UNIVERSITY OF OTTAWA

Estimating the Gender Earnings Gap in Brazil

Ghadeer Daham

2285832

Major Paper presented to the Department of Economics of the University of Ottawa in partial fulfillment of the requirements of the M.A. Degree

Supervisor: Professor Fernanda Estevan

1. Introduction

Gender discrimination is a major area of research in social sciences including economics. Economists attempt at empirically measuring the effect of gender discrimination in a major part of society: the labour market. Labour economists pursue the study of gender discrimination by investigating the male-female wage differentials and seek to identify whether systemic discrimination contributes to some of the wage differential. Some societies in the world have a common the perception that women come 'second' to men when it comes to major responsibilities including participation in the labour force. However, the role of males as 'bread earners' and females as 'child bearer' is changing overtime. Female labour participation rate have increased worldwide even in less developed economies along with a decline in fertility rates (United Nations, 2010).

On a global spectrum, the 'The World's Women 2010' report by the United Nations show that women's labour participation rate increased in the 1980s, but remained stable in the last two decades from 1990 to 2010 lingering around 52%. Conversely, the global male participation rate declined steadily from 81% to 77% during the same period. Women labour participation varies in different sub-regions showing increases in some regions and declines in others. In particular, during the 1990 to 2010 period, the Caribbean and Latin America regions showed noteworthy increases in women's labour participation from a little below 40% to more than 50%. This can be compared to approximately 40% and 30% in Southern and Western Asia and Northern Africa for the same period, respectively.

Females are evidently more attached to the labour force than before. They also outperform males in certain levels of education and have similar experience and training (Drolet, 2001; United Nations, 2010). However, women still receive lower wages than men. The International Labour Organization (ILO) reports that in 2008, women earned on average 70% to 90% of men's wage in most countries. In 30 countries from the European Union, for example, the wage differential is between 15% and 25% and the bigger gender pay gap for women is found at the senior occupational level and for women with higher educational backgrounds. On average and across all occupations, women earn 68%, 85%, and 81% relative to males, in the Republic of Korea, United Kingdom, and Brazil, respectively. As in Europe, the wage differential is larger in the high paid occupations in Brazil and United Kingdom (ILO, 2009).

Economists with an interest in examining race and gender wage differentials developed decomposition techniques to determine the source of the persistent wage gap. A fundamental and frequently used approach decomposes the wage gap by isolating a component related to observable characteristics, which includes human capital skills and individual attributes. The remaining component of the wage differential constitutes of characteristics that were not observed in the study and may be partially considered systemic discrimination against women (Oaxaca, 1973; Neumark, 1988; Oaxaca and Ransom, 1999; Baker and Drolet, 2010). These two components in the wage differential are referred to 'explained' and 'unexplained', respectively, by most researchers in the field.

The focus of this paper is to identify the sources of earnings differential between males and females in a developing country context. I conduct a male-female earnings differential analysis using recent cross-sectional Brazilian data for December 2011 and compare my results to previous research to determine whether the wage differential has declined overtime. To examine the change in wage differential, I utilize standard decomposition techniques that are prominent in the economics of discrimination literature (Oaxaca, 1973; Neumark, 1988). The measurement will decompose the wage differential into two components: the difference of the average group productivity characteristics which is referred to the 'explained' part. The productivity characteristics are observable factors that include human capital and social variables. The second component accounts for the difference in the returns of those characteristics and factors not included in the model, which is referred to the 'unexplained' part and maybe an attribute of discrimination.

In this study, I use a linear regression model that measures the impact of wage determining characteristics on male and female monthly effective earnings. For both females and males, I find that wage determining characteristics, which I also refer to 'observable characteristics' such as education, potential experience, employment status, tenure, and firm size are positively correlated with monthly effective earnings. In addition, the results are statistically and economically significant for a wide range of factors. Consequently, I use the returns to observable characteristics to measure the difference between females and males leading to the second set of results. Using the Oaxaca decomposition, I find that 133%, 113%, and 110% cannot be explained in the gender earnings differential, across all specifications, respectively.[1] This means that the entire earnings differential cannot be explained by the observable characteristics chosen in my study and the explained component arises as negative due to females exceeding males in the averages of observable characteristics.

In addition, when using the Neumark decomposition, I find that 145%, 121%, and 117.3% cannot be explained across the specifications. For the correction of self-selectivity bias, I use the Heckman procedure (1979) along with the Oaxaca (1973) decomposition technique. I find that the earnings differential reduces when correcting for self-selection and the unexplained component reduces to 123%, 108.5% and 104% across the three specifications. The different decomposition methods and correction for potential biases have an effect on the magnitude of the explained and unexplained component. However, all the methods result in the entire earnings differential as an attribute of discrimination.

The remaining of the paper is structured as follows. Section 2 provides the literature review documenting some of the main findings in this field. Section 3 outlines the data and sample used in the study and its sources. In Section 4, I describe the mean earnings differentials between males and females. Section 5 introduces the econometric model. Section 6 lays out the results of the impact of certain characteristics on male-female earnings. Section 7 provides the log earnings decomposition of the wage differential, and finally, section 8 contains concluding remarks.

---

[1] The specifications used in this study are linear earnings regressions that measure the impact of selected observable characteristics on earnings. The first specification includes age, education, experience, race, dependence, school attendance. The second specification adds regions in Brazil, firm size, tenure, employment status, type of contract, and social security. The third specification only includes occupation to the second specification.

2. Literature Review

Although studies on the female-male wage differential analyze different productivity related characteristics, demographics and regions, they all confirm that the gender wage gap is persistent overtime. Furthermore, they all support the conclusion that the gender wage gap that cannot be explained by observable characteristics included in their models accounts for a large amount in the wage differential. In other words, the component in the wage differential that is attributed to unobservable characteristics, which may be partially discrimination, outweighs the component that explains the difference in earnings that result from males surpassing females in productivity or other characteristics deemed important.

Becker (1957) and Arrow (1972a) are considered the pioneers of theoretically modelling the economics of discrimination, particularly employers' discriminatory behaviour. Becker (1957) introduces the first competitive model of discrimination showing that employers may have 'discriminatory tastes' and prefer not to employ individuals from minority groups. Therefore, minority employees may have to be more productive than average or receive lower wages for identical productivity characteristics related to other groups. Arrow (1972a) examines the causes of the wage differential that may stem from nepotism towards the superior group or discrimination against the minority group. The authors develop assumptions thus, generating utility functions to demonstrate employer's discriminatory behaviour.

Under particular assumptions about the employers' discriminatory tastes introduced by Becker (1957), Oaxaca (1973) and Blinder (1973) simultaneously derive a decomposition technique which has become a core method when conducting empirical studies.[2] Since both studies emerged at the same time, many researchers refer to the decomposition technique as the Blinder-Oaxaca (1973) approach. Using two separate wage regressions for males and females, the wage differential is calculated. Subsequently, the wage differential is decomposed into two components: the difference in earnings of the average individual and human capital characteristics between two groups and the return to coefficients of the same characteristics, which he refers to the latter as the discrimination or residual component. This means that the effects of discrimination describe the residual that is deducted from effects of all the observable characteristics included in the regression from the overall wage differential. One strong assumption that Oaxaca (1973) makes when using the estimator of wage discrimination is the choice of the reference group before decomposing the wage differential. When using this method, one has to choose a female or male wage structure as the 'no-discrimination wage structure' so that to identify by how much does the female or male group are disadvantaged.

In general terms, the Oaxaca decomposition technique provides a comparison of wage structures for any two groups by estimating two wage regressions. The researcher chooses one group and their corresponding wage structure as the leading group that

[2] Becker (1957) assumes that employers have a 'taste for discrimination', a racial preference, for example. In a competitive market, discrimination arises as a result of deriving employers' utility functions from profits of the firm and also from the value of the composition of the labour force (superior and minority groups discriminated against). In order to allow for flexibility in conducting empirical work, the types of labour of both groups are perfectly substitutable. In partial equilibrium, disadvantaged employees must compensate employers by working harder or accepting lower wages relative to the superior group. In general equilibrium, the employers compensate for the value of their distaste and the cost of hiring the employees that are from the minority group.

predominates in the absence of discrimination. The researcher considers the chosen group as the 'reference group' or the "no-wage discrimination" structure that does not suffer from discrimination and, subsequently, compares it with the group that is assumed to suffer from discrimination. When the technique was first introduced, researchers were interested in identifying whether discrimination exists against disadvantaged groups that include females and black individuals. Therefore, the chosen reference groups for the empirical work when using the techniques were males and white individuals, respectively. The component which may be partially an attribute of discrimination in the wage differential is the difference between the current wage structure and the "no-wage discrimination" wage structure. The remaining part of the wage differential is an attribute of the differences in productivity related to characteristics included in the wage equation.

