1996-1999		Total change	
	Formal	Informal	Formal	Informal	Formal	Informal
Duncan & Duncan	0.6111	0.7838	0.5710	0.7465	-0.0402	-0.0373
	(0.0016)	(0.0016)	(0.0017)	(0.0017)	(0.0024)	(0.0023)
Gini	0.4581	0.6214	0.4165	0.6254	-0.0416	0.0041
	(0.0017)	(0.0021)	(0.0017)	(0.0021)	(0.0024)	(0.0030)
Karmel & MacLachlan	0.2065	0.3048	0.1970	0.2877	-0.0095	-0.0171
	(0.0008)	(0.0011)	(0.0008)	(0.0011)	(0.0011)	(0.0015)

Estimates based on an ad hoc classification with 23 occupation outcomes in the formal segment and 16 outcomes in the informal segment (see Table 2.1, above) for workers aged 18 and 65 years. Bootstrapped standard errors in parentheses.

The same data provides persuasive evidence to suggest that occupational segregation by gender has exhibited a substantial reduction in urban Colombia for both formal and informal workers between the two selected periods 1986-1989 and 1996-1999. Besides the conventional dissimilarity index proposed by Duncan and Duncan (1955), we also computed the Gini coefficient of the distribution of jobs by gender proposed by Deutsch et al. (1994) and the index of segregation proposed by Karmel and MacLachlan (1988). All three indices suggest a reduction in the degree of dissimilarity of occupational distributions by gender amongst formal workers across the 23 occupation categories defined in this study, which are statistically different from zero according to the standard errors obtained through 500 bootstrapped iterations. In the case of informal workers, all indices suggest a similar reduction except in the case of Gini whose change is not statistically different from zero (see Table 2.3). We also observe that all segregation indices are higher for the informal sector in both periods, despite the fact that the number of occupation categories in this segment is smaller, as this excludes professional and managerial jobs in our definition.

As in Chapter 1, we implement also a decomposition of changes in the segregation indices presented above using a methodology proposed by (Deutsch et al., 2006) in order to understand the reasons behind the aforementioned reductions. According to this methodology, changes in segregation indices may be the result of variations in three components, namely, (i) occupation weights, (ii) female labour participation and, (iii) the internal gender structure of particular occupations (see Section 1.5.2 in Chapter 1 for details). Changes in the first two components have the potential of biasing the overall result of segregation indices but do not imply, necessarily, that women are less segregated within particular occupations. The results for this decomposition confirm that changes in the internal structure represent between one halve (in the case of the Duncan and Duncan index) and thre- quarters (in the case of the Karmel and McLachlan index) of the reduction amongst formal workers while they explain more than 100 per cent of the total variation amongst informal workers according to all indices reported here (see Table 2.4). The observed changes in segregation measures between 1986-1989 and 1996-1999 are not only statistically significant but can also be related perhaps to the emergence of a more egalitarian composition of occupations by gender. In other words, the decomposition results suggests the gender composition of occupation categories became more balanced over the years examined here.[24]

[24] These results, however, are not strictly comparable to those presented for the Shapley decomposition in the first chapter as the number of occupation categories in this present chapter is substantially smaller.

Table 2.4: Shapley decomposition of changes in segregation indices in urban Colombia: 1986-1989 and 1996-1999

Segregation index	*(1) Occupation weights*	*(2) Female labour participation*	*(3) Margins = (1) + (2)*	*(4) Internal structure*	*(5) Total change*
Formal workers (i)					
Duncan & Duncan	-0.0098	-0.0105	-0.0203	-0.0198	-0.0402
	24.4%	26.3%	50.6%	49.4%	100.0%
Gini	-0.0083	-0.0095	-0.0178	-0.0238	-0.0416
	19.9%	22.9%	42.8%	57.2%	100.0%
Karmel & McLachlan	0.0002	-0.0023	-0.0021	-0.0074	-0.0095
	-2.1%	24.7%	22.6%	77.4%	100.0%
Informal workers (ii)					
Duncan & Duncan	0.0036	0.0040	0.0077	-0.0449	-0.0373
	-9.7%	-10.9%	-20.5%	120.5%	100.0%
Gini	0.0077	0.0083	0.0160	-0.0120	0.0041
	189.3%	204.8%	394.2%	-294.2%	100.0%
Karmel & McLachlan	0.0015	0.0017	0.0032	-0.0203	-0.0171
	-8.8%	-9.8%	-18.6%	118.6%	100.0%

Source: household survey microdata for labour force aged 18 and 65 years. (i): based on an ad-hoc classification of 23 occupational categories (see Table 2.1 for details). (ii): based on an ad-hoc classification of 16 occupational categories which excludes professionals, managers and directives.

Counterfactual estimates based on $\hat{P}^{*}$ indicate that segregation measures in the formal sector would be reduced by around 80 per cent in 1986-1989 and by 78 per cent in 1996-1999 in the hypothetical scenario that female observable characteristics were treated as male characteristics for the purposes of job allocation in the labour market. In the case of the informal sector, the same set of indices would be reduced by around 91 and 93 per cent, respectively, for the same periods of years. This might suggests that occupational segregation is largely attributable to a differentiated treatment of observable characteristics in the labour market. But these counterfactual estimates should be interpreted with a degree of caution as they imply a drastic modification to the overall structure of workers across occupations in the labour market with obvious general equilibrium implications.

As a prelude to our empirical results, we present some differences on average characteristics between men and women in the formal and informal sectors (see Tables 2.5a and 2.5b). First and foremost, there is a sizeable reduction of the gender log wage hourly gap between 1986-1989 and 1996-1999, particularly amongst formal workers as it fell by almost three-quarters to just 0.02 log points over these years compared to a reduction of one-fifth to 0.21 log points amongst informal workers. Educational levels as measured by spline variables with two knots at 11 and 16 years of formal schooling indicate, not surprisingly, higher schooling levels amongst formal workers compared to their informal counterparts. The inclusion of education in the spline form aims at a consistent characterisation of the labour market across different occupational categories over the two periods 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, the duration of 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 for job access in the formal sector. The knot at 16 years marks the end of college education and additional years of schooling are presumably on postgraduate education or a second college degree. As indicated in section 2.2, above, they also show that women have higher schooling levels than men in the formal sector while the converse is observed amongst informal workers over all years examined here. Conversely, average potential experience (i.e., age – years of formal schooling – 6) is higher for all workers in the informal sector compared to that observed amongst formal workers. This can be explained by two aspects, first, because the average number of years of formal schooling is higher amongst formal workers (therefore reducing the life span for potential experience) and, second, because informal workers tend to be older than those in the formal sector.[25] Potential experience is also higher for men compared to women in both formal and informal sectors over the two periods of years under study. In terms of marital status, most men in the labour force are either married or in free union while the proportion of women in that state is substantially lower. A possible rationale for these differences is a gender-specific pattern of labour force participation in which women in marital relationships are less likely than their male partners to engage in paid work outside the household. In contrast, the proportion of divorced or widowed female workers is substantially higher than amongst male workers in the same marital state for both segments of employment over all the years reviewed here.

Sample proportions indicate that the composition of employment by industries is also highly differentiated in terms of gender and segments of employment. As is the case in other developing countries, informal employment in urban Colombia is highly concentrated in service sector activities. In the case of women, we observe that more than half of those included in our dataset are in the service sector; as mentioned above, nearly 50 per cent of those women work in housekeeping occupations. An important proportion of female employment in the informal sector, 25 per cent in 1986-1989 and 30 per cent in 1996-1999, is employed in retail trade activities including street sales, restaurants and small shops. In the case of male informal employment, two-thirds of

[25] Some research in Colombia (Florez 2002) suggest the existence of a 'informal-formal-informal' pattern of labour force participation whereby young workers start their labour life in informal occupations such as family workers without remuneration, then they move into the formal sector as waged workers and, after accumulating some experience and capital, they finally move back to the informal economy as owners of small firms or own-account workers.

those surveyed are concentrated in retail trade for both periods of years. The same figures indicate that one in five informal male workers in our sample are in the (mainly personal) services sector while the sample proportions in the transport sector, construction and manufacturing are considerably higher than those observed amongst the female subsample. In contrast, the employment structure in the formal sector is clearly less concentrated and differentiated in terms of gender. One out of three formal sector women and one out of four sector formal men in our sample are in the services' sector, which in this case includes government, education and health workers. And about one-tenth of our sample of formal workers from both genders groups for 1996-1999 is in the financial sector, while around a quarter are employed in manufacturing.

Table 2.5a: Mean sample values, informal workers (own-account workers - except professionals- and domestic servants), urban Colombia: 1986-1989 and 1996-1999

		1986-1989		***1996-1999***	
	Variables	***Men***	***Women***	***Men***	***Women***
	Log hourly wage	7.829	7.555	7.704	7.493
		(0.00346)	(0.00409)	(0.00357)	(0.00375)
Personal characteristics	Experience	26.53	23.60	26.03	24.99
		(0.0564)	(0.0555)	(0.0583)	(0.0576)
	spline: 0-11 yrs formal schooling	6.012	5.178	6.771	6.250
		(0.0133)	(0.0130)	(0.0143)	(0.0146)
	spline: 12-16 yrs formal schooling	0.145	0.0687	0.193	0.137
		(0.00309)	(0.00220)	(0.00375)	(0.00320)
	spline: more than 16 yrs formal schooling	0.00210	0.00115	0.00262	0.00168
		(0.000205)	(0.000161)	(0.000319)	(0.000253)
Marital status (base category: single)	Married/free union	0.726	0.360	0.701	0.431
		(0.00185)	(0.00204)	(0.00200)	(0.00219)
	Divorced/widowed	0.0544	0.247	0.0723	0.279
		(0.000942)	(0.00183)	(0.00113)	(0.00198)
Industries (base category: services and utilities: electricity, gas and water)	Agriculture	0.0154	0.00211	0.0134	0.00242
		(0.000511)	(0.000195)	(0.000503)	(0.000217)
	Mining and Quarrying	0.00269	7.22e-05	0.00264	9.76e-05
		(0.000215)	(3.61e-05)	(0.000224)	(4.36e-05)
	Manufacturing	0.108	0.165	0.105	0.146
		(0.00129)	(0.00158)	(0.00134)	(0.00156)
	Construction	0.103	0.000325	0.150	0.00117
		(0.00126)	(7.65e-05)	(0.00156)	(0.000151)
	Wholesale and retail trade, restaurants…	0.426	0.254	0.338	0.303
		(0.00205)	(0.00185)	(0.00207)	(0.00203)
	Transport, Storage and Communications	0.116	0.00258	0.170	0.00632
		(0.00133)	(0.000215)	(0.00164)	(0.000350)
	Financial, insurance and real state	0.0172	0.00363	0.0232	0.00997
		(0.000540)	(0.000255)	(0.000659)	(0.000439)
	Observations	57,982	55,426	52,312	51,229

Sample proportion for cities ommited. For sample proportions of occupational categories see Table 2.2a.

Table 2.5b: Mean sample values, formal workers (professionals, managers, employers and waged workers), urban Colombia: 1986-1989 and 1996-1999

		1986-1989		***1996-1999***	
	Variables	***Men***	***Women***	***Men***	***Women***
	Log hourly wage	8.151	8.079	8.161	8.142
		(0.00173)	(0.00213)	(0.00229)	(0.00255)
Personal characteristics	Experience	19.32	15.67	19.14	16.60
		(0.0290)	(0.0341)	(0.0344)	(0.0373)
	Experience^2	526.1	360.6	511.0	392.5
		(1.455)	(1.484)	(1.687)	(1.607)
	spline: 0-11 yrs formal schooling	7.587	8.643	8.512	9.388
		(0.00741)	(0.00934)	(0.00860)	(0.00895)
	spline: 12-16 yrs formal schooling	0.689	0.878	0.946	1.278
		(0.00382)	(0.00558)	(0.00526)	(0.00697)
	spline: more than 16 yrs formal schooling	0.0186	0.0153	0.0533	0.0556
		(0.000385)	(0.000461)	(0.000988)	(0.00116)
Marital status (base category: single)	Married/free union	0.625	0.364	0.640	0.410
		(0.00114)	(0.00153)	(0.00137)	(0.00170)
	Divorced/widowed	0.0321	0.179	0.0456	0.201
		(0.000413)	(0.00122)	(0.000596)	(0.00138)
Industries (base category: services and utilities: electricity, gas and water)	Agriculture	0.0209	0.00773	0.0189	0.00675
		(0.000336)	(0.000279)	(0.000389)	(0.00028)
	Mining and Quarrying	0.00757	0.00241	0.00474	0.00135
		(0.000203)	(0.000156)	(0.000196)	(0.00012)
	Manufacturing	0.289	0.279	0.252	0.233
		(0.00106)	(0.00143)	(0.00124)	(0.00146)
	Construction	0.0970	0.0120	0.0905	0.0126
		(0.000694)	(0.000346)	(0.000820)	(0.00038
	Wholesale and retail trade, restaurants…	0.181	0.273	0.188	0.269
		(0.000903)	(0.00142)	(0.00112)	(0.00153)
	Transport, Storage and Communications	0.0890	0.0201	0.0936	0.0255
		(0.000668)	(0.000447)	(0.000832)	(0.00054)
	Financial, insurance and real state	0.0798	0.0859	0.103	0.100
		(0.000636)	(0.000891)	(0.000870)	(0.00104)
	Observations	181,630	98,924	122,486	84,150

Sample proportion for cities omitted. For sample proportions of occupational categories see Table 2.2b.

2.5 Empirical results

2.5.1 Wage equations by gender and sector of employment

Now our attention turns to the interpretation of the wage equations estimates by gender outlined in equations (2.1) and (2.2) in which the dependent variable is the logarithm of hourly wages expressed in constant December 2008 prices, using the consumer price index for each one of the seven cities included in our sample as a deflator. Our log wage specification is austere in terms of human capital and productivity characteristics as the Colombian data do not include explicit information on labour market experience and characteristics related to specific types of education and abilities possessed by the individual in all waves. Hence, we have to rely only on those personal characteristics noted above: potential labour experience and its quadratic, dummies for two marital status in which the 'singles' category is the base group (including free union or married and widowed or divorced) and the number of years of formal schooling in the form of a piece-wise linear spline function with one knot at 11 years of formal schooling and another at 16 years. We include also geographic controls for the set of Colombian cities. The specification also features dummies for all 23 occupation categories in the case of formal workers and 16 categories in the case of informal workers (as professional are all coded as formal workers). The models are estimated without constant terms given the full set of occupations is used in both cases.

Estimates from log hourly wage equations for the informal and formal segments are presented in tables 2.6a and 2.6b, respectively. The significance of both components of labour market experience indicates that both, men and women in informal employment tend to reach a maximum on hourly log wages at a certain age and then decline thereafter. In the case of informal workers, estimates for both periods of years indicate that women tend to reach their maximum log hourly earnings after 33-34 years of finishing their schooling while men tend to do so at 39-40 years, on average and *ceteris paribus*. For those in the formal sector, their maximum returns from potential experience come some years later in all cases. In the case of men, their log hourly wages are maximised at 43 years of potential experience in 1986-1996 and 46 years in 1996-1999 while women in formal employment do so at 42 years in the former and 52 years in the latter case.

The results for our spline specification for years of schooling reveal that the coefficients are well determined amongst formal workers from both gender groups in both the 1986-1989 and the 1996-1999 periods. In the case of informal workers, they are statistically significant in all cases except for years of postgraduate or second college degree (more than 16 years of formal schooling). A possible rationale for this result is that our definition of informal workers excludes professional and managerial jobs where academic credentials are relevant. Thus, within the spectrum of occupations in the informal segment postgraduate education is probably not a pertinent determinant of hourly wages as confirmed by our spline specification estimates. The same coefficients also suggest three important policy implications. Firstly, that returns from education are higher amongst men in both sectors and time periods, as suggested by the spline coefficients which are statistically different from zero in all cases. Secondly, that the returns from secondary, college and postgraduate education in the formal sector have increased in urban Colombia for both genders between 1986-1989 and 1996-1999 while they seem to decrease for women and men in the informal sector.[26]

---

[26] T-tests of differences performed amongst these coefficients, either between women and men or across time periods, reveal that they are statistically different from zero in all cases with probability values well below one per cent in all cases. This concluding evidence drawn from

To some extent, this might reflect the effects of higher premiums to education derived from a skilled biased technological change which has been noted in the literature on this country as one of the driving forces for the increase of wage differentials between skilled and unskilled workers during 1990s (see Attanasio, et al. 2004; Cárdenas and Gutierrez, 1997; Fanjzylber and Maloney, 2005; Goldberg and Pavcnik, 2005, Isaza and Meza, 2006).

In regard to marital status, we find in the informal sector that men in a marital relationship (either married or in free union) earn higher wages not only with respect to all other male workers but also compared to all women, as indicated by the dummy coefficients from the two periods of years included here, on average and everything else the same. The results for the informal segment indicate the highest wage advantage for divorced/widowed women in 1986-1989 while, incidentally there seems to be a penalty for women in a marital relationship in 1996-1999 as their hourly wages are, on average, 3.5 per cent lower than those of single women, *ceteris paribus*. This may reflect the fact that in this sector marriage is used as a negative productivity signal for women but a positive one for men. Our results for the formal sector suggest both men and women in marital relationship tend to earn, on average, the highest wages of this group of workers in both time periods. They also indicate smaller gender differences than those found amongst informal workers although men in a marital relationship enjoy also the highest wage advantage.

---

these test (not reported here) is explained by the small size of the standard errors which are originated from large datasets.

Table 2.6a: Robust log wage equations, informal workers (own-account workers -except professionals- and domestic servants), urban Colombia: 1986-1989 and 1996-1999

		1986-1989		***1996-1999***	
	Variables	***Men***	***Women***	***Men***	***Women***
Personal characteristics	Experience	0.0222**	0.0227**	0.0200**	0.0166**
		(0.00112)	(0.00135)	(0.00109)	(0.00118)
	Experience^2	-0.00028**	-0.00034**	-0.00026**	-0.00025**
		(1.84e-05)	(2.33e-05)	(1.83e-05)	(2.04e-05)
	spline: 0-11 yrs formal schooling	0.0844**	0.0657**	0.0704**	0.0539**
		(0.00125)	(0.00169)	(0.00126)	(0.00145)
	spline: 12-16 yrs formal schooling	0.112**	0.156**	0.114**	0.129**
		(0.00524)	(0.00845)	(0.00466)	(0.00613)
	spline: more than 16 yrs formal schooling	-0.0549	-0.318*	0.00928	0.120
		(0.0849)	(0.137)	(0.0706)	(0.121)
Marital status (base category:	Married/free union	0.143**	0.0351**	0.166**	-0.0348**
		(0.00869)	(0.0116)	(0.00876)	(0.00927)
	Divorced/widowed	0.0173	0.0457**	0.0503**	-0.0153
		(0.0160)	(0.0117)	(0.0147)	(0.00979)
Industries (base category: services and utilities: electricity, gas and water)	Agriculture	0.0999	-0.246	0.0265	-0.567**
		(0.0521)	(0.256)	(0.0520)	(0.207)
	Mining and Quarrying	-0.198**	0.548*	-0.176**	0.0722
		(0.0644)	(0.272)	(0.0668)	(0.142)
	Manufacturing	0.0795**	-0.00633	0.0343**	-0.0388
		(0.0117)	(0.0379)	(0.0130)	(0.0359)
	Construction	0.0250*	0.185	0.0183	0.142
		(0.0106)	(0.119)	(0.0111)	(0.0908)
	Wholesale and retail trade, restaurants…	0.0514*	-0.0773**	-0.0326	-0.0517*
		(0.0211)	(0.0288)	(0.0220)	(0.0237)
	Transport, Storage and Communications	0.261**	0.732**	0.158**	0.739**
		(0.0112)	(0.0744)	(0.0111)	(0.0618)
	Financial, insurance and real state	0.247**	0.305**	0.134**	0.344**
		(0.0332)	(0.0889)	(0.0287)	(0.0566)

Table 2.6a: (Continuation)

		1986-1989		1996-1999	
	Variables	***Men***	***Women***	***Men***	***Women***
	Bookkeepers, cashiers and computing	6.960**	7.281**	7.026**	7.056**
		(0.0643)	(0.108)	(0.0572)	(0.0565)
	Clerical workers	6.841**	6.949**	7.135**	7.234**
		(0.0476)	(0.0763)	(0.0487)	(0.0687)
	Transport and communication workers	6.685**	6.693**	6.553**	6.890**
		(0.116)	(0.202)	(0.0592)	(0.161)
	Wholesale and Retail Trade workers	7.020**	7.060**	7.112**	7.146**
		(0.0295)	(0.0399)	(0.0315)	(0.0354)
	Sales workers	7.286**	7.620**	7.539**	7.533**
		(0.0374)	(0.0784)	(0.0435)	(0.0662)
	Shop assistants	6.649**	6.836**	6.917**	7.236**
		(0.0285)	(0.0382)	(0.0311)	(0.0354)
	Catering and lodging workers	6.864**	7.124**	7.057**	7.248**
		(0.0411)	(0.0499)	(0.0500)	(0.0495)
	Housekeeping workers	6.963**	7.093**	7.028**	7.187**
Occupations		(0.0528)	(0.0199)	(0.0404)	(0.0220)
(intercepts)	*Other personal services workers*	6.696**	7.049**	6.871**	7.265**
		(0.0272)	(0.0302)	(0.0278)	(0.0293)
	Farm and related workers	6.962**	7.575**	6.977**	7.408**
		(0.0455)	(0.244)	(0.0462)	(0.197)
	Spinners, Weavers, Knitters, Dyers	6.635**	6.828**	6.757**	6.870**
		(0.0823)	(0.0534)	(0.0957)	(0.0599)
	Food and Beverage Processers	6.838**	6.972**	6.947**	7.030**
		(0.0367)	(0.0652)	(0.0410)	(0.0552)
	Tailors, Dressmakers, Sewers	6.706**	6.871**	6.902**	7.140**
		(0.0297)	(0.0465)	(0.0361)	(0.0449)
	Shoemakers and Leather Goods Makers	6.622**	6.813**	6.758**	7.070**
		(0.0256)	(0.0648)	(0.0296)	(0.0590)
	Material-Handling Equipment Operato	6.594**	6.930**	6.793**	6.694**
		(0.0267)	(0.182)	(0.0270)	(0.142)
	Other blue collar workers	6.819**	6.975**	6.944**	7.087**
		(0.0198)	(0.0481)	(0.0224)	(0.0433)
	Observations	57,982	55,426	52,312	51,229
	Unadjusted R-squared	0.4661	0.3029	0.4355	0.3427

Robust standard errors in parentheses. ** p<0.01, * p<0.05. Dummy coefficients for cities omitted (see Table A3a in Appendix 1).

Table 2.6b: Robust log wage equations, formal workers (professionals, managers, employers and waged workers), urban Colombia: 1986-1989 and 1996-1999

		1986-1989		***1996-1999***	
	Variables	***Men***	***Women***	***Men***	***Women***
Personal characteristics	Experience	0.0290**	0.0268**	0.0230**	0.0197**
		(0.000493)	(0.000651)	(0.000586)	(0.000656)
	Experience^2	-0.00033**	-0.00032**	-0.00025**	-0.00019**
		(9.19e-06)	(1.50e-05)	(1.14e-05)	(1.54e-05)
	spline: 0-11 yrs formal schooling	0.0628**	0.0581**	0.0595**	0.0530**
		(0.000629)	(0.000992)	(0.000789)	(0.00113)
	spline: 12-16 yrs formal schooling	0.124**	0.102**	0.143**	0.124**
		(0.00149)	(0.00151)	(0.00178)	(0.00160)
	spline: more than 16 yrs formal schooling	0.0965**	0.0812**	0.154**	0.150**
		(0.0106)	(0.0141)	(0.00612)	(0.00657)
Marital status (base category: single)	Married/free union	0.0880**	0.0854**	0.0895**	0.0781**
		(0.00342)	(0.00408)	(0.00434)	(0.00449)
	Divorced/widowed	0.00948	0.0312**	0.0210*	0.0228**
		(0.00873)	(0.00532)	(0.00906)	(0.00562)
Industries (base category: services and utilities: electricity, gas and water)	Agriculture	0.0428*	-0.0166	-0.0233	-0.0632*
		(0.0169)	(0.0276)	(0.0204)	(0.0296)
	Mining and Quarrying	0.357**	0.358**	0.321**	0.158*
		(0.0178)	(0.0351)	(0.0298)	(0.0783)
	Manufacturing	0.0279**	-0.0408**	-0.0302**	-0.0885**
		(0.00398)	(0.00661)	(0.00545)	(0.00725)
	Construction	-0.0472**	0.0118	-0.0278**	0.0176
		(0.00496)	(0.0148)	(0.00707)	(0.0161)
	Wholesale and retail trade, restaurants…	-0.0601**	-0.106**	-0.115**	-0.144**
		(0.00511)	(0.00594)	(0.00652)	(0.00636)
	Transport, Storage and Communications	0.0293**	0.101**	-0.0221**	0.0390**
		(0.00564)	(0.0127)	(0.00740)	(0.0131)
	Financial, insurance and real state	-0.0116*	0.0363**	-0.0759**	0.0128
		(0.00554)	(0.00671)	(0.00672)	(0.00750)

Table 2.6b: (Continuation)

		1986-1989		1996-1999	
	Variables	***Men***	***Women***	***Men***	***Women***
Occupations (intercepts)	*Engineers, technicians*	7.616**	7.590**	7.703**	7.759**
		(0.0147)	(0.0240)	(0.0186)	(0.0281)
	Medical workers and life	7.531**	7.503**	7.652**	7.640**
		(0.0179)	(0.0197)	(0.0210)	(0.0203)
	Social sciences and humanities	7.409**	7.527**	7.480**	7.582**
		(0.0145)	(0.0193)	(0.0176)	(0.0210)
	Accountants	7.372**	7.520**	7.471**	7.608**
		(0.0169)	(0.0222)	(0.0225)	(0.0237)
	Jurists	7.534**	7.675**	7.696**	7.847**
		(0.0189)	(0.0239)	(0.0236)	(0.0256)
	Teachers	7.388**	7.396**	7.438**	7.463**
		(0.0134)	(0.0154)	(0.0165)	(0.0175)
	Managers and directors	7.538**	7.514**	7.654**	7.613**
		(0.0118)	(0.0174)	(0.0155)	(0.0192)
	Bookkeepers, cashiers and	7.231**	7.269**	7.342**	7.331**
		(0.0110)	(0.0150)	(0.0143)	(0.0167)
	Clerical workers	7.197**	7.249**	7.275**	7.321**
		(0.0106)	(0.0140)	(0.0139)	(0.0159)
	Transport and communication	7.057**	7.167**	7.137**	7.217**
		(0.0102)	(0.0224)	(0.0135)	(0.0219)
	Wholesale and Retail Trade	7.701**	7.689**	7.789**	7.715**
		(0.0173)	(0.0300)	(0.0223)	(0.0328)
	Sales workers	7.497**	7.523**	7.550**	7.466**
		(0.0156)	(0.0224)	(0.0185)	(0.0203)
	Shop assistants	7.098**	7.122**	7.209**	7.189**
		(0.0105)	(0.0143)	(0.0139)	(0.0163)
	Catering and lodging workers	7.472**	7.526**	7.526**	7.512**
		(0.0284)	(0.0411)	(0.0377)	(0.0433)
	Housekeeping workers	7.048**	7.076**	7.082**	7.097**
		(0.0142)	(0.0132)	(0.0159)	(0.0144)
	Other personal services workers	7.050**	7.200**	7.153**	7.273**
		(0.00915)	(0.0134)	(0.0125)	(0.0151)
	Farm and related workers	7.336**	7.300**	7.327**	7.420**
		(0.0206)	(0.0414)	(0.0241)	(0.0480)
	Spinners, Weavers, Knitters,	7.266**	7.134**	7.216**	7.129**
		(0.0150)	(0.0187)	(0.0193)	(0.0272)
	Food and Beverage Processers	7.109**	7.087**	7.196**	7.168**
		(0.0128)	(0.0240)	(0.0168)	(0.0263)
	Tailors, Dressmakers, Sewers	7.018**	7.043**	7.119**	7.104**
		(0.0143)	(0.0142)	(0.0176)	(0.0163)
	Shoemakers and Leather Goods	6.932**	6.975**	7.075**	7.065**
		(0.0109)	(0.0171)	(0.0163)	(0.0194)
	Material-Handling Equipment	7.105**	7.108**	7.192**	7.156**
		(0.0105)	(0.0157)	(0.0140)	(0.0181)
	Other blue collar workers	7.104**	7.075**	7.181**	7.135**
		(0.00845)	(0.0154)	(0.0116)	(0.0173)
	Observations	181,630	98,924	122,486	84,150
	Unadjusted R-squared	0.6482	0.6199	0.6709	0.6694

Robust standard errors in parentheses. ** p<0.01, * p<0.05. Dummy coefficients for cities omitted (see Table A3a in Appendix 1).

A unique feature of our econometric specification is the inclusion of the full set of occupations as equation intercepts - 23 in the case of formal workers and 16 in the case of informal workers, all of which are statistically significant at the one percent level. Most of them appear higher for women in informal employment for both periods but, given that the mean wage differences between men and women in this segment of employment are sizeable, drawing any comparisons between gender groups on occupational returns is difficult. Beyond this caveat, intercept effects for informal workers reveal that *Sales workers* are the occupation with the highest intercept effect in the informal sector for both genders over the two time periods reviewed in this chapter. The same results for the formal sector indicate that *Wholesale and Retail Trade workers* followed by *Engineers and technicians*, are the two occupation categories with the highest returns for the formal subsample. However, the same results also suggests that gender wage differences in urban Colombia tend to be associated with higher male returns from observable productivity characteristics such as education and potential labour experience.