Oaxaca (1973) uses his model to examine the male-female wage differential utilizing the male wage structure as the base group. Using the 1967 American Survey of Economic Opportunity, he finds that the residual in the wage differential accounts for 77.7% and 93.6% for white and black individuals, respectively. Thus, differences in observable characteristics fail to explain almost the entire wage differential between the black males and females, leading to the conclusion that black females may be experiencing the most discrimination, and white females are not far behind black females.

Oaxaca (1973) also discusses that setting up the decomposition technique with a different reference group will yield different results, which he refers to the so-called 'index number problem'. In his study, he finds that when using the female wage structure as the "no-discrimination" wage structure, the residual component is 64% whereas when the male wage structure is used, the estimate reduces to 53%.

An alternative measurement that has been widely used is the Blinder (1973) decomposition technique. The slight difference between the Blinder (1973) from the Oaxaca (1973)'s model is the breakdown of the residual component of the wage differential into shift coefficients (intercepts) and the return to coefficients. Using the 1967 Michigan Survey, Blinder (1973) finds that the wage differential is an attribute of discrimination entirely. The breakdown of the discrimination component into the returns to coefficients and intercepts shows no significance leading to the same end results as the Oaxaca (1973) technique.

Jones (1983) discusses some limitations of the decomposition approach. In particular, he empirically shows that the residual cannot be broken into differences in coefficients and intercepts, due to the choice of base groups in the specifications. Oaxaca and Ransom (1999) reaffirm and formalize Jones critical claims of Blinder's method by conducting a comparative study using different base groups in two wage specifications. Using the 1989 American Survey of College and University Faculty, they find that the overall discrimination and productivity related characteristics in both specifications are the same, 0.90 and 0.176, respectively. However, the intercept terms along with the return to coefficients in the discrimination part differ in magnitude in both specifications. It is interesting to note that the Blinder approach is still widely used and can be found in Labour Economics textbooks as a standard method.[3]

Neumark (1988) explores ways to treat the index number problem and argues that a reference group should not be chosen arbitrarily because the type of discrimination can

[3] I do not use the Blinder-Oaxaca (1973) method in this study that breakdowns the intercept from the returns to coefficients component.

change the nature of decomposition, thus, leading to different results. He provides the example of gender discrimination that can stem from nepotism toward men which leads to overpayment for men while women are paid the non-discriminatory wage. Alternatively, discrimination against women can cause employers to pay men the competitive wage (non-discriminatory wage) but underpay women. In addition, employers can practice nepotism and discrimination at the same time. To tackle this issue, Neumark (1988) extends the original employers' discrimination models developed by Becker (1957) and Arrow (1972) and derives a no-wage discrimination structure, which is the result of a pooled wage structure defined as the 'weighted average of the male and female wage structure'. The no-discrimination wage structure is combined with the Oaxaca (1973) decomposition technique and has been known as the 'Neumark decomposition' technique.

Neumark (1988) takes the alternative decomposition technique and examines the female-male wage differential, using the 1980 American National Longitudinal Survey of Young Men and Women (NLS). He finds that when using the male and female wage structure as the no-wage discrimination structure, the residual component becomes 70% and 69%, respectively. However, the discrimination component reduces to 57% when the Neumark decomposition is used. Neumark (1988) has also faced some criticism. The pooled wage structure may result in an econometric bias because in some cases, it could amplify the explained component in the differential by transferring the unexplained component to the explained.

Another problem that researchers face in the application of the decomposition technique is the self-selection bias. Heckman (1976, 1979) describes the bias as a result of the non-randomly selected sample and treats this issue by considering it as an ordinary 'specification error' or 'omitted variables' bias when estimating wage equations. The bias occurs when individuals select themselves in a group which does not represent the population as a whole. For example, if 'ability' is an unobservable variable in the wage equation, and there is self-selection on ability, the average ability of individuals in the sample will differ from the average ability in the population. In general, wage equations estimated on selected samples do not estimate random samples drawn from the population. Problems can also arise when data is not randomly selected due to missing observations or attrition conducted by the researchers. This occurs in longitudinal studies and during the course of the study; researchers are forced to drop observations that belong to non-respondents. Heckman (1979) explains that this has the same composition effect as self-selection.

Heckman (1979) provides a treatment to tackle self-selection as a specification error arising from missing observations and attrition or an omitted variables bias. Heckman (1979) proposes a consistent two stage step method to correct for the bias. In the two gender wage equations, researchers only estimate wage observations for males and females who choose to work. This means individuals who work are selected non-randomly from the population. Estimating the wage determining characteristics from the selected population may lead to a bias because the dependent variable is observed only from a non-random sample. In the first stage, Heckman (1979) introduces a model based on economic theory for the probability to work which is represented in a probit regression. The second stage is when self-selection is corrected taking the transformed estimated individual probabilities as additional regressors, known as the 'inverse Mills ratio'. The method has become the standard estimation method in equations with specification errors known as the 'two step' or the 'limited maximum likelihood (LIML)' method. Generally, the procedure will then take out the selection effects from the overall wage differential then use the standard decomposition technique to determine the effect of unexplained component and the other component that can be explained by chosen characteristics. Jann (2008) provides an example correcting for selectivity using the Heckman (1979) procedure. He finds that the corrected wage differential reduces from 0.173 to 0.165. This is as a result from the adjustment of wages for females that were slightly downward biased. Consequently, the unexplained component reduced to 0.073 from 0.082 in the overall wage differential.

The above-mentioned issues have become a gateway for more research in this area. Economists have criticized some of the proposed treatments and developed alternative methods to correct for self-selection and the index number problem, thereby creating extensions of the Oaxaca (1973) model (Duan et al, 1983, 1984, 1985; Olsen, 1980; Neumark, 1988; J. Cotton (1988); Oaxaca and M. Ransom, 1999; Juhn et al, 1991; Puhani, 2000).

Canadian empirical studies using these decomposition techniques have been conducted in the last decades to examine whether the gender pay gap is persistent overtime (Baker et al, 1995; Baker and Fortin, 2001; Baker and Drolet, 2010). These papers identify the wage differential that can be compared over time by using the standard Oaxaca decomposition technique with methodological extensions to accommodate for the question of interest. Baker and Drolet (2010) investigate the gender wage gap using time series data over the last 30 years. The dataset is a combination of surveys between 1981 and 2008 containing wage data.[4] They use a one wage equation and decompose the wage differential with the standard Blinder-Oaxaca (1973) decomposition technique. The overall unexplained component of the log wage differential rises from below 70% in 1981 to approximately 80% in 2008.

Drolet (2001) argues that the choice of dataset is crucial for the purpose of obtaining accurate results and uses the 1997 Survey of Labour and Income Dynamics (SLID). Unlike most empirical studies in wage discrimination, she uses data that provides actual labour market experience instead of establishing a proxy that measures potential experience. Her results show that the use of proxies such as age or potential experience leads to larger gender wage gaps.[5] Drolet (2001) uses two different samples to measure earnings and to determine the gender differential: annual earnings and hourly wage rates. She highlights that hourly wage rates is a better measure since it avoids problems in annual earnings.[6] Using the Blinder-Oaxaca (1973) decomposition technique, she finds that 51% to 63% in the gender wage gap cannot be explained by observable characteristics, similar to other research of female/male differentials.[7]

[4] Baker and Drolet (2010) use the following surveys: Survey of Work History (SWH) for 1981, Survey of Union Membership (SUM) for 1984, the Labour Market Activity Survey (LMAS) for the 1986-1990 period, Survey of Labour Income Dynamics (SLID) for the 1993-1996 period, and Labour Force Survey (LFS) for 1997-2008 period

[5] As a proxy for actual labour market experience, economists use age or potential experience by taking age and subtracting number of years of schooling and an additional 6 years. Similarly, I use this proxy for experience since the actual experience variable is not available in the survey.

[6] Using annual earning, restrictions to full-year full time workers are needed due to discrepancies in the number of hours worked per week and number of weeks worked per year. Also, hourly wage rates give accurate results when controlling for job characteristics such as occupation or tenure. The annual earnings data contains the combination of wages and salaries from all jobs and does not distinguish particular job characteristics, which does not allow measuring the impact of a certain job characteristic.

[7] Gunderson (1998) find the unexplained component in the gender wage gap account for 53% to 70%; Baker et al (1995) find the unexplained component, particularly in the 1990s to account for 84%.