2.5.2 The effects of segregation on the gender wage gap

In the case of informal workers, the hourly wage gap fell from 0.27 log points to 0.21 log points over the two time periods reviewed in this study. The decomposition results suggest that the main forces behind this reduction are better progress on schooling levels amongst the female informal labour force (*Endowments: levels of education*) as well as higher returns that women receive within particular occupations once the effects from all other variables included in the model are accounted for (*Treatments within occupations*). As mentioned above, the vast majority of the 16 intercept coefficients in the informal sector are higher for women than men over all years reviewed here though this observation does not take into account the fact that male wages are higher than female wages on average. However, most of the wage disadvantage against women in the informal sector is still sourced in a presumably discriminatory treatment of most of the personal characteristics included in the log wage specification, namely, education, potential experience, marital status, etc.[27]

As explained in section 2.3, above, the role of occupational segregation on the gender wage gap is analysed through the inclusion of the last two terms in the decomposition presented in Table 2.7 which capture the effects of the *explained* and *unexplained* differences in the *allocation of workers* across occupational categories. In the case of informal workers, the *explained allocation of workers* component contributes less than six per cent of the log hourly wage gap in 1986-1989 compared to almost one fifth in 1996-1999. In contrast, the unexplained allocation of workers (or the 'pure' segregation) component actually helps to reduce between one quarter and one fifth of the log wage hourly gap in the informal sector over the years reviewed here. Although the extent of disadvantage against women in terms of both wages and segregation by occupations appears to be highest in the informal sector, these results suggests that the segregated distribution of jobs across gender actually helps to attenuate the overall wage penalty on female workers in this segment of employment. A closer look at both coefficients and subsample proportions discussed above reveals that most of the

[27] We acknowledge that the specification for the informal sector may be biased due to relevant omitted variables in the determination of wages in this segment of the labour market. In particular, we are not accounting for the effects of non-human capital characteristics (i.e., physical assets) which may be relevant in the income determination process amongst own-account workers which comprise the majority of those classified as 'informal' workers in this study. This omitted variable problem may exert an upward bias in our return estimates from observed personal characteristics amongst informal workers, particularly human capital measures.

positive effect from this segregation component is driven by one single occupation, *Housekeeping workers*. This occupation represents about one half of female informal workers (see Table 2.5a, above) and its intercept effects represent, on average, about 14 percentage points of wage advantage in favour of women indicating that, once we control for the effects of other covariates, female housekeeping workers are actually enjoying an hourly pay advantage in performing such occupation in relation to other job options available to them in the informal economy of urban Colombia (see Table 2.6a, above). Thus, the combined effect from both aspects is actually helping to reduce the gender wage gap in this segment of the labour market. In other words, if those women working as *Housekeeping workers* in the informal sector of urban Colombia were working in other occupations of this segment of employment, as implied by the counterfactual scenario of the distribution of jobs by gender, our decomposition results indicate that they would secure lower wages.[28] One possible explanation for these results is that women workers in the informal sector of urban Colombia are actually finding in housekeeping jobs a better remuneration alternative to other occupations available to them in the informal economy. We believe this is an interesting finding that deserves further research in its own right.

[28] It is worth to remember that, according to results presented in Table 2.2a, the counterfactual estimates of the proportion of jobs occupied by women for the housekeeping category indicate that they would be reduced from 53.1 per cent of the female informal jobs to 2.8 per cent in 1986-1989 and from 47.6 per cent to 1.9 per cent in 1996-1999.

Table 2.7: Decomposition of log hourly wage gaps, informal and formal workers, urban Colombia: 1986-1989 and 1996-1999

	1986-1989		***1996-1999***	
	Informal	***Formal***	***Informal***	***Formal***
Treatments: returns from education	0.094**	0.059**	0.101**	0.085**
	(0.011)	(0.009)	(0.012)	(0.012)
Treatments: returns from experience	0.032	0.030**	0.079**	0.032**
	(0.022)	(0.000)	(0.020)	(0.000)
Treatments: other observables	0.159**	0.04**	0.138**	-0.009
	(0.018)	(0.006)	(0.020)	(0.008)
Endowments: levels of education	0.079**	-0.089**	0.043**	-0.100**
	(0.001)	(0.001)	(0.001)	(0.001)
Endowments: levels of experience	0.021**	0.051**	0.005**	0.029**
	(0.001)	(0.001)	(0.000)	(0.000)
Endowments: other observables	0.102**	0.024**	0.073**	0.016**
	(0.005)	(0.001)	(0.004)	(0.002)
Treatments within occupations	-0.151**	-0.039**	-0.190**	-0.020
	(0.038)	(0.014)	(0.036)	(0.018)
Explained allocation of workers	0.008**	-0.008**	0.004**	-0.018**
	(0.001)	(0.000)	(0.001)	(0.000)
Unexplained allocation of workers	-0.068**	0.005**	-0.042**	0.004*
	(0.022)	(0.002)	(0.017)	(0.002)
Hourly log wage gap	0.274**	0.073**	0.211**	0.019**
	(0.005)	(0.003)	(0.005)	(0.003)

Standard errors in parentheses. ** p<0.01, * p<0.05. See Appendix 2 for details on formulas for standard errors.

The results for the formal sector indicate not only substantially lower levels of gender hourly wage differences compared to those observed amongst their informal counterparts but also a sizeable reduction of the log wage gap, from 0.07 log points in 1986-1989 to just 0.02 in 1996-1999. The main force behind this reduction is a decline in the gender differential on returns from other observable characteristics such as marital status and industry affiliation (*Treatments: other observables*). Increasing educational levels amongst the female labour force also contributed towards reducing the hourly wage disadvantage against women in the formal sector (*Endowments: levels* *of education*). In fact, higher average schooling levels amongst female formal workers are found to reduce the wage gap over both periods of years reviewed here. The same decomposition results suggest that most of the remaining wage penalty on women in the formal sector could be explained in terms of a discriminatory treatment of both schooling (*Treatments: returns from education*) and potential labour force experience (*Treatments: returns from experience*) which receive higher returns amongst the male formal subsample over all years reviewed here. On this we should observe that the use of potential labour force experience for women may lead to an upward bias in the unequal treatment effect alluded in our decomposition results of gender wage differentials. This is because our labour force experience measure is poorly correlated with actual labour force experience for women given their labour force interruption pattern mainly due to childbearing. For a detailed discussion on this issue, see Wright and Ermisch (1991).

Both the explained and unexplained components of the allocation of workers by gender appear to contribute towards reducing the gender wage gap in the formal sector between 1986-1989 and 1996-1999, particularly for the former which accounts for about one fifth of the change over these years. Their effects on the wage differential structure between female and male formal workers appear to be the opposite of what is found in the informal sector: while the *explained allocation of workers* component appear to reduce the hourly gender wage gap, the *unexplained allocation of workers* fraction contributes to almost seven per cent of it in 1986-1986 and 19 per cent in 1996-1999.

In summary, although the gender wage gap has exhibited a substantial reduction for both formal and informal workers, most of it can still be attributed to some sort of unequal treatment in which male characteristics tend to be better rewarded than the female ones. We find that the unequal distribution of women and men across different occupations in the labour force actually helps to reduce the gender wage gap in both segments of employment over all years reviewed here. In the case of informal workers, it is the *unexplained allocation of workers* component which may be considered as a 'pure' segregation effect that reduces the wage gap between men and women. Conversely, the *Explained allocation of workers* component helps to attenuate the wage disadvantage against women in the formal segment. Thus, the effects of what may be deemed as a result of segregation (*unexplained allocation of workers*) are found to reduce the gender hourly wage gap in the informal sector, although conversely they seem to explain some of it in the formal sector. This small contribution of occupational segregation to the gender hourly wage gap in urban Colombia is in line with the findings from similar studies for Ireland (Reilly, 1991), rural China (Meng and Miller, 1995), United States (Brown et al., 1980) and United Kingdom (Miller, 1987). In general, this literature suggests that occupational segregation plays a marginal role in explaining the magnitude of gender wage differences. In an addition to the existing literature, we find that for informal women occupational segregation actually acts to reduce unequal treatment.

2.5.3 Feminisation of occupations and gender wage differences

One of the empirical regularities from the literature on the effects of occupational segregation on the gender wage gap is that occupations with a high concentration of women tend to offer lower wages for both gender groups compared to male dominated occupations (see, for example, Baker and Fortin, 2001, Bayard et al., 2003, Jurajda, 2003). We now extend our empirical analysis in order to investigate this hypothesis in the case of urban Colombia. In order to do this, we estimated wage equations by gender for the informal and formal segments of the labour market with the same controls for personal characteristics, cities and industries. In our first specification, we include the ratio of women over the total number of workers in each one of the 82 occupation categories included in the original classification of occupations available in the Colombian dataset for urban areas. It should be noted here that we are not using just the 23 and 16 occupation categories defined above for, respectively, the formal and informal segments of employment. Instead, we prefer to use the original classification with a larger (and more precisely defined) number of job categories in order to fully exploit the variance within the original information. Therefore, the gender wage equations (2.1) and (2.2) could be reformulated as

$$W^m = X^m \beta^m + FEM_j \delta_j^m + \mu^m \qquad (2.6)$$

$$W^f = X^f \beta^f + FEM_j \delta_j^f + \mu^f \qquad (2.7)$$

where *FEM* is a vector with the ratios of female workers with respect to the total number of workers in occupation *j*, δ_j^m and δ_j^f are the corresponding vectors of coefficients, while μ^m and μ^f represent the error terms from male and female equations.[29] Clearly this framework assumes that the distribution of workers across occupations is exogenously determined and there is no need to model its distribution. In order to control for the effects of occupation characteristics other than the proportion of female workers, *FEM*, our second specification includes dummy variables for each of the *ad hoc* 23 occupational categories defined earlier this paper for the

[29] This follows the approach of Baker and Fortin (2001) who implement a similar specification with data for Canada and US workers. They also use a two-step approach in which dummy coefficients for each occupation are obtained and then, regressed on a set of occupation characteristics. Their dataset is richer in terms of firm, sector and occupation characteristics which are not available for urban Colombia for the whole period covered in this study. For this reason, we limited the analysis to the one-step approach proposed by them.

formal sector (and 16 for the informal). The idea behind this strategy is that occupation category dummies should capture occupation fixed effects other than those arising from the gender composition of particular occupations. Thus, the wage equations are reformulated as

$$W^m = X^m\beta^m + FEM_j\delta_j^m + P_k\rho_k^m + \mu^m \tag{2.8}$$

$$W^f = X^f\beta^f + FEM_j\delta_j^f + P_k\rho_k^f + \mu^f \tag{2.9}$$

where P is a vector of occupational dummies for the k occupational categories, (23-1=) 22 in the formal sector and (16-1=) 15 in the informal sector and all other variables are defined as above.

The results from our first specification for the informal sector reveal that there is a wage disadvantage from female occupational intensity only for men in 1986-1989 with an implied elasticity at the average percentage female intensity of less than (-0.09*0.20=) -0.02 but the sign of this result reverses in 1996-1999.[30] In the case of formal workers, the results from our first specification clearly indicate a wage disadvantage from female occupational intensity for female workers in both periods of years with an implied elasticity of -0.11 compared to a small wage advantage for male workers with an elasticity of 0.01. An evident drawback from the first specification is that it does not control for relevant occupation characteristics which are highly correlated with wages such as physical demands, hazards and specific training requirements as demonstrated by Baker and Fortin (2001); as explained above, this type of data are not readily available for urban Colombia. Other omitted variables in our specification of determinants of wages at the level of occupation are, for instance, specific information on major degree, availability to work part time or full time, and specific occupation abilities such as people oriented skills, quantitative skills, verbal

[30] The implied elasticity is given by $\frac{\partial \ln(w)}{\partial FEM} * \overline{FEM}$.

skills and so on.[31] Our second specification attempts to address this problem by including dummies for occupation fixed effects as per specifications (2.8) and (2.9). The results for this specification for the formal sector confirm those from the first model without controls for occupation characteristics (in which female occupation intensity is associated to a wage premium for male workers and a wage penalty on female workers), but they also indicate smaller wage effects arising from female occupation intensity in most cases. The implied elasticity estimates in 1996-1999 are 0.04 for the formal male subsample and -0.07 for their female counterparts. In the case of informal workers, the coefficients for female occupation intensity suggest a negative impact on female wages in both years and a positive impact on male wages only in 1996-1999 (see Table 2.8).

Estimates of the wage penalty effect originated in female occupation intensity reported in this Chapter are not strictly comparable to those from existing studies due to differences in both the econometric specification and the structure of data collected in other countries. In addition, the division of results along the formal/informal classification of workers presented here may be inappropriate to the labour markets in industrialised economies for which most of the estimates in the literature can be compared to. Beyond these objections, some comparisons can still be made for reference. Jurajda and Harmgart (2007) report for West Germany that female occupation intensity has a positive effect on male wages with a coefficient of 0.37 and a negative effect on female wages with an estimate of -0.05. To some extent, these results are somehow similar to those for the male and female subsamples in urban Colombia,

[31] See Shauman, 2006 for an empirical application on the determinants of wages and gender sorting amongst occupations.

particularly in 1996-1999 where most of our specifications yield a positive coefficient for the former and a negative coefficient for latter.[32]

Some studies suggest that female occupational intensity entails a wage disadvantage but differences in methodology make somehow difficult to relate these results from those presented here. For instance, estimates for Czech Republic and Slovak Republic presented in Jurajda (2005) are based on pooled regressions for both female and male workers and they confirm a wage disadvantage from female occupational intensity with coefficients of -0.13 in the former and -0.1 in the latter. Using also a pooled sample approach, estimates from Bayard et al (2003) indicate a slightly larger wage penalty (-0.14) in the United States. To some extent, the findings in these and other studies using a similar approach (Groshen, 1991, Jurajda, 2003, Killingsworth et al., 1986) also confirm a female wage disadvantage from female occupational intensity.

[32] In the case of East Germany, the same authors find a positive effect according to this coefficient for both female (0.10) and male subsamples (0.12). These estimates control not only for personal characteristics but also for firm characteristics. The positive effect of female occupational intensity found in East and West Germany by Jurajda and Harmgart (2007) is difficult to rationalise in the same terms as in Urban Colombia given the structural differences prevailing between the two countries. Massive layoffs in the eastern part of Germany during the integration process leaded a selection of the most qualified women into the labour market of that part of the country. This process entailed an increase of average productive characteristics amongst the female labour force which, according to Jurajda and Harmgart (2007), explains the wage premium endorsed to the female share of occupations in that country.

Table 2.8: Occupational gender composition coefficients, informal and formal workers, urban Colombia: 1986-1989 and 1996-1999

	1986-1989		*1996-1999*	
	Men	***Women***	***Men***	***Women***
Informal workers				
Specification 1: only female shares by ISCO 68	-0.0943**	-0.00528	0.0770**	-0.00354
	(0.0206)	(0.0298)	(0.0207)	(0.0273)
Specification 2: average characteristics by	0.167**	-0.139	0.151**	-0.447**
occupation	(0.0552)	(0.0891)	(0.0581)	(0.0774)
Specification 3: dummies for 16 occupational	-0.0259	-0.201*	0.117**	-0.242**
categories	(0.0397)	(0.0862)	(0.0380)	(0.0681)
Formal workers				
Specification 1: only female shares by ISCO 68	0.0386**	-0.185**	0.0923**	-0.205**
	(0.00688)	(0.00707)	(0.00828)	(0.00890)
Specification 2: average characteristics by	-0.590**	-0.394**	-0.461**	-0.545**
occupation	(0.0192)	(0.0238)	(0.0286)	(0.0288)
Specification 3: dummies for 23 occupational	0.101**	-0.140**	0.135**	-0.115**
categories	(0.0122)	(0.0134)	(0.0154)	(0.0152)

Robust standard errors in parentheses. ** p<0.01, * p<0.05.

Overall, evidence from the literature just reviewed above is, to some extent, in line with the findings from the regression models presented above which suggest that occupational segregation of women as measured by female occupational intensity is associated with lower wages for women in the formal sector over all years reviewed here. When controlling for occupation fixed effects, a negative effect from female occupation intensity is also confirmed for women in the informal sector. Interestingly, the share of women in occupations appears to exert a positive effect on male wages for both formal and informal workers. To some extent, this might be indicative that gender wage discrimination is not ascribed to lower remuneration of typically female dominated occupations. On the contrary, this suggests that male workers in female dominated occupations enjoy some sort of wage premium. This may be because men in these occupations are situated at higher grades within these occupations. This finding deserves by itself further investigation through the examination of vertical discrimination within occupations. However, the results reported here could be taken to broadly provide confirmation that most of the persistent wage disadvantage suffered by women in urban Colombia is explained by their lower returns for observable characteristics such as education and potential labour market experience (see section 2.5.2, above). In other words, the fact that female occupational intensity entails a wage disadvantage only for women suggests that gender wage discrimination in this country is a generalised phenomenon not confined to particular occupational categories.

2.6. Final remarks

In this chapter we attempted to contribute to the existing knowledge about the relationship between occupational segregation and the gender wage gap with an empirical application to urban Colombia. An interesting feature of this case study is the existence of a highly segmented labour market in which a clear divide in terms of employment conditions, regulations and compliance with existing labour legislation can be delineated between formal and informal workers.

We confirm that informal workers are not only more segregated in terms of gender than their formal counterparts but also that the magnitude of the wage gap between women and men is at its widest and is persistent. This finding is to be anticipated given the unregulated nature of informal employment where the enforcement of labour standards and gender equality legislation is presumably weak. However, both segments of the labour force have experienced reductions not only in their levels of occupational segregation but also in terms of the wage differentials observed between women and men over the years examined here.

As was found in a few similar applications for other countries, our results indicate that differentiated returns from observable characteristics explain most of the gender wage gap for both formal and informal workers in urban Colombia. At the same time, improvements in educational levels amongst female workers acted to reduce the magnitude of wage differentials between men and women, particularly in the formal segment where the most educated tend to work. Our decomposition technique makes use of a counterfactual distribution of female employment across occupations, using multinomial logit coefficients from the male sub-samples of workers. This feature allowed us to differentiate between the explained and unexplained portions of the wage gap attributed to the unequal distributions of jobs across gender. We found that the explained portion of occupational segregation contributes towards reducing the wage gap amongst formal workers, a result that is congruent with the increasing educational levels of the female labour force in this country. Conversely, we find that the unexplained portion of occupational segregation (or what may be regarded as unjustified or 'pure' segregation) actually helps to reduce the hourly wage gap in the informal sector.

Finally, we provided some evidence suggesting that female occupation intensity is associated with lower wages for women in the formal sector of urban Colombia, a result that is consistent with the empirical findings from similar applications in other countries. Estimates for the informal sector suggest that the penalty on female wages associated with the proportion of women in occupations tends to be lower than that of the formal sector. Interestingly, our most recent estimates suggest that female occupational intensity is associated with higher male wages for both formal and informal workers. There is certainly substantial scope for further research on this matter. In particular, incorporating information on other job characteristics would provide the basis for a useful and informative exercise.

More generally, the investigation on the effects of occupational segregation on gender wage differences could be easily extended with more recent household surveys from this country. There is also some scope for improvement in the specification of the wage equations for the informal sector by the inclusion of working place conditions and other productivity related characteristics that are available from some of the waves in the household surveys conducted in urban Colombia. In addition, the issue of identification for the purposes of modelling selection effects is an issue that needs addressing and can only be done once better datasets become available. We regard these foregoing issues as important for the development of the research agenda on this topic in the future.

Appendix 2.1

Table A2.1.1a: Multinomial logit coefficients of occupational attainment, formal workers (professionals, managers, employers and waged workers), urban Colombia: 1986-1989

	Engineers, technicians and physical scientist	*Medical workers and life scientist*	*Social sciences and humanities*	*Accountants*	*Jurists*	*Teachers*	*Managers and directors*	*Bookkeepers, cashiers and computing machine operators*	*Clerical workers*	*Transport and communication workers*	*Wholesale and Retail Trade workers*
Experience	0.0422**	0.0334**	0.0266**	0.0448**	0.0893**	0.1231**	0.0785**	0.0057	0.0067	-0.1021**	0.0889**
	(0.0066)	(0.0087)	(0.0053)	(0.0092)	(0.01)	(0.0066)	(0.004)	(0.0049)	(0.0045)	(0.0048)	(0.0063)
Experience^2	0	0.0009**	0.0001	0.0003	-0.0002	-0.0013**	-0.0006**	0.0001	0.0002**	0.0016**	-0.0006**
	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
spline: 0-11 yrs formal schooling	0.5466**	0.2775**	0.2822**	0.7152**	0.7557**	0.8423**	0.3009**	0.506**	0.3477**	0.1252**	0.2364**
	(0.0226)	(0.0344)	(0.0089)	(0.048)	(0.0802)	(0.0236)	(0.0058)	(0.0083)	(0.0066)	(0.0065)	(0.0078)
spline: 12 or more yrs formal schooling	1.1276**	1.6092**	0.7701**	1.2099**	1.4246**	0.9016**	0.7067**	0.3064**	0.3047**	0.0321	0.5023**
	(0.0133)	(0.0244)	(0.0116)	(0.0181)	(0.023)	(0.0117)	(0.0103)	(0.012)	(0.0124)	(0.0185)	(0.0142)
household head	0.1233**	0.0732	-0.0567	0.259**	0.0695	-0.1044*	0.3584**	-0.0134	-0.0394	-0.4748**	1.0466**
	(0.0474)	(0.0663)	(0.0415)	(0.0674)	(0.0721)	(0.0457)	(0.0317)	(0.0347)	(0.0335)	(0.0353)	(0.0604)
Number of infants (<2 yrs) in household	-0.2027**	-0.3271**	-0.1306**	-0.14*	-0.2984**	-0.2283**	-0.1623**	-0.1175**	-0.107**	-0.1187**	-0.1043*
	(0.0466)	(0.0685)	(0.04)	(0.0641)	(0.0727)	(0.0454)	(0.0286)	(0.0326)	(0.0311)	(0.0317)	(0.0426)
Number of children (2-5 yrs) in household	-0.1354**	-0.2083**	-0.1716**	-0.0027	-0.1169**	-0.156**	-0.1587**	-0.129**	-0.1418**	-0.1274**	-0.1348**
	(0.0308)	(0.0439)	(0.027)	(0.0407)	(0.0449)	(0.0293)	(0.0187)	(0.022)	(0.0209)	(0.0213)	(0.0276)
average adult schooling yrs –rest adults in hh	0.1108**	0.1074**	0.0817**	0.079**	0.0823**	0.0977**	0.1108**	0.08**	0.076**	0.0332**	0.1226**
	(0.0056)	(0.0071)	(0.0053)	(0.0072)	(0.0074)	(0.0054)	(0.004)	(0.0048)	(0.0047)	(0.0051)	(0.0061)
Medellin	-0.0969	0.2714**	-0.6034**	-0.0825	-0.362**	0.2768**	0.3326**	-0.3053**	-0.1581**	-0.0996**	-0.0526
	(0.0519)	(0.0729)	(0.0461)	(0.0726)	(0.0851)	(0.0503)	(0.03)	(0.0387)	(0.0357)	(0.0362)	(0.0491)
Barranquilla	-0.2428**	0.1022	-0.7562**	-0.2601**	-0.0588	-0.0092	-0.2749**	-0.0609	-0.1316**	0.1395**	0.4335**
	(0.0568)	(0.0779)	(0.0571)	(0.078)	(0.0784)	(0.058)	(0.0408)	(0.0434)	(0.0434)	(0.0434)	(0.0508)
Manizales	0.1169	1.0162**	-0.3679**	0.1992	0.4721**	1.3327**	0.3366**	0.3251**	0.1559*	-0.0789	0.0118
	(0.0981)	(0.113)	(0.0894)	(0.1346)	(0.131)	(0.0717)	(0.0561)	(0.0634)	(0.0642)	(0.0725)	(0.0929)
Pasto	-0.8253**	0.443**	-0.6491**	-0.5241**	0.2766**	1.0407**	-0.4677**	0.1108	0.1554**	-0.1917**	0.0037
	(0.1005)	(0.1007)	(0.0814)	(0.125)	(0.1012)	(0.0642)	(0.0623)	(0.0601)	(0.0569)	(0.0695)	(0.0788)
Bucaramanga	0.2436**	0.8982**	-0.3915**	-0.1104	0.1535	0.5604**	0.2661**	-0.1951**	0.0178	-0.0403	0.2327**
	(0.0659)	(0.0843)	(0.0636)	(0.1028)	(0.1013)	(0.0644)	(0.0427)	(0.0556)	(0.0495)	(0.0512)	(0.0637)
Cali	-0.3071**	-0.0011	-0.4269**	-0.1385	-0.3374**	-0.0502	-0.0744*	-0.1209**	-0.1832**	-0.0191	-0.0716
	(0.0568)	(0.0768)	(0.0494)	(0.0738)	(0.0819)	(0.0579)	(0.0372)	(0.0428)	(0.0422)	(0.0423)	(0.0564)
Constant	-10.7736**	-11.401**	-6.0446**	-13.8442**	-15.7291**	-13.8821**	-6.7781**	-7.1082**	-5.5427**	-1.9289**	-8.0362**
	(0.2487)	(0.3509)	(0.1199)	(0.5195)	(0.8639)	(0.2672)	(0.0854)	(0.1099)	(0.0943)	(0.0931)	(0.1318)
Observations	5,630	4,518	7,614	3,230	2,968	14,392	14,920	12,805	25,343	7,194	4,990

Base category: other blue collar workers. Own estimates based on household survey microdata for male workers aged 18 and 65 years old in seven main metropolitan areas. Dummies for quarters omitted.

Table A2.1.1a: continuation

	Sales workers	*Shop assistants*	*Catering and lodging workers*	*Housekeeping workers*	*Other personal services workers*	*Farm and related workers*	*Spinners, Weavers, Knitters, Dyers*	*Food and Beverage Processers*	*Tailors, Dressmakers, Sewers*	*Shoemakers and Leather Goods Makers*	*Material-Handling Equipment Operators*
Experience	0.0369**	-0.0539**	0.0645**	-0.0349**	-0.0225**	0.0015	0.0466**	-0.0436**	-0.0153	-0.0504**	-0.0399**
	(0.0069)	(0.0034)	(0.012)	(0.0079)	(0.0031)	(0.0063)	(0.0078)	(0.0067)	(0.0085)	(0.0056)	(0.0047)
Experience^2	0	0.0009**	-0.0002	0.0008**	0.0006**	0.0006**	-0.0007**	0.0005**	0.0001	0.0005**	0.0006**
	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)
spline: 0-11 yrs formal schooling	0.4542**	0.093**	0.2163**	0.0039	0.0606**	-0.1047**	0.106**	-0.0955**	0.0552**	-0.0721**	-0.0436**
	(0.0116)	(0.0045)	(0.015)	(0.0104)	(0.0039)	(0.0088)	(0.0093)	(0.0084)	(0.0106)	(0.007)	(0.0059)
spline: 12 or more yrs formal schooling	0.4774**	0.1837**	0.4654**	0.0324	0.0193	0.6358**	-0.1087*	-0.1076*	-0.1352*	-0.1855**	-0.212**
	(0.0138)	(0.013)	(0.0275)	(0.0474)	(0.0173)	(0.021)	(0.0516)	(0.0528)	(0.054)	(0.0477)	(0.0419)
household head	0.2912**	-0.2888**	0.7775**	0.016	0.3135**	-0.2336**	0.2403**	0.1432**	-0.1452*	-0.3316**	-0.0392
	(0.0539)	(0.0246)	(0.1099)	(0.0598)	(0.0231)	(0.0492)	(0.0543)	(0.0462)	(0.0593)	(0.0381)	(0.033)
hh. Number of infants (<2 yrs)	-0.2056**	-0.0857**	-0.1646	-0.0853	0.0813**	-0.0124	-0.0652	0.0583	-0.035	0.0594	0.0447
	(0.0503)	(0.022)	(0.0869)	(0.0525)	(0.0188)	(0.0431)	(0.0486)	(0.0379)	(0.0522)	(0.0322)	(0.0275)
hh. Number of children (2-5 yrs)	-0.1519**	-0.0611**	-0.1034	-0.0154	-0.0022	-0.0793**	-0.1687**	-0.0584*	-0.0499	-0.0452*	-0.0142
	(0.0329)	(0.0143)	(0.054)	(0.0328)	(0.0124)	(0.0278)	(0.0328)	(0.0257)	(0.0342)	(0.0217)	(0.0182)
average adult schooling yrs –rest adults in hh.	0.1293**	0.0265**	0.0591**	-0.0417**	-0.0222**	-0.0176*	0.0256**	-0.0045	0.0108	-0.0427**	-0.0186**
	(0.007)	(0.0037)	(0.0118)	(0.0087)	(0.0033)	(0.0069)	(0.008)	(0.0071)	(0.009)	(0.006)	(0.0051)
Medellín	0.0543	-0.2986**	0.4099**	-0.1787**	-0.1461**	0.1885**	2.026**	0.1482**	0.0349	-0.8007**	0.3262**
	(0.0599)	(0.0273)	(0.0903)	(0.0643)	(0.0235)	(0.0609)	(0.0622)	(0.0489)	(0.0589)	(0.0448)	(0.0344)
Barranquilla	0.3364**	0.248**	-0.1197	0.4188**	0.03	0.3515**	-0.9861**	0.2557**	-0.1189	-0.8093**	0.5094**
	(0.0634)	(0.0302)	(0.1262)	(0.0689)	(0.0288)	(0.0723)	(0.1618)	(0.0604)	(0.0797)	(0.0615)	(0.0411)
Manizales	0.5377**	-0.1698**	0.4941**	-0.2379	-0.0041	2.1579**	0.885**	0.4546**	-0.628**	-1.1552**	0.1853**
	(0.0971)	(0.0531)	(0.1572)	(0.1297)	(0.0439)	(0.061)	(0.116)	(0.0809)	(0.1526)	(0.1124)	(0.066)
Pasto	0.2482**	-0.0553	0.1635	-0.4715**	-0.3335**	0.8721**	-1.5697**	0.1919*	-0.231*	-0.2051**	-0.5223**
	(0.0919)	(0.0462)	(0.1545)	(0.1286)	(0.044)	(0.0781)	(0.3074)	(0.0815)	(0.1164)	(0.0679)	(0.0796)
Bucaramanga	0.9166**	-0.0185	0.3256*	-0.4602**	-0.3191**	0.8902**	-1.3601**	-0.1674*	-0.4106**	0.198**	-0.1327*
	(0.0645)	(0.0365)	(0.127)	(0.1018)	(0.0355)	(0.0663)	(0.2212)	(0.0751)	(0.0995)	(0.0466)	(0.0541)
Constant	-9.0709**	-1.5454**	-8.8571**	-3.1005**	-1.6523**	-3.2566**	-5.5435**	-1.842**	-3.5485**	-0.9353**	-1.4599**
	(0.1596)	(0.0666)	(0.2767)	(0.1635)	(0.0612)	(0.1357)	(0.1531)	(0.1253)	(0.1649)	(0.109)	(0.0906)
Observations	4,004	26,079	1,393	9,038	27,232	3,739	3,978	4,100	10,438	7,031	8,906

Base category: other blue collar workers. Own estimates based on household survey microdata for male workers aged 18 and 65 years old in seven main metropolitan areas. Dummies for quarters omitted.

Table A2.1.1b: Multinomial logit coefficients of occupational attainment, informal workers (own-account workers -except professionals- and domestic servants), urban Colombia: 1986-1989

	Bookkeepers, cashiers and computing machine operators	*Clerical workers*	*Transport and communication workers*	*Wholesale and Retail Trade workers*	*Sales workers*	*Shop assistants*	*Catering and lodging workers*	*Housekeeping workers*
Experience	0.0236	0.0214	-0.0687	-0.0039	0.0443**	-0.0412**	-0.0056	-0.0868**
	(0.0261)	(0.0195)	(0.038)	(0.0045)	(0.0105)	(0.0037)	(0.0142)	(0.0186)
Experience²	0.0004	0.0002	0.0012*	0.0003**	-0.0002	0.0006**	0.0001	0.0011**
	(0.0005)	(0.0003)	(0.0006)	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0003)
spline: 0-11 yrs formal schooling	0.5499**	0.3054**	-0.0814	0.0813**	0.1892**	-0.0531**	0.0458**	-0.2969**
	(0.0461)	(0.0265)	(0.0534)	(0.0055)	(0.0131)	(0.0047)	(0.0172)	(0.0254)
spline: 12 or more yrs formal schooling	0.1448**	0.1473**	-0.2262	0.2223**	0.3223**	0.0845**	0.2092**	-0.2927
	(0.0542)	(0.0521)	(0.4065)	(0.0156)	(0.0261)	(0.0189)	(0.0438)	(0.1765)
household head	-0.0431	-0.1737	-1.0498**	0.5609**	0.2233**	0.081**	0.7889**	-1.6654**
	(0.2053)	(0.1529)	(0.2866)	(0.0371)	(0.086)	(0.0277)	(0.1252)	(0.1481)
Number of infants (<2 yrs) in household	-0.0714	0.0719	-0.4523	-0.1099**	-0.13	0.082**	-0.2505*	-0.325*
	(0.2061)	(0.1438)	(0.337)	(0.0314)	(0.0793)	(0.0234)	(0.1067)	(0.1474)
Number of children (2-5 yrs) in household	-0.2963*	-0.1364	0.1485	-0.142**	-0.1068*	0.017	-0.2586**	-0.3419**
	(0.1467)	(0.0999)	(0.1577)	(0.02)	(0.0496)	(0.0147)	(0.0677)	(0.0943)
average adult schooling yrs –rest adults household	0.1055**	0.0282	-0.048	0.0645**	0.0663**	-0.0446**	0.0137	0.1839**
	(0.0266)	(0.02)	(0.0464)	(0.0045)	(0.0104)	(0.0039)	(0.0137)	(0.0208)
Medellín	-0.7414**	0.1828	1.9617**	-0.3374**	0.6468**	0.3006**	-0.1199	-0.46*
	(0.2735)	(0.1703)	(0.3595)	(0.0378)	(0.0879)	(0.0357)	(0.1137)	(0.197)
Barranquilla	-1.3035**	-1.8465**	-0.8892	-1.0469**	-0.8099**	0.7605**	-1.0379**	0.579**
	(0.2988)	(0.3318)	(0.6615)	(0.0421)	(0.1269)	(0.0305)	(0.1454)	(0.1411)
Manizales	-0.8765	0.0626	1.5123*	0.0271	0.7188**	0.4297**	-0.0713	-0.543
	(0.7183)	(0.3938)	(0.5961)	(0.0746)	(0.1715)	(0.0708)	(0.2411)	(0.4629)
Pasto	-0.9011*	0.3738	1.0484*	-0.8544**	-0.006	0.0677	-0.7871**	-0.1252
	(0.424)	(0.2181)	(0.5191)	(0.0634)	(0.1466)	(0.0517)	(0.2105)	(0.2737)
Bucaramanga	-0.4317	-0.0434	-0.3503	-0.755**	1.6238**	0.9503**	-1.154**	-0.8593**
	(0.3097)	(0.2361)	(0.7768)	(0.0551)	(0.0824)	(0.0384)	(0.2229)	(0.3211)
Cali	-0.1675	0.2823	0.4857	-0.6945**	-0.5346**	0.4055**	-0.2755*	-0.6057**
	(0.2039)	(0.1592)	(0.4767)	(0.0411)	(0.1227)	(0.035)	(0.1167)	(0.2269)
Constant	-10.5225**	-7.0945**	-4.1287**	-2.0445**	-6.8746**	-0.0431	-4.362**	-1.8641**
	(0.6199)	(0.4109)	(0.7996)	(0.0924)	(0.2578)	(0.0784)	(0.303)	(0.4103)
Observations	240	458	84	13674	1411	22251	1253	29790

Base category: other blue collar workers. Own estimates based on household survey microdata for male workers aged 18 and 65 years old in seven main metropolitan areas. Dummies for quarters omitted.

Table A2.1.1b: -continuation

	Other personal services workers	*Farm and related workers*	*Spinners, Weavers, Knitters, Dyers*	*Food and Beverage Processers*	*Tailors, Dressmakers, Sewers*	*Shoemakers and Leather Goods Makers*	*Material-Handling Equipment Operators*
Experience	-0.0388**	-0.0226*	-0.0491	-0.0121	0.0307**	-0.0068	-0.0592**
	(0.0089)	(0.0097)	(0.0375)	(0.0148)	(0.0118)	(0.009)	(0.0108)
Experience^2	0.0008**	0.0008**	0.0009	0	-0.0003	0.0003	0.0006**
	(0.0001)	(0.0001)	(0.0006)	(0.0002)	(0.0002)	(0.0001)	(0.0002)
spline: 0-11 yrs formal schooling	-0.0229*	-0.0815**	-0.0231	-0.035*	0.0654**	-0.0266*	-0.2457**
	(0.0116)	(0.013)	(0.0516)	(0.0176)	(0.0138)	(0.0111)	(0.0142)
spline: 12 or more yrs formal schooling	0.0836	0.4664**	0.135	-0.0203	-0.1365	-0.104	-0.1383
	(0.0475)	(0.0304)	(0.1355)	(0.0796)	(0.0714)	(0.0659)	(0.1436)
household head	-0.1821**	-0.2382**	-0.3628	0.4727**	0.0621	-0.1399*	-0.1199
	(0.0671)	(0.0765)	(0.2971)	(0.114)	(0.0882)	(0.0664)	(0.0746)
Number of infants (<2 yrs) in household	-0.0662	-0.1046	-0.3027	-0.0262	-0.0481	-0.0559	0.0783
	(0.0626)	(0.0713)	(0.3118)	(0.0926)	(0.0786)	(0.0598)	(0.0611)
Number of children (2-5 yrs) in household	-0.0648	-0.106*	-0.1291	-0.0173	-0.084	0.0298	0.1117**
	(0.0388)	(0.0439)	(0.1846)	(0.0575)	(0.0493)	(0.0355)	(0.0368)
average adult schooling yrs –rest adults household	-0.0825**	0.0194	0.0686	-0.0081	-0.0117	-0.0525**	-0.1066**
	(0.0096)	(0.0103)	(0.0416)	(0.0145)	(0.0115)	(0.0092)	(0.0117)
Medellín	0.4364**	0.2315*	-0.216	-0.0587	-0.0819	-0.0229	0.8471**
	(0.0944)	(0.0974)	(0.3322)	(0.1237)	(0.0982)	(0.0815)	(0.091)
Barranquilla	0.7169**	0.6148**	-1.5726**	-0.4666**	-0.6055**	-0.3783**	0.2143*
	(0.0825)	(0.0827)	(0.5325)	(0.1286)	(0.1072)	(0.0831)	(0.0958)
Manizales	1.4836**	0.0872	-27.1787	0.8635**	-0.1073	0.2119	0.7326**
	(0.1256)	(0.2097)	(511608.9)	(0.1775)	(0.2154)	(0.156)	(0.1726)
Pasto	0.0827	0.8121**	0.4733	-0.3968*	0.3542**	0.6919**	1.5682**
	(0.1435)	(0.1076)	(0.3642)	(0.1967)	(0.1166)	(0.0876)	(0.0951)
Bucaramanga	0.4089**	0.6538**	-0.1428	0.3138*	-0.3976**	0.1211	-0.0674
	(0.1155)	(0.1042)	(0.4194)	(0.1334)	(0.142)	(0.0975)	(0.1413)
Cali	1.1061**	0.035	-0.6552	-0.4837**	-0.1004	0.141	0.1547
	(0.0807)	(0.1027)	(0.3966)	(0.1436)	(0.0979)	(0.0775)	(0.111)
Constant	-2.0636**	-3.1099**	-4.3763**	-3.2583**	-3.9416**	-2.087**	-0.5876**
	(0.1873)	(0.2163)	(0.7339)	(0.295)	(0.2384)	(0.1811)	(0.2114)
Observations	4442	1406	896	986	7487	1893	1280

Base category: other blue collar workers. Own estimates based on household survey microdata for male workers aged 18 and 65 years old in seven main metropolitan areas. Dummies for quarters omitted.

Table A2.1.1c: Multinomial logit coefficients of occupational attainment, formal workers (professionals, managers, employers and waged workers), urban Colombia: 1996-1999

	Engineers, technicians and physical scientist	*Medical workers and life scientist*	*Social sciences and humanities*	*Accountants*	*Jurists*	*Teachers*	*Managers and directors*	*Bookkeepers, cashiers and computing machine operators*	*Clerical workers*	*Transport and communication workers*	*Wholesale and Retail Trade workers*
Experience	0.003	0.0056	-0.0268**	0.0454**	0.0602**	0.0334**	0.0399**	-0.0428**	-0.0088	-0.0664**	0.0631**
	(0.0072)	(0.0096)	(0.0054)	(0.0114)	(0.0116)	(0.0065)	(0.0046)	(0.0054)	(0.005)	(0.0049)	(0.0075)
Experience^2	0.0005**	0.0009**	0.001**	-0.0001	0.0002	0.0007**	-0.0001	0.0009**	0.0004**	0.001**	-0.0003*
	(0.0002)	(0.0002)	(0.0001)	(0.0003)	(0.0003)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
spline: 0-11 yrs formal schooling	0.508**	0.2748**	0.2426**	0.8717**	1.1121**	0.7174**	0.2161**	0.4464**	0.2999**	0.1311**	0.0991**
	(0.0326)	(0.0496)	(0.0112)	(0.1419)	(0.2531)	(0.0298)	(0.0077)	(0.0123)	(0.0086)	(0.0074)	(0.0099)
spline: 12 or more yrs formal schooling	1.0808**	1.4474**	0.7689**	1.2514**	1.3265**	1.0565**	0.7126**	0.3951**	0.3318**	0.0525**	0.468**
	(0.0137)	(0.0201)	(0.0119)	(0.0204)	(0.0213)	(0.0124)	(0.0108)	(0.0123)	(0.0129)	(0.0186)	(0.0154)
household head	-0.0724	-0.1141	-0.2091**	-0.113	-0.0395	-0.2512**	0.2684**	-0.2426**	-0.2441**	-0.4769**	0.9295**
	(0.0494)	(0.0663)	(0.0424)	(0.0727)	(0.0773)	(0.0455)	(0.0359)	(0.0394)	(0.037)	(0.0372)	(0.0678)
Number of infants (<2 yrs) in household	-0.0426*	-0.0453*	-0.0204	-0.0396	-0.0525*	0.0084	-0.0574**	-0.0465**	-0.0197	-0.0221*	-0.0731**
	(0.0165)	(0.0226)	(0.0131)	(0.0241)	(0.0248)	(0.014)	(0.0109)	(0.0121)	(0.0111)	(0.0105)	(0.0164)
Number of children (2-5 yrs) in household	-0.0129	-0.0548**	-0.0188*	-0.003	-0.0354**	-0.0073	-0.0276**	-0.0215**	-0.0165**	-0.0204**	-0.0418**
	(0.009)	(0.0124)	(0.0073)	(0.0131)	(0.0135)	(0.0077)	(0.006)	(0.0067)	(0.0062)	(0.006)	(0.0091)
average adult schooling yrs –rest adults household	0.1677**	0.1499**	0.1317**	0.0951**	0.1394**	0.0656**	0.1771**	0.1232**	0.0957**	0.0436**	0.2107**
	(0.0083)	(0.0104)	(0.0076)	(0.0111)	(0.0117)	(0.0073)	(0.0062)	(0.0074)	(0.007)	(0.0073)	(0.0098)
Medellín	0.1267*	0.4049**	-0.2812**	-0.1768	-0.2603*	0.353**	0.5521**	-0.2005**	0.0167	0.0671	0.0549
	(0.0644)	(0.0949)	(0.0553)	(0.1037)	(0.1114)	(0.0665)	(0.0464)	(0.0529)	(0.0527)	(0.0505)	(0.0838)
Barranquilla	-0.0002	0.7173**	-0.4647**	0.3322**	0.4492**	0.2309**	0.0811	-0.0278	0.1099*	0.1274*	1.2974**
	(0.0628)	(0.0841)	(0.0589)	(0.0859)	(0.0874)	(0.0657)	(0.0519)	(0.0539)	(0.0557)	(0.0559)	(0.0732)
Manizales	-0.2717**	0.4653**	-0.3991**	-0.0726	-0.0548	0.6287**	0.2443**	-0.3922**	0.2125**	-0.2282**	-0.0275
	(0.0781)	(0.0997)	(0.0646)	(0.1102)	(0.113)	(0.0702)	(0.0541)	(0.0634)	(0.057)	(0.0616)	(0.0955)
Pasto	-0.578**	0.2704**	-0.6076**	-0.7792**	0.0423	0.7682**	-0.059	-0.2578**	0.1016	-0.3235**	0.4966**
	(0.0779)	(0.0955)	(0.068)	(0.1198)	(0.1)	(0.0671)	(0.0584)	(0.0636)	(0.061)	(0.0674)	(0.0873)
Bucaramanga	0.2716**	0.8255**	-0.0052	0.0411	0.3855**	0.7801**	0.7026**	0.1143	0.2298**	0.2338**	0.5783**
	(0.0764)	(0.1044)	(0.0654)	(0.1221)	(0.1164)	(0.074)	(0.0552)	(0.0628)	(0.0632)	(0.0598)	(0.0938)
Cali	-0.254**	0.0799	-0.3403**	0.109	-0.2072	0.3009**	0.4028**	-0.1615**	0.1441*	0.1429*	0.6284**
	(0.078)	(0.1117)	(0.0644)	(0.1061)	(0.1204)	(0.0743)	(0.0537)	(0.0603)	(0.0587)	(0.0565)	(0.0861)
Constant	-10.8263**	-10.9786**	-5.9355**	-16.0375**	-20.3059**	-12.6236**	-6.7732**	-6.9154**	-5.7015**	-2.7151**	-7.9689**
	(0.3603)	(0.5351)	(0.1447)	(1.5518)	(2.7765)	(0.3331)	(0.1095)	(0.1531)	(0.1213)	(0.1104)	(0.1686)
Observations	4,627	4,471	7,177	2,696	2,538	13,557	11,840	10,838	18,417	5,727	3,580

Base category: other blue collar workers. Own estimates based on household survey microdata for male workers aged 18 and 65 years old in seven main metropolitan areas.

Table A2.1.1c: -continuation

	Sales workers	*Shop assistants*	*Catering and lodging workers*	*Housekeeping workers*	*Other personal services workers*	*Farm and related workers*	*Spinners, Weavers, Knitters, Dyers*	*Food and Beverage Processers*	*Tailors, Dressmakers, Sewers*	*Shoemakers and Leather Goods Makers*	*Material-Handling Equipment Operators*
Experience	-0.0006	-0.0539**	0.0489**	-0.0009	-0.0345**	0.0354**	0.0607**	-0.0352**	-0.025**	-0.0175*	-0.0472**
	(0.0072)	(0.0038)	(0.0144)	(0.0077)	(0.0033)	(0.0079)	(0.0128)	(0.0073)	(0.0091)	(0.0075)	(0.005)
Experience^2	0.0004**	0.0008**	0.0001	0.0002	0.0008**	0.0001	-0.0007**	0.0003*	0.0004*	0.0001	0.0008**
	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0001)
spline: 0-11 yrs formal schooling	0.4305**	0.0653**	0.1278**	-0.0138	0.0944**	-0.147**	0.1427**	-0.053**	0.0409**	-0.0639**	-0.0117
	(0.0159)	(0.0054)	(0.0195)	(0.0099)	(0.0045)	(0.0105)	(0.0154)	(0.0093)	(0.0123)	(0.0094)	(0.0068)
spline: 12 or more yrs formal schooling	0.4875**	0.1888**	0.4756**	-0.0429	-0.0414*	0.6352**	-0.283**	-0.0934*	-0.2827**	-0.1243*	-0.126**
	(0.0139)	(0.0133)	(0.028)	(0.0457)	(0.0164)	(0.0229)	(0.0967)	(0.0453)	(0.0704)	(0.0527)	(0.0328)
household head	0.1056*	-0.3273**	0.8111**	-0.1694**	0.1753**	-0.4491**	0.1316	0.1037*	-0.2083**	-0.2664**	-0.1654**
	(0.0531)	(0.0283)	(0.1335)	(0.0558)	(0.025)	(0.0569)	(0.0897)	(0.0522)	(0.0671)	(0.0519)	(0.0375)
Number of infants (<2 yrs) in household	-0.0654**	-0.0326**	-0.0461	-0.0174	-0.005	-0.0238	-0.0123	0.0258	0.004	-0.0304*	0.0041
	(0.0163)	(0.008)	(0.0317)	(0.0153)	(0.0067)	(0.0157)	(0.0274)	(0.0137)	(0.0192)	(0.014)	(0.0104)
Number of children (2-5 yrs) in household	-0.0168	-0.0076	-0.0866**	0.0003	-0.0033	0.0011	0.0031	-0.0029	-0.0059	0.005	0.0013
	(0.0089)	(0.0045)	(0.0181)	(0.0086)	(0.0038)	(0.0087)	(0.0154)	(0.0079)	(0.0109)	(0.0077)	(0.0058)
average adult schooling yrs –rest adults household	0.1847**	0.0709**	0.1479**	-0.0508**	0.007	0.0043	0.0259	-0.0004	0.0268	-0.0378**	-0.0167*
	(0.0098)	(0.0057)	(0.0194)	(0.0114)	(0.0049)	(0.0112)	(0.0168)	(0.0106)	(0.0137)	(0.0111)	(0.0077)
Medellín	0.4402**	-0.2023**	0.153	0.1827*	-0.25**	-0.1764	1.7042**	0.1846*	0.4091**	0.0001	0.2361**
	(0.0754)	(0.0413)	(0.1418)	(0.0845)	(0.0342)	(0.1075)	(0.1272)	(0.0796)	(0.0817)	(0.0876)	(0.0518)
Barranquilla	0.4442**	0.5409**	0.4698**	1.1778**	0.1586**	-0.0582	-0.6405**	0.3992**	-0.222*	0.0393	0.3863**
	(0.0767)	(0.0408)	(0.1425)	(0.082)	(0.0369)	(0.1253)	(0.2165)	(0.0884)	(0.1073)	(0.1031)	(0.0579)
Manizales	0.059	-0.1966**	-0.0056	0.0471	-0.1397**	1.5773**	0.0328	0.5791**	-1.0881**	-0.5706**	-0.2375**
	(0.0901)	(0.0475)	(0.1608)	(0.0975)	(0.0385)	(0.0922)	(0.175)	(0.083)	(0.141)	(0.116)	(0.0649)
Pasto	0.1622	-0.2085**	-0.0073	-0.6401**	-0.3405**	0.497**	-2.2716**	0.2711**	-0.7151**	0.5799**	-0.7086**
	(0.0892)	(0.0503)	(0.1705)	(0.1258)	(0.0426)	(0.1082)	(0.4634)	(0.0914)	(0.1322)	(0.0915)	(0.079)
Bucaramanga	1.0364**	0.2885**	0.564**	-0.5814**	0.0459	0.8667**	-1.0463**	0.0719	-0.6793**	1.2616**	-0.0563
	(0.0814)	(0.0459)	(0.1627)	(0.1242)	(0.0398)	(0.1032)	(0.2786)	(0.0957)	(0.1317)	(0.0823)	(0.0652)
Cali	0.4426**	0.1267**	-0.1982	0.2345*	0.0723	0.028	-0.6395**	0.2981**	-0.1803	0.5141**	0.1797**
	(0.0841)	(0.0445)	(0.1825)	(0.0951)	(0.0374)	(0.1205)	(0.2215)	(0.0885)	(0.1059)	(0.0903)	(0.0594)
Constant	-9.3015**	-1.964**	-8.6483**	-3.0683**	-1.8727**	-3.5922**	-6.5223**	-2.4104**	-3.6311**	-2.3225**	-1.5707**
	(0.2057)	(0.083)	(0.3236)	(0.1672)	(0.0708)	(0.1809)	(0.2638)	(0.1526)	(0.1958)	(0.1589)	(0.108)
Observations	4,568	19,197	1,038	7,711	21,591	2,250	1,260	2,992	7,333	3,552	5,990

Base category: other blue collar workers. Own estimates based on household survey microdata for male workers aged 18 and 65 years old in seven main metropolitan areas. Dummies for quarters omitted.

Table A2.1.1d: Multinomial logit coefficients of occupational attainment, informal workers (own-account workers -except professionals- and domestic servants), urban Colombia: 1996-1999

	Bookkeepers, cashiers and computing machine operators	*Clerical workers*	*Transport and communication workers*	*Wholesale and Retail Trade workers*	*Sales workers*	*Shop assistants*	*Catering and lodging workers*	*Housekeeping workers*
Experience	-0.1564**	-0.0297	-0.0792**	-0.0169**	0.035**	-0.0215**	-0.0391*	-0.0703**
	(0.0186)	(0.0173)	(0.0201)	(0.005)	(0.013)	(0.0038)	(0.0162)	(0.0149)
Experience²	0.0031**	0.0011**	0.0012**	0.0005**	0.0003	0.0004**	0.001**	0.0009**
	(0.0004)	(0.0003)	(0.0003)	(0.0001)	(0.0002)	(0.0001)	(0.0003)	(0.0002)
spline: 0-11 yrs formal schooling	0.5277**	0.3322**	-0.0886**	0.0323**	0.2295**	-0.0446**	0.081**	-0.2267**
	(0.0572)	(0.0295)	(0.0287)	(0.0064)	(0.018)	(0.0048)	(0.0229)	(0.02)
spline: 12 or more yrs formal schooling	0.3843**	0.2511**	-0.0446	0.2293**	0.3515**	0.1615**	0.2465**	-0.0046
	(0.0353)	(0.0406)	(0.1156)	(0.0156)	(0.0251)	(0.0157)	(0.0441)	(0.0747)
household head	-0.3014	-0.2032	-1.0742**	0.5509**	0.2356*	0.1392**	0.3334*	-0.8248**
	(0.1589)	(0.1378)	(0.165)	(0.0398)	(0.1007)	(0.0279)	(0.1347)	(0.1132)
Number of infants (<2 yrs) in household	-0.0588	-0.0634	-0.0146	-0.0206*	-0.0822**	0.0101	-0.0408	-0.0514
	(0.0451)	(0.0419)	(0.0395)	(0.0102)	(0.0285)	(0.007)	(0.0352)	(0.0316)
Number of children (2-5 yrs) in household	0.0167	-0.022	-0.0036	-0.0139*	-0.0236	-0.0081*	-0.0162	-0.0046
	(0.0242)	(0.023)	(0.0225)	(0.0058)	(0.0155)	(0.004)	(0.0197)	(0.0176)
average adult schooling yrs –rest adults household	0.0829**	-0.0157	0.0266	0.0629**	0.1379**	-0.0385**	0.0564*	0.2039**
	(0.0295)	(0.0248)	(0.0356)	(0.0067)	(0.0167)	(0.0056)	(0.0234)	(0.0227)
Medellín	-0.279	0.254	1.3861**	-0.5029**	1.0959**	1.0798**	-0.0807	0.2289
	(0.2792)	(0.1977)	(0.2998)	(0.0506)	(0.1466)	(0.0599)	(0.1801)	(0.1888)
Barranquilla	-0.79**	-1.9471**	0.2754	-2.0284**	-0.5039**	1.4099**	-1.2499**	0.0747
	(0.2259)	(0.2808)	(0.3003)	(0.0581)	(0.1616)	(0.0536)	(0.1909)	(0.1639)
Manizales	0.1809	0.3921	0.7627*	-0.6383**	1.2141**	1.2207**	0.3457	0.027
	(0.2814)	(0.2125)	(0.3444)	(0.0604)	(0.1534)	(0.0634)	(0.1786)	(0.225)
Pasto	0.7448**	0.4239*	-0.0882	-0.2603**	1.0932**	-0.024	-0.5622**	-0.607*
	(0.2085)	(0.1956)	(0.3821)	(0.0495)	(0.1495)	(0.0725)	(0.2134)	(0.2432)
Bucaramanga	0.1619	-0.1573	-0.1941	-2.921**	1.0621**	1.4174**	-1.6329**	-0.368
	(0.2377)	(0.224)	(0.3833)	(0.1221)	(0.15)	(0.0579)	(0.3143)	(0.2206)
Cali	-0.5646	0.0297	-0.2987	0.0132	-0.0377	-0.2685**	0.2727	-0.1886
	(0.294)	(0.2043)	(0.3824)	(0.0448)	(0.1843)	(0.0729)	(0.1616)	(0.2062)
Constant	-8.5669**	-7.1462**	-4.4379**	-1.7722**	-8.4519**	-1.3515**	-5.211**	-3.0979**
	(0.6794)	(0.4548)	(0.6699)	(0.1069)	(0.3127)	(0.0942)	(0.3871)	(0.38)
Observations	588	521	244	12155	1088	17953	919	24920

Base category: other blue collar workers. Own estimates based on household survey microdata for male workers aged 18 and 65 years old in seven main metropolitan areas. Dummies for quarters omitted.

Table A2.1.1d: -continuation

	Other personal services workers	*Farm and related workers*	*Spinners, Weavers, Knitters, Dyers*	*Food and Beverage Processers*	*Tailors, Dressmakers, Sewers*	*Shoemakers and Leather Goods Makers*	*Material-Handling Equipment Operators*
Experience	-0.0303**	-0.0221*	-0.0023	-0.0168	0.0491**	-0.0262**	-0.0423**
	(0.0077)	(0.0105)	(0.0463)	(0.0145)	(0.0137)	(0.01)	(0.0085)
Experience²	0.0006**	0.0008**	-0.0001	0.0002	-0.0003	0.0006**	0.0004**
	(0.0001)	(0.0002)	(0.0008)	(0.0002)	(0.0002)	(0.0002)	(0.0001)
spline: 0-11 yrs formal schooling	-0.0382**	-0.1468**	0.0501	-0.0574**	0.0807**	-0.0249	-0.165**
	(0.01)	(0.0138)	(0.0585)	(0.0179)	(0.0162)	(0.0128)	(0.0106)
spline: 12 or more yrs formal schooling	0.09*	0.4144**	-0.0316	0.0998	0.0064	-0.0957	-0.2079*
	(0.0366)	(0.0361)	(0.2094)	(0.0606)	(0.0548)	(0.0647)	(0.0952)
household head	-0.1499**	-0.3028**	-0.1773	0.4771**	-0.076	0.0276	-0.0715
	(0.0564)	(0.0752)	(0.3297)	(0.1116)	(0.0944)	(0.0748)	(0.0587)
Number of infants (<2 yrs) in household	-0.013	0.0139	-0.0113	0.0312	-0.0331	0.0172	0.009
	(0.0144)	(0.0188)	(0.085)	(0.0262)	(0.0247)	(0.0179)	(0.0142)
Number of children (2-5 yrs) in household	0.0086	0.0011	0.0096	-0.0082	0.0335*	0.0012	0.0169*
	(0.0081)	(0.0107)	(0.0479)	(0.015)	(0.0135)	(0.0102)	(0.0081)
average adult schooling yrs –rest adults household	-0.0492**	-0.0391*	-0.0766	-0.0114	0.0245	-0.0521**	-0.1648**
	(0.0116)	(0.0152)	(0.0639)	(0.0197)	(0.018)	(0.0148)	(0.0126)
Medellín	0.6377**	0.0298	-1.5223*	-0.1603	-0.062	-0.1717	0.4375**
	(0.1008)	(0.1534)	(0.6418)	(0.158)	(0.1341)	(0.1277)	(0.115)
Barranquilla	-0.1638	0.2425	-1.4202**	-1.123**	-0.981**	-0.6096**	0.2212*
	(0.0977)	(0.1288)	(0.4387)	(0.1601)	(0.1325)	(0.114)	(0.1046)
Manizales	1.0634**	0.5376**	-1.4815	0.3206*	-0.3322*	-0.2851	0.4914**
	(0.1019)	(0.1506)	(0.7634)	(0.1571)	(0.1648)	(0.149)	(0.124)
Pasto	-0.142	0.5181**	0.1577	0.2479	-0.1006	0.3272**	0.7858**
	(0.12)	(0.1421)	(0.3865)	(0.1463)	(0.1404)	(0.1165)	(0.1106)
Bucaramanga	0.1436	0.3946**	-1.1159*	-0.8623**	-0.6329**	0.4047**	-0.3653**
	(0.109)	(0.1419)	(0.5337)	(0.1906)	(0.1572)	(0.1121)	(0.1308)
Constant	-2.0138**	-2.6973**	-4.9011**	-3.1433**	-4.8301**	-2.561**	-0.3391
	(0.183)	(0.2618)	(0.9493)	(0.3128)	(0.2901)	(0.2319)	(0.1945)
Observations	5514	1150	439	1254	5763	1385	1743

Base category: other blue collar workers. Own estimates based on household survey microdata for male workers aged 18 and 65 years old in seven main metropolitan areas. Dummies for quarters omitted.

Table A2.2a: Dummy coefficients for cities, robust log wage equations, informal workers, urban Colombia: 1986-1989 and 1996-1999

		1986-1989		***1996-1999***	
	Variables	***Men***	***Women***	***Men***	***Women***
Cities (base category: Bogotá)	Medellin	0.0281**	-0.0377**	-0.178**	-0.151**
		(0.0101)	(0.0126)	(0.0146)	(0.0156)
	Barranquilla	0.0119	-0.0563**	-0.197**	-0.225**
		(0.00862)	(0.0129)	(0.0120)	(0.0142)
	Manizales	-0.298**	-0.313**	-0.443**	-0.399**
		(0.0209)	(0.0228)	(0.0164)	(0.0165)
	Pasto	-0.365**	-0.481**	-0.433**	-0.501**
		(0.0137)	(0.0154)	(0.0145)	(0.0152)
	Bucaramanga	0.0773**	-0.162**	-0.0288*	-0.128**
		(0.0116)	(0.0124)	(0.0138)	(0.0154)
	Cali	-0.0165	-0.154**	-0.213**	-0.201**
		(0.00978)	(0.0122)	(0.0141)	(0.0157)
	Observations	57982	55426	52312	51229

Robust standard errors in parentheses. ** p<0.01, * p<0.05.

Table A2.2b: Dummy coefficients for cities, robust log wage equations, formal workers, urban Colombia: 1986-1989 and 1996-1999

		1986-1989		***1996-1999***	
	Variables	***Men***	***Women***	***Men***	***Women***
Cities (base category: Bogotá)	Medellin	-0.00631	-0.0126**	-0.0631**	-0.0169**
		(0.00364)	(0.00454)	(0.00587)	(0.00627)
	Barranquilla	-0.0786**	-0.144**	-0.135**	-0.104**
		(0.00455)	(0.00596)	(0.00634)	(0.00696)
	Manizales	-0.183**	-0.159**	-0.251**	-0.170**
		(0.00681)	(0.00817)	(0.00681)	(0.00739)
	Pasto	-0.301**	-0.256**	-0.320**	-0.275**
		(0.00667)	(0.00940)	(0.00713)	(0.00803)
	Bucaramanga	-0.00156	-0.0704**	-0.0259**	-0.0381**
		(0.00501)	(0.00609)	(0.00678)	(0.00701)
	Cali	-0.0946**	-0.119**	-0.116**	-0.106**
		(0.00416)	(0.00536)	(0.00678)	(0.00726)
	Observations	181630	98924	122486	84150

Robust standard errors in parentheses. ** p<0.01, * p<0.05.