Many studies investigate the gender earnings gap in developing countries, particularly in African countries. In these studies, most economists adapt the standard decomposition methodologies, but include some modifications to correct for selectivity biases. Appleton et al (1999) conduct a comparative empirical study investigating the gender wage gap as well as its determinants in three African countries: Ethiopia, Uganda, and Cote d'Ivoire. They combine the Oaxaca and Neumark techniques, along with a sectoral decomposition technique to identify discrimination within the public and private sector. The latter decomposition reveals the residual component within sectors that cannot be explained by differences in characteristics. Using a number of surveys, they find that the gender wage gap is significant in Uganda and Ethiopia and relatively smaller in Cote d'Ivoire.[8]

Using Neumark's (1988) analysis on the type of discrimination, Appleton et al (1999) divide the unexplained component in the wage differential to determine the portion related to nepotism towards men and discrimination against women, referring to both parts as male and female deviations, respectively. In Uganda, the unexplained component is broken into the return to male and female deviations which account for 23% and 56%, respectively. This suggests that discrimination against women is larger in the unexplained component. Ethiopia's unexplained component of the wage differential is 110% equally divided for male and female returns. Contrary to Uganda and Ethiopia, wage differential for Cote d'Ivoire is entirely attributed to observable characteristics. Looking at the wage decomposition by sector, Ethiopia has a larger wage gap in the private sector than in the public sector. Conversely, Cote d'Ivoire has a larger wage in the public sector than in the

[8] In Ethiopia, they use the Survey of Adolescent Fertility, Reproductive Behaviour and Employment Status of the Youth Population in Urban Ethiopia for the year 1990. In Uganda, the Integrated Survey of Uganda for 1992 is used. For Cote d'Ivoire, the Livings Standards Survey for 1985, 1986, and 1987.

private sector. In addition, Uganda has almost the same wage gap in the public and private sector.

Glick and Sahn (1997) also use the Neumark decomposition technique and examine gender earnings differentials in Guinea within three sectors: public sector, private sector, and self-employment. They find that in 1990, 55% is due to discrimination for self-employment. The authors argue that in self-employment, discrimination does not reflect employers' discriminatory behaviour but women have less experience in entrepreneurship or less access to funding or credit to establish a small business. Therefore, the discrimination component is relatively small in gender wage differential with respect to self-employment. In the public sector, the discrimination component accounts for 75% and observable characteristics did not account much in the wage differential since women are slightly more educated than men in this sector. The private sector shows a similar outcome of that in the public sector with a negative difference due to productivity characteristics.

The remaining African studies use the standard Oaxaca (1973) decomposition using the male wage structure as the reference group. Knight and Sabot (1982) look into labour market discrimination across gender and race in a poor urban country in Africa, using the 1971 Tanzanian manufacturing firm's survey, and a simple Oaxaca decomposition technique; they find that 17% of the wage differential is a product of gender discrimination. The authors claim that the Tanzanian government oppose gender and race discrimination which may be the reason behind the small discrimination component in the overall wage differential. The findings could also be a result of poor measurement or inaccuracy of data, and not related to policy implementation.

There are a rising number of studies investigating gender wage differentials in Latin America. Psacharopoulos and Tzannatos (1992) review 21 studies for 15 Latin American countries and East Asia, and find that on average, the unexplained component account for 88%. A more recent study in Brazil that focuses on measuring discrimination after economic reform and stabilization find interesting results. Arabsheibani et al (2003) investigate the gender wage differential, using Pesquisa Nacional por Amostra de Domicilios (PNAD), over the period 1988 to 1998. First, they look at whether women improved their presence in the labour market during trade and financial liberalization between 1988 and 1992. Subsequently, they measure the effect of price stabilization in the four years following the 'Real Plan', introduced in 1994 by the Brazilian Finance Minister, Fernando Henrique Cardoso.

Using the Juhn, Murphy and Pierce decomposition extending Blinder-Oaxaca decomposition to take account of distributional changes over time, they conclude that women's status in the workforce benefited due to the improvement of their productivity characteristics as well as the reduction of the discrimination component of the wage differential over that period. In 1988, men were getting paid more than women by 300%, decreasing to 162% in 1992, and after the implementation of the 'Real Plan', the gap reduced to 33% in 1998. Arabsheibani et al (2003) suggest that the reduction in the gender wage gap could be the result of a change in the relative distribution of wages, improvement of women's endowments raising their wages thus shrinking the wage gap, and successful implementation of anti-discrimination policy brought by the reform. In addition, they find the total effect of women's endowments as well as observed and unobserved prices narrowed the gender gap by 33.1%.

Another Brazilian study took a slightly different approach during the same period of economic reform measuring race and gender discrimination for the rural and urban regions. Loureiro et al (2004) use the Oaxaca-Blinder (1973) with the Heckman procedure (1974, 1979, 1980) to correct for selectivity bias. Using PNAD data for the urban labour market, the discrimination component is around 52% in 1992 increasing to 59% in 1998. There was an increase in the male-female wage differential during this period in the urban and rural labour markets and the authors concur that the increase is due to a more competitive economic environment. During a turbulent period following the economic reform in 1994, the increase in unemployment combined with worker turnover may have resulted in more competition for jobs and since males occupy more managerial jobs and can influence earnings, this drives up discrimination in the labour market.

Thus, the results on the gender wage differentials worldwide differ due to the variation of methodology, choice of datasets, human capital and demographic variables, as well as the selection of a base group. However, the component in the wage differential that cannot be explained, which may be partially an attribute of discrimination is consistent among the literature ranging from 50% to 90%.

3. Data and Sample

This paper uses a cross-sectional data collected by the Brazilian Institute of Geography and Statistics (IBGE) which is a public foundation of the Brazilian federal administration. Inaugurated in 1934, the digital repository for statistical, economic and geographic data contains databases and surveys on labour, including the Monthly Employment Survey (EMS). Established in 1980 and recognized by the International Labour Organization, the EMS has comprehensive data on the labour force and accompanying information on employment conditions, family related characteristics, income, demographics, and classes of workers. The advantage of this survey is that it is conducted on a monthly basis only covering metropolitan regions in Brazil so that to achieve homogenous characteristics across the sample. The population of interest are the civilian, non-institutionalized population 10 years of age or older.[9] Surveying 45,267 households as collection units, the EMS includes all employed individuals during the reference week (interview date), the unemployed, and workers that are working but not receiving income. Total observations in my sample survey for the year 2011 in the month of December are 98,783.[10]

I restrict the sample to individuals between the ages of 15 to 75 who are paid employees and worked 15 hours or more during the reference week in their main job only.[11] I also drop individual observations with missing information from the selected variables, and exclude the self-employed and employers. First, the age restriction stems from the fact that most individuals who are older than 75 are retired and therefore earn some of their income through retirement plans. This reasoning also applies to the exclusion of the self-employed or employers. They are not part of the study mainly because their earnings are not from a wage system established by public and private institutions.. After the restrictions are in place, the total number of observations in this study reduces to 24,877 employed individuals – 14,958 male and 9,919 female observations.

[9] Individuals living on military reserves, correctional institutions, hospital patients, school boarders, and military employees, residents in nursing homes, orphanages, members of religious orders in convents, monasteries, etc, are excluded from the survey.
[10] More than one individual are surveyed in each household.
[11] The main job is defined as the individual's job holding the greatest number of actual hours worked.

Table 1 provides the descriptive statistics by gender of observable characteristics used in this study. In my sample, on a monthly average, males earn 433.5 Brazilian Real (BRL) more than females, consistent with previous similar studies. In Canadian dollars, this means that males earn $235.45 more than females a month. The female average earnings in this study is 1305.86 BRL which gives an average earnings gap of 0.85. Similarly, the United Nations find that in 2010, the male-female earnings ratio is 0.81 in Brazil.

In the male sample, 53.4% who claim they are the main providers in the household, while this is the case for only 33.4% of females. Furthermore, females outperform males in the average years of education with 9.10 years on average against 8.46 years for males. By level of education, males and females are generally balances, except that 65.5% of females have, on average, eleven or more years of schooling followed by males with 59.2%. Since income has been positively correlated with educational attainment in previous human capital empirical studies, females outdoing males in education will impact how much it will be accounted for in the gender earning gap.

The survey does not include information on the actual level of experience for individuals, and therefore, I use a conventional proxy referred to potential work experience by researchers. The proxy can be calculated by using age minus schooling minus 6.[12] This measurement has been criticized by some researchers because it overestimates actual experience, but however, due to data limitations, many researchers resort to developing proxies for age or experience. Drolet (2001) discusses the disadvantages of the proxy for experience particularly the fact that it does not take into account work interruptions of

[12] In this study, the potential experience variable is demonstrated as a categorical variable to measure the impact of experience on earnings at different levels of experience. This approach is consistent with other economics of discrimination literature that use a proxy for experience (Baker et al., 1995; Gunderson, 1998; Loureiro et al., 2004)

females when they take on family related responsibilities including children bearing, thus overestimating actual experience. In the sample, females have an average of 18.37 years of experience followed by males with an average of 18.07 years.

Moreover, other researchers in the field find that using the variable experience in the earnings regressions may lead to imprecise results. Corcoran and Duncan (1979) argue that males acquire continuous work experience building job-related skills, tenure and on the job training. Females face interrupted periods of work due to child bearing and household responsibilities, which in turn affects females' earnings in many ways. First, they accumulate less work experience, tenure, and seniority. Second, human capital skills may depreciate during time out of the labour force which will impact earnings. Third, females who leave the labour force temporarily will delay training which will also affect their earning.