Appendix 2.2: Derivation of standard errors in the decomposition of gender log hourly wage gaps

The derivation of standard errors for the different components in the decomposition outlined in equation (2.4) above is derived as follows. The standard error of the *treatments component* is defined as

$$\sqrt{VAR[\overline{\boldsymbol{X}}^{f}(\widehat{\boldsymbol{\beta}}^{m}-\widehat{\boldsymbol{\beta}}^{f})]}=\sqrt{\overline{\boldsymbol{X}}^{f'}\left[VAR(\widehat{\boldsymbol{\beta}}^{m})+VAR(\widehat{\boldsymbol{\beta}}^{f})\right]\overline{\boldsymbol{X}}^{f}} \tag{A2.1}$$

where $\overline{\boldsymbol{X}}^{f}$ is a vector of mean characteristics for the female subsample and $\widehat{\boldsymbol{\beta}}^{m}$ and $\widehat{\boldsymbol{\beta}}^{f}$ represent, respectively, the set of estimated male and female coefficients. In turn, $VAR(\widehat{\boldsymbol{\beta}}^{m})$ and $VAR(\widehat{\boldsymbol{\beta}}^{f})$ symbolize a subset of the variance-covariance matrices estimated from the male and female subsamples by Ordinary Least Squares which corresponds to the set of variables included in the model except those representing the intercepts of specific occupations. In the case of the *endowments* component, its standard error is defined as

$$\sqrt{VAR\left[(\overline{\boldsymbol{X}}^{m}-\overline{\boldsymbol{X}}^{f})\widehat{\boldsymbol{\beta}}^{m}\right]}=\sqrt{(\overline{\boldsymbol{X}}^{m}-\overline{\boldsymbol{X}}^{f})'VAR(\widehat{\boldsymbol{\beta}}^{m})(\overline{\boldsymbol{X}}^{m}-\overline{\boldsymbol{X}}^{f})} \tag{A2.2}$$

where $\overline{\boldsymbol{X}}^{m}$ is a vector of mean characteristics for the male subsample and all other terms are defined as above. The standard error for the *treatments within occupations* component is

$$\sqrt{VAR\left[\overline{\boldsymbol{P}}_{j}^{f}\left(\widehat{\boldsymbol{\gamma}}_{j}^{f}-\widehat{\boldsymbol{\gamma}}_{j}^{m}\right)\right]}=\sqrt{\overline{\boldsymbol{P}}_{j}^{f'}\left[VAR(\widehat{\boldsymbol{\gamma}}_{j}^{m})+VAR\left(\widehat{\boldsymbol{\gamma}}_{j}^{f}\right)\right]\overline{\boldsymbol{P}}_{j}^{f}} \tag{A2.3}$$

where $\overline{\boldsymbol{P}}_{j}^{f}$ represents the observed sample proportions of female workers across *j* occupations (= 16 in the informal sector and 23 in the formal one) and $VAR(\widehat{\boldsymbol{\gamma}}_{j}^{m})$ and

$VAR\left(\hat{\boldsymbol{\gamma}}_j^f\right)$ denote the subsectors of the variance-covariance matrix defined by the *j* parameters for equal number of occupation controls. In order to facilitate estimation, we assume that the sample proportions of occupations are constant so, variances in that case are approximated to their squared values. Thus, the standard error for the *explained allocations of workers* component is

$$\sqrt{VAR\left[\left(\overline{\boldsymbol{P}}_j^m - \overline{\boldsymbol{P}}_j^*\right)\hat{\boldsymbol{\gamma}}_j^m\right]} = \sqrt{\left(\overline{\boldsymbol{P}}_j^m - \overline{\boldsymbol{P}}_j^*\right)'\left[VAR\left(\hat{\boldsymbol{\gamma}}_j^m\right)\right]\left(\overline{\boldsymbol{P}}_j^m - \overline{\boldsymbol{P}}_j^*\right)} \qquad \text{(A2.4)}$$

while the standard error for the *unexplained allocation of workers* component is

$$\sqrt{VAR\left[\left(\overline{\boldsymbol{P}}_j^* - \overline{\boldsymbol{P}}_j^f\right)\hat{\boldsymbol{\gamma}}_j^m\right]} = \sqrt{\left(\overline{\boldsymbol{P}}_j^m - \overline{\boldsymbol{P}}_j^*\right)'\left[VAR\left(\hat{\boldsymbol{\gamma}}_j^m\right)\right]\left(\overline{\boldsymbol{P}}_j^m - \overline{\boldsymbol{P}}_j^*\right)}. \qquad \text{(A2.5)}$$

Chapter 3: Female intensity, trade reforms and capital investments in Colombian Manufacturing Industries: 1981-2000

3.1. Introduction

The process of trade liberalisation in developing countries has taken place at the same time that their labour markets witnessed an increase in female labour force participation to historically unprecedented levels. The effects of trade as well as other economic policies are expected to have a differentiated effect on women due to asymmetries in the distribution of rights over economic resources, as well as segregated roles in relation to both the market economy and within the household (Fontana, 2003). Although the increase in female employment over the last decades is the result of long-term development trends pertaining to demographic and cultural change, there is also a concern in the literature to understand the effects of trade reforms and other economic policies on labour market outcomes from a gender perspective.

An increasing body of economic literature has emerged in which the interactions between trade and gender differences in the labour market have been explored. From an economic perspective, trade liberalisation might affect employment dynamics by gender in at least four different ways. First, the opportunities for increasing exports, as well as competition in the form of imported goods, have both the potential of changing gender differences in the labour market if women are concentrated in sectors more exposed to trade (Collier, 1994). Second, trade liberalisation may change the relative prices of imported technology and capital goods in developing countries. Some studies indicate strong complementarities between technology and female labour (Galor and Weil, 1996, Weinberg, 2000, Welch, 2000). Third, according to the "taste for discrimination hypothesis" formulated by Becker (1957), any policy measure towards increased competition is likely to reduce the extent of discrimination against women and ethnic minorities in the labour market. A number of empirical studies have tried to identify the effects of trade policies on the unexplained portion of the gender wage gap that can be attributed to discrimination (Artecona and Cunningham, 2002, Black and Brainerd, 2004, Oostendorp, 2009, Reilly and Vasudeva-Dutta, 2005). Fourth and last, as a counterpart to Becker's hypothesis, increasing competition arising from trade liberalisation might weaken the bargaining position of women in female-intensive industries (see: Williams and Kenison, 1996, Williams, 1987, Darity and Williams, 1985). Berik et al. (2004) found in the case of Korea and Taiwan supportive evidence for this hypothesis.

Most of this literature has focused on the effects of trade on gender wage differences while the effects on the gender composition of employment have received less attention. The experience from developed economies indicates that both trade and industrialisation are closely interrelated to the gender composition of economic activities. For instance, Goldin (quoted in Galor and Weil (1996)) indicates that the necessity for fine motor skills in textiles during the industrialisation in the United Kingdom and the United States, and more recently in the electronics industry in Asian economies, represent examples of absolute and comparative advantage of female over male labour along the pathway of economic development. However, there is still a vacuum in the existing knowledge on how trade liberalisation, as well as the industrialisation process, is affecting the gender composition of employment across manufacturing activities in developing countries.

This chapter provides an empirical application to identify the effects of trade on the gender composition of employment across manufacturing industries in Colombia. In particular, we exploit a natural experiment of trade liberalisation which took place in this country at the beginning of the 1990s to assess its possible effects on the gender composition of the workforce across industrial activities. In order to account for the effects of changes in capital technology, our empirical strategy controls for different types of average stock of capital per worker (namely, machinery, office equipment and transport equipment) across manufacturing industries. We implement a panel data strategy based on fixed-effects instrumental variables (FE-IV, hereafter) in order to address potential endogeneity problems on some of the regressors. Our findings confirm that increasing levels of trade openness in the terms of both, import penetration and export orientation tend to be associated to higher shares of female employment although this effect appears to be differentiated in terms of skill level. Equally we find that manufacturing industries with higher levels of industry concentration tend to have lower female shares of jobs. Our variables for different types of the stock of capital per worker suggest that machinery and office equipment are associated with higher shares of female jobs, particularly in the white-collar workers category. The remainder of this paper is organised as follows. The following section presents the literature review and a third provides some background information for the country describing the data used for this empirical application. The fourth reports the econometric results in the light of the existing literature. The fifth and last section offers some concluding remarks.

3.2. Literature review

Trade theory provides some explanations for the effects of increased foreign competition on employment patterns between men and women. In particular, the Stolper-Samuelson theorem within the Heckscher-Ohlin-Samuelson trade model indicates that trade liberalisation increases the demand for, and the returns to, the most abundant factor of production. Thus, if women constitute the abundant factor in exporting industries boosted by trade, it is possible that their returns will grow faster than those of male workers and, in this way, the gender wage gap will be reduced. Wood (1991) provides one of the first studies to survey the relationship between trade and the gender composition of the labour force in developing countries. The author investigated the effects of trade on female employment ratios in manufacturing for a sample of countries and found that increasing exports to industrialised economies are associated with higher relative demand for female intensive goods from developing countries. But at the same time, Wood (1991) found that trade flows of manufacturing goods from the 'South', which to a great extent are intensive in female labour, were not associated with reductions in relative demand for female labour in manufacturing industries from developed countries.

In a more recent study, Chamarbagwala (2006) studied the effects of trade liberalisation on the gender (and skill) wage gap in India using a non parametric methodology developed by Katz and Murphy (1992). In addition to the Stolper-Samuelson proposition just mentioned above, this work tests the four "Skill Enhancing Trade (SET)" hypotheses proposed by Robbins (1996 –referenced in Chamarbagwala, 2006, see note 4) which indicate that trade liberalisation promotes, through different channels, the demand for and wages of skilled workers in developing countries. Chamarbagwala (2006) finds increasing skill premiums and diminishing gender wage gaps in India, the former being consistent with "skill-biased technical change" (*cfr.*, Acemoglu, 2002) and the latter due to a relocation process of female and male workers between traded and non-traded sectors.

From a theoretical point of view, trade liberalisation might affect employment dynamics by gender in at least four different ways. First, as long as women and men are imperfect substitutes in production, increased trade may affect the relative demand (as well as relative wages) of one gender group with respect to another. New opportunities arising from increasing exports, as well as more competition from imported goods, have the potential for both changing gender differences in the labour market if women are concentrated in sectors more exposed to trade (Collier, 1994). Second, trade liberalisation may change the relative prices of imported technology and capital goods in developing countries. For instance, the introduction of more capital intensive production processes in semi-industrialised economies might open new employment opportunities for women as physical strength becomes less relevant. In this sense, some studies indicate strong complementarities between technology and female labour (Galor and Weil, 1996, Weinberg, 2000, Welch, 2000). Third, according to the "taste for discrimination hypothesis" formulated by Becker (1957), any policy measure inducing increased competition is likely to reduce the extent of discrimination against women and ethnic minorities in the labour market. As long as gender discrimination is costly, increasing competition from imported goods and services is likely to increase competitive forces and reduce the scope for non-competitive behaviour in the form of discrimination (Artecona and Cunningham, 2002, Black and Brainerd, 2004). Fourth and lastly, as a counterpart to Becker's hypothesis, increasing competition arising from trade liberalisation might weaken the bargaining position of women in female-intensive industries (see: Williams and Kenison, 1996, Williams, 1987, Darity and Williams, 1985). Local entrepreneurs might respond to increasing imports with cost-cutting strategies to reduce labour costs and this might affect women if they are more concentrated in formerly protected industries. In what follows in this section, we review this literature with respect to these four hypothetical effects of trade on women.

3.2.1 Men and women as imperfect substitutes

Trade may have a differentiated effect in terms of gender because women and men may be imperfect substitutes. A recent article by Qian (2008) on the impact of tea prices and gender imbalance in China illustrates how female workers in this country have a comparative advantage in the production of that crop as "picking requires the careful plucking of whole tender leaves [which] gives adult women absolute and comparative advantages over children and men". In this case, women's comparative advantage is magnified by the fact that both the price and quality of tea leaves increases significantly with leaf tenderness. In another study for India, Rosenzweig (2004 –quoted in Duflo (2005)) documents how the choice of language instruction for boys and girls during school instruction in Mumbai entailed skill differences which became highly valuable after economic liberalisation in India over the 1990s. According to this study, low caste girls were more likely to attend English speaking schools while boys were more likely to attend Marathi-speaking schools. With the increase of service industries such as telemarketing and software as a result of economic liberalisation, the labour market returns of possessing English as a second language skill exhibited a dramatic increase. As a result, low-caste women enjoy a comparative advantage in the export-oriented service sector of Mumbai with respect to their male counterparts, with more possibilities for better wages and, to some extent, more opportunities for social mobility. Another example of imperfect substitution between men and women is provided by Goldin (quoted in Galor and Weil (1996)) who argues that the process of industrialisation is responsible for the increase in demand

(and thereby, wages) of female labour. The necessity for fine motor skills in textiles during the industrialisation in the United States and United Kingdom and, more recently in the electronics industry in Asian economies, provide examples of absolute and comparative advantage of female over male labour along the pathway of economic development.

3.2.2 The role of technology and women

Trade liberalisation has the potential of bringing about technological change or, at least, reconversion towards more capital-intensive production processes in semi-industrialised countries as imported machinery and equipment become cheaper. In the same vein, the increase in the number of foreign-owned firms might lead to the introduction of more capital-intensive production processes compared to local firms. In both cases, the question is whether the increase in capital per worker enhances the participation of women in the labour market.

Galor and Weil (1996) formalise a microeconomic model in which women and men are imperfect labour substitutes. The model has multiple steady-state equilibriums, one in which the economy has low capital per worker, high fertility rates, low female labour participation and low wages; at the other extreme, there is another equilibrium characterised by high capital per worker, low fertility rates and high relative female wages. The authors argue that countries might converge to a development trap of high fertility, low capital per worker and low productivity in which low female wages induce women to a low labour participation/high fertility outcome which in turn further reduces capital per worker. As the process of economic development allows some increase in the capital per worker, physical strength becomes less relevant and there is more scope for female labour participation. Increasing labour demand (and wages) for nonphysical strength skills, which can be supplied by women, entail an opportunity cost to childbearing as well as an incentive for reduced fertility. This in turn permits the accumulation of more capital per worker and this reinforces a cycle of higher demand for female labour, higher female wages, higher female labour participation and, ultimately, lower fertility.

In the case of the United States, Welch (2000) reviews the trends in relative female/male wages as well as wage inequality among men. His evidence is persuasive in favour of the hypothesis according to which women enjoy an advantage in cognitive skills. He finds that behind both the increasing trend in women's relative wages and growing income inequality among men in the United States there is a common factor: a growing demand for intellectual skills. Compared to average men, male workers at the top of the income distribution as well as women in general are relatively more intensive in such skills. Thus, the increase in demand for skilled labour shifts the income distribution in favour of these two groups. In the case of women, increasing schooling levels, as well as less frequent temporal withdrawals from the labour force due to maternity, might explain not only the improvement in female relative wages but also their higher work force share in a number of occupations.

In another study for the United States, Weinberg (2000) finds that the increasing use of computers accounts for about one half of the increase in demand for female workers, a finding that is in line with the hypothesis of imperfect substitution between female and male work noted above. He also proposes a microeconomic model in which the introduction of computers not only increases the share of female employment in a number of industries but also favours their demand in non-computer jobs by changing production processes in ways that are both less physically demanding and less hazardous. Based on his empirical findings, Weinberg (2000) concludes that a substitution process between highly skilled women and less skilled men, as previously documented in other studies, might be explained by the increase in computer use.

3.2.3 Trade, competition and gender discrimination

In 1957, Becker formulated an influential hypothesis in relation to labour market discrimination known as 'the taste for discrimination'. According to this hypothesis, discriminating employers and their employees are willing to sacrifice part of their income or rents in order to avoid working with people possessing some characteristics (Becker, 1957). The implication of Becker's hypothesis is, therefore, that the scope for non-competitive behaviour of firms can only be afforded through some sort of monopolistic rents which permit them to exercise their taste for discrimination against minorities in the labour market. In this sense, any policy measure towards enhanced competition should lead to the elimination of these rents and, therefore, to a reduction in the scope for costly discrimination.

There is a growing body of empirical literature in which Becker's formulation has been tested by linking trade liberalisation and gender outcomes in the labour market. This literature has focused on the effects of increased competition from trade on the magnitude of the inter-industry gender wage gaps while the effects on the gender composition of employment across economic activities have merited little attention. Two studies with a similar econometric strategy, Artecona and Cunningham (2002) and Black and Brainerd (2004), investigated the effects of increasing trade and the degree of industry concentration on the 'residual gender wage gap'.[33] The former study

[33] The residual gender wage gap is estimated as "the gender wage gap that remains after one controls for differences in education and potential labour market experience" (Black & Brainerd 2004: 544).

used data from Mexico while the latter used data from the United States. Both studies find evidence that the residual gender wage gap fell more in industries with high degree of concentration which were exposed to increased levels of foreign competition. In the same vein, Reilly and Vaseudeva (2005) investigated the relationship between trade-related measures (i.e., tariffs and imports and exports shares) on the inter-industry gender wage gap with microdata for India and found some evidence that more open sectors in that country tend to report lower levels of wage discrimination against women. In another application for Mexico, Aguayo-Téllez et al. (2010) found that trade liberalisation in this country favoured the creation of female employment in export-oriented industries at the same time that labour reallocation across sectors explains about two fifths of the increase in the female wage bill share. One of the few studies using cross sectional data is Oostendorp (2009), who investigates the effects of trade and foreign direct investment (FDI) on the gender pay gap across 161 occupations in 83 countries.[34] This study suggests that the occupational gender wage gap tends to decrease with log GDP per capita, trade and net inflows of FDI but only for richer countries while the effect on poorer countries is not statistically significant. These findings lead Oostendorp (2009) to conclude that this evidence is in line with Boserup's (1970) hypothesis according to which gender discrimination is inversely related to the level of economic development.

As noted above, the effects of trade on the gender composition of particular occupations have not yet been extensively surveyed. Most of the empirical literature has focused on the effects of trade on the gender wage gap while the implications in

[34] The dataset used in this study is the *ILO October Inquiry*, collected annually by the International Labour Organisation. It contains information on wages, earnings, and hours of work for occupations defined along the International Standard Classification of Occupations of 1968 at four digits.

terms of gender based industry segregation is yet to receive the same empirical attention. In this context, we should note Becker's assertion that

> If an individual has "taste for discrimination" he must act as if he were willing to pay something, either directly or in the form of a reduced income, to be associated with some persons instead of others (Becker, 1957: 14p.).

Here we find a segregation dimension in which discrimination not only involves a monetary cost in terms of "reduced income" but also encompasses a compositional dimension of the labour force which should be reflected in a disproportionately smaller share of women (or minority) workers in discriminating industries. In other words, as the economy becomes more liberalised, gender industry segregation should decrease in formerly protected sectors as their rents to indulge in gender discrimination shrink.

3.2.4 Trade and the bargaining position of women in the labour market

There are also alternative interpretations for the effects of trade on gender discrimination in the labour market. In a study for Korea and Taiwan, Berik et al. (2004) find a positive association between gender wage discrimination and increased levels of foreign competition in concentrated industries. The authors indicate their evidence supports a non-neoclassical hypothesis (see: Williams and Kenison, 1996, Williams, 1987, Darity and Williams, 1985) according to which increased levels of trade competition push employers to cost-cutting strategies that lessen the bargaining position of female and ethnic minority workers. A similar proposition is put forward by Seguino (2000) who argues that, in the case of semi-industrialised countries, "gender inequality has a positive effect on technical progress and growth" as low female wages provide a comparative advantage for export industries to succeed and earn the foreign currency to purchase imported capital goods, intermediate inputs and technology.

These causation links subsequently lead to reinforcing and self-fulfilling cycles of export growth, technical progress and, ultimately, economic growth. Her econometric estimates from a panel of semi-industrialised middle-income countries provide evidence of a positive relationship between gender income inequality and economic growth via two channels: (i) increased exports, technological change and growth and (ii) more investment. It should be noted that although Berik et al. (2004) and Senguino (2000) are implicitly assuming an opposite direction in the causality relationship between trade and gender wage discrimination, they ultimately concur in the notion that increasing competition arising from globalisation weakens the bargaining position of female workers in export oriented industries.

3.3. Background and data: trade liberalisation and labour markets in Colombia

3.3.1 Female share of jobs in manufacturing industries

As in other developing countries, Colombia has experienced a remarkable increase in female labour participation over the last decades. Between 1990 and 2004, female labour participation for the seven largest urban areas rose from 43.3% to 55.9% (Isaza et al., 2007). A number of factors have been cited in the literature to explain this trend. First, demographic change coupled with a smaller number of children per household has increased female labour participation in this country (Arango and Posada, 2002, Tenjo and Ribero, 1998). Second, increased educational levels amongst the female population have not only increased their probability of labour participation (Arango and Posada, 2002) but have also influenced female aspirations in terms of professional success (Gilbert, 1997). Lastly, the third factor is economic change (more closely associated with the reforms), which according to Farné, (cited in Gilbert, 1997) has encouraged the development of new occupations that fit both the skills and the social role of women. There is also some agreement in the Colombian literature that the growing labour force participation of secondary family members during the 1990s (mainly women) was motivated by an added worker effect exacerbated by adverse circumstances in the economy at the end of this decade (Isaza, 2002, Isaza, 2006, Santa María and Rojas, 2001, Tenjo and Ribero, 1998).[35]

Employment estimates of the female share of jobs across manufacturing industries for this empirical application are based on data from the Annual Manufacturing Survey (AMS hereafter) administered by the National Statistical Administrative Department (DANE, from its initials in Spanish). The survey can be considered as a census in the sense that it is gathered annually amongst nearly all manufacturing enterprises with more than ten workers since 1975. The economic classification under which the survey was collected from 1981 to 2000 is the International Standard Industrial Classification –ISIC, Rev. 2. Data for subsequent years were gathered using the ISIC Rev.3 which renders unfeasible comparisons with previously collected data. Figure 3.1a displays the total number of both, female and male workers across two broad categories, white collar and blue collar. This broad characterisation, on which we base subsequent

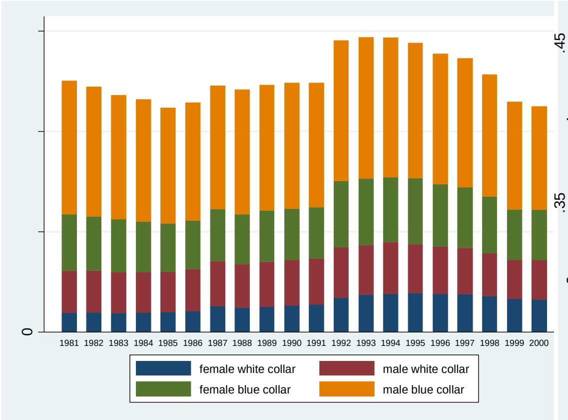
[35] It is noted that the long term trend of increasing real wages may have played an important role in the increasing female labour participation reported in urban Colombia. According to figures from Isaza *et al.* (2007), mean labour incomes rose 21.3% among men and 8.8% among women in the seven largest cities of this country between 1990 and 2004. Although it has not been found specific research on this regard for urban areas of this country, growing female earnings in the labour market may have entailed higher opportunity costs to households' fertility and, thus, increased the participation of women in the labour market. This interpretation is in line with the formulation given by Welch (2000) for the United States where the growing demand for intellectual skills explains not only the improvement in female wages but also their higher female labour force participation.

analyses, is preferred to other dis-aggregations of the labour force as the AMS was subject to changes in the questionnaires over the years analysed here regarding the classification of workers. It should be observed that other divisions of the labour force, namely by skill, hierarchical and contractual status, are not possible for the whole time period from 1981 to 2000. From the figures presented in Figure 3.1a, we observe a stagnation pattern in the employment dynamics of Colombian manufacturing industries for all groups analysed here where only in the case of female white collar workers is there an absolute increase in the number of jobs between the beginning of the 1980s and the end of 1990s. This sluggish pattern in employment growth could be attributed to a number of factors including an increased exit rate of plants after the introduction of trade liberalisation reforms introduced in 1990 (Eslava et al., 2009), weaker demand for Colombian manufactured goods internally due to a severe economic downturn at the end of the 1990s, as well as a less competitive position of Colombian manufacturing exports due to the appreciation of the Colombian currency for most of that decade (Ocampo et al., 2004). Goldberg and Pavcnik (2003) argue that labour market rigidities (rather than trade liberalisation) are also a major factor contributing to the informalisation of urban employment –and thus, the stagnation of formal employment in manufacturing firms over the 1990s.

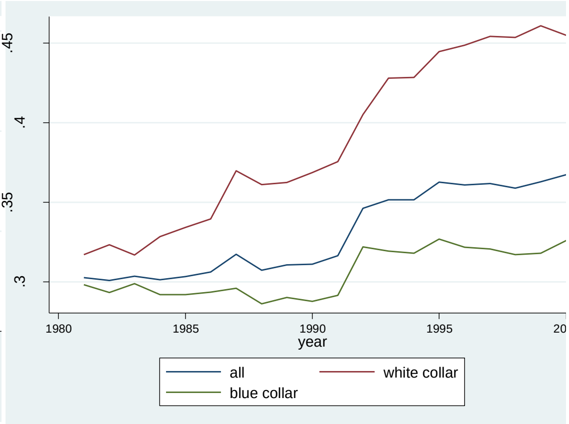
The same figures provide the basis for the calculation of the percentage of female jobs by skill level in manufacturing (see Figure 3.1b). They indicate that the female share of jobs for all workers rose from around 30 per cent at the beginning of the 1990s to more than 36 per cent from 1995 onwards. This increase was more pronounced amongst white collar workers as their share of female jobs increased from 31.7 percent in 1981 to 45.5 per cent in 2000 compared to a more modest rise from 29.8 per cent to 32.6 per cent in the case of blue collar workers over the same years. These trends are in line with the findings in the literature reviewed in section 3.2, above, according to which increasing female labour force participation is concomitant with the process of economic development.

Figure 3.1: Number of jobs and gender composition of employment across white and blue collar workers and gender in all manufacturing industries, Colombia: 1981-2000

a) Number of jobs



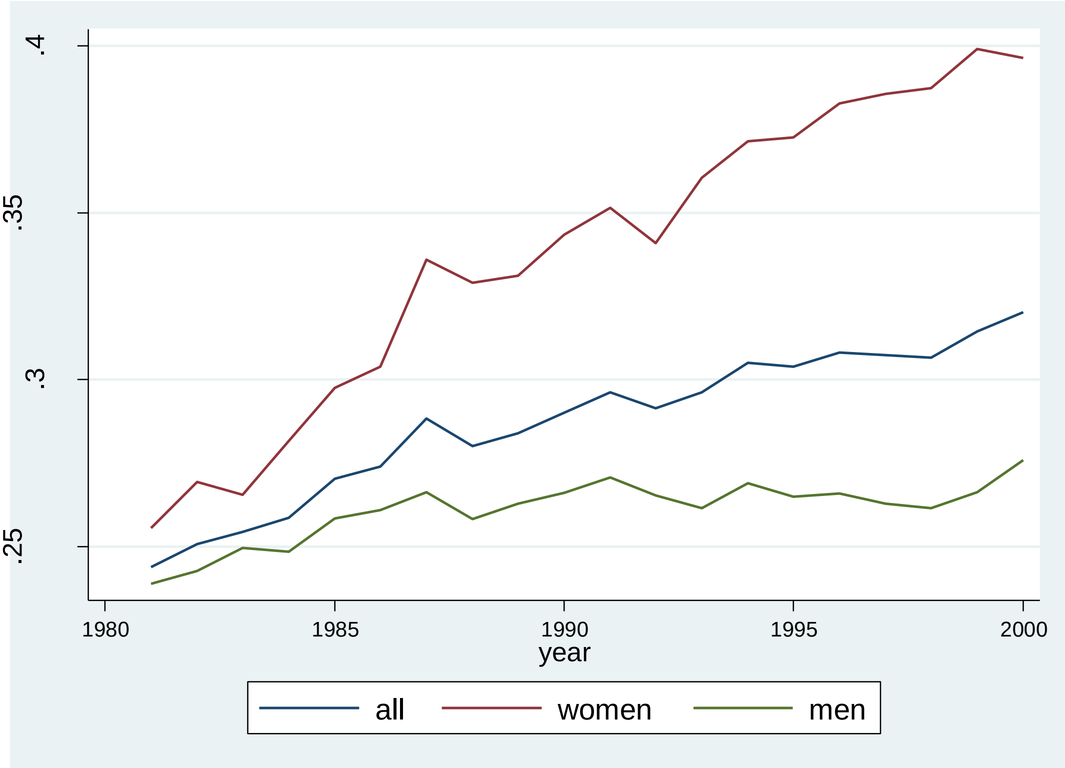
b) % of female jobs



Own estimates based on Annual Manufacturing Survey microdata.

The structure of manufacturing employment in Colombia has also experienced a structural transformation in terms of the skill composition of the labour force over the years analysed here. Employment figures from the AMS indicate that the percentage of white collar jobs has grown for both men and women although, this increase has been more pronounced amongst the latter (see Figure 3.2). These trends suggest that the process of economic development in Colombia has favoured a structural transformation of the manufacturing employment composition by skill level in which the increasing proportion of white collar workers is benefiting on the margin the incorporation of more women into the manufacturing labour force. This finding could be rationalised in terms of the literature reviewed in section 3.2.2, above (Galor and Weil, 1996, Welch, 2000), according to which the incorporation of technology in production processes is complementary to both, the demand of skilled workers and female labour.

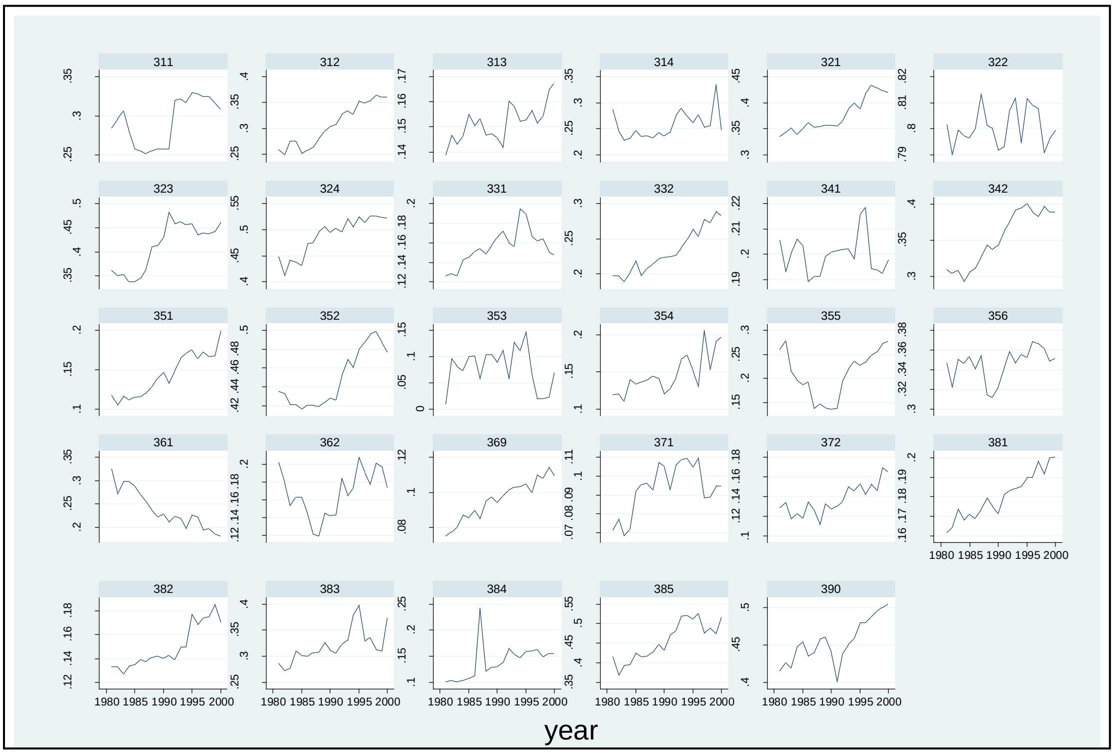
Figure 3.2: % white collar jobs by gender in all manufacturing industries, Colombia: 1981-2000



Own estimates based on Annual Manufacturing Survey microdata.

The increasing proportion of women reported in Figure 3.1b above, can also be plotted across 29 manufacturing industries using the ISIC Rev. at three digits (see Figure 3.3). With the exception of *353- Petroleum refineries* and *361- Pottery, china and earthenware*, all other manufacturing industries have increased the share of female workers within their labour force over these years. They indicate also that most of the industries with the highest female intensity over most years are those related to the textile-clothing-footwear production chain, this is, *322- Wearing apparel, except footwear*, *324- Footwear*, *323- Leather and products of leather* and, *321- Textiles*. These could be characterised as light industry in which production processes are intensive in both female labour and fine motor skills. Other industries have also experienced important increases in the female share of jobs. This is the case of *385- Measuring & controlling equipment*, *312- Food for animals*, and *342- Printing, publishing and allied* *industries* where most of the increment in the proportion of women workers took place in the form of more jobs into the white-collar category.

Figure 3.3: Proportion of female jobs across manufacturing industries, Colombia: 1981-2000



Own estimates based on Annual Manufacturing Survey microdata.

3.3.2 Tariffs and trade

Trade reforms in Colombia at the beginning of the 1990s evolved around two elements. The first one was the signing of trade agreements with México and Chile, on the one hand, and with the Andean countries of Venezuela, Ecuador, Peru and Bolivia, on the other. The second element was a reduction of the protective structure. According to Attanasio et al. (2004), Goldberg and Pavcnik (2005b, 2005a) and Jaramillo and Tovar (2006), one of the interesting features of Colombia is that this country did not participate in the GATT negotiations for the reduction of trade tariffs, so the level of protection was very high before the reforms. The removal of trade barriers was started in 1990 with the idea of a gradual approach over a time horizon of more than three years including the elimination of non-tariff barriers and reductions in both the number and level of import tariffs which were assumed to be complemented with a policy of exchange rate depreciation. Macroeconomic circumstances such as high inflation and a dramatic increase in the inflow of foreign capital, besides a reduction in trade flows (both, imports and exports), compounded a scenario in which Colombian authorities decided to speed up the liberalisation process. Thus, the initial liberalisation schedule for 1994 was completed in terms of non-tariff barriers and import tariffs by the end of 1991 (Edwards, 2001).

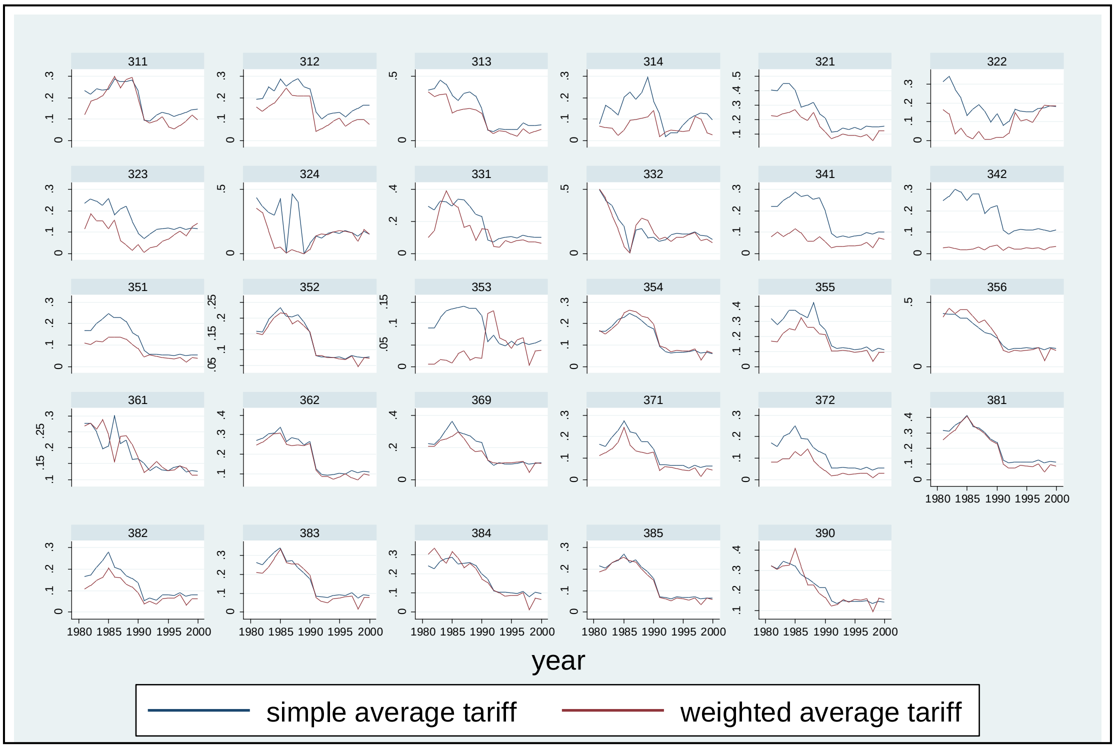
In order to measure the degree of trade openness in Colombia, we use in this empirical application a number of trade measures including import tariff data from the National Planning Department. Import tariffs were originally reported at eight-digit level according to the Nandina[36] classification. For expositional purposes of this analysis, we collapsed these data into 29 sectors defined by the ISIC Rev.2 at three-digit level in order to match it with the employment data presented in section 3.3.1, above (see Figure 3.4). According to these estimates, weighted average import tariffs for all manufacturing industries fell from 16.9 per cent in 1981-1984 to 6.4 per cent in 1997-2000.[37] The largest reductions on weighted tariffs over these years (all of which were more than 20 percentage points) were reported on *356- Plastic products*, *313- Beverage industries*, *384- Transport equipment*, *381- Fabricated metal products* and, *332- Furniture and fixtures*. Some studies for this country suggest that industries with a high intensity of unskilled labour were more protected before the reforms and thus, experienced the largest reductions in tariffs during the liberalisation period (Attanasio

[36] This is a harmonised trade classification for Andean countries.

[37] Weights are based on imports value in US dollars.

et al., 2004, Goldberg and Pavcnik, 2003, Goldberg and Pavcnik, 2005b, Goldberg and Pavcnik, 2005a, Jaramillo and Tovar, 2006).

Figure 3.4: Simple and weighted average tariffs across manufacturing industries, Colombia: 1981-2000

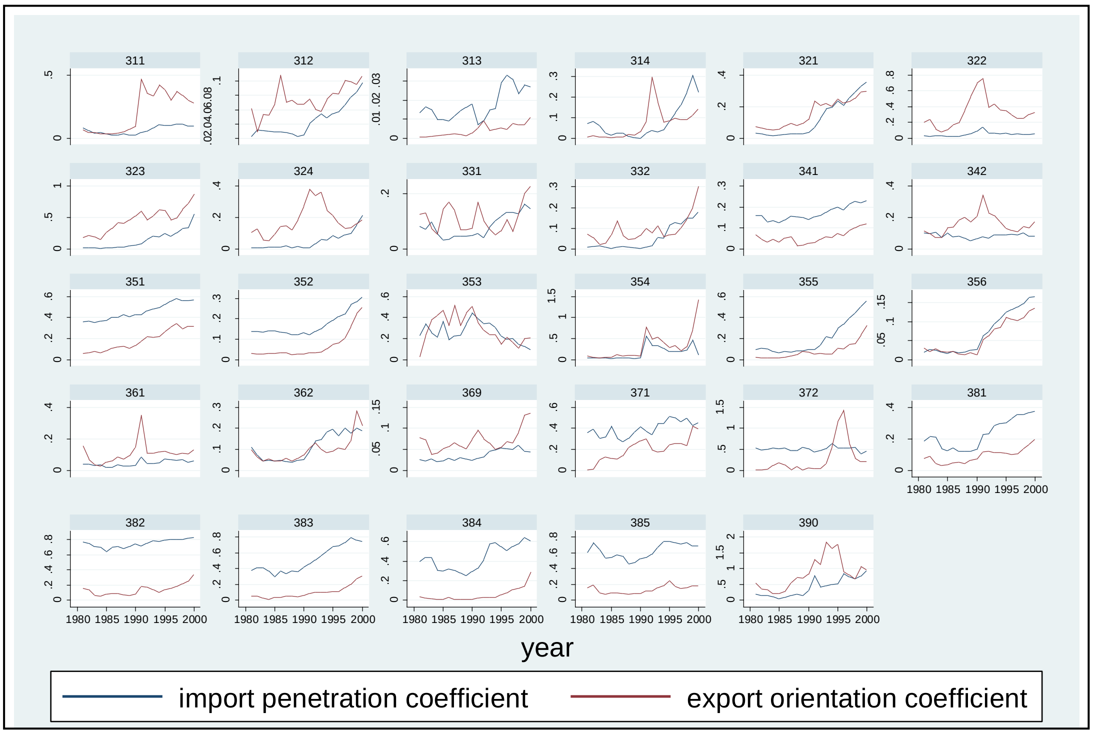


Own estimates based on tariff data from National Planning Department -DNP. Weights are based on import values in Col Pesos.

It should also be remarked that the process of tariff removal in Colombia was initiated in some industries in the early 1980s from which *332- Furniture and fixtures*, *322- Wearing apparel, except footwear* and, *321- Textiles* experienced reductions of more than ten percentage points over the pre-reform period (1985-1989) so, their reductions during the reform period (1990-1994) were more modest compared to other manufacturing industries. As a result of this process, the manufacturing industries with the lowest level of import tariffs over the post-reform period were mainly producers of intermediate goods such as *372- Non-ferrous metal basic* *industries*, *351- Industrial chemicals*, *353- Petroleum refineries*, *371- Iron and steel basic industries*, *354- Products of petroleum and coal* and, *352- Other chemical products*.

Some studies have previously used tariff data in order to assess the effects of trade policy on employment outcomes in Colombia (Attanasio et al., 2004, Goldberg and Pavcnik, 2003, Goldberg and Pavcnik, 2005b, Jaramillo and Tovar, 2006). In particular, Jaramillo and Tovar (2006) claim that tariff rates are "the most direct measure of trade policy available" in the Colombian case. But other important direct measures of trade policy such as Non-tariff barriers (NTBs hereafter), on the other hand, are only available after 1991 and, therefore, tariff rates provide a just a partial picture of trade policy in Colombia. For this reason, we focus our analysis on two commonly used indicators of trade policy, import penetration coefficient (*IPC)* and export orientation coefficient (*EOC*) that are readily available from the National Planning Department at three-digit level of the ISIC Rev.2. We believe that these measures represent superior indicators of trade policy as they display changes in trade flows, which are the ultimate objective of changes in the trade regime. The *IPC* measures the share of the domestic market in a given industry that is supplied with imports while the *EOC* indicates the percentage of domestic production in a given industry that is exported to other countries and thus, provides a crude measure of comparative advantage. The results for these trade measures are presented in Figure 3.5 and provide convincing evidence that most of Colombian manufacturing industries became more open in terms of both import penetration and export orientation. The *IPC* indicates that imported goods represented 18.9 per cent of the internal demand of all manufacturing goods in 1981-1985 and 32.4 per cent in 1996-2000. In general, only two out of 30 manufacturing industries examined here (*353- Petroleum refineries* and *342- Printing, publishing and allied industries*) report a reduction in this coefficient after trade liberalisation in 1991. The same figures indicate that the industries with the largest increments in import penetration over these years were *390- Other Manufacturing Industries*, *355- Rubber* *products*, *383- Electrical machinery apparatus, appliances*, *354- Products of petroleum and coal*, *323- Leather and products of leather* and, *321- Textiles*. In turn, *EOC* suggests that while 6.9 per cent of the domestic manufacturing product of traded goods in 1981-1984 was exported, this proportion grew to 21.3 per cent in 1996-2000. According to this coefficient, all manufacturing industries, except *353- Petroleum refineries*, became more export-oriented over these years. The largest increments in the *EOC* over this period were reported by *390- Other Manufacturing Industries*, *354- Products of petroleum and coal*, *323- Leather and products of leather* and, *372- Non-ferrous metal basic industries*, all of which experienced increases of more than 30 percentage points. It is worth to mention that *323- Leather and products of leather* and *321- Textiles*, the two sectors with the highest proportion of female workers (see 3.3.1 section, below), reported large increments in both export orientation and import penetration.

Figure 3.5: Import Penetration and Export Orientation coefficients across manufacturing industries, Colombia: 1981-2000



Source: National Planning Department -DNP.

3.3.3 Concentration, market power and trade reforms

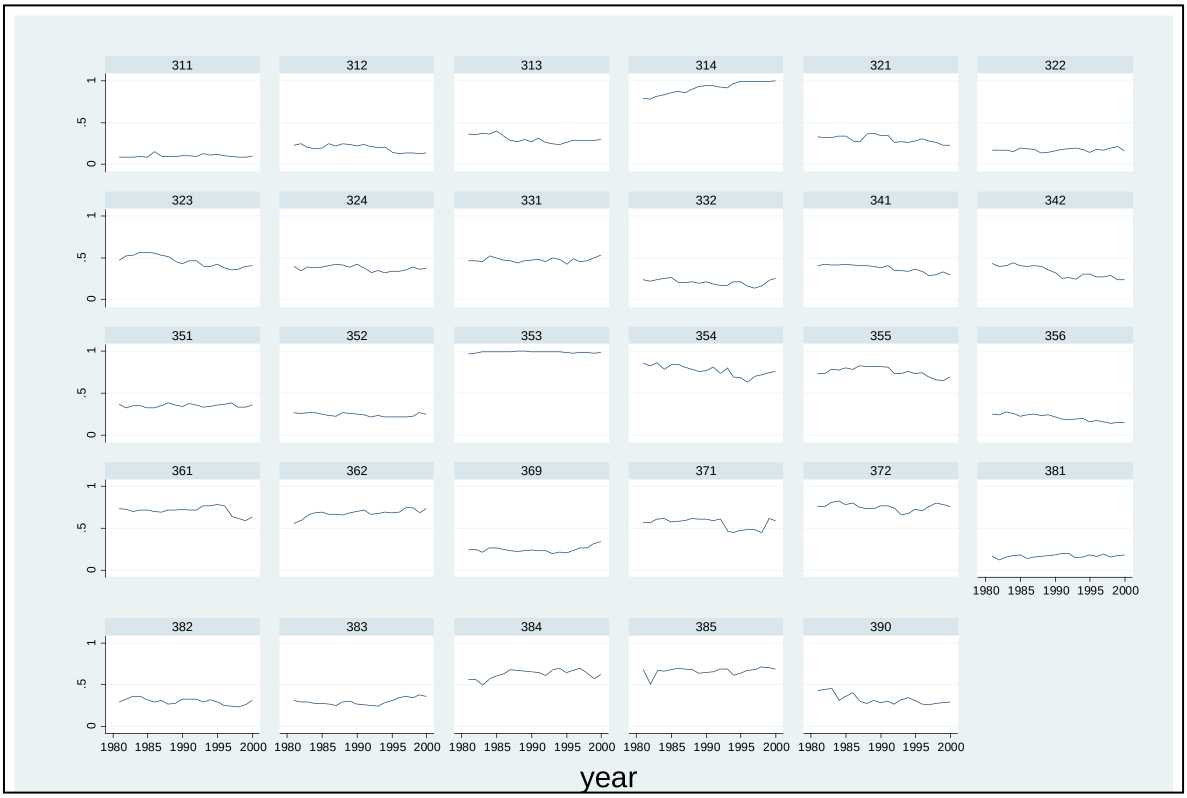
As explained in section 2.3 above, trade liberalisation has the potential to bring about more competition in the form of increased imports which, in turn, might reduce the scope for costly gender discrimination. On the other hand, section 2.4 suggests the possibility that increasing competition from imports may reinforce the bargaining position of local firms in the labour market as the number of employers is being reduced and workers have fewer options for employment within a given industry.

In order to control for the effects of market structure, we compute a conventional four-firm concentration ratio (CR_4) across industries based on the ratio between the gross product value from the four largest firms within a given industry and the total gross product value for the same industry as follows

$$CR_4 = \sum_{i=1}^{4} S_i \qquad (3.1)$$

where S_i denotes the gross product share of the i firm in the total gross product of a given industry. According to this index, there has been a slight reduction in the degree of concentration along the two decades defined in this study, from an average of 0.452 in 1981 to 0.439 in 2000. Figure 3.6 displays this concentration ratio for each of the 29 ISIC sectors along the years defined in this study. We plotted concentration ratios on an identical scale in order to display the high degree of stability in the ranking of the most (and less) concentrated sectors. Thus, *353- Petroleum refineries*, *314- Tobacco manufactures*, *354- Products of petroleum and coal*, *372- Non-ferrous metal basic industries*, *355- Rubber products* and, *361- Pottery, china and earthenware* emerge as the most concentrated ones in which the value of production for the top four firms represents more than 70 per cent of their corresponding industry. In contrast, *311- Food products*, *381- Fabricated metal products* and, *322- Wearing apparel, except* *footwear* appear as the least concentrated industries over the years reviewed here as their concentration index ranks, on average, below 20 per cent.

Figure 3.6: Concentration Indices (based on Gross Product Values) across manufacturing industries, Colombia: 1981-2000



Own estimates based on Annual Manufacturing Survey microdata.

3.3.4 Capital equipment

The interaction of trade with employment dynamics by gender has multiple dimensions. As explained by Galor and Weil (1996), the process of economic development allows increases in the availability of capital per worker which make physical strength less relevant and, thus, may lead to increased female labour participation. Since trade liberalisation facilitates the access to imported technology, there is the possibility of significant interactions with employment dynamics by gender.

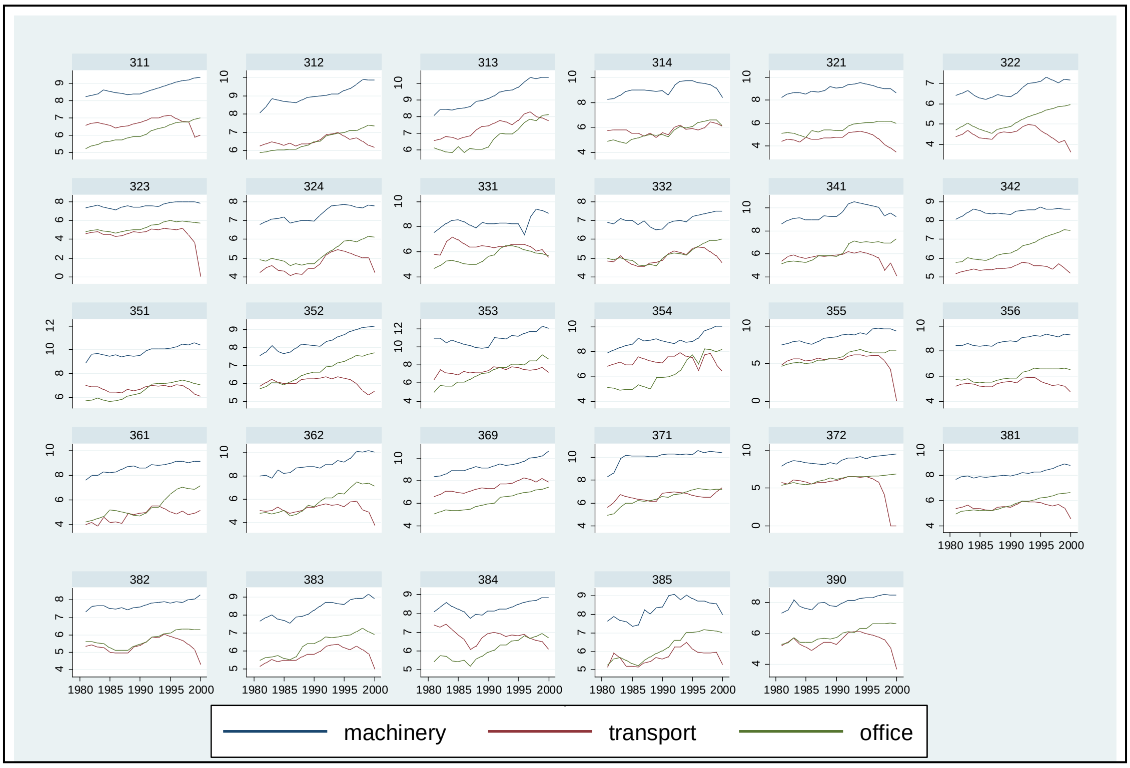
In order to test these possible relationships between employment dynamics by gender and trade, we also investigated the changes in capital investment across manufacturing industries. For this purpose, we computed the stock of three different types of capital over the fiscal year using AMS microdata. These are (i) machinery and equipment, (ii) transport equipment, and (iii) office equipment. In order to control for scale differences, we computed separately capital stocks per worker in natural logarithms expressed in constant 1999 Colombian Pesos. Capital stocks were estimated using a perpetual inventories approach according to the following expression:

$$K_{it} = K_{it-1} + I_{it} + (K_{it-1} + I_{it}) * D_i \tag{3.2}$$

where K denotes the capital stock of industry i at the beginning of year t, I represents the gross investment of industry i and, D depicts the observed depreciation rate of industry i estimated by Pombo (1999) at the ISIC Rev.2, 3-digit level industries. Figure 3.7 displays our estimates for the logarithm of the capital stock per worker across the 29 manufacturing industries defined in this study from 1981 to 2000. Capital stocks per worker of both machinery equipment and office equipment reported net increases between 1981-1985 and 1996-2000 for all manufacturing industries reviewed here. Contrastingly, transport equipment per worker reported net increases only in 14 out of 29 manufacturing sectors over the same time period. The largest increases in the stock of machinery equipment per worker between 1981-1985 and 1996-2000 were reported by *313- Beverage industries*, *362- Glass and glass products*, *355- Rubber products* and, *369- Other non-metallic mineral products*. In the case of transport equipment, the largest increases were found in *313- Beverage industries*, *369- Other non-metallic mineral products*, *361- Pottery, china and earthenware*, *324- Footwear* and, *371- Iron and steel basic industries*. Finally, the largest increases in office equipment per worker were recorded by *354- Products of petroleum and coal*, *353- Petroleum refineries*, *362- Glass and glass products*, *361- Pottery, china and earthenware*, *369-*

Other non-metallic mineral products and*, 313- Beverage industries*. From this, it is evident that *313- Beverage industries* was the most dynamic sector in terms of investments of all three types of capital equipment reviewed here, followed by *362- Glass and glass products*, a complementary sector of the former. A similar remark could be made for industries dedicated to the production of non-metallic mineral manufactures such as *361- Pottery, china and earthenware*, and *369- Other non-metallic mineral products* with some of the largest increments in their stock of the three types of capital per worker examined here.

Figure 3.7: Capital Equipment (Machinery, Transport and Office) per Worker across manufacturing industries, Colombia: 1981-2000



Own estimates based on Annual Manufacturing Survey microdata.

3.4. Econometric analysis

3.4.1 Methodology

In order to explain the effects of trade policy on the gender composition of the workforce across manufacturing industries, we implement different panel data models including fixed-effects instrumental variables (FE-IV). As technological changes are also likely to affect the share of female jobs across manufacturing industries over a time span of two decades, our empirical strategy also incorporates the three explanatory variables for the capital stock per worker (in logarithms) explained above in section 3.3.4, namely, machinery equipment, transport equipment and, office equipment. In addition, we control for the effects of changes in market structure with the inclusion of a concentration index based on expression 3.1 in Section 3.3.3, above.

The FE-IV approach adopted here is based on an individual industry effects model

$$y_{it} = x'_{it}\beta + \alpha_i + \varepsilon_{it} \qquad (3.3)$$

where y_{it} represents the female share of jobs in industry i at time t, x'_{it} is a set of explanatory variables and β depicts the coefficients to be estimated. The structure of the error component in (3.3) assumes the existence of unobserved time-invariant factors across the cross-section units depicted by α_i plus a conventional random component ε_{it}. Provided the existence of adequate instruments, z_{it}, FE-IV provide consistent estimates of β even in cases where the regressors contained in x_{it} are correlated with the random component ε_{it}. The key characteristic of such instruments is that they are uncorrelated to the error term ε_{it} so,

$$E(\varepsilon_{it}|\alpha_i, z_{i1}, \dots, z_{it}, \dots, z_{iT}) = 0. \qquad (3.4)$$

Under the assumption that (3.2) is upheld by the data, FE-IV provides consistent estimates. As it is normally the case with panel data, if the assumptions for the idiosyncratic error term notably $\varepsilon_{it} \sim (0, \sigma_\varepsilon^2)$ are not satisfied, conventionally computed standard errors are inaccurate. According to Cameron and Trivedi (2009), this assumption can be relaxed by the use of standard errors that allow for intergroup correlation. This is achieved with the estimation of a variance-covariance matrix that is adjusted with a clustered sandwich estimator. Chapter 8 of Angrist and Pischke (2008) describe this and other procedures for robust covariance matrix estimation in panel data applications whose observations are correlated within groups.[38] The estimation of FE-IV models presented in this application is performed using the ***xtivreg2*** Stata command developed by Schaffer and Stillman (2010) which allows for this type of cluster-robust standard errors. In the case of models without instruments, cluster-robust standard errors can be estimated with the conventional ***xtreg*** Stata command.

[38] Chapter 10 in Cameron and Trivedi (2009) provides also a review of different estimates for the variance-covariance matrix including the cluster-robust procedure. More formally, the cluster-robust standard errors approach implemented in this application is a generalization of White's (1980) procedure for the estimation of a robust covariance matrix of the following form:

$$\hat{\Omega}_{cl} = \left(X'X\right)^{-1} \left(\sum_g X_g \hat{\Psi}_g X_g \right) (X'X)^{-1}$$

where $\hat{\Psi}_g = a\hat{e}_g\hat{e}'_g = a \begin{bmatrix} \hat{e}_{1g}^2 & \cdots & \hat{e}_{1g}\hat{e}_{n_g g} \\ \vdots & \ddots & \vdots \\ \hat{e}_{1g}\hat{e}_{n_g g} & \cdots & \hat{e}_{n_g g}^2 \end{bmatrix}$,

X_g is the matrix of regressors for g groups, $\hat{e}_{ig}$ are the estimated residuals clustered around g groups of data and a is a factor adjustment which makes a degrees of freedom correction. See ANGRIST, J. D. & PISCHKE, J.-S. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion,* Princeton, New Jersey.: 312-313p.

3.4.2 Results

As a departure point, Table 3.1 describes the variables included in the models presented in this section while Table 3.2 reports their variance decomposition of them. All variables have no missing values and are within the expected range. To facilitate interpretation and estimation under different methods, all our variables are continuous measures within the 0 to 1 range, except in the case of capital per worker variables as they are expressed in logs in Colombian Pesos at constant 1999 prices. For all variables but the log of office equipment per worker variable (*lnkpw_office*), most of the variation occurs *between* manufacturing industries rather than *within* manufacturing industries.

Table 3.1 Variable definitions

label	variable	definition
femshare	female share of jobs: all workers	female share of jobs in industry i at time t amongst all workers
wc_femshare	female share of jobs: white-collar workers	female share of jobs in industry i at time t amongst white collar workers
bc_femshare	female share of jobs: blue-collar workers	female share of jobs in industry i at time t amongst blue collar workers
ipc	import penetration coefficient	$ipc_{it} = \frac{M_{it}}{Y_{it} + M_{it} - X_{it}}$ where Y, M and X denote, respectively, the gross product, imports and exports of industry *i* at time *t*.
eoc	export orientation coefficient	$eoc_{it} = \frac{X_{it}}{Y_{it}}$ where X and Y denote, respectively, exports and the gross product of industry *i* at time *t*.
CIGP	Concentration index	See expression (3.1) in text and details on it.
lnkpw_mach	ln(capital equipment per worker: machinery)	See expression (3.2) in text and details on it.
lnkpw_trans	ln(capital equipment per worker: transport)	
lnkpw_office	ln(capital equipment per worker: office equipment)	

To begin with, we want to test whether there is a relationship between the female share of jobs, on the one hand, and two selected trade variables on the other. The trade variables are the import penetration coefficient –*ipc* and the export orientation coefficient –*eoc*. These models are presented in Tables 3.3 and 3.4, from top to bottom, for *all workers*, *white-collar workers* and *blue collar workers*. All the reported specifications use clustered-robust standard errors as described in the preceding section.