Other work related observable characteristics used in this study that influence earning are tenure, employment status, type of contract, occupation, firm size and social security. Large number of empirical literature shows that real income of an individual has a positive relationship with tenure and firm size (Pavan, 2011 & Morissette, 1991). The average number of working hours for males exceeds that of females by 2.40 hours which is not very large. In the sample, individuals who are working more than 44 hours in the reference week are considered working overtime. The proportion of females and males working more than 44 hours in this study are 81.5% and 91%, respectively. These large proportions of individuals working overtime are common in the workforce within the Brazilian economy. Telles (2003) measure the labour market segmentation using 1980 Brazilian census data. He reports the proportions of females and males working between

40 to 48 hours are 71% and 66%, respectively. In addition, 29% females and 15% males work more than 48 hours. Telles (2003) describes 40 or more hours worked as full time work in an urban economy.

The gender composition of occupation varies with more males working in occupations such as mining, electricity & gas with 21.9% compared to females with 13.2%. Not surprisingly, females are over-represented in occupations in the field of public administration, education, health, and social services with 31.8% compared to 13.5% males. In addition, females slightly outnumber males in professional occupations including financial intermediation and real estate with 18.7% compared to 18.0% males. Drolet (2001) finds the same pattern in gender composition when it comes to major field of study, having more females in health and education relative to males and more males in applied science technologies, and trade fields relative to females. Drolet (2001) mentions that the choice of each major field of study gap and in this study, occupation, will account for in the explained portion of the earning differential thus influence the gender earning gap.

4. Mean Gender Earnings Differentials

Table 2 to 4 provide the mean gender earnings differentials corresponding with their demographic, human capital, and work related characteristics. The findings in this study, particularly the male-female earnings ratios, are consistent with the Brazilian study conducted by Arabsheibani et al (2003), indicating that the gender wage or earnings gap ratios has increased after the economic reform in 1994. In this study, earnings ratio tends to be higher when the pay gap is small, and conversely a large pay gap describes a smaller earnings ratio between both genders.

As highlighted in Table 2, the earnings ratio is the highest for having 11 years of schooling or more as well as graduates and above showing 0.73 and 0.76 ratios, respectively. On the contrary, individuals having only four to seven years of schooling or with adult literacy have the smallest ratio of 0.67 and 0.63, respectively. Graduates or individuals with high number of years of schooling have similar skill sets, ability, and labour force attachment, which may explain the smaller pay gap ratio between genders.

Table 2 also shows the mean earnings differential between males and females by experience. The mean earnings ratio for individuals with one to ten years of experience is the largest in the range showing 0.81. The average earning gap increases monotonically with the increase in the number of years in experience, showing an average earnings ratio of 0.69 for over 20 years of experience. Females and males with low levels of experience as well as other similar characteristics enter the labour force with almost the same earning. Individuals with a high number of years in experience tend to have diverse characteristics such as work interruptions for females leading to skill depreciation. In other words, other factors that are attributed to females increase the average earning gap in this category.

In Table 3, the average earnings ratio ranges from 0.80 for the oldest age group, 55 to 75, to 0.89 for a younger age group, 15 to 34. The younger age group have similar skills, education, and level of experience for females and males and they both simultaneously enter the labour force after some time spent in school. Women in the older age group may not have the same educational attainment than the younger women which drives their wage lower relative to their male counterpart.

The gender earnings gap is the highest among the main providers in the household (0.66) and the lowest amongst the son, daughter, and relative (0.95). This may be related to the fact that the latter group are single individuals who were never married and are from a younger age group. Consequently, they have low household responsibilities and are more attached to the labour force. Females who are the main individuals in the households, especially those who have dependents, juggle between their careers and family responsibilities. Polachek (1979) argues that females under those circumstances choose to work part-time or accept lower wages in exchange for shorter or flexible hours, or jobs that are closer to home.

In the sample, the proportion of white males and females' individuals outnumber having a mixed race, blacks, and Asians that follow in ordinal order. Although Asians are the highest paid individuals, the male-female earnings ratio is the lowest showing a large gender earnings gap of 0.70 relative to their counterparts. In contrast, blacks are the lowest paid individuals with the smallest gender earnings gap showing 0.78. Cotton (1988) reports that white males on average have more years of schooling and experience, as well as the return to education and experience is significantly high relative to black males. Females and males with low levels of experience and education have homogeneous characteristics and are often seen in low paying jobs. It was reported by other researchers that females tend to have lower salaries relative to males in executive positions or with supervisory salaries (Baker and Drolet, 2010; Drolet, 2001).

Table 4 provides the male-female ratios of work related characteristics that have several implications on the gender earning gap. For tenure of current or past employer, the gender earning gap is 0.94 for up to thirty days of tenure and decreases as the number of years of tenure increases (0.75 for more than two years). Empirical literature that examines the relationship between tenure and wage shows that as tenure increases so will wage. One would assume that an increase in the female tenure will give the employer an impression that she has a high commitment to her job, thus will drive her wage higher relative to males. However, Drolet (2001) mentions that the gender earning gap does not necessarily shrink as tenure increases.

In my sample, females earn about 0.74 of the average earning of males in the public sector, and 0.81 in the private sector. As seen previously, the gender earnings gap is larger among older cohorts, and those who have more than two years of tenure which may be clustered in the public sector thus leading to such result. Foguel et al (2000) examines the wage differential between the public and private sector in Brazil and finds a significant wage gap between the two sectors with the public sector paying more to their employees than the private sector. They also find that workers in the public sector are more educated, older and have longer tenure with their present employer. Another empirical study by Saldanha et al (1988) indicates that there is a high degree of wage heterogeneity among the public sector segments in Brazil which may also impact earning differences among genders as well.

Moreover, occupations also carry interesting implications. Females earn about 0.75 of the average earning of males in mining/electricity, gas and water, 0.67 in public administration/education and health, and 0.69 in services. The earning gap is smaller in trade/automotive and oil & gas showing 0.80 earnings ratio and working in a company office, females on average are paid higher than males do showing 1480.30 and 1305.53, respectively. In heavy duty related jobs, women probably perform administrative tasks and therefore receive a relatively large wage relative to male workers in these sectors. Similarly, Knight and Sabot (1982) report that females receive higher returns of 69% working clerical jobs compared to males who only receive 48% in the manufacturing sector.

5. Econometric Model

This paper uses an econometric model to examine the effect of a change in endowments on monthly earnings for males and females to decompose the wage differential into explained and unexplained parts. The general linear model that estimates this effect for both females and males is the following:

$$lnw^m_i = X^m_i\beta^m + \varepsilon^m_i \quad i=1,..,n \qquad (1)$$

$$lnw^f_i = X^f_i\beta^f + \varepsilon^f_i \quad i=1,..,n \qquad (2)$$

Where lnw_i represents the logarithm of effective monthly earning for individual i, and X_i represents a vector of observable characteristics for each individual. The subscripts in the equations, namely f and m, represent females and males, respectively. The error terms are represented by ε_i. Assuming that the error terms are not correlated with the explanatory variables, $(\varepsilon_i \mid X_i) = 0$. The equations for males and females become the following:

$$E(lnw^m_i \mid \text{female}=0, X^m) = E(X^m_i)\,\beta^m \quad (3)$$

$$E(lnw^f_i \mid \text{female}=1, X^f) = E(X^f i)\,\beta^f \qquad (4)$$

I use the method of ordinary least squares (OLS) for estimating the coefficients in the linear regression model. By the properties of OLS, I can also write (3) and (4) as:

$$ln\overline{w}^f = \overline{X}^f \hat{\beta}^f \qquad (6)$$

$$ln\overline{w}^m = \overline{X}^m \hat{\beta}^m \qquad (5)$$

Where $\overline{w}^m$ and $\overline{w}^f$ are the average monthly earnings for males and females, respectively. $\hat{\beta}^m$ and $\hat{\beta}^f$ are the estimated coefficients and the vector of mean values of the observable characteristics are represented by $\overline{X}^m$ and $\overline{X}^f$ for males and females, respectively. I further analyze the log of earnings differential by using the standard method of decomposition (Oaxaca, 1973). This method will identify impact of differences in the coefficient of endowments to the overall result of earnings differential. The Oaxaca decomposition technique takes equations (5) and (6) and presents the wage differential of males and females in the following manner:

$$ln\overline{w}^m - ln\overline{w}^f = (\overline{X}^m - \overline{X}^f)\hat{\beta}^m + (\hat{\beta}^m - \hat{\beta}^f)\,\overline{X}^f \qquad (7)$$

The term on the left-hand side represents the average earnings differential between males and females. The right-hand side divides the earning differential into two parts. The first term is the earnings difference in the group average productivity characteristics (endowment effect) between males and females which is typically referred to as the explained part. The second term measures the contribution of the difference in the returns to those characteristics (coefficients) which can be referred to the unexplained part of the average earnings differential.

Equation (7) calculates the decomposition of earning differential assuming that males have the non-discriminatory wage structure. In other words, the first term of equation (7) describes how much females would earn if they face a male wage structure. Alternatively, the above representation can be expressed with having a female's non-discriminatory wage structure:

$$ln\overline{w}^m - ln\overline{w}^f = (\overline{X}^m - \overline{X}^f)\hat{\beta}^f + (\hat{\beta}^m - \hat{\beta}^f)\overline{X}^m \qquad (8)$$

The first term of equation (8) can be interpreted as how much males would earn if they face a female wage structure. Both expressions yield different results reflecting the index number problem, however, since average female wages are lower than male wages across the world, the common perception and empirical findings indicate that females are discriminated against in the labour force. Hence, I will only show the wage differential decomposition using the male wage structure as non-discriminatory, and show the decomposition when I correct for selectivity biases using the Heckman (1979) procedure.

In addition, I also use the Neumark decomposition:

$$ln\overline{w}^m - ln\overline{w}^f = (\overline{X}^m - \overline{X}^f))\hat{\beta}^* + (\hat{\beta}^m - \hat{\beta}^*)\overline{X}^m \qquad (9)$$

Where the first term in the expression is the 'explained' component of the earning differential, and the latter terms are the 'unexplained' component capturing discrimination and the unobservable characteristics. Equation (9) encompasses the 'pooled' wage structure, $\hat{\beta}^*$, which is also referred to the no-discrimination coefficient vector. As highlighted by Oaxaca and Ransom (1994), the pooled wage structure is the weighted average of the male and female wage structures:

$$\hat{\beta}^{*} = \Pi\hat{\beta}^{m} + (1 - \Pi)\hat{\beta}^{f} \qquad (10)$$

With the weighting matrix comprising of:

$$\Pi = (X'_m X_m + X'_f X_f)^{-1}(X'_m X_m) \qquad (11)$$

In equation (11), the first term represents the observable characteristics' matrix for the pooled sample whereas the second term is the corresponding matrix for males.

6. Returns to the Earning Functions

Table 5 presents OLS regressions examining the effect of a change in a number of observable characteristics on earnings. In all specifications, the dependent variable is the logarithm of monthly earnings. Each coefficient is the percentage change in monthly earnings concurring with a change in observable characteristics.

The first specification includes the standard human capital and social variables such as education and potential experience level, age, a quadratic in age and potential experience, family status determining the income flow in the household, race, and whether the individual is attending school or not. The second specification adds control for region, firm size, tenure, employment status (whether worker is in fulltime, part time, or overtime), type of employment contract, and contribution to social security. The third specification adds occupation as control.

Researchers in the field of economics of discrimination all agree that including occupation or industry in the regression could lead to underestimating the discrimination component of the decomposition. This occurs mainly due to the fact that discrimination can occur through occupation and industry and using the two controls could take away some of the unexplained (which is considered discrimination) and transfer it to the explained as productivity related characteristics causing some of the wage gap. The discrimination can occur because females must accept to work in certain occupations because they have no access to others, thus, they are forced in certain occupations paying lower wages. In addition, females also face the difficulty of getting promoted to high executive levels in their occupations (Baker and Drolet, 2010). Therefore occupation controls are included in the third specification only for comparison purposes.

Like other empirical studies, education is positively correlated with log of monthly earning in all specifications.[13] However, the magnitudes of the coefficients corresponding with females and males differ, showing that the education return for females is larger than for males. Compared against no education to 3 years of education, females' returns of four to seven years of education exceedingly leads males with around 20%% compared to males' returns of around 10% for all specifications.

On the contrary, eight to ten years of education with respect to no education to 3 years of education does not provide results consistent with empirical studies examining the relationship of earnings and education. For females (males), results show an increase of 3.0% (5.3%) in the first specification, a decrease in returns of 0.9% (1.2%) in the second specification an increase of 4.0% for females whereas a reduction of 0.4% for males.. The findings of the returns of eight to ten years of education are relatively smaller compared to four to seven years of education. A possibility of this result can be that individuals who have lower levels of education attended schools that teach practical skills

[13] The only exception is at the level of eight to ten years of education showing a negative correlation in the second specification and the third specification for males only and they are not significant.

and perform certain hands on jobs that generate higher returns compared with eight to ten years of education.

Furthermore, eleven years or more of education for females increases monthly earning by 35.8%, 24.1%, and 25.1% in the first, second, and third specifications respectively with respect to no education to 3 years of education. Their male counterparts follow with 32.8%, 20.9%, and 21.6%.[14]

Looking at gender differences in the returns of education, these results are consistent with Canadian data with the exception of the eight ten years of education category. In a large number of Canadian empirical studies, females receive more returns in majority of categories in higher education relative to males, while they are worse off with less than high school education (Baker et al, 1995; Drolet, 2001; Baker and Drolet, 2010).

The coefficients for potential experience for both females and males have a positive sign and are statistically significant at the 1% level. In comparison with the base group of twenty one years and over of experience, one to ten years of experience, one additional year of potential experience increases the monthly earnings of females (males) by 9.4 % (16.3%), 4.8% (12.7%), and 3.5% (11.9%) in all three specifications. Although females on average have more years of experience, there is an economically significant gender difference in the returns to experience benefiting males. The returns to potential experience for both males and females converge at a decreasing rate as the number of years of experience increase.[15] The quadratic of potential experience is negatively correlated with earnings for both genders in all specifications at the 1% significance level.

---

[14] The positive correlations are all statistically significant at the 1% level.

[15] One additional year of eleven to twenty years of potential experience increases the monthly earnings of females (males) by 4.8 % (7.8%), 2.6% (6.5%), and 2.0% (6.0%)

Earnings levels are higher across genders increasing monotonically with eleven or more employees in firm size, more than two years of tenure, and employees working overtime. All are positively correlated at the 1% significance level but differ for both genders having males earn higher returns than females for the abovementioned coefficients. With respect to a firm size of six to ten employees, females (males) who work in a firm that has eleven or more employees earn 10.6% (13.6%) and 7.6% (11.2%) in the second and third specification, respectively. The tenure coefficient deduces consistent results having males earn higher returns than females for each level of tenure. For more than two years of tenure, earning increases by 18.9% -19.0% (26.7%-27.4%) for females (males) with respect to less than thirty days. Both coefficients are economically and statistically significant at the 1% significance level. The longer the period of tenure with both genders' current employer, the higher wages individuals are offered. This correlation is prominent in the labour economics literature and not surprisingly, females' earnings in the same tenure level falls behind men' earning, which will be reflected in the decomposition.

There is a positive correlation between earnings and an increase in the number of hours worked a week for both genders as one would expect. Comparable with part time status, females (males) having full time status earnings increase by 3.8%-3.6% (4.9%-5.2%) in the second and third specifications. These findings are statistically significant at the 10% level. There is a significant increase of earnings returns for both females and females when working overtime is compared against part time status. The returns to earnings for females (males) increase by 19.3%-22.0% (21.2%-21.8%) for the second and third specification, with 1% significance level. There is considerable amount of individuals working overtime across genders, and this could be linked to the increase in their hourly earnings across genders implying that employers reward individuals who spend approximately more than 8 hours a day in the workplace.[16]

The occupation coefficients show a negative correlation between certain occupations and earnings and the majority of the statistical inferences are significant at the 1% confidence level comparable with respect to working in a company's office base group. The downward relationship can be the result of the comparable base group since it has been seen in Table 4 that females earn more than males working at a company office. Galvez (2006) finds that although females are concentrated in low paying industries, men earn much more than females in industrial jobs, and even in services. Although females are more or less equally present in the labour force participating in such occupations, their returns are significantly lower with respect to an occupation in a company's office. Hence, this is why many empirical researchers in discrimination are reluctant to include occupations as part of the regression.

As we have seen the returns to earnings for females and males differ quite significantly in some areas, and males returns are higher comparable to females. This result will outweigh the unexplained relative to the explained component of the earning differential since I will be using the males wage structure as the reference group.

[16] Holding a full time status is approximately one fifth of the increase in earnings when working overtime.

7. Log Earnings Decompositions

We have seen that females and males have different returns to observable characteristics as well as mean proportions. These findings will be reflected when using the decomposition technique. However, those differences may not be able to explain the entire wage differential between both genders. Explaining the wage gap can be divided into two parts using the Oaxaca decomposition (1973) approach using a male wage structure as non-discriminatory. The Oaxaca decomposition will also be represented using the Heckman procedure (1979) to correct for self-selection. Furthermore, and a Neumark decomposition (1988) is also used with a pooled wage structure as non-discriminatory.