Table 3.2: Panel summary statistics: within and between variation

Variable		Mean	Std. Dev.	Min	Max	Observations
isic	overall	349.897	25.064	311	390	N = 580
	between		25.486	311	390	n = 29
	within		0.000	349.897	349.897	T = 20
year	overall	1990.5	5.771	1981.0	2000.0	N = 580
	between		0.000	1990.5	1990.5	n = 29
	within		5.771	1981.0	2000.0	T = 20
femshare	overall	0.2701	0.1598	0.0096	0.8135	N = 580
	between		0.1596	0.0785	0.8007	n = 29
	within		0.0296	0.1824	0.3727	T = 20
wc_femshare	overall	0.3761	0.1019	0.0364	0.6704	N = 580
	between		0.0861	0.1597	0.5987	n = 29
	within		0.0567	0.2528	0.6486	T = 20
bc_femshare	overall	0.2295	0.1889	0.0032	0.8697	N = 580
	between		0.1899	0.0224	0.8516	n = 29
	within		0.0281	0.1200	0.3487	T = 20
CIGP	overall	0.4429	0.2462	0.0836	0.9990	N = 580
	between		0.2463	0.0985	0.9894	n = 29
	within		0.0442	0.2799	0.5702	T = 20
ipc	overall	0.2189	0.2218	0.0005	0.9456	N = 580
	between		0.2023	0.0176	0.7511	n = 29
	within		0.0980	-0.1337	0.7527	T = 20
eoc	overall	0.1717	0.2237	0.0006	1.8409	N = 580
	between		0.1636	0.0041	0.8421	n = 29
	within		0.1555	-0.4653	1.3417	T = 20
lnkpw_mach	overall	8.6275	0.9976	6.2089	12.3023	N = 580
	between		0.8888	6.7036	10.9253	n = 29
	within		0.4807	6.8639	10.0045	T = 20
lnkpw_trans	overall	5.8333	1.0772	0.0000	8.2848	N = 580
	between		0.9121	4.4789	7.4737	n = 29
	within		0.5964	0.5224	7.0625	T = 20
lnkpw_office	overall	6.0298	0.8434	4.2137	9.1494	N = 580
	between		0.5027	5.1884	7.1754	n = 29
	within		0.6833	3.8663	8.0038	T = 20

In Table 3.3, Column 1 reports pooled OLS regression estimates featuring only *ipc* as a regressor. The coefficients for manufacturing employment disaggregated by broad skill types are poorly determined as their statistical significance lies outside the 10 per cent level. However, there is a remarkable gain in efficiency as well as an increase in the magnitude of the *ipc* coefficient when we control for fixed effects using the (within) FE estimator in Column 2. In this case we find a positive and well determined relationship between import penetration and the female share of jobs; the size of the coefficients suggests that this effect is stronger amongst *white collar workers*. This relationship is confirmed in Column 3 for *all workers* and *white collar workers* when we include a trend variable while it turns out statistically insignificant for blue-collar workers. We also check in Column 4 whether this relationship holds when we lag the trade variable as the presumed effects of import penetration in manufacturing industries on their female share of jobs might exhibit some persistence over time. The estimates in Column 4 are quite similar in terms of both size and statistical significance to those from the FE with no trend in Column 2. The inclusion of a time-trend variable in addition to the lagged *ipc* variable in Column 5 yields a sizeable reduction in the size of the coefficients while standard errors are slightly larger so the statistical significance is consequently reduced, particularly amongst *blue collar workers*. Finally, Column 6 features coefficients based on a first-difference estimator. As the variables are in differences while the *ipc* variable is lagged one period, there is a reduction in the number of observations with respect to the FE models based on the mean-difference estimator in Columns 2 and 3. First differencing reduces the size of the coefficients dramatically and they are well determined only when the dependent variable is the female share of jobs for *all workers*.

Table 3.3 Female share equations, trade variable: import penetration coefficient (ipc)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**OLS**	**FE**	**FE+trend**	**FE: ipc_{t-1}**	**FE: ipc_{t-1} + trend**	**Differences: D.Y = f(D.ipc_{t-1})**
			All workers			
ipc	-0.0799	0.1445***	0.0728**			
	(0.1192)	(0.0306)	(0.0339)			
trend			0.0021***		0.0023***	
			(0.0007)		(0.0007)	
L.ipc				0.1499***	0.0694**	
				(0.0299)	(0.0334)	
LD.ipc						0.0334***
						(0.0088)
Constant	0.2876***	0.2384***	0.2320***	0.2390***	0.2304***	
	(0.0423)	(0.0067)	(0.0050)	(0.0064)	(0.0050)	
			White-collar workers			
ipc	0.0515	0.3334***	0.1075**			
	(0.0583)	(0.0556)	(0.0472)			
trend			0.0066***		0.0071***	
			(0.0008)		(0.0009)	
L.ipc				0.3219***	0.0784	
				(0.0553)	(0.0504)	
LD.ipc						0.0022
						(0.0365)
Constant	0.3648***	0.3031***	0.2828***	0.3109***	0.2848***	
	(0.0226)	(0.0122)	(0.0092)	(0.0118)	(0.0104)	
			Blue-collar workers			
ipc	-0.1323	0.0673**	0.0682			
	(0.1428)	(0.0310)	(0.0433)			
trend			-0.0000		0.0002	
			(0.0008)		(0.0008)	
L.ipc				0.0729**	0.0647	
				(0.0315)	(0.0427)	
LD.ipc						0.0220
						(0.0175)
Constant	0.2585***	0.2148***	0.2149***	0.2136***	0.2127***	
	(0.0499)	(0.0068)	(0.0069)	(0.0067)	(0.0069)	
Observations	580	580	580	551	551	522

Cluster-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.4 provides a similar set of econometric results with respect to those commented above but this time the trade variable is represented by the export orientation coefficient –*eoc*. Both OLS and FE estimates in Columns 1 and 2 indicate that manufacturing industries with higher levels of export orientation tend to have larger shares of female jobs. The coefficients for the *eoc* variable are statistically significant at the 1 per cent level for all dis-aggregated measures of the labour force.

With the inclusion of a time trend variable in Columns 2 and 3 the *eoc* coefficient still yields a positive coefficient in all cases but the size and the statistical significance is drastically reduced. A similar outcome is observed in Columns 4 and 5 with the incorporation of a one-lag version for this explanatory variable either with or without a trend control. The first-differenced results reported in Column 6 suggest that changes in export orientation might be positively associated with changes in the female share of jobs in the case of *all workers* and *blue collar workers* while they exert no independent effect amongst *white collar workers*. Notwithstanding, this positive effect amongst the *blue collar workers* is statistically significant only at the 10 per cent level.

The preceding findings from models featuring only one explanatory trade variable (plus a time trend in some cases) deserve some reflection. Estimates from the FE models using the mean-difference estimator suggest that manufacturing industries with high levels of both import penetration and export orientation tend to have a larger share of jobs occupied by women. The use of the first-difference estimator yields slightly less convincing evidence in favour of trade as a positive explanation for the growing proportion of female jobs in manufacturing industries. At best, these results suggests that the effects of increased trade in the gender composition of employment of manufacturing industries in urban Colombia are unevenly distributed across the two categories of jobs defined in this study. While changes in import penetration might be associated with a larger share of female jobs amongst *white collar workers*, changes in export orientation might be associated with increasing shares of jobs amongst *blue collar workers*. More importantly, the poor significance of the trade coefficients in some specifications suggests that other variables may have played a role in the incorporation of women in manufacturing. So far, we have implicitly assumed that the trade variables are uncorrelated to the error term ε_{it}. In other words, we have not dealt yet with any potential endogeneity problems that may contaminate these estimates.

Table 3.4 Female share equations, trade variable: export orientation coefficient (eoc)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**OLS**	**FE**	**FE+trend**	**FE: eoc_{t-1}**	**FE: eoc_{t-1} + trend**	**Differences: D.Y = f(D.eoc_{t-1})**
			All workers			
eoc	0.1960**	0.0594***	0.0178			
	(0.0730)	(0.0212)	(0.0152)			
trend			0.0026***		0.0028***	
			(0.0006)		(0.0006)	
L.ipc				0.0641**	0.0224	
				(0.0241)	(0.0168)	
LD.ipc						0.0190**
						(0.0093)
Constant	0.2364***	0.2599***	0.2395***	0.2604***	0.2369***	
	(0.0246)	(0.0036)	(0.0052)	(0.0040)	(0.0057)	
			White-collar workers			
eoc	0.1594***	0.1396***	0.0211			
	(0.0409)	(0.0309)	(0.0192)			
trend			0.0075***		0.0075***	
			(0.0008)		(0.0008)	
L.ipc				0.1419***	0.0283*	
				(0.0335)	(0.0149)	
LD.ipc						0.0076
						(0.0124)
Constant	0.3487***	0.3521***	0.2942***	0.3562***	0.2920***	
	(0.0147)	(0.0053)	(0.0072)	(0.0055)	(0.0084)	
			Blue-collar workers			
eoc	0.1940**	0.0201	0.0122			
	(0.0877)	(0.0178)	(0.0180)			
trend			0.0005		0.0007	
			(0.0006)		(0.0006)	
L.ipc				0.0269	0.0167	
				(0.0206)	(0.0197)	
LD.ipc						0.0241*
						(0.0122)
Constant	0.1962***	0.2260***	0.2221***	0.2248***	0.2190***	
	(0.0293)	(0.0031)	(0.0061)	(0.0034)	(0.0066)	
Observations	580	580	580	551	551	522

Cluster-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

For these reasons, we now implement the FE-IV approach by incorporating additional explanatory variables in our modelling strategy, namely, a concentration index of the gross product described in section 3.3.3 (*CIGP*), and the three measures of the stock of capital equipment per worker detailed on section 3.3.4 (*lnkpw_mach*, *lnkpw_trans* and, *lnkpw_office* –see Table 3.1 for definitions). Under this framework, we control for endogeneity problems through the use of instruments for both trade measures already incorporated in the models presented in Tables 3.3 and 3.4 and the concentration index variable (*CIGP*) discussed in Section 3.3.3, above. We base our decision on which variables to instrument on a version of the Hausman test of endogenous regressors developed in Stata™ by Schaffer and Stillman (2010) that is robust to violations of conditional homoskedasticity. The results for this test, under different specifications, are presented in the Statistical Appendix 3.1 of this chapter (see Tables A3. 1 and A3.2); they indicate that the null hypothesis that a given set of regressors is exogenous can be safely rejected in the case of the concentration index variable (*CIGP*) and the two trade measures (*ipc* and *eoc*).[39] Thus, we instrumented *CIGP* with the logarithm of the number of firms, *ipc* with average tariffs (see section 3.3.2, above) and, *eoc* with a conventional relative trade balance measure (*RTB*) constructed as follows:

$$RTB_{it} = \frac{X_{it} - M_{it}}{X_{it} + M_{it}} \tag{3.5}$$

where X_{it} and M_{it} denote the exports and imports, respectively, from industry *i* at time *t*.

The rationality for the use of these instruments is justified not only on the grounds that they are highly correlated to the endogenous variables (we test formally this below) but also on their theoretical validity. In the case of the import penetration, we argue that average tariffs represent an appropriate instrument measure of trade policy as they are aimed at moderating import flows. On this it should be mentioned that some empirical applications dealing with the effects of trade on labour market outcomes in Colombia have directly relied on tariffs as a proxy measure of trade policy (Attanasio et al., 2004, Goldberg and Pavcnik, 2003, Goldberg and Pavcnik, 2005b, Jaramillo and

[39] See notes at Tables 3.A1 and 3.A2 for details on the structure of this test.

Tovar, 2006).[40] We believe that using tariffs instead of import penetration as a variable to control for the impact of trade policy on the labour market is problematic as it omits the effects of other trade policy measures such as import licences and import quotas. Contrastingly, import penetration provides an outcome measure of the effects of trade policy on the competitive environment in which local firms have to operate. Tariffs instead provide a good instrument for import penetration as they embody a trade policy measure aimed specifically at moderating import flows into the domestic economy. In the case of the export orientation coefficient, we use a relative trade balance measure described in expression (3.5) as it represents a reasonable estimate of the competitive position of manufacturing industries with rich variation across sectors and over time. We also instrument the concentration index of gross product (CIGP) variable with the natural logarithm of the corresponding number of firms for each combination of industries and years based on the assumption that more competitive industries (i.e., with a lower concentration index) have, on average, a larger number of firms.

In Table 3.5 we test formally the association between the endogenous regressors and the selected instruments incorporated in subsequent FE-IV models presented below. According to these results, we can reasonably be confident that our instruments are highly correlated with the endogenous regressors not only in terms of the FE within estimator (see Columns 1, 3 and 5) but also in terms of the first-differences specification (see Columns 2, 4 and 6). As in other models presented along this chapter, the standard errors reported in Table 3.5 are robust for cluster correlation. On these

[40] On these papers, Attanasio et al (2004) use tariffs at the beginning of the 1980s interacted world coffee prices as instruments for tariffs while Goldberg and Pavcnik (2005b) perform an identical strategy. Jaramillo and Tovar (2006) also use tariffs at the beginning of the 1980s interacted with annual exchange rates.

results we verify a negative association between import penetration (*ipc*) and average tariffs (*a_tariffs*) as can be seen in the regression coefficients in Columns 1 and 2 which are statistically significant at the one per cent level in the case of the FE estimator and, at the five percent level in the case of the first-differences estimator. We confirm also a negative association between the concentration index of gross product (*CIGP*) and the natural logarithm of the number of plants (*ln_noplants*) as can be inferred from the estimated coefficients in Columns 3 and 4 of Table 3.5. Lastly, we corroborate a positive relationship with statistically significant coefficients at the one per cent level between export orientation (*eoc*) and the relative trade balance measure (*rtb*) presented in expression (3.5), above.

Table 3.5: Testing the relevance of instruments: fixed-effects and first-differences estimates

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
VARIABLES	**Ipc**	**D.ipc**	**CIGP**	**D.CIGP**	**eoc**	**D.eoc**
a_tariffs	-0.5688***					
	(0.1193)					
D.a_tariffs		-0.1196**				
		(0.0593)				
ln_noplants			-0.1075**			
			(0.0399)			
D.ln_noplants				-0.1022***		
				(0.0240)		
rtb					0.2237**	
					(0.0816)	
D.rtb						0.2995***
						(0.0824)
Constant	0.3189***	0.0073***	0.9762***	-0.0007***	0.1933***	0.0096***
	(0.0210)	(0.0024)	(0.1978)	(0.0000)	(0.0079)	(0.0005)
Observations	580	551	580	551	580	551
R-squared	0.2586	0.0078	0.1702	0.0620	0.1223	0.2061
Number of isic	29	29	29	29	29	29

Cluster-robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: (1) features *ipc* as a dependent variable against average tariffs (*a_tariffs*) as a single explanatory variable while (2) features the same variables in differences. (3) features *CIGP* as a dependent variable against the logarithm of the number of firms (*ln_noplants*) as a single explanatory variable while (4) features the same variables in differences. (5) features *eoc* as a dependent variable with the relative trade balance (*rtb*) as a single explanatory variable while (6) features the same variables in differences.

Results for our FE-IV estimates for the effects of import penetration on the female share of jobs are presented in Table 3.6. In order to check the robustesness of our FE-IV estimates, we also estimate the same female share equations with instruments derived from their lagged values. Standard errors for FE-IV models presented on Table 3.6 are robust for cluster serial autocorrelation (see Section 3.4.1, above). To further check these results, we present in the Statistical Appendix 3.2, estimates using the Generalised Method of Moments approach developed by Arellano and Bover (1995) and Blundell and Bond (1998).

As a natural reference point, Column 1 on Table 3.6 shows conventional FE with no instrumental variables. The trade variable, *ipc*, shows well determined coefficients for *all workers*, *white collar workers* and *blue collar workers* pointing towards a positive relationship between import penetration and the female share of jobs, a finding that confirms our previous results from Table 3.3. The use of instruments presented under different specifications in Columns 2 to 7confirm this result for both, *all workers* and *white collar workers*. In the case of *blue collar workers*, the choice of instruments affects the statistical significance of this variable and this casts some doubt on the effects of import penetration in the female share of jobs amongst this category. Results for the *ipc* variable using the linear dynamic panel data procedure presented in the Statistical Appendix 3.2 confirm that its effect on the female share of jobs is both negative and statistically different from zero only in the case of *white collar workers*. These results suggest that import penetration has a differentiated effect in the female share of jobs across the labour force categories defined in this study suggesting that some of the presumably positive effects of increased import penetration tend to favour the insertion of women mainly into the *white collar workers* category.

Table 3.6 Fixed-effects IV estimation of female share equations; trade variable: import penetration coefficient (ipc)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					All workers				
ipc	0.0892***	0.1600***	0.1247***	0.1752***	0.1505***	0.1698**	0.1102***	0.1708**	0.1065***
	(0.0128)	(0.0551)	(0.0157)	(0.0362)	(0.0144)	(0.0742)	(0.0165)	(0.0701)	(0.0166)
CIGP	-0.0787***	-0.4516***	-0.1096***	-0.4496***	-0.1221***	-0.4519***	-0.0961***	-0.4490***	-0.0968***
	(0.0249)	(0.0951)	(0.0368)	(0.0957)	(0.0377)	(0.0961)	(0.0368)	(0.0946)	(0.0368)
lnkpw_mach	0.0005	0.0019	0.0119***					0.0026	0.0030
	(0.0038)	(0.0048)	(0.0027)					(0.0047)	(0.0039)
lnkpw_trans	-0.0033*			-0.0016	-0.0011			-0.0018	-0.0029
	(0.0018)			(0.0022)	(0.0018)			(0.0026)	(0.0018)
lnkpw_office	0.0124***					0.0001	0.0106***	-0.0011	0.0095***
	(0.0028)					(0.0052)	(0.0021)	(0.0055)	(0.0030)
					White-collar workers				
ipc	0.1626***	0.6532***	0.2336***	0.7203***	0.3390***	0.5719***	0.1691***	0.5461***	0.1666***
	(0.0196)	(0.1008)	(0.0258)	(0.0701)	(0.0256)	(0.1230)	(0.0257)	(0.1129)	(0.0261)
CIGP	-0.0621	0.1353	-0.1097*	0.1177	-0.1551**	0.1330	-0.0520	0.1256	-0.0542
	(0.0382)	(0.1740)	(0.0606)	(0.1856)	(0.0673)	(0.1593)	(0.0575)	(0.1524)	(0.0575)
lnkpw_mach	0.0075	0.0155*	0.0467***					0.0041	0.0073
	(0.0058)	(0.0087)	(0.0045)					(0.0076)	(0.0062)
lnkpw_trans	-0.0018			0.0101**	0.0054			0.0057	-0.0017
	(0.0028)			(0.0043)	(0.0033)			(0.0042)	(0.0028)
lnkpw_office	0.0412***					0.0183**	0.0434***	0.0169*	0.0398***
	(0.0043)					(0.0086)	(0.0032)	(0.0089)	(0.0047)

Table 3.6 (Continuation)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Blue-collar workers				
ipc	0.0721***	0.0579	0.0863***	0.0033	0.0782***	0.0914	0.0868***	0.0959	0.0883***
	(0.0141)	(0.0560)	(0.0168)	(0.0366)	(0.0150)	(0.0753)	(0.0179)	(0.0714)	(0.0182)
CIGP	-0.0416	-0.3885***	-0.0253	-0.3848***	-0.0212	-0.3869***	-0.0270	-0.3893***	-0.0260
	(0.0275)	(0.0967)	(0.0395)	(0.0967)	(0.0394)	(0.0975)	(0.0401)	(0.0964)	(0.0402)
lnkpw_mach	-0.0067	-0.0097**	-0.0038					-0.0045	-0.0034
	(0.0041)	(0.0049)	(0.0029)					(0.0048)	(0.0043)
lnkpw_trans	0.0006			-0.0013	0.0006			0.0011	0.0011
	(0.0020)			(0.0023)	(0.0019)			(0.0026)	(0.0020)
lnkpw_office	0.0012					-0.0096*	-0.0023	-0.0076	-0.0006
	(0.0031)					(0.0053)	(0.0022)	(0.0056)	(0.0033)
Observations	580	580	551	580	551	580	551	580	551
Instruments									
- tariffs		Yes		yes		yes		yes	
- ln(number of plants)		Yes		yes		yes		yes	
- One lag			Yes		Yes		yes		yes

Cluster-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Constant omitted. Column (1) displays conventional FE with no instrumental variables. Columns (2) to (9) display FE-IV estimates; see bottom of table for chosen instruments. Import penetration coefficient (ipc) instrumented with either average tariffs or its own lag. Concentration index based on gross product (CIGP) instrumented with either the natural logarith of the number of firms or its own lag. FE-IV with cluster-robust stantandard errors estimated with the **xtivreg2** Stata command developed by Schaffer and Stillman (2010).

Models presented in Table 3.6 also investigate the effects of other variables commented on in the literature review in Section 3.2 above. Our measure of the degree of market concentration (*CIGP*) discussed in Section 3.3.3, above, is negative and statistically significant at the one per cent level for *all workers* and *blue collar workers* and performs poorly in the case of *white collar workers*. Results from dynamic panel data presented in Appendix 3.2 also suggest that the degree of market concentration is inversely correlated with the female share of jobs for all employment groupings analysed here, with well determined coefficients in most cases. Overall, the econometric evidence presented in both the main text and Appendix 3.2 is in line with the segregation dimension implicit in Becker's hypothesis of labour market discrimination in the sense that increased levels of market competition should erode monopolistic rents to discriminate against women. Although we do not have any evidence of reduced gender discrimination, we do observe that more competitive industries tend to have, on average, higher female shares of jobs. At least, this is what we would expect according to Becker's hypothesis in terms of the gender composition of the labour force as a result of increasing competition. In any case, we remain agnostic on whether this inverse relationship between market concentration and the female share of jobs across manufacturing industries is in any extent related to lower levels of gender discrimination. The same could be said regarding the results for the *ipc* variable commented above which could be rationalised in terms of the increased levels of market competition induced by increasing import penetration.

The results in Table 3.6 also feature the effects of the stock of capital investments per worker (in natural logarithms of Col Pesos of 1999) under the three categories discussed in Section 3.3.4, above. Columns 2 to 7 display the effects of these variables one by one using either average tariffs + the number of firms in logs (Columns 2, 4 and 6) or lagged values (Columns 3, 5 and 7) as instruments for both, the trade variable (*ipc*) and the concentration index variable (*CIGP*). It is worth reiterating that we could not find evidence indicating the necessity to instrument our capital equipment variables based on the version of the Hausman test of endogenous regressors explained above (see Appendix 3.1). Compared to the baseline specification with no instruments (Column 1), only our office equipment variable (*lnkpw_office*) is statistically significant for *all workers* and *white collar workers* while it tends to perform poorly for *blue collar workers*. This relationship is confirmed by our dynamic panel data estimates presented in Table A3.2.1 in Appendix 3.2. On this we should recall the discussion presented in Section 3.2.2 according to which increases in the availability of capital per worker enhances the participation of women in the labour market. In particular, the positive relationship between the capital stock of office equipment and the female share of jobs observed in our results is consistent with the hypothesis supported by some of the studies reviewed above (Galor and Weil, 1996, Weinberg, 2000, Welch, 2000) which suggest that women enjoy a comparative advantage in cognitive skills.

The fact that the estimates for our office equipment variable is not statistically significant for blue collar workers (a result that is also confirmed by dynamic panel data estimates in Table A3.2.1 in Appendix 3.2) might in a way be interpreted as a confirmation that the investments in office equipment are complementary to skilled female labour which tend to be concentrated in the *white collar category*. This interpretation is, to some extent, in line with the formulation proposed by Weinberg (2000) who argues that, in the case of the United States, a substitution process between highly skilled women and less skilled men might be explained by the increase in computers use which, on the margin, tends to favour the former. Figures presented in Figure 3.1a, above, suggest that this phenomenon might also be happening in Colombian manufacturing industries as female *white collar workers* were the only group of the labour force which shows an absolute increase of employment levels between 1981 and 2000. Contrastingly, male *blue collar workers* were the group with the largest reduction in manufacturing employment over these years in both absolute and relative terms.

In regard to the other two capital equipment variables reported in Table 3.6, we observe less clear cut results. Coefficients for the machinery equipment variable (*lnkpw_mach*) in Columns 2 and 3 suggest a positive and statistically significant relationship in the case of *white collar workers* when this variable is instrumented with its lagged values but this result turns out statistically insignificant when all capital regressors are simultaneously included in the model as can be seen in Column 7. Our dynamic panel data estimates presented in Table A3.2.1 of Appendix 3.2 indicate that this relationship is well determined only for *all workers*. In the same vein, the transport equipment variable shows up statistically significant at the 5 per cent level only in the case of *white collar workers* when we instrument both *ipc* and *CIGP* with average tariffs and the log of the number of firms, respectively (see Column 4 in Table 3.6). This finding is also confirmed by our dynamic panel data estimates from Table A3.2.1 in Appendix 3.2. When we switch our IV strategy to lagged values, our results presented in the main text indicate that this coefficient is not statistically different from zero. As in the case of the machinery equipment, the transport equipment variable loses its statistical significance when all capital regressors are simultaneously included in our FE-IV models in Table 3.6, a result that is also confirmed by our dynamic panel data estimates in Table A3.2.1 in Appendix 3.2.

Table 3.7 Fixed-effects IV estimation of female share equations; trade variable: export orientation coefficient (eoc)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					All workers				
eoc	0.0292***	-0.0015	0.0503***	0.0162	0.0675***	0.0023	0.0430***	-0.0026	0.0427***
	(0.0075)	(0.0242)	(0.0106)	(0.0229)	(0.0106)	(0.0230)	(0.0103)	(0.0233)	(0.0103)
CIGP	-0.0946***	-0.5540***	-0.1410***	-0.6239***	-0.1727***	-0.5210***	-0.1124***	-0.5240***	-0.1104***
	(0.0255)	(0.0774)	(0.0387)	(0.0795)	(0.0405)	(0.0794)	(0.0385)	(0.0792)	(0.0384)
lnkpw_mach	0.0003	0.0142***	0.0177***					0.0041	0.0022
	(0.0039)	(0.0039)	(0.0027)					(0.0048)	(0.0041)
lnkpw_trans	-0.0048***			-0.0032	-0.0026			-0.0051**	-0.0046**
	(0.0018)			(0.0025)	(0.0020)			(0.0023)	(0.0018)
lnkpw_office	0.0168***					0.0113***	0.0155***	0.0100**	0.0148***
	(0.0028)					(0.0029)	(0.0019)	(0.0040)	(0.0029)
					White-collar workers				
eoc	0.0543***	-0.1012***	0.1029***	-0.0050	0.1564***	-0.0850***	0.0786***	-0.0955***	0.0780***
	(0.0116)	(0.0383)	(0.0179)	(0.0386)	(0.0197)	(0.0326)	(0.0163)	(0.0334)	(0.0163)
CIGP	-0.0908**	-0.4268***	-0.1606**	-0.7453***	-0.2641***	-0.2339**	-0.0671	-0.2370**	-0.0665
	(0.0395)	(0.1223)	(0.0655)	(0.1338)	(0.0756)	(0.1127)	(0.0608)	(0.1139)	(0.0608)
lnkpw_mach	0.0069	0.0729***	0.0569***					0.0123*	0.0056
	(0.0060)	(0.0062)	(0.0046)					(0.0069)	(0.0065)
lnkpw_trans	-0.0045			0.0036	0.0019			-0.0057*	-0.0042
	(0.0028)			(0.0042)	(0.0037)			(0.0033)	(0.0029)
lnkpw_office	0.0492***					0.0610***	0.0502***	0.0555***	0.0476***
	(0.0043)					(0.0042)	(0.0030)	(0.0057)	(0.0047)

Table 3.7 (continuation)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Blue-collar workers				
eoc	0.0156*	-0.0170	0.0286***	-0.0226	0.0294***	-0.0191	0.0261**	-0.0176	0.0264**
	(0.0082)	(0.0243)	(0.0110)	(0.0215)	(0.0105)	(0.0239)	(0.0111)	(0.0242)	(0.0111)
CIGP	-0.0575**	-0.4505***	-0.0526	-0.4350***	-0.0541	-0.4535***	-0.0460	-0.4542***	-0.0444
	(0.0279)	(0.0776)	(0.0405)	(0.0744)	(0.0401)	(0.0825)	(0.0413)	(0.0824)	(0.0414)
lnkpw_mach	-0.0066	-0.0041	0.0008					-0.0030	-0.0036
	(0.0042)	(0.0039)	(0.0028)					(0.0050)	(0.0044)
lnkpw_trans	-0.0006			-0.0013	-0.0001			-0.0009	-0.0004
	(0.0020)			(0.0023)	(0.0020)			(0.0024)	(0.0020)
lnkpw_office	0.0052*					-0.0025	0.0021	-0.0008	0.0041
	(0.0031)					(0.0031)	(0.0020)	(0.0041)	(0.0032)
Observations	580	580	551	580	551	580	551	580	551
Instruments									
- relative trade balance		Yes		yes		yes		yes	
- ln(number of plants)		Yes		yes		yes		yes	
- One lag			yes		yes		yes		yes

Cluster-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Constant omitted. Column (1) displays conventional FE with no instrumental variables. Columns (2) to (9) display FE-IV estimates; see bottom of table for chosen instruments. Export orientation coefficient (eoc) instrumented with either relative trade balance or its own lag. Concentration index based on gross product (CIGP) instrumented with either the natural logarith of the number of firms or its own lag. FE-IV with cluster-robust stantandard errors estimated with the **xtivreg2** Stata command developed by Schaffer (2010).

The econometric results presented in Table 3.7 are intended to investigate the effects of an alternative trade variable, the export orientation coefficient –*eoc*. In this case, increased levels of trade in the form of export orientation tend to be statistically different from zero in a number of specifications. Nevertheless, the sign of the coefficient proves to be sensitive to the choice of instruments in this case. When we base our IV strategy on lagged values of endogenous regressors, the coefficient for *eoc* is both, positive and statistically significant at the one per cent level in all specifications (and for all breakdowns of the manufacturing employment) analysed here. This result is well supported by our dynamic panel data estimates presented in Table A3.2.2, Appendix 3.2, particularly in the case of blue collar workers. To a lesser extent, our first differences estimates from Table 3.4 point to a similar relationship. The simultaneous use of tariffs and the number of firms (in logs) as instruments (in Columns 2, 4, 6 and 8) yields less convincing results indicating that the *eoc* coefficient turns either statistically insignificant (for *all* and *blue collar workers*) or negative (for *white collar workers*). Overall, these results suggest that export orientation in manufacturing industries may be associated to larger shares of female workers in employment and some of the coefficients imply that this effect might be stronger amongst *blue collar workers*, a result that is somehow evident by comparing estimates from FE-IV and dynamic panel data. In the case of *white collar workers*, estimates from different methods provide a less coherent picture in terms of sign, size of coefficients and, statistical significance. From a conservative point of view, these results prove inconclusive in the case of *white collar workers* while they also suggest that export orientation might be associated to higher female shares of jobs in the *blue collar* category as indicated by the majority of our FE-IV and dynamic panel data estimates presented in Table 3.7 and Appendix 3.2, respectively.

The results in Table 3.7 also reveal the effects of other variables on the female share of jobs in manufacturing industries. In the case of our market concentration variable

(*CIGP*), there is strong evidence of its negative association with the female share of jobs for all labour force groupings analysed here. Coefficients for this variable are well determined in most cases, particularly for *all workers* and *white collar workers* where their statistical significance lies at the one per cent level in most cases. We observe also that the size and statistical significance of the coefficients tend to be reduced by the joint use of tariffs and the number of firms in logs as instruments of, respectively, *eoc* and *CIGP*. The negative association between the female share of jobs and our market concentration variable is better supported by our dynamic panel data results from Table A3.2.2 in Appendix 3.2 in which this variable is well determined in all cases.