The mean earnings differential between males and females, in logarithm terms, is 0.179. In Brazil, Arabsheibani et al (2003) show the decline in the log of mean earnings differential from 0.335 in 1988 to 0.147 in 1998. In addition, another study in Brazil finds the log of mean wage differential between both genders is 0.177 in 1998 (Loureiro et al, 2004). Although, there are variations in the earnings differential in the Brazilian literature, there is a substantial difference in the magnitude of the wage differential between past decades and recent times showing improvement in the earnings of females relative to males. Brazilian literature explores whether the economic reform has contributed to the decline in the wage or earnings differential, but looking at other studies from an international level, other countries are also facing significant reduction in the earnings differential during the same periods.[17]

[17] Baker and Drolet (2010) find that the log wage gap reduces from 0.268 in 1981 to 0.166 in 2008 in Canada. Baker et al (1995) measure gender discrimination in the United States showing that the log wage gap reduces from 0.508 in 1970 to 0.402 in 1990.

Table 6 presents the log earnings decompositions using the Oaxaca (1973) method for all the specifications used in this study. The unexplained component is as high as 133% in the first specification followed by 113% in the second specification. In other words, the entire earnings functions of observable characteristics do not explain anything of the wage differential of 17.9%.[18] Adding occupation as a control in the third specification reduces the unexplained component to 110%. It is quite ambiguous whether the occupation control takes away some of what is perceived as discrimination occurring through the type of job performed thus reducing the unexplained component, or it may really explain a portion of the earnings differential. Observable characteristics including age, firm size, tenure, employment status, and occupation can explain a very small portion of the earnings differential.

Moreover, correcting for self-selection bias using the Heckman (1979) procedure, the log earnings differential slightly reduces from 0.179 to 0.175. This results from the increase of the logarithm of monthly earnings for females indicating that there was a downward bias in the original earnings differential. As a result, the log earnings decomposition reduces the unexplained component to 123%, 108% and 104% in the first, second, and third specification respectively. The correction for self-selection did not reduce the earnings differential or the unexplained component significantly, thus remaining above 100%.

In addition, Table 7 presents the log earnings decomposition using the Neumark (1988) method and the results are consistent with the Oaxaca (1973) decomposition. However, the unexplained component is substantiated in all specification showing 145.3% in the

[18] The explained components show -33%, -13.4%, and -10.5% in the first, second, and third specifications.

first specification. Adding more controls results in 121.2% of the unexplained component in the second specifications and reducing to 117.3% in the third specification when adding occupation. Looking at the breakdown of the explained component across specifications, it appears that the gender earnings differential is partially due to age, work related characteristics including firm size, all levels of tenure, employment status, type of contract across all specifications, and occupation in the third specification. The only two occupations that explain a small portion of the gender earnings gap are in public administration, health services, and administration as well as general services. All the other industrial occupations added to the unexplained component. The jobs that have a role in explaining some of the gender earnings gap are where females are mostly concentrated.

In Table 7, the Neumark decomposition can be divided to determine whether the differences in coefficients accounts for higher returns for males (nepotism) or lower returns to females (discrimination against females) compared to a pooled wage structure. It turns out that in all specifications, there are lower returns to females and can be interpreted as discrimination against females accounting for the entire earnings differential.[19]

These results also have other implications that need to be considered. The decomposition shows negative explained components across all specifications in both methods. The mean proportions for females exceed male proportions in sizable amounts in characteristics including age, number of average years of schooling, potential experience, full time status, having temporary contracts and some levels in tenure. Females are

[19] In the unexplained component, the female returns compared to the pooled wage structure accounts for earnings differential across all specifications, which implies as interpreted by Neumark's study as discrimination against females.

succeeding in catching up, if not having an edge on males, in characteristics influencing returns to earnings, thus shrinking the gender earnings gap related to these characteristics and overstating the discrimination component. This in turn drives to the result of a negative explained component when using the standard discrimination decompositions such as Oaxaca (1973) and Neumark (1988).

The choice of observable characteristics can influence how much of the overall earnings differential can be explained and cannot be unexplained. There is no universal approach in choosing observable characteristics to be part of the earnings functions; however, most researchers seem to prefer the conventional human capital controls as well as other standard characteristics that they consider to influence earnings. In addition to the above findings with respect to the large unexplained components, the literature that deals with time series data and controls for the same observable characteristics over the period in question, the unexplained component steadily grows in the overall earnings differential with a simultaneous decline in the unexplained component (Baker and Drolet, 2010). Some researchers have alternate explanations to the growing unexplained component. (Loureiro et al, 2004; Frijters, 1998)

Frijters (1998) argues that the uncertainties that are within the labour force generate high employment expectations and discrimination. In other words, discrimination can occur if there is uncertainty for employers and employees on whether they will retain their current jobs in the future. This can happen due to the prevalence of limited jobs available. The proportion of females in occupations with authority and decision making is very small and females are often seen in the lower ends of the hierarchy.[20] Males occupy managerial, legislation and senior official positions and they can influence earnings of females or enforce the 'glass ceiling' in order to gain job security. Laureiro et al (2004) adds that an increase in the discrimination component in the Brazilian economy could be stemming from a more competitive economic environment especially that was the result of the economic reform that was implemented in 1994 for the same abovementioned reasons.

8. Conclusion

In this paper, I investigate the gender earnings gap using the Employment Monthly Survey (LFS) for Brazil in the month of December 2011. On average, I find that female earnings are 85% of the males' average. In all three specifications I use in this study, the earnings determining factors including education, age, potential experience, employment status, tenure, are generally positively correlated with statistical and economic significance. The returns to earnings coefficient for males in most observable characteristics exceed female coefficients, explaining some of the gender earnings gap. However, these observable characteristics did not contribute substantially in the gender earnings gap. In the log earnings decompositions using the Oaxaca method and correcting for self-selection as well as the Neumark method, I find that the entire earnings differential cannot be explained by the observable characteristics ranging from 104% to 145% across specifications. Using the different methods to correct for potential biases such as self-selection and the index number problem does not reduce the unexplained

[20] United Nations (2010) report the proportion of females in occupations with authority in overall total of females employed over the period of 2004 to 2008 are 10% in Northern Africa, 30% in Latin America and the Carribean, 30% in Asia and 40% in more developed countries.

component of the overall earnings differential. Hence, the entire earnings differential cannot be explained by observable characteristics and may be partially due to discrimination.

This finding and others in past literature vary depending on the dataset, choice of controls and base groups as well as using a proxy for actual experience. During the last two decades, most American and Canadian literature report 50% to 80% unexplained component which may be an attribute of discrimination.[21] For the same period, in Latin America, it has been reported that the unexplained component is on average 88% and in some African countries as we have seen previously, the unexplained component accounts for the entire wage differential. This may be due to the fact that the developing region is yet behind in implementing legislations that fights against discrimination. In Brazil, Arabsheibani et al (2003) reports that the gender wage gap reduced substantially after the implementation of the 'Real Plan' in 1994. Laureiro et al (2004) also shows that the unexplained component was steadily growing in the overall wage differential between 1992 and 1998, and predicting that this may be as a result of economic competition intensifying discrimination. Future research in Brazil should focus on measuring the direct relationship between economic competition and gender discrimination.

[21] American literature reports an unexplained component in the higher scale of 50% to 90%.

References

[1] ABRAMO, L., AND M. VALENZUELA (2005): "Women's Labour Force Participation Rates in Latin America," *International Labour Review*, 144 (4), 369-400.

[2] APPLETON, S., J. HODDINOTT, AND P. KRISHNAN (1999): "The Gender Gap in Three African Countries," *Economic Development and Cultural Change*, 47(2), 289-312.

[3] ARABSHEIBANI, G., F. CARNEIRO, AND A. HENLEY (2003): "Gender Wage Differentials in Brazil: Trends over a Turbulent Era," *World Bank Policy Research*, Working Paper 3148.

[4] ARROW, K. (1972a): "Models of Job Discrimination," *Racial Discrimination in Economic Life*, Lexington: D.C. Heath, 83-101.

[5] BAKER, M., AND N. FORTIN (2001): "Occupational Gender Composition and Wages in Canada, 1987-1990," *The Canadian Journal of Economics*, 34 (2), 345-376.

[6] BAKER, M., AND M. DROLET (2010): "A New View of the Male/Female Pay Gap," *Canadian Public Policy*, 36(4), 429-464.

[7] BECKER, G. (1957): The Economics of Discrimination, *University of Chicago Press, Chicago, IL.*

[8] BLAU, F., AND L. KAHN (1997): "Swimming Upstream: Trends in the Gender Wage Differential in the 1980s," *Journal of Labor Economics*, 15(1), 1-42.

[9] BLAU, P. (1977): Inequality and Heterogeneity- A Primitive Theory of Social Structure, *United States, Free Press.*

[10] BLINDER, A. (1973): "Wage Discrimination: Reduced Form and Structural Estimates," *The Journal of Human Resources*, 8(4), 436-455.

[11] COTTON, J. (1988): "On the Decomposition of Wage Differentials," *The Review of Economics and Statistics*, 70(2), 236-243.

[12] Department of Economic and Social Affairs (2010): "The World's Women 2010, Trends and Statistics," *United Nations*, ST/ESA/STAT/SER.K/19.

[13] DROLET, M. (2001): "The Persistent Gap: New Evidence on the Canadian Gender Wage Gap," *Analytical Studies Branch - Research Paper Series*, Statistics Canada, Catalogue no. 11F0019MIE2001157.

[14] FAIRLIE, R.W. (1999): "The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment," *Journal of Labor Economics*, 17(1), 80-108.

[15] FAIRLIE, R.W. (2005): "An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models," *Journal of Economic and Social Measurement*, 30(4), 305-316.

[16] FRIJTERS, P. (1998): "Discrimination and Job Uncertainty," *Journal of Economic Behaviour and Organization*, 36(4), 433-446.

[17] GLICK, P. AND D. E. SAHN (1997): "Gender and Education Impacts on Employment and Earnings in West Africa: Evidence from Guinea," *Economic Development & Cultural Change*, 45(4), 793-824.

[18] GUNDERSON, M. (1998): "Females and the Canadian Labour Market: Transition Towards the Future," *Census Monograph Series*, Statistics Canada, Catalogue Number 96-321-MPE No. 2.

[19] HECKMAN, J.J. (1979): "Sample Selection Bias as a Specification," *Econometrica*, 47(1), 153-161.

[20]HOFFMANN, R., AND A.L. KASSOUF (2005): "Deriving Conditional and Unconditional Marginal Effects in Log Earnings Equations Estimated by Heckman's Procedure," *Applied Economics*, 37(11), 1303-1311.

[21] International Labour Office (2008a): "Global Wage Report 2008/2009, Geneva: ILO.

[22] JANN, B. (2008): "The Blinder-Oaxaca Decomposition for Linear Regression Models," The STATA Journal, 8(4), 453-479.

[23] JONES, F. L. (1983): "On Decomposing the Wage Gap: A Critical Comment on Blinder's Methods," *Journal of Human Resources*, 18(1), 126-130.

[24] KNIGHT, J. AND R. SABOT (1982): "Labor Market Discrimination in a Poor Urban Economy," *Journal of Development Studies*, 19(1), 67-87.

[25] LOUREIRO, P., A. SACHSIDA, AND F. CARNEIRO (2004): "Race and Gender Discrimination in the Labor Market: An Urban and Rural Sector Analysis for Brazil," *Journal of Economic Studies*, 31(2), 129-143.

[26] MORISSETTE, R. (1991): "Canadian Jobs and Firm Size: Do Small Firms Pay Less?," *Analytical Studies Branch Research Paper Series*, No. 35, Statistics Canada.

[27] NEUMARK, D. (1988): "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination," *The Journal of Human Resources*, 23(3), 279-295.

[28] NEUMARK, D., TROSKE, K., HELLERSTEIN, J., AND K. BAYARD (1999): "New Evidence on Sex Segregation and Sex Differences in Wages from Matched Employee-Employer Data," NBER Working Paper No. 7003.

[29] OAXACA, L. R., AND M. RANSOM (1994): "On Discrimination and the Decomposition of Wage Differentials," *Journal of Econometrics*, 61(1), 5-21.

[30] OAXACA, L. R. (1973): "Male-Female Wage Differentials in Urban Markets," *International Economic Review*, 14(3), 693-709.

[31] OAXACA, L. R. AND M.R. RANSOM (1999): "Identification in Detailed Wage Decompositions," *Review of Economics and Statistics*, 81(1), 154-157.

[32] PAVAN, R. (2011): "Career Choice and Wage Growth," *Journal of Labor Economics*, 29(3), 549-587.

[33] POLACHEK, S. (1975): "Potential Biases in Measuring Male-Female Discrimination," *Journal of Human Resources*, 10(2), 205-229.

[34] PUHANI, P. (2002): "The Heckman Correction for Sample Selection and Its Critique," *Journal of Economic Surveys*, 14(1), 53-68.

[35] SANBORN, H. (1964): "Pay Differences between Men and Women," *Industrial and Labor Relations Review*, XVII, 534-550.

[36] TELLES, E. (1993): "Urban Labor Market Segmentation and Income in Brazil," *Economic Development and Cultural Change*, 41(2), 231-249.

Table 1
Descriptive Statistics

Variable	Mean		Std. Dev.	
	Male	Female	Male	Female
Wage	1739.36	1305.86	2354.8	1638.94
Age	36.10	36.82	12.35	12.00
Years of education	8.46	9.10	6.94	6.92
Years of education:				
Zero to three years	.042	.040	.199	.655
Four to seven years	.176	.155	.381	.361
Eight to ten years	.190	.150	.392	.357
Eleven or more years	.592	.655	.492	.475
Family status:				
Main provider	.534	.334	.498	.473
Spouse	.119	.350	.324	.477
Son/daughter/relative	.345	.312	.476	.463
Race:				
White	.474	.506	.499	.499
Black	.116	.114	.319	.318
Asian	.005	.005	.067	.068
Mixed	.405	.374	.490	.484
Number of potential years of experience	18.07	18.37	13.44	13.368
Average number of working hours per week	43.26	40.87	7.53	8.19
Employment status:				
Full Time	.034	.064	.224	.325
Part Time	.053	.120	.183	.245
Overtime	.911	.815	.283	.387
Tenure (conditional on woring)				
Up to thirty days	.018	.021	.134	.145
Month to a year	.232	.246	.422	.431
One to two years	.138	.139	.345	.346
More than two years	.612	.593	.487	.491
Firm size:				
Two to five employees	.092	.110	.289	.314
Three to ten	.051	.079	.220	.270
Eleven and up	.856	.809	.350	.392
Type of Contract:				
Temporary	.941	.956	.234	.180
Permanent	.058	.043	.234	.180

Table 1 (Continued) Descriptive Statistics				
Variable	**Mean**		**Std. Dev.**	
	Male	**Female**	**Male**	**Female**
Contribution to social security:				
Contribute to social security	.852	.848	.354	.358
Not contribute to social security	.147	.151	.354	.358
Occupation:				
Mining and distribution. Electricity/gas and water.	.219	.132	.413	.339
Company's office	.110	.021	.310	.120
Trade/automotive/oil & gas	.176	.180	.381	.390
Financial intermediation/real estate/rents	.182	.187	.385	.389
Public administration/defense Education/health/social services	.135	.318	.341	.465
Services	.179	.160	.384	.367
Number of Observations	14958	9,919		

Note: The sample used in this table is restricted to paid workers in their main jobs that were working in the reference week, aged 15 to 75, not self-employed or employers, and with effective monthly earning.

Table 2
Gender differentials in mean monthly earning by education and experience

	Monthly earning		
	Female	Male	Ratio
Years of Study			
Zero to Three years	580.14	840.72	0.69
Four to Seven years	625.42	936.55	0.67
Eight to ten years	685.36	996.62	0.68
Eleven years or more	1517.09	2065.36	0.73
Educational Attainment			
Adult Literacy	713.89	1128.18	0.63
Less than high school	643.74	1092.28	0.59
High school to degree	808.93	1145.10	0.71
Graduate and above	5757.64	7553.95	0.76
School Status			
Attending school	987.78	1246.96	0.79
Not attending school	1243.23	1649.89	0.75
Experience			
One to ten years	776.93	963.80	0.81
Eleven to twenty years	936.39	1232.53	0.76
Twenty one and over	1571.14	2280.13	0.69

Table 3
Gender differentials in mean monthly earning by demographic characteristics

	Monthly earning		
	Female	Male	Ratio
Age			
15-34 years	11.652	13.177	0.89
35 -44 years	21.140	25.260	0.84
45 -54 years	21.431	26.487	0.81
55-75 years	19.828	24.671	0.80
Dependence			
Main provider in the household	1312.86	1974.614	0.66
Spouse	1301.849	1523.465	0.85
Son/daughter/relative	1022.446	1080.319	0.95
Race			
White	1487.50	2015.32	0.74
Black	883.98	1139.90	0.78
Asian	2200.46	3149.70	0.70
Mixed	941.231	1254.58	0.75
Region			
Recife	919.28	1203.81	0.76
Salvador	1145.31	1486.04	0.77
Belo Horizonte	1114.40	1622.03	0.69
Rio de Janeiro	1359.59	1782.65	0.76
Sao Paulo	1389.27	1772.63	0.78
Porto Alegre	1202.31	1535.97	0.78

Table 4
Gender differentials in mean monthly earning by work related characteristics