Regarding our stock of capital per worker variables, results from Table 3.7 also indicate that both machinery (*lnkpw_mach*) and office equipment (*lnkpw_office*) exhibit a positive association with the female share of jobs for *all workers* and *white collar workers*. These results are equally confirmed by our dynamic panel data results from Table A3.2.2 in Appendix 3.2, according to which the coefficients for these two variables are well determined for the same the labour force groupings. In contrast, our measure of the stock of transport equipment per worker (*lnkpw_trans*) tends to be statistically insignificant in most cases, except for *all workers* when it is included simultaneously with the other two capital per worker variables just mentioned above.[41]

[41] It should be highlighted that the sign of the coefficient for this variable in this case is negative. This result is just partially replicated by our dynamic panel data coefficients reported in Table A3.2.2 of the Statistical Appendix 3.2 where this variable appears statistically significant only for *blue collar workers*. These results suggest that manufacturing industries with a high intensity in the use of transport equipment tend to have lower proportions of jobs occupied by women, an interpretation that might be plausible if we take into account that occupations related to the operation of transport equipment tend to be performed almost exclusively by men in urban Colombia. This interpretation is supported by the household survey microdata used in the previous two chapters according to which around 98 per cent of those working as

We also find strong evidence that the same three capital measures are uncorrelated with the female share of jobs amongst *blue collar workers*, as indicated by their coefficients in all specifications for this labour group. Compared to our dynamic panel estimates from Table A3.2.2 in Appendix 3.2, evidence of a relationship between the female share of jobs and the stock of capital equipment can only be confirmed in the case of *white collar workers* for machinery equipment and office equipment variables.

3.5. Concluding remarks

This chapter provides new evidence on the relationship between trade reforms and employment outcomes by gender with an empirical application to Colombian manufacturing industries. Given some data limitations discussed below, our empirical approach had to innovate by looking at the effects of trade liberalisation on the gender composition of employment in manufacturing industries. Although the evidence presented in this chapter does not formally test whether women are more (or less) discriminated in the labour market, our empirical results suggests that trade liberalisation, as well as some of the structural transformations in terms of the degree of market competition and the capital intensity of economic activities, are somehow related to the gender composition of employment in Colombian manufacturing industries.

We found convincing evidence that increased levels of import penetration are positively associated with higher female shares of jobs in manufacturing industries. Different econometric techniques presented in this chapter point towards a similar

---

"Transport Equipment Operators" between 1984 and 2004 are men, indicating that this occupation ranks as one of the most segregated in the labour market of this country.

conclusion and they indicate that this effect was probably stronger amongst *white collar workers*. Increasing levels of export orientation suggest a similar pattern although this result might be even more pronounced in the case of *blue collar workers*. Likewise, we found persuasive evidence that higher levels of market concentration as measured by our concentration index of gross product (*CIGP*) are negatively associated with the female share of jobs in manufacturing industries, indicating that more competitive environments are more likely to incorporate larger shares of female employment. So far, this is what we expected to find from the literature review presented in section 3.2.3 in relation to the segregation dimension implicit in Becker's hypothesis of labour market discrimination. As our dependent variable is the female share of jobs, we remain agnostic as to whether the effects of increased competition, either in the form of import penetration or in the form of market concentration, have any effect on the extent of gender pay discrimination. These results, however, suggest that increasing levels of competition are associated with higher shares of female employment and this is the type of result we would expect to encounter on the gender composition of the labour force if increased trade has an effect on gender discrimination. We should stress that increasing levels of female employment in manufacturing industries could occur with or without improvements in the bargaining position of women in the labour market. On this we should remember that higher levels of trade might also push employers to cost-cutting strategies that lessen the bargaining position of women, as suggested by the study of Berik et al. (Berik et al., 2004) for east Asian economies discussed in Section 3.2.4, above. In that Section, we reviewed also a study by Seguino (2000) who argues that low female wages might encourage the hiring of women workers in export oriented industries. For all these reasons, our findings are only suggestive of some of the positive effects of trade on gender differences in the labour market and further research is needed to establish whether the participation of women in Colombian manufacturing industries was accompanied by a reduction of gender discrimination.

We could also verify some complementarities between female labour and the use of some types of capital equipment. Our estimates under different panel data techniques are suggestive that the increasing use of office equipment is concomitant with higher shares of female employment in manufacturing industries of urban Colombia. This result was robust even in cases were other types of capital equipment were simultaneously controlled for. These findings provide further support of the hypothesis that the increasing use of technology favours the incursion of women in the labour market as they enjoy a comparative advantage in cognitive skills (Galor and Weil, 1996, Weinberg, 2000, Welch, 2000). This finding is further supported by the fact that the presumably positive effect derived from the increasing use of office equipment is econometrically stronger amongst the *white collar* group where the most qualified women tend to be concentrated. In the same vein, the fact that increasing female shares of jobs are also positively associated with increases in the use of machinery equipment suggests that the growing demand for intellectual skills not only favours the relative demand for female labour but also that this might entail an incentive for fertility decline as implied by Galor and Weil (1996).

We attempted to reconcile results from different econometric techniques, including FE-IV. The appropriateness of instruments and their validity in terms of both economic and statistical theory was assessed by comparing results drawn from different econometric techniques. The use of different methods to verify the relationships between the female share of jobs and some variables related to the economic development process provides a sound basis for statistical inference. We were fortunate to verify that many of these relationships were robust to the use of different instruments. From an empirical point of view, we believe that the results outlined along this chapter are well supported by a number of methods pointing in the same direction.

The findings encountered along this chapter also provide an interesting picture from an economic development perspective. To some extent, the evidence presented here is suggestive that the incorporation of women in manufacturing industries is concomitant not only with increased levels of trade, but also with capital intensification (in terms of machinery and office equipment) of productive processes in a number of industries. As a whole, our findings are consistent with Boserup's (1970) hypothesis according to which gender discrimination is inversely related to the level of economic development. This claim, however, deserves some qualifications as we could observe along this chapter that these effects are highly differentiated across the labour market groupings defined in our data. In this sense, the selection of women into the *white collar workers* category appears to be more successful than the case of the blue collar workers category and this differentiated pattern appears to be biased in favour of the most qualified (and, presumably, better off) women.

Finally, it should be remarked that this investigation in its present state could be further developed in a number of ways. As explained above, it would be desirable to verify the effects of trade on gender wage differences. Employment data used in this chapter come from the Annual Manufacturing Survey which does not provide disaggregated information on wages and labour costs by gender. This limitation in the availability of data rendered impossible the further investigation of the effects of trade policies on labour market outcomes from a gender perspective, particularly in regard to the paramount issue of wage differences. An alternative to this problem would be to use household survey microdata, which are available in Colombia on a regular basis since 1984. Based on statistical analyses not presented in this study, we found that this type of data has some limitations in terms of the accuracy in the recording of the information related to the ISIC codes to describe the economic activity of household respondents in the workforce, which is based on the supply side of the labour market, as opposed to the Annual Manufacturing Survey. Therefore, our attempts to verify a relationship between trade measures and gender wage gaps were inconclusive using household survey data but we believe this issue remains an important avenue for further research.

Appendix 3.1

Table A3. 1a Fixed-effects IV estimation of female share equations, trade variable: import penetration coefficient (ipc) –Endogeneity test of endogenous regressors

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
						All workers						
ipc	0.3163***	0.2887***	0.3375***	0.0662***	0.0782***	0.0595***	0.1600***	0.1752***	0.1698**	0.4281	0.1183***	0.0921***
	(0.0488)	(0.0276)	(0.0729)	(0.0169)	(0.0173)	(0.0168)	(0.0551)	(0.0362)	(0.0742)	(0.5911)	(0.0230)	(0.0221)
CIGP	-0.0046	-0.0121	-0.0032	-0.5102***	-0.5645***	-0.4990***	-0.4516***	-0.4496***	-0.4519***	-0.2955	-0.1166***	-0.0842***
	(0.0365)	(0.0327)	(0.0388)	(0.0832)	(0.0858)	(0.0857)	(0.0951)	(0.0957)	(0.0961)	(0.3754)	(0.0427)	(0.0286)
lnkpw_mach	-0.0042			0.0090***			0.0019			-0.1666		
	(0.0048)			(0.0031)			(0.0048)			(0.3242)		
lnkpw_trans		0.0003			-0.0025			-0.0016			-0.0346	
		(0.0021)			(0.0023)			(0.0022)			(0.0258)	
lnkpw_office			-0.0048			0.0075***			0.0001			0.0111**
			(0.0054)			(0.0024)			(0.0052)			(0.0053)
Endogeneity test of endogenous regressors												
χ^2 (1 or 2):	30.704	55.951	20.004	44.259	53.755	42.196	61.707	85.728	51.935	2.933	2.703	0.195
P-val =	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0868	0.1002	0.6592
						White-collar workers						
Ipc	0.6273***	0.7110***	0.5430***	0.2153***	0.2889***	0.1664***	0.6532***	0.7203***	0.5719***	0.2857	0.2878***	0.0900***
	(0.0853)	(0.0549)	(0.1120)	(0.0224)	(0.0251)	(0.0208)	(0.1008)	(0.0701)	(0.1230)	(0.3198)	(0.0326)	(0.0331)
CIGP	0.0614	0.0818	0.0555	-0.1385	-0.3930***	-0.0400	0.1353	0.1177	0.1330	-0.1820	-0.0981	-0.0239
	(0.0638)	(0.0651)	(0.0596)	(0.1101)	(0.1246)	(0.1063)	(0.1740)	(0.1856)	(0.1593)	(0.2031)	(0.0605)	(0.0428)
lnkpw_mach	0.0166**			0.0489***			0.0155*			0.0050		
	(0.0084)			(0.0041)			(0.0087)			(0.1754)		
lnkpw_trans		0.0100**			0.0062*			0.0101**			0.0081	
		(0.0042)			(0.0034)			(0.0043)			(0.0365)	
lnkpw_office			0.0191**			0.0453***			0.0183**			0.0658***
			(0.0083)			(0.0030)			(0.0086)			(0.0079)
Endogeneity test of endogenous regressors												
χ^2 (1 or 2):	44.896	111.032	20.326	0.084	4.839	0.038	46.380	111.640	21.804	0.085	0.150	6.178
P-val =	0.0000	0.0000	0.0000	0.7723	0.0278	0.8458	0.0000	0.0000	0.0000	0.7709	0.6988	0.0129

Table A3. 1a (continuation)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
							Blue-collar workers					
ipc	0.1966***	0.0983***	0.2387***	0.0375**	0.0252	0.0424**	0.0579	0.0033	0.0914	0.4767	0.0508**	0.0934***
	(0.0461)	(0.0256)	(0.0697)	(0.0171)	(0.0163)	(0.0173)	(0.0560)	(0.0366)	(0.0753)	(0.7382)	(0.0242)	(0.0247)
CIGP	0.0081	-0.0183	0.0071	-0.4013***	-0.3588***	-0.4078***	-0.3885***	-0.3848***	-0.3869***	-0.3113	-0.0754*	-0.0658**
	(0.0344)	(0.0303)	(0.0371)	(0.0842)	(0.0809)	(0.0880)	(0.0967)	(0.0967)	(0.0975)	(0.4687)	(0.0449)	(0.0320)
lnkpw_mach	-0.0151***			-0.0082***			-0.0097**			-0.2336		
	(0.0045)			(0.0031)			(0.0049)			(0.4049)		
lnkpw_trans		0.0003			-0.0011			-0.0013			-0.0379	
		(0.0020)			(0.0022)			(0.0023)			(0.0271)	
lnkpw_office			-0.0140***			-0.0063***			-0.0096*			-0.0091
			(0.0052)			(0.0025)			(0.0053)			(0.0059)
Endogeneity test of endogenous regressors												
χ^2 (1 or 2):	9.183	2.558	7.708	27.542	22.804	25.879	31.23	22.879	28.768	3.991	1.933	1.311
P-val =	0.0024	0.1097	0.0055	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0457	0.1644	0.2523

Clustered robust standard errors in parentheses estimated with the **xtivreg2** Stata command developed by Schaffer (2010). The endogeneity test incorporated in this command is robust to heteroskedasticity and is compatible with clusted-robust standard errors. The test statistic is distributed as a Chi-squared with degrees of freedom equal to the number of tested regressor and is defined as a difference between two Sargan-Hansen tests from two models, one where the concerning variables are treated as endogenous and another where these variables are treated as exogenous.

*** p<0.01, ** p<0.05, * p<0.1. N = 580 and t = 20 in all cases.

(1) to (3): import penetration coefficient (ipc) is instrumented with average tariffs.

(4) to (6): concentration index based on gross product (CIGP) instrumented with the natural logarithm of the number of firms.

(7) to (9): both, ipc and CIGP instrumented as described above.

(10) to (12): capital per worker variables instrumented with the natural logarithms of net investment per worker.

Table A3. 1b Fixed-effects IV estimation of female share equations (pooled capital variables), trade variable: import penetration coefficient (ipc) - Endogeneity test of endogenous regressors

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	All workers				**White-collar workers**				**Blue-collar workers**			
ipc	0.3360***	0.0556***	0.1708**	-4.1142	0.5166***	0.1649***	0.5461***	2.7724	0.2443***	0.0424**	0.0959	-5.7164
	(0.0686)	(0.0169)	(0.0701)	(139.8278)	(0.1025)	(0.0210)	(0.1129)	(93.3871)	(0.0659)	(0.0175)	(0.0714)	(193.8449)
CIGP	-0.0033	-0.4971***	-0.4490***	3.4727	0.0460	-0.0337	0.1256	-2.3203	0.0110	-0.4117***	-0.3893***	4.8274
	(0.0381)	(0.0846)	(0.0946)	(118.7512)	(0.0569)	(0.1054)	(0.1524)	(79.3106)	(0.0366)	(0.0875)	(0.0964)	(164.6262)
lnkpw_mach	-0.0013	0.0035	0.0026	-4.6197	0.0048	0.0073	0.0041	3.0729	-0.0080*	-0.0041	-0.0045	-6.4093
	(0.0049)	(0.0046)	(0.0047)	(155.5796)	(0.0073)	(0.0058)	(0.0076)	(103.9073)	(0.0047)	(0.0048)	(0.0048)	(215.6818)
lnkpw_trans	0.0015	-0.0040*	-0.0018	4.0828	0.0051	-0.0017	0.0057	-2.7432	0.0040	0.0001	0.0011	5.6662
	(0.0027)	(0.0022)	(0.0026)	(137.9452)	(0.0040)	(0.0028)	(0.0042)	(92.1298)	(0.0026)	(0.0023)	(0.0026)	(191.2351)
lnkpw_office	-0.0042	0.0063*	-0.0011	1.6270	0.0175**	0.0416***	0.0169*	-0.9845	-0.0103*	-0.0041	-0.0076	2.2281
	(0.0058)	(0.0036)	(0.0055)	(54.0780)	(0.0086)	(0.0045)	(0.0089)	(36.1172)	(0.0055)	(0.0038)	(0.0056)	(74.9689)
Endogeneity test of endogenous regressors												
χ^2 (1 or 2):	23.099	42.619	54.813	4.331	20.239	0.084	21.878	15.844	9.230	27.374	31.239	3.797
P-val =	0.0000	0.0000	0.0000	0.2279	0.0000	0.7722	0.0000	0.0012	0.0024	0.0000	0.0000	0.2842

Clustered robust standard errors in parentheses estimated with the **xtivreg2** Stata command developed by Schaffer (2010). The endogeneity test incorporated in this command is robust to heteroskedasticity and is compatible with clusted-robust standard errors. The test statistic is distributed as a Chi-squared with degrees of freedom equal to the number of tested regressor and is defined as a difference between two Sargan-Hansen tests from two models, one where the concerning variables are treated as endogenous and another where these variables are treated as exogenous.

*** p<0.01, ** p<0.05, * p<0.1. N = 580 and t = 20 in all cases.

(1), (5), (9): import penetration coefficient (ipc) instrumented with average tariffs.

(2), (6), (10): concentration index based on gross product (CIGP) instrumented with the natural logarithm of the number of firms.

(3), (7), (11): both, ipc and CIGP instrumented as described above.

(4), (8), (12): capital per worker variables instrumented with the natural logarithms of net investment per worker.

Table A3.2a Fixed-effects IV estimation of female share equations, trade variable: export orientation coefficient (eoc)
Endogeneity test of endogenous regressors

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
						All workers						
eoc	-0.0562**	-0.0321	-0.0469**	0.0194**	0.0276***	0.0168*	-0.0015	0.0162	0.0023	-0.3140	0.0682***	0.0324***
	(0.0231)	(0.0217)	(0.0210)	(0.0097)	(0.0103)	(0.0094)	(0.0242)	(0.0229)	(0.0230)	(1.1172)	(0.0193)	(0.0100)
CIGP	-0.1633***	-0.2002***	-0.1234***	-0.5222***	-0.6004***	-0.5001***	-0.5540***	-0.6239***	-0.5210***	0.6267	-0.1398***	-0.1157***
	(0.0310)	(0.0325)	(0.0288)	(0.0844)	(0.0863)	(0.0871)	(0.0774)	(0.0795)	(0.0794)	(2.3912)	(0.0467)	(0.0322)
lnkpw_mach	0.0262***			0.0126***			0.0142***			0.4888		
	(0.0034)			(0.0031)			(0.0039)			(1.5157)		
lnkpw_trans		-0.0017			-0.0033			-0.0032			-0.0500*	
		(0.0022)			(0.0024)			(0.0025)			(0.0303)	
lnkpw_office			0.0217***			0.0106***			0.0113***			0.0103*
			(0.0023)			(0.0024)			(0.0029)			(0.0056)
Endogeneity test of endogenous regressors												
χ^2 (1 or 2):	23.458	21.458	18.684	37.928	49.759	36.211	60.205	70.14	54.523	6.003	4.82	1.741
P-val =	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0143	0.0281	0.1870
						White-collar workers						
Eoc	-0.1246***	-0.0487	-0.0960***	0.0737***	0.1123***	0.0576***	-0.1012***	-0.0050	-0.0850***	-0.1992	0.1404***	0.0363**
	(0.0410)	(0.0415)	(0.0341)	(0.0134)	(0.0160)	(0.0120)	(0.0383)	(0.0386)	(0.0326)	(0.9146)	(0.0284)	(0.0146)
CIGP	-0.2596***	-0.3615***	-0.1457***	-0.1614	-0.5048***	-0.0277	-0.4268***	-0.7453***	-0.2339**	0.4298	-0.1752**	-0.0529
	(0.0549)	(0.0621)	(0.0467)	(0.1170)	(0.1340)	(0.1116)	(0.1223)	(0.1338)	(0.1127)	(1.9577)	(0.0687)	(0.0469)
lnkpw_mach	0.0780***			0.0598***			0.0729***			0.4373		
	(0.0060)			(0.0043)			(0.0062)			(1.2409)		
lnkpw_trans		0.0049			0.0035			0.0036			-0.0331	
		(0.0041)			(0.0038)			(0.0042)			(0.0446)	
lnkpw_office			0.0633***			0.0533***			0.0610***			0.0647***
			(0.0038)			(0.0030)			(0.0042)			(0.0082)
Endogeneity test of endogenous regressors												
χ^2 (1 or 2):	39.13	26.556	30.496	0.001	4.579	0.339	39.154	30.773	30.882	1.668	1.833	1.694
P-val =	0.0000	0.0000	0.0000	0.9766	0.0324	0.5606	0.0000	0.0000	0.0000	0.1965	0.1758	0.1931

Table A3. 2a (continuation)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
							Blue-collar workers					
eoc	-0.0653***	-0.0592***	-0.0641***	0.0030	-0.0006	0.0032	-0.0170	-0.0226	-0.0191	-0.3561	0.0305	0.0242**
	(0.0237)	(0.0216)	(0.0227)	(0.0098)	(0.0095)	(0.0097)	(0.0243)	(0.0215)	(0.0239)	(1.1910)	(0.0187)	(0.0112)
CIGP	-0.1051***	-0.1141***	-0.0896***	-0.4202***	-0.3898***	-0.4213***	-0.4505***	-0.4350***	-0.4535***	0.7178	-0.0844*	-0.1009***
	(0.0318)	(0.0323)	(0.0311)	(0.0851)	(0.0795)	(0.0898)	(0.0776)	(0.0744)	(0.0825)	(2.5492)	(0.0451)	(0.0359)
lnkpw_mach	0.0066*			-0.0055*			-0.0041			0.4995		
	(0.0035)			(0.0031)			(0.0039)			(1.6158)		
lnkpw_trans		-0.0001			-0.0013			-0.0013			-0.0444	
		(0.0022)			(0.0022)			(0.0023)			(0.0293)	
lnkpw_office			0.0070***			-0.0037			-0.0025			-0.0093
			(0.0025)			(0.0024)			(0.0031)			(0.0063)
Endogeneity test of endogenous regressors												
χ^2 (1 or 2):	16.911	17.08	16.642	26.243	24.782	24.216	42.321	41.182	40.571	6.034	2.546	3.265
P-val =	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0140	0.1106	0.0708

Clustered robust standard errors in parentheses estimated with the **xtivreg2** Stata command developed by Schaffer (2010). The endogeneity test incorporated in this command is robust to heteroskedasticity and is compatible with clusted-robust standard errors. The test statistic is distributed as a Chi-squared with degrees of freedom equal to the number of tested regressor and is defined as a difference between two Sargan-Hansen tests from two models, one where the concerning variables are treated as endogenous and another where these variables are treated as exogenous.

*** p<0.01, ** p<0.05, * p<0.1. N = 580 and t = 20 in all cases.

(1) to (3): export orientation coefficient (eoc) instrumented with relative trade balance (see text for details).

(4) to (6): concentration index based on gross product (CIGP) instrumented with the natural logarithm of the number of firms.

(7) to (9): both, ipc and CIGP instrumented as described above.

(10) to (12): capital per worker variables instrumented with the natural logarithms of net investment per worker for each type of capital: machinery, transport and office equipment.

Table A3. 2b Fixed-effects IV estimation of female share equations (pooled capital variables), trade variable: export orientation coefficient (eoc). Endogeneity test of endogenous regressors

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	All workers				**White-collar workers**				**Blue-collar workers**			
eoc	-0.0514**	0.0158*	-0.0026	-0.1514	-0.1058***	0.0565***	-0.0955***	0.1776	-0.0624***	0.0034	-0.0176	-0.2380
	(0.0214)	(0.0094)	(0.0233)	(0.6215)	(0.0350)	(0.0121)	(0.0334)	(0.4718)	(0.0230)	(0.0097)	(0.0242)	(0.8677)
CIGP	-0.1259***	-0.4980***	-0.5240***	0.3708	-0.1528***	-0.0230	-0.2370**	-0.2691	-0.0878***	-0.4247***	-0.4542***	0.5318
	(0.0290)	(0.0859)	(0.0792)	(1.4861)	(0.0474)	(0.1105)	(0.1139)	(1.1282)	(0.0311)	(0.0893)	(0.0824)	(2.0749)
lnkpw_mach	0.0029	0.0034	0.0041	-0.5158	0.0121*	0.0064	0.0123*	0.3742	-0.0042	-0.0038	-0.0030	-0.7318
	(0.0043)	(0.0047)	(0.0048)	(1.8036)	(0.0070)	(0.0060)	(0.0069)	(1.3692)	(0.0046)	(0.0049)	(0.0050)	(2.5182)
lnkpw_trans	-0.0055***	-0.0049**	-0.0051**	0.4066	-0.0058*	-0.0045	-0.0057*	-0.3161	-0.0013	-0.0007	-0.0009	0.5767
	(0.0020)	(0.0022)	(0.0023)	(1.4670)	(0.0033)	(0.0028)	(0.0033)	(1.1136)	(0.0022)	(0.0023)	(0.0024)	(2.0482)
lnkpw_office	0.0212***	0.0094**	0.0100**	0.2087	0.0579***	0.0504***	0.0555***	-0.0601	0.0095***	-0.0015	-0.0008	0.2689
	(0.0033)	(0.0037)	(0.0040)	(0.6700)	(0.0053)	(0.0047)	(0.0057)	(0.5087)	(0.0035)	(0.0038)	(0.0041)	(0.9355)
Endogeneity test of endogenous regressors												
χ^2 (1 or 2):	20.164	36.878	56.300	3.870	33.093	0.434	33.638	11.723	15.767	25.511	40.733	4.127
P-val =	0.0000	0.0000	0.0000	0.2759	0.0000	0.5102	0.0000	0.0084	0.0001	0.0000	0.0000	0.2481

Clustered robust standard errors in parentheses estimated with the **xtivreg2** Stata command developed by Schaffer (2010). The endogeneity test incorporated in this command is robust to heteroskedasticity and is compatible with clusted-robust standard errors. The test statistic is distributed as a Chi-squared with degrees of freedom equal to the number of tested regressor and is defined as a difference between two Sargan-Hansen tests from two models, one where the concerning variables are treated as endogenous and another where these variables are treated as exogenous.

*** p<0.01, ** p<0.05, * p<0.1. N = 580 and t = 20 in all cases.

(1), (5), (9): export orientation coefficient (eoc) instrumented with relative trade balance (see text for details).

(2), (6), (10): concentration index based on gross product (CIGP) instrumented with the natural logarithm of the number of firms.

(3), (7), (11): both, eoc and CIGP instrumented as described above.

(4), (8), (12): capital per worker variables instrumented with the natural logarithms of net investment per worker.

Appendix 3.2

The discussion about the validity of instruments in the context panel data has been widely documented in the literature. In order to have a yardstick of comparison for our FE-IV estimates, we also implement a dynamic panel data system strategy based on the GMM developed by Arellano and Bover (1995) and Blundell and Bond (1998). This GMM procedure consists of a simultaneous estimation of two equations, one in levels and another in differences with a set of instruments used in each equation. In principle, the general model can be expressed as:

$$y_{it} = \delta y_{i,t-1} + x'_{i,t}\beta + \rho_i + \varepsilon_{i,t}\ . \qquad \text{(A3.1)}$$

where $y_{i,t}$ represents the share of female jobs in the total number of jobs of industry i at year $t, x_{i,t}$ is a set of explanatory variables (in this case, a trade variable, a concentration index and, a capital stock per worker measure in logs – either machinery, transport or office equipment), ρ_i depicts a vector of industry fixed effects and, $\varepsilon_{i,t}$ is an *i.i.d.* random component. First differencing of (A3.1) allows the elimination of the industry fixed effects as follows,

$$y_{i,t} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i,t-2}\right) + \beta\left(x_{i,t} - x_{i,t-1}\right) + \left(\varepsilon_{i,t} - \varepsilon_{i,t-1}\right).$$

(A3.2)

In this specification, the choice of instruments aimed to solve endogeneity problems amongst the explanatory variables is performed in such a way that present realisations on the explanatory variables are influenced by past realisations of the dependent variable. Thus, instead of assuming strict ortogonality in the regressors, a less restrictive assumption of *weak exogeneity* is adopted. Under the two assumptions of (i) no serial autocorrelation in the residuals and, (ii) *weak exogeneity*, the following moment conditions apply:

$$E\left[y_{i,t-\dot{s}}\left(\varepsilon_{i,t} - \varepsilon_{i,t-1}\right)\right] = 0 \qquad \forall \qquad s \geq 2; t = 3, \dots, T. \qquad \text{(A3.3)}$$

$$E\left[x_{i,t-\dot{s}}\left(\varepsilon_{i,t}-\varepsilon_{i,t-1}\right)\right]=0 \qquad \forall \qquad s\geq 2; t=3,\ldots,T. \qquad \text{(A3.4)}$$

Moment conditions (A3.3) and (A3.4) represent the basis for the GMM estimator of differences. This differences estimator is, however, characterised by low asymptotic precision and small sample biases and that is why it should be complemented with the regression equation in levels. Furthermore, when the lagged dependent and explanatory variables are persistent over time they represent weak instruments for the regression equation in differences (Blundell and Bond, 1998). According to Griliches and Hausman (1986), another problem is that the differences estimator is biased due to decreasing signal-to-noise ratios. For all of this, Arellano and Bover (1995) system estimator reduces potential biases by incorporating simultaneously the estimation of equations (A3.1) and (A3.2). Industry-specific effects at this stage ought to be controlled with instrumental variables for which lagged differences represent adequate instruments for the regression in levels. Even though, industry-specific effects may be correlated with right-hand side variables, there is no correlation between them when they are expressed in differences. Under these circumstances, the following stationarity property should hold,

$$E[y_{i,t+p}\cdot\rho_i]=E[y_{i,t+q}\cdot\rho_i]\,; \quad E[x_{i,t+p}\cdot\rho_i]=E[x_{i,t+q}\cdot\rho_i]\,; \quad \forall\, p \; and \; q \quad \text{(A3.5)}$$

from which the additional moment conditions for this part of the system are given by

$$E\left[\left(y_{i,t-s}-y_{i,t-s-1}\right)\times\left(\rho_i+\varepsilon_{i,t}\right)\right]=0 \qquad for\; s=1, \qquad \text{(A3.6)}$$

$$E\left[\left(x_{i,t-s}-x_{i,t-s-1}\right)\times\left(\rho_i+\varepsilon_{i,t}\right)\right]=0 \qquad for\; s=1. \qquad \text{(A3.7)}$$

Conditions (A3.3) to (A3.7) provide the basis for the GMM procedure to generate consistent estimates of the parameters of interest in which the weighting matrix can be any symmetric, positive definite matrix (Arellano and Bover, 1995). From these matrices, the most efficient GMM estimator is generated by applying the weighting matrix based on the variance-covariance matrix for the moment conditions. Consistency of this GMM estimator relies on whether the validity of the lagged explanatory variables as adequate instruments holds in practice. According to Arellano and Bond (1991) and Arellano and Bover (1995), two tests can be implemented to verify the validity of such instruments, the Sargan test for over-identifying restrictions and, the second-order serial correlation test. The former is expressed as follows:

$$s = \hat{\epsilon}^{\prime} Z\left(\frac{1}{N}\sum_{i=1}^{N} Z_i^{\prime}\,\hat{\epsilon}_i\hat{\epsilon}_i^{\prime} Z_i\right)^{-1} Z_i^{\prime}\,\hat{\epsilon} \qquad \text{(A3.8)}$$

where $\hat{\epsilon}_{jt}$ are the estimated residuals and Z represents the set of valid instruments in the differenced equation. Under the null hypothesis that instruments are exogenous, S follows a χ^2_{m-r} distribution where $m - r$ is the number of instruments minus the number of exogenous variables. The Sargan test evaluates the overall validity of the instruments by assessing the sample analogue of the moment conditions used in the estimation process in which failure to reject the null hypothesis gives support to our model.

The second test examines the hypothesis of no serial correlation in the error term. In particular, we test whether the residuals from the regression in differences are first- and second-order serially correlated. Following Arellano and Bond (1991) and Arellano and Bover (1995), when this test fails to reject the null hypothesis of no second-order serial correlation, we conclude that the original error term is serially uncorrelated in accordance to the moment conditions set above.