	Monthly earning		
	Female	Male	Ratio
Job tenure with last/present employer			
Up to thirty days	804.83	859.51	0.94
Month to one year	791.69	1022.02	0.77
One to two years	983.05	1247.39	0.79
More than two years	1463.23	1938.56	0.75
Employment Status			
Full Time	1096.21	1376.24	0.80
Part Time	1007.45	1328.41	0.76
Overtime	1276.26	1642.05	0.78
Firm Size			
One to five employees	775.64	853.07	0.90
Six to ten employees	895.94	1012.35	0.89
Eleven or more employees	1228.43	1518.89	0.81
Type of Contract			
Permanent	815.99	903.56	0.90
Temporary	1201.49	1479.42	0.81
Social Security			
Contribute to social security	1371.24	1726.51	0.79
Not contribute to social security	627.73	890.81	0.70
Industry			
Private Sector	1152.84	1428.9	0.81
Public Sector	2169.64	2948.35	0.74
Occupation			
Mining and distribution. Electricity/gas and water.	1288.69	1711.90	0.75
Company's office	1480.30	1305.53	1.13
Trade/automotive/oil & gas	896.96	1111.68	0.80
Financial intermediation/real estate/rents	1422.74	1628.74	0.87
Public administration/defense Education/health/social services	1785.62	2638.92	0.67
Services	979.39	1406.99	0.69

Table 5
OLS Regression Results

Dependent Variable: Log monthly earning						
Variable	(1)		(2)		(3)	
	Female	Male	Female	Male	Female	Male
Age						
Age	.027 (.005)***	.034 (.003)***	-.001 (.005)	.008 (.003)**	-.004 (.005)	.007 (.003)*
Age squared	.0001 (.0001)	.0001 (.0001)	.0005 (.0001)***	.0005 (.00006)***	.0005 (.0001)***	.0005 (.0001)***
Potential Experience						
One to ten years	.094 (.043)**	.163 (.035)***	.048 (.039)	.127 (.033)***	.035 (.038)	.119 (.032)***
Eleven to twenty years	.048 (.025)**	.078 (.020)***	.026 (.025)	.065 (.019)***	.020 (.023)	.060 (.019)***
Experience squared	-.001 .0001***	-.0005 (.00005)***	-.0007 (.00008)***	-.0007 (.00005)***	-.0007 (.0001)*	-.0007 (.0001)***
Dependence						
Main provider	-.0231 (.015)	.140 (.015)***	.002 (.014)	.097 (.013)***	.002 (.014)	.094 (.014)***
Relative	-.066 (.016)***	-.071 (.017)***	-.058 (.015)***	-.083 (.016)***	-.057 (.014)***	-.084 (.016)***
Race						
Black	-.327 (.021)***	-.338 (.015)***	-.261 (.019)***	-.269 (.015)***	-.258 (.019)***	-.272 (.015)***
Asian	.496 .083***	.248 (.071)***	.405 (.077)***	.229 (.066)***	.417 (.076)***	.228 (.066)***
Mixed	-.302 (.013)***	-.273 (.010)***	-.230 (.013)***	-.213 (.011)***	-.223 (.013)***	-.214 (.011)***
Years of education						
Four to seven years	.196 (.056)***	.076 (.033)**	.214 (.052)***	.134 (.031)***	.203 (.051)***	.122 (.032)***
Eight to ten years	.030 (.030)**	.053 (.019)**	-.009 (.027)	-.012 (.017)	.004 (.027)	-.004 (.017)
Eleven or more years	.358 (.042)***	.328 (.028)***	.241 (.039)***	.209 (.026)***	.251 (.038)***	.216 (.026)***
School Attendance						
Not attending school	-.009 (.022)	-.010 (.019)	-.064 (.020)***	-.066 (.018)***	-.049 (.020)**	-.063 (.018)***
Region						
Recife	-	-	-.308 (.022)***	-.270 (.017)***	-.310 (.022)***	-.274 (.017)***
Salvador	-	-	-.155 (.022)***	-.155 (.018)***	-.155 (.022)***	-.161 (.017)***
Belo Horizonte	-	-	-.099 (.018)***	.049 (.015)***	-.105 (.019)***	.042 (.014)**
Sao Paulo	-	-	.168 (.017)***	.193 (.014)***	.164 (.018)***	.186 (.014)***
Porto Alegre	-	-	-.102 (.019)***	-.020 (.016)	-.099 (.020)***	-.026 (.016)

Table 5 (Continued)
OLS Regression Results

Dependent Variable: Log monthly earning						
Specification	(1)		(2)		(3)	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Firm size						
One to five employees	-	-	-.060 (.026)**	-.073 (.025)**	-.058 (.026)**	-.088 (.025)***
Eleven or more employees	-	-	.106 (.021)***	.136 (.020)***	.076 (.021)***	.112 (.021)***
Tenure						
Month to year	-	-	-.059 (.060)	.077 (.055)	-.054 (.060)	.083 (.055)
One to two years	-	-	.039 (.062)	.131 (.056)**	.038 (.061)	.135 (.055)**
More than two years	-	-	.189 (.061)***	.267 (.055)***	.190 (.060)***	.274 (.055)***
Employment status						
Fulltime	-	-	.038 (.028)*	.049 (.033)*	.036 (.028)*	.052 (.033)*
Overtime	-	-	.193 (.018)***	.212 (.023)***	.220 (.019)***	.218 (.024)***
Type of Contract						
Temporary contract	-	-	.044 (.020)**	.113 (.034)***	.048 (.035)	.131 (.034)***
Social Security						
Contribute to social security	-	-	.257 (.018)***	.167 (.015)***	.253 (.019)***	.165 (.015)***
Occupation						
Mining and distribution. Electricity/gas and water.	-	-	-	-	-.214 (.043)***	-.035 (.016)**
Trade/automotive/oil & gas	-	-	-	-	-.354 (.042)***	-.146 (.016)***
Financial /real estate/rents	-	-	-	-	-.157 (.042)***	-.059 (.017)***
Public admin/defence Education/health/social services	-	-	-	-	-.180 (.042)***	-.001 (.027)
Services	-	-	-	-	-.247 (.043)***	-.084 (.016)***
R-squared	0.22	0.29	0.35	0.38	0.36	0.38
F-Test	203.96	437.15	189.54	326.56	169.59	282.33
Number of observations	9,919	14958	9,919	14958	9,919	14958

Robust standard errors are in parentheses.
* t-test undertaken at the 10% significance level, ** 5% significance level, and *** 1% significance level.

Table 6 Decomposition of log monthly earnings using the Oaxaca (1973) Method and correction for self-selection			
Log of monthly earning for males	**7.07**		
Log of monthly earning for females	**6.89**		
Difference	**.179**		
		Specification	
	(1)	**(2)**	**(3)**
Explained			
Age	.037	.009	.004
Age squared	.009	.048	.003
Dependence	.034	.020	.012
Race	-.020	-.016	-.014
Years of Education	-0.05	-.032	-.045
Not attending school	-.0002	-.001	-.009
Years of potential experience	-.01	-.008	-.004
Experience Squared	-.060	-.081	-.094
Region	-	-.001	-.004
Firm Size	-	.006	.007
Tenure	-	.007	.005
Employment Status	-	.018	.012
Temporary contract	-	.001	-.002
Social Security contribution	-	~	.001
Occupation	-	-	.110
Total difference in average observable characteristics (explained)	**-.059 (-33)**	**-.024 (-13.4)**	**-.018 (-10.05)**
Total difference in the returns to observable characteristics (unexplained)	**.238 (133)**	**.203 (113.4)**	**.197 (110.05)**
Correction for self-selection using the Oaxaca (1988) method and Heckman (1979) procedure			
Log of monthly earning for males	**7.05**		
Log of monthly earning for females	**6.88**		
Difference	**.175**		
Total difference in average observable characteristics (explained)	**-.040 (-23)**	**-.015 (-8.5)**	**-.007 (-4)**
Total difference in the returns to observable characteristics (unexplained)	**.216 (123)**	**.190 (108.5)**	**.182 (104)**

Percentages are in parentheses

~Too small to report

Table 7
Decomposition of log monthly earnings using the Neumark (1988) Method

Log of monthly earning for males	**7.07**		
Log of monthly earning for females	**6.89**		
Difference	**.179**		
	Specification		
	(1)	**(2)**	**(3)**
Explained			
Age	.037	.006	.004
Age squared	.003	.043	.044
Dependence	.017	.014	.014
Race	-.019	-.015	-.015
Years of Education	-.056	-.037	-.039
Not attending school	-.0002	-.001	-.001
Years of potential experience	-.009	-.006	-.006
Experience Squared	-.054	-.073	-.082
Region	-	-.001	-.001
Firm Size	-	.006	.005
Tenure	-	.007	.009
Employment Status	-	0.018	.019
Temporary contract	-	.001	.0007
Social Security contribution	-	~	~
Occupation	-	-	.017
Total difference in average observable characteristics (explained)	**-.081** **(-45.25)**	**-.037** **(-21.22)**	**-.031** **(-17.31)**
Total difference in the returns to observable characteristics (unexplained)	**.260** **(145.25)**	**.217** **(121.22)**	**.210** **(117.31)**
Female returns compared to the pooled wage structure	**.260** **(145.25)**	**.217** **(121.22)**	**.210** **(117.31)**
male returns compared to the pooled wage structure	**-**	**-**	**-**

Percentages are in parentheses
~Too small to report