Results for the GMM procedure outlined above are presented in Tables A3.2.1 and A3.2.2 in this Appendix. Models presented in Table A3.2.1 feature *ipc* as the trade explanatory variable whereas models in Table A3.2.2 use *eoc* as a trade variable. The layout of results in these two tables is divided along the breakdowns of the labour force outlined along this chapter, namely, all workers (Columns 1 to 4), white collar workers

(Columns 5 to 8) and, blue collar workers (Columns 9 to 12). For each of the labour force breakdowns, the three capital variables are introduced, first, one by one and then, simultaneously. Arellano-Bond test for first and second order serial correlation is presented at the bottom ob tables, followed by the Sargan test for the overall validity of the instruments. These results are used for reference to comment other models in the main text.

Table A3.2.1: Arellano-Bover/Blundell-Bond linear dynamic panel-data estimates: female share of jobs across manufacturing industries, 1981-2000. Trade variable: import penetration coefficient (ipc)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	***All workers***				***White-collar workers***				***Blue-collar workers***			
Lagged Dep. Var.	0.7854***	0.7900***	0.7740***	0.7729***	0.6146***	0.6424***	0.5291***	0.5146***	0.8217***	0.8159***	0.8211***	0.8161***
	(0.0304)	(0.0304)	(0.0310)	(0.0316)	(0.0297)	(0.0295)	(0.0313)	(0.0322)	(0.0285)	(0.0282)	(0.0283)	(0.0286)
ipc	0.0076	0.0207**	0.0014	0.0006	0.0771***	0.1182***	0.0430**	0.0485**	0.0165	0.0119	0.0216*	0.0173
	(0.0114)	(0.0098)	(0.0128)	(0.0129)	(0.0210)	(0.0197)	(0.0206)	(0.0209)	(0.0115)	(0.0102)	(0.0127)	(0.0129)
CIGP	-0.0622***	-0.0587***	-0.0581***	-0.0599***	-0.1850***	-0.1519***	-0.1754***	-0.1592***	-0.0353**	-0.0332**	-0.0374**	-0.0355**
	(0.0147)	(0.0146)	(0.0145)	(0.0147)	(0.0194)	(0.0179)	(0.0170)	(0.0189)	(0.0148)	(0.0146)	(0.0149)	(0.0150)
lnkpw_mach	0.0054**			0.0026	0.0175***			0.0117**	-0.0017			0.0014
	(0.0024)			(0.0037)	(0.0038)			(0.0057)	(0.0025)			(0.0041)
lnkpw_trans		-0.0011		-0.0018		0.0061**		0.0013		-0.0039**		-0.0036**
		(0.0016)		(0.0016)		(0.0028)		(0.0027)		(0.0017)		(0.0017)
lnkpw_office			0.0049**	0.0036			0.0248***	0.0320***			-0.0024	-0.0023
			(0.0021)	(0.0033)			(0.0031)	(0.0048)			(0.0021)	(0.0036)
Constant	0.0392*	0.0861***	0.0588***	0.0555**	0.0637**	0.1454***	0.0995***	0.1465***	0.0675***	0.0768***	0.0675***	0.0769***
	(0.0219)	(0.0154)	(0.0149)	(0.0245)	(0.0286)	(0.0205)	(0.0164)	(0.0320)	(0.0242)	(0.0142)	(0.0162)	(0.0269)
Observations	551	551	551	551	551	551	551	551	551	551	551	551
Arellano-Bond test (p-values)												
First order	0.0012	0.0009	0.0013	0.0012	0.0024	0.0021	0.0033	0.0031	0.0022	0.0027	0.0022	0.0024
Second order	0.1996	0.2191	0.213	0.2049	0.225	0.2269	0.2338	0.2398	0.4506	0.4876	0.466	0.4681
***Sargan test:* χ^2(188)**	92.541	93.108	90.995	90.109	25.817	26.229	26.245	25.799	50.367	49.493	49.828	49.172
p-values	0.0520	0.0479	0.0647	0.0731	0.2135	0.1978	0.1972	0.2142	0.1050	0.1211	0.1147	0.1274

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A3.2.2: Arellano-Bover/Blundell-Bond linear dynamic panel-data estimates: female share of jobs across manufacturing industries, 1981-2000. Trade variable: export orientation coefficient (eoc)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	***All workers***				***White-collar workers***				***Blue-collar workers***			
Lagged Dep. Var.	0.7865***	0.7968***	0.7726***	0.7711***	0.6472***	0.7014***	0.5382***	0.5206***	0.8306***	0.8215***	0.8327***	0.8231***
	(0.0299)	(0.0297)	(0.0310)	(0.0317)	(0.0291)	(0.0286)	(0.0314)	(0.0326)	(0.0273)	(0.0273)	(0.0274)	(0.0279)
eoc	0.0069	0.0139**	0.0062	0.0053	0.0165	0.0387***	0.0102	0.0154	0.0144*	0.0122*	0.0152**	0.0136*
	(0.0072)	(0.0066)	(0.0072)	(0.0073)	(0.0132)	(0.0130)	(0.0123)	(0.0124)	(0.0076)	(0.0070)	(0.0077)	(0.0077)
CIGP	-0.0624***	-0.0577***	-0.0597***	-0.0612***	-0.1958***	-0.1537***	-0.1792***	-0.1626***	-0.0328**	-0.0313**	-0.0337**	-0.0323**
	(0.0147)	(0.0146)	(0.0144)	(0.0146)	(0.0195)	(0.0186)	(0.0171)	(0.0190)	(0.0147)	(0.0146)	(0.0148)	(0.0148)
lnkpw_mach	0.0053**			0.0023	0.0216***			0.0117**	-0.0017			0.0002
	(0.0023)			(0.0037)	(0.0038)			(0.0058)	(0.0024)			(0.0041)
lnkpw_trans		-0.0012		-0.0019		0.0041		-0.0000		-0.0041**		-0.0039**
		(0.0016)		(0.0016)		(0.0029)		(0.0027)		(0.0017)		(0.0017)
lnkpw_office			0.0044**	0.0034			0.0271***	0.0346***			-0.0017	-0.0010
			(0.0018)	(0.0029)			(0.0029)	(0.0047)			(0.0019)	(0.0032)
Constant	0.0401*	0.0867***	0.0620***	0.0603**	0.0344	0.1552***	0.0915***	0.1459***	0.0657***	0.0763***	0.0610***	0.0798***
	(0.0210)	(0.0156)	(0.0139)	(0.0250)	(0.0288)	(0.0218)	(0.0163)	(0.0332)	(0.0232)	(0.0143)	(0.0147)	(0.0271)
Number of obs.	551	551	551	551	551	551	551	551	551	551	551	551
Arellano-Bond test (p-values)												
First order	0.0014	0.0012	0.0015	0.0014	0.0035	0.0031	0.0043	0.004	0.0032	0.0032	0.0032	0.0032
Second order	0.1996	0.2022	0.2154	0.2087	0.2281	0.24	0.2358	0.2469	0.4407	0.4717	0.446	0.4562
***Sargan test:* χ^2(188)**	91.778	92.128	91.251	19.214	27.426	27.947	27.849	27.565	50.229	48.909	49.404	49.028
p-values	0.0580	0.0552	0.0624	0.5714	0.1572	0.1417	0.1445	0.1529	0.1074	0.1329	0.1228	0.1304

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Conclusions, Limitations and Agenda for Further Research

Some of the findings of this doctoral thesis research indicate that gender differences in the labour market of urban Colombia have evolved in a positive way. According to the estimates presented in Chapters 1 and 2, segregation measures have exhibited a statistically significant decrease over a time span of two decades. Likewise, gender wage differentials presented in Chapter 2 have experienced sizeable reductions over the same years for workers in both the formal or waged employment and informal or own-account employment. Taking into account the increase of GDP per capita and other welfare indicators recorded by this country over most of the 1980s and 1990s, these trends are suggestive of Boserup's (1970) assertion that gender discrimination (either in the form of occupational segregation or wage discrimination) should decrease with the pace of economic development.

The breakdown of segregation measures across different groups of the labour force suggests that this positive picture of the evolution of gender differences in urban Colombia is a fate not shared by everyone. In Chapter 1 we found marked differences in the extent of occupational segregation whereby dissimilarity indices have increased (not decreased!) for older workers and those with secondary or less education. According to the results for the Shapley decomposition implemented in Chapter 1, only in the case of workers with university education and those in government positions was there a change in the gender composition of individual occupations towards a less segregated pattern in the gender distribution of jobs. In fact, dissimilarity measures indicate that gender-based occupational segregation remains at very high levels despite the statistical reductions observed over the years examined here. Thus, most of the reduction in segregation measures observed between the mid-1980s and 2004 was driven by the increasing proportion of women in the overall composition of the labour force. In other words, these results show that, despite the fact that women represent an increasing proportion of the workforce, most occupations have remained equally segregated across all years reviewed in this study, particularly for those with the lowest educational levels and those in both middle age and of older years.

We have also investigated the extent to which occupational segregation explains gender wage differences in urban Colombia. The results for Chapter 2 indicate that informal workers are not only more segregated in terms of gender than their formal counterparts, but also that the magnitude of the pay gap between women and men is the widest. As in Chapter 1, these results are suggestive that institutions do play a role in determining gender differences in the labour market. However, the same results point towards a substantial reduction in hourly wage differentials between women and men in urban Colombia since the mid-1980s.

Improvements in educational levels amongst female workers appear as the main driving force behind the reduction in the magnitude of wage differentials between men and women, particularly in the formal segment where the most educated tend to work. Surprisingly, we find that occupational segregation does not play an influential role in determining gender wage differences in the labour market of urban Colombia. Our decomposition results, using an innovative framework in which the explained and unexplained portions of occupational segregation are accounted for, are suggestive that the way women are sorted out into informal occupations actually helps to reduce gender pay differences in that segment of the labour market. In this sense, given their low levels of human capital and other valuable characteristics, informal women do better by working as domestic servants compared to other male dominated occupations in the informal economy where their labour incomes would presumably be lower. This issue raises an interesting research question for the future in looking at this domestic help sector in its own right. For those in the formal employment segment, we found that the explained portion of occupational segregation contributes towards reducing the wage gap amongst formal workers, a result that is concurrent with increasing educational levels of the female labour force in this country.

Even though occupational segregation does not exert a considerable role on gender pay differences, we find that female occupation intensity is associated with lower wages, particularly in the formal sector where the penalty associated with the proportion of women in occupations is substantially higher than that in the informal sector. This result is in line with similar findings obtained for other countries such as United States, Czech Republic, and Slovak Republic. Interestingly, the same econometric results suggest that men enjoy a wage advantage in female dominated occupations, indicating that gender wage discrimination is not ascribed to lower remuneration of typically female dominated occupations. This finding represents an avenue for future research in which information on job characteristics could be linked to household survey data in order to explain more comprehensively the roots of female wage disadvantage (as well as men's wage advantage) in female-dominated occupations.

As explained in the introduction to this thesis, Colombian economic authorities implemented a comprehensive package of market oriented reforms at the beginning of the 1990s. At the core of the reforms, trade liberalisation was assumed to play a central role in order to enhance competition and a more efficient allocation of productive resources of the economy. In Chapter 3 we examined the effects of trade liberalisation in the incorporation of women in manufacturing industries. In contrast to the type of data used in the previous two chapters, we conducted this piece of research with aggregate data from manufacturing industries on employment by gender and skill level, the stocks of different types of capital equipment, and information on the degree of market competition. Our measure of market concentration was based on a constructed index in which the ratio of the production value from the largest four firms within each industry and year is expressed as a proportion of the total production value in a given industry. We found convincing evidence that increased levels of import penetration are positively associated with higher female shares of jobs in manufacturing industries, a result that could be tested under different instrumental variable techniques. Increasing levels of export orientation suggest a similar pattern although this result might be even more pronounced in the case of *blue collar workers*. In the same chapter, we found persuasive evidence that higher levels of market concentration are negatively associated with the female share of jobs in manufacturing industries, indicating that more competitive environments are more likely to incorporate larger shares of female employment. In general, this is what we expected to find in relation to the segregation dimension implicit in Becker's hypothesis of labour market discrimination. We could also verify some complementarities between female labour and the use of some types of capital equipment. In particular, our estimates are suggestive that the increasing use of office equipment is correlated with higher shares of female employment in the manufacturing industries of urban Colombia. This finding provides additional evidence to the existing literature in which the increasing use of technology provides new labour opportunities for women.

The empirical strategy adopted in the third chapter of the thesis was dictated by major difficulties in relation to the information available to estimate gender wage gaps across manufacturing industries from household survey microdata. Ideally, an assessment of the effects of trade on gender employment patterns should rely on wage differentials by gender. Our efforts to link changes in trade flows originated in the economic liberalisation of the Colombian economy to wage patterns by gender from household survey microdata were inconclusive. Unfortunately, the manufacturing surveys data do not report information of labour costs or wages by gender. Statistical analyses not reported in this thesis from household survey microdata suggested that this source has limitations in terms of the accuracy in the recording of the information related to the ISIC codes (at two-digit level) to describe the economic activity of household respondents in the workforce, which is based on the supply side of the labour market, as opposed to the Annual Manufacturing Survey. Therefore, our attempts to verify a relationship between trade measures and gender wage gaps were inconclusive using the household survey data. We still believe this issue remains as an important avenue for further research. As a result of these limitations, we remain agnostic as to whether the effects of increased competition, either in the form of import penetration or in terms of market concentration, exert any effect on the degree of gender pay discrimination. For all these reasons, our findings are only suggestive of some of the positive effects of trade on gender differences in the labour market and further research is needed to establish whether the participation of women in Colombian manufacturing industries was accompanied by a reduction in gender discrimination.

Looking at the results from all three empirical chapters presented in this thesis, there is a common pattern that dominates the story of gender differences in the labour market of urban Colombia. In Chapters 1 and 2 we could see how the situation of women in informal or precarious working conditions is substantially different from that of women in formal or waged employment. Although the type of data in Chapter 3 refers exclusively to formal employment, it also reveals that the beneficial effects of trade and increasing competition on the incorporation of women in manufacturing tend to be concentrated in the white-collar workers category who are, presumably, the better rewarded in manufacturing employment.

All of the foregoing suggests that gender differences in the labour market are subject to a great degree of heterogeneity between different groups of the labour force. The story from urban Colombia suggests that the most educated females are making good progress in reducing their differences with respect to their male counterparts. They also tend to be less discriminated against in terms of pay and have more options to choose from different types of jobs, mainly in the formal sector where labour institutions offer better working conditions. At the other side of the labour market, sizeable numbers of women are still facing difficult working conditions with no contract or labour guarantees, and their income disparities with respect to men are substantially wider. Gender differences amongst so-called informal employment in Colombia are also decreasing but at a substantially slower rate. The economic reforms are providing Colombians with both new opportunities and challenges. In sum, gender differences in the labour market of urban Colombia have been shortened substantially since the mid-1980s, but their pace of improvement is progressing to a large extent slower for the most disadvantaged women.

References

ACEMOGLU, D. 2002. Technical Change, Inequality, and the Labor Market. *Journal of Economic Literature,* 40**,** 7-72.

AGUAYO-TELLEZ, E., AIROLA, J. & JUHN, C. 2010. Did Trade Liberalization Help Women? The Case of Mexico in the 1990s. *Working Paper Series 16195.* Cambridge, MA: National Bureau of Economic Research.

AKTER, S. 2005. Occupational Segregation, Wage Discrimination, and Impact on Poverty in Rural Bangladesh. *The Journal of Developing Areas,* 39**,** 15-39.

ANGRIST, J. D. & PISCHKE, J.-S. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion,* Princeton, New Jersey.

ANKER, R. 1997. Theories of occupational segregation by sex: An overview. *International Labour Review,* 136**,** 315.

ANKER, R., MALKAS, H. & KORTEN, A. 2003. Gender-based occupational segregation in the 1990's. *In:* PEEK, P. & SHAHEED, Z. (eds.) *Declaration Working Paper No.16.* Geneva: ILO.

ANKER, R. & MELKAS, H. 1997. Occupational segregation by sex in Nordic countries: An empirical investigation. *International Labour Review,* 136**,** 341.

APPLETON, S., HODDINOTT, J. & KRISHNAN, P. 1999. The Gender Wage Gap in Three African Countries. *Economic Development and Cultural Change,* 47**,** 289-312.

ARANGO, L. E. & POSADA, C. E. 2002. La participación laboral en Colombia. *Borradores de Economía No.217* [Online]. Available: http://www.banrep.gov.co/docum/ftp/borra217.pdf [Accessed August, 2002].

ARELLANO, M. & BOND, S. 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies,* 58, 277-297.

ARELLANO, M. & BOVER, O. 1995. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics,* 68, 29-51.

ARTECONA, R. & CUNNINGHAM, W. 2002. Effects of Trade Liberalization on the Gender Wage Gap in Mexico. *Policy Research Report on Gender and Development.* Washington: The World Bank, Development Research Group.

ARULAMPALAM, W., ALISON, L. B. & BRYAN, M. L. 2007. Is There a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution. *Industrial and Labor Relations Review,* 60, 163-186.

ATTANASIO, O., GOLDBERG, P. K. & PAVCNIK, N. 2004. Trade reforms and wage inequality in Colombia. *Journal of Development Economics,* 74, 331-366.

BAKER, M. & FORTIN, N. M. 2001. Occupational Gender Composition and Wages in Canada, 1987-1988. *The Canadian Journal of Economics / Revue canadienne d'Economique,* 34, 345-376.

BAUNACH, D. M. 2002. Trends in Occupational Sex Segregation and Inequality, 1950 to 1990. *Social Science Research,* 31, 77-98.

BAYARD, K., HELLERSTEIN, J., NEUMARK, D. & TROSKE, K. 2003. New Evidence on Sex Segregation and Sex Differences in Wages from Matched Employee-Employer Data. *Journal of Labor Economics,* 21, 887-922.

BECKER, G. 1957. *The Economics of Discrimination* Chicago, Ill, University of Chicago Press.

BECKER, G. 1971. *The Economics of Discrimination* Chicago, University of Chicago Press.

BERIK, G., VAN DER MEULEN RODGERS, Y. & ZVEGLICH, J. E. 2004. International Trade and Gender Wage Discrimination: Evidence from East Asia. *Review of Development Economics,* 8, 237-254.

BLACK, S. E. & BRAINERD, E. 2004. Importing Equality? The Impact of Globalization on Gender Discrimination. *Industrial and Labor Relations Review,* 57, 540-559.

BLACKBURN, R. M., BROOKS, B. & JARMAN, J. 2001. Occupational Stratification: The Vertical Dimension of Occupational Segregation. *Work, Employment & Society,* 15, 511-538.

BLACKBURN, R. M. & JARMAN, J. 2005. Segregation and Inequality. *GeNet Working Paper,* 3, 15.

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

BLUNDELL, R. & BOND, S. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics,* 87, 115-143.

BOOTH, A. L., FRANCESCONI, M. & FRANK, J. 2003. A sticky floors model of promotion, pay, and gender. *European Economic Review,* 47, 295-322.

BORGHANS, L. & GROOT, L. 1999. Educational presorting and occupational segregation. *Labour Economics,* 6, 375-395.

BOSERUP, E. 1970. *Woman's Role in Economic Development,* New York, St Martin's Press.

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.

CAMERON, C. & TRIVEDI, P. K. 2009. *Microeconometrics Using Stata,* College Station, Texas, StataCorp LP.

CHAMARBAGWALA, R. 2006. Economic Liberalization and Wage Inequality in India. *World Development,* 34**,** 1997-2015.

COLLIER, P. 1994. A Theoretical Framework and the Africa Experience: Gender Aspects of Labor Allocation During Structural Adjustment. *In:* HORTON, S. K., R., AND MAZUMDAR, D. (ed.) *Labor Markets in an Era of Adjustment.* Washington: The International Bank for Reconstruction and Development/The World Bank.

DANE 2000. Encuesta Nacional de Hogares, Ciudades y Areas Metropolitanas, Manual de Crítica y Codificación, Etapa 108. Dirección Técnica de Estadísticas Básicas, DANE, Bogotá, D.C.

DANE 2004. Metodología de la Encuesta de Hogares. *In:* BÁSICAS, D. T. D. E. (ed.). Bogotá D.C.: Departamento Administrativo Nacional de Estadística.

DANE 2010. Boletín de Prensa: Medición del Empleo Informal, Trimestre Enero - Marzo de 2010. *In:* BÁSICAS), D. A. N. D. E. D. T. D. E. (ed.). Bogotá.

DARITY, W. A., JR. & WILLIAMS, R. M. 1985. Peddlers Forever?: Culture, Competition, and Discrimination. *The American Economic Review,* 75**,** 256-261.

DEMING, W. E. & STEPHAN, F. F. 1940. On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known. *The Annals of Mathematical Statistics,* 11**,** 427-444.

DEUTSCH, J., FLUCKIGER, Y. & SILBER, J. 1994. Measuring occupational segregation : Summary statistics and the impact of classification errors and aggregation. *Journal of Econometrics,* 61**,** 133-146.

DEUTSCH, J., FLÜCKIGER, Y. & SILBER, J. 2006. The Concept of Shapley Decomposition and the Study of Occupational Segregation -paper presented at the international conference on. *New Frontiers in the Field of Segregation Measurement and Analysis.* Monte Verita, Switzerland.

DEUTSCH, R., MORRISON, A., PIRAS, C. & ÑOPO, H. 2002. Occupational Segregation by Gender in Three Latin American Countries. *In:* DEPARTMENT, S. D. (ed.) *Technical Papers Series.* Washington, D. C.: Inter-American Development Bank.

DUFLO, E. 2005. Gender Equality in Development. *BREAD Policy Paper No. 011.* Bureau for Research in Economic Analysis of Development - Massachusetts Institute of Technology.

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.

ESLAVA, M., HALTIWANGER, J., KUGLER, A. & KUGLER, M. 2009. Trade Reforms and Market Selection: Evidence from Manufacturing Plants in Colombia. *Discussion Paper No. 4256.* Bonn, Germany: The Institute for the Study of Labor (IZA).

FLÓREZ, C. E. 2003. Migration and the Urban Informal Sector in Colombia. *Conference on African Migration in Comparative Perspective,.* Johannesburg, South Africa.

FONTANA, M. 2003. The Gender Effects Of Trade Liberalisation In Developing Countries: A Review Of The Literature. *Discussion Paper Series -Center for Economic Policy Research.* Brighton: University of Sussex.

GALOR, O. & WEIL, D. N. 1996. The Gender Gap, Fertility, and Growth. *The American Economic Review,* 86, 374-387.

GILBERT, A. 1997. Employment and Poverty during Economic Restructuring: The Case of Bogotá, Colombia. *Urban Studies,* 34, 1047-1070.

GOLDBERG, P. & PAVCNIK, N. 2003. The response of the informal sector to trade liberalization. *Journal of Development Economics,* 72, 463-496.

GOLDBERG, P. K. & PAVCNIK, N. 2005a. The effects of the Colombian trade liberalization on urban poverty. *NBER Working Paper 11081.* Cambridge, MA: National Bureau of Economic Research.

GOLDBERG, P. K. & PAVCNIK, N. 2005b. Trade, wages, and the political economy of trade protection: evidence from the Colombian trade reforms. *Journal of International Economics,* 66, 75-105.

GRAZIER, S. & SLOANE, P. J. 2007. Accident risk, gender, family status and occupational choice in the UK. *Labour Economics,* 15, 968-957.

GROSHEN, E. L. 1991. The Structure of the Female/Male Wage Differential: Is It Who You Are, What You Do, or Where You Work? *The Journal of Human Resources,* 26, 457-472.

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.

HANRATTY, D. M. & MEDITZ, S. W. 1988. *Colombia: A Country Study,* Washington.

HECKMAN, J. J. 1979. Sample Selection Bias as a Specification Error. *Econometrica,* 47, 153-161.

ISAZA CASTRO, J. G. 2002. Women Workers in Bogotá's Informal Sector: Gendered Impact of Structural Adjustment Policies in The 1990s. *Planeación y Desarrollo,* XXXIII**,** 411-468.

ISAZA, J. G. 2002. Women Workers in Bogotá's Informal Sector: Gendered Impact of Structural Adjustment Policies in The 1990s. *Planeación y Desarrollo,* XXXIII**,** 411-468.

ISAZA, J. G. 2006. Trade Liberalization and Gender Effects: A Literature Review for Colombia. *Equidad y Desarrollo,* 4.

ISAZA, J. G., ACOSTA, C. E. & MEZA, C. A. 2007. Género, Pobreza Y Dinámicas Del Mercado De Trabajo: Un Análisis De Microsimulaciones Entre 1990 Y 2004 Para Las Áreas Urbanas De Colombia. *In:* RAMIREZ, L. F. & ISAZA, J. G. (eds.) *Empleo, pobreza y desigualdad: Una mirada a partir de la investigación Universitaria* Bogotá: Ediciones Unisalle.

ISAZA, J. G., ACOSTA, C. E. & MEZA, C. A. 2010. *Gender, Poverty and Labour Market Dynamics: A Microsimulations Analysis for Urban Colombia between 1990 and 2004,* Saarbrücken, Germany, LAP Lambert Academic Publishing.

JARAMILLO, C. & TOVAR, J. 2006. Trade and labor markets: evidence from the Colombian trade liberalization process. *DOCUMENTO CEDE No.28.* Bogotá: Universidad de Los Andes.

JURAJDA, S. 2003. Gender wage gap and segregation in enterprises and the public sector in late transition countries. *Journal of Comparative Economics,* 31**,** 199-222.

JURAJDA, S. & HARMGART, H. 2007. When do female occupations pay more? *Journal of Comparative Economics,* 35**,** 170-187.

JURAJDA, S. N. 2005. Gender segregation and wage gap: an east-west comparison. *Journal of the European Economic Association,* 3**,** 598-607.

KARMEL, T. & MACLACHLAN, M. 1988. Occupational Sex Segregation--Increasing or Decreasing? *Economic Record,* 64**,** 187.

KATZ, L. F. & MURPHY, K. M. 1992. Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *The Quarterly Journal of Economics,* 107**,** 35-78.

KILLINGSWORTH, M. R., HECKMAN, J. J. & ORLEY, C. A. A. R. L. 1986. Chapter 2 Female labor supply: A survey. *Handbook of Labor Economics.* Elsevier.

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.

LUCIFORA, C. & REILLY, B. 1990. Wage Discrimination and Female Occupational Intensity. *Labour,* 4**,** 147-168.

MACPHERSON, D. A. & HIRSCH, B. T. 1995. Wages and Gender Composition: Why do Women's Jobs Pay Less? *Journal of Labor Economics,* 13**,** 426-471.

MAIN, B. G. M. & REILLY, B. 1992. Women and the Union Wage Gap. *The Economic Journal,* 102**,** 49-66.

MELKAS, H. & ANKER, R. 1997. Occupational segregation by sex in Nordic countries: An emprical investigation. *International Labour Review,* 136.

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.

MULEKAR, M. S., KNUTSON, J. C. & CHAMPANERKAR, J. A. 2007. How useful are approximations to mean and variance of the index of dissimilarity? *Computational Statistics & Data Analysis,* In Press, Corrected Proof.

NEUMAN, S. & SILBER, J. G. 1996. Wage discrimination across ethnic groups: Evidence from Israel. *Economic Inquiry,* 34**,** 648.

OAXACA, R. 1973. Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review,* 14**,** 693-709.

OCAMPO, J. A., SÁNCHEZ, F. J., HERNÁNDEZ, G. A. & PRADA, M. F. 2004. Crecimiento de las exportaciones y sus efectos sobre el empleo, la desigualdad y la pobreza en Colombia. *DOCUMENTO CEDE***,** 34.

OOSTENDORP, R. H. 2009. Globalization and the Gender Wage Gap. *The World Bank Economic Review,* 23**,** 141-161.

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

POMBO, C. 1999. Productividad industrial en Colombia: una aplicación de números índices. *Revista de Economía de la Universidad del Rosario,* 2**,** 107-139.

QIAN, N. 2008. Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance. *The Quarterly Journal of Economics,* 123**,** 1251-1285.

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

REILLY , B. & VASEUDEVA, P. 2005. The Gender Pay Gap and Trade Liberalisation: Evidence for India. *PRUS Working Paper No.32.* Brighton: Poverty Research Unit, Department of Economics, University of Sussex.

REILLY , B. & VASUDEVA-DUTTA, P. 2005. The Gender Pay Gap and Trade Liberalisation: Evidence for India. *PRUS Working Paper,* 32**,** 1-52.

ROBERTS, M. J. & SKOUFIAS, E. 1997. The Long-Run Demand for Skilled and Unskilled Labor in Colombian Manufacturing Plants. *The Review of Economics and Statistics,* 79**,** 330-334.

ROBERTS, M. J. & TYBOUT, J. R. 1997. The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs. *The American Economic Review,* 87**,** 545-564.

SANTA MARÍA, M. & ROJAS, N. 2001. La participación laboral: qué ha pasado y qué podemos esperar. *ARCHIVOS DE ECONOMÍA.* Bogotá: Departamento Nacional de Planeación.

SCHAFFER, M. E. & STILLMAN, S. 2010. xtoverid: Stata module to calculate tests of overidentifying restrictions after xtreg, xtivreg, xtivreg2 and xthtaylor. http://ideas.repec.org/c/boc/bocode/s456779.html.

SEGUINO, S. 2000. Gender Inequality and Economic Growth: A Cross-Country Analysis. *World Development,* 28**,** 1211-1230.

SEMYONOV, M. & JONES, F. 1999. Dimensions of Gender Occupational Differentiation in Segregation and Inequality: A Cross-National Analysis. *Social Indicators Research,* 46**,** 225-247.

SHAUMAN, K. A. 2006. Occupational sex segregation and the earnings of occupations: What causes the link among college-educated workers? *Social Science Research,* 35**,** 577-619.

SILBER, J. G. 1989. On the measurement of employment segregation. *Economics Letters,* 30, 237-243.

TENJO, J. & RIBERO, R. 1998. Participación, desempleo y mercados laborales en Colombia. *ARCHIVOS DE ECONOMÍA.* Bogotá: Departamento Nacional de Planeación.

TZANNATOS, Z. 1999. Women and Labor Market Changes in the Global Economy: Growth Helps, Inequalities Hurt and Public Policy Matters. *World Development,* 27, 551-569.

WEINBERG, B. A. 2000. Computer Use and the Demand for Female Workers. *Industrial and Labor Relations Review,* 53, 290-308.

WELCH, F. 2000. Growth in Women's Relative Wages and in Inequality among Men: One Phenomenon or Two? *The American Economic Review,* 90, 444-449.

WHITE, H. 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica,* 48, 817-838.

WILLIAMS, R. M. 1987. Capital, Competition, and Discrimination: A Reconsideration of Racial Earnings Inequality. *Review of Radical Political Economics,* 19, 1-15.

WILLIAMS, R. M. & KENISON, R. E. 1996. The Way We Were?: Discrimination, Competition, and Inter-industry Wage Differentials in 1970. *Review of Radical Political Economics,* 28, 1-32.

WOOD, A. 1991. North-South Trade and Female Labour in Manufacturing: An Asymmetry. *Journal of Development Studies,* 27, 168.

WRIGHT, R. E. & ERMISCH, J. F. 1991. Gender Discrimination in the British Labour Market: A Reassessment. *The Economic Journal,* 101, 508-522.