LUND UNIVERSITY
School of Economics and Management

Master in Economic Development and Growth

Women in the Workforce: A Decomposition of Female Racial Wage Gaps in Brazil

Laura Giles Álvarez

laura.giles_alvarez.824@student.lu.se

Abstract: This paper analyses how racial wage gaps amongst women in the Brazilian urban labour market have evolved in response to a strong increase in workers' skills between 2001 and 2011. For this purpose, Melly (2005, 2006) quantile decomposition frameworks will be applied using PNAD data. In so doing we acknowledges a gap in the literature by focusing on women instead of men, extending wage decomposition frameworks applied to Brazil and updating wage gap estimates by using the latest 2011 data. Our conclusions thus coincide with Salardi (2013) in that increased female labour force skills are the strongest factor determining the reduction, albeit maintenance of glass ceiling effects. However contrasting previous findings by Foguel & Acevedo (2006) we find that the combination of higher skills, greater returns to those skills and changing unobserved factors drive, in conjunction, the lowering of racial wage gaps around the median of the earnings distribution. Finally our key contribution to the existing empirical literature in Brazil reveals that is has been mostly the combined effects of higher skills and unobserved economic wage structures which mostly explain the unexpected rise of sticky floor effects amongst women in the workforce.

Key words: Race, women, wages, inequality, labour market, skills.

Acknowledgements

Contents

List of Tables and Figures

Part I. Introduction

Racial wage discrimination is known to be the primary source of inequality in the Brazilian labour market today (Leite, 2005). While gender wage gaps have reduced considerably since 2001, having caught much public and political interest, racial earnings differentials have not fallen as much (Soares, 2000). Indeed, it was not until recently that discrimination in terms of skin colour was even acknowledged as a problem in a country where non-white workers, comprising 50.7% of the population today (Census, 2010), earn around 45% less than their white counterparts (Atal, Ñopo & Winder, 2009 and The Economist, 2012). This is a cause of concern; as it has been proven that racial wage gaps constrain prospective economic growth by producing inefficiencies in the market when moving resources from the most productive to less productive agents (Lang & Lehmann, 2011).

Traditionally, racial wage discrimination has been defined as the differing labour returns across individuals, due to their race, who are equally productive and employed in the same sector and/or region (Loureiro, 2003). However this concept is now thought to be too narrow. The most recent strands in the literature hence reveal much more complex structures behind racial disparities in the labour market (Bailey, Loveman & Muniz, 2012). They account for this form of discrimination as the result of a combination of differing skill levels across workers (a Composition Effect), social associations due to skin colour stemming from the historical background (a Price Effect) and other wage and labour market structures (a Residual Effect) (Juhn , Murphy & Pierce, 1993). O'Neill (1990) and Juhn et al. (1993) theorized that differing human capital externalities in the form of these three effects underpin the basic mechanics of discrimination, principally so for women (Neal & Johnson, 1996).

In the particular case of Brazil between 2001 and 2011, and despite the prevailing yet lowering racial wage gaps, the country has undergone one of the most radical economic and social transformations in its history. Inequality levels have decreased to unprecedented rates below 0.55 Gini Index values; welfare levels have improved beyond any previous record by doubling GNI per capita[1] and the country has established itself as the 6th largest world economy (World Bank, 2000-2011). Most importantly in the light of O'Neill's (1990) ideas, there has been a great improvement in skills: On the one hand higher life expectancy, decreasing fertility rates and higher quality of life (Lee, 2003) have increased age expectations and consequently years of work experience amongst the population. On the other hand school attendance for under 5 year olds has doubled, 83% school attendance for 15-18 year olds has been achieved and access to

---

[1] GNI per capita levels from 6.82 to 10.18 at current PPP level.

tertiary education for whites has doubled, whilst tripling for non-whites between 2001 and 2011 (Síntese de Indicadores Sociais (SIS), 2012).

In concurrence with these increases in skills, racial wage gaps for women have decreased in Brazil between 2001 and 2011, as perceived in Figure 1. However, the interesting fact is that this reduction has not been homogenous across the whole wage distribution, which could explain why racial wage gaps are not falling as fast as gender wage disparities (Soares, 2000). For example we corroborate previous findings by Salardi (2013) and Foguel & Acevedo (2006) of falling, yet prevalence of glass ceiling effects amongst women A glass ceiling effect in this context is understood as the disadvantages encountered by a sector of the population to reach occupations higher up the job hierarchy (Morrison, White& The Centre of Creative Leadership, 1987) : as perceived in Figure 1, there are rising wage differentials throughout the top half of the wage distribution. These differentials decrease considerably more for the 9th decile than for any other part of the wage distribution between 2001 and 2011. Moreover, from Figure 1 we also perceive the growth of sticky floor effects by 2011: despite the fact that sticky floor effects exist in 2001, their reduction is smaller across time in comparison to the rest of the wage distribution. Therefore they become more pronounced across time (also visible in Figures 5 & 6). We define sticky floor effects as the situation arising when non-white workers are appointed at the bottom of the pay scale and whites further up that same pay scale; even thought they are otherwise an identical workforce (Arulamplam, Booth & Bryan, 2006).

Figure 1. Female Hourly Log Wage Racial Differentials by Year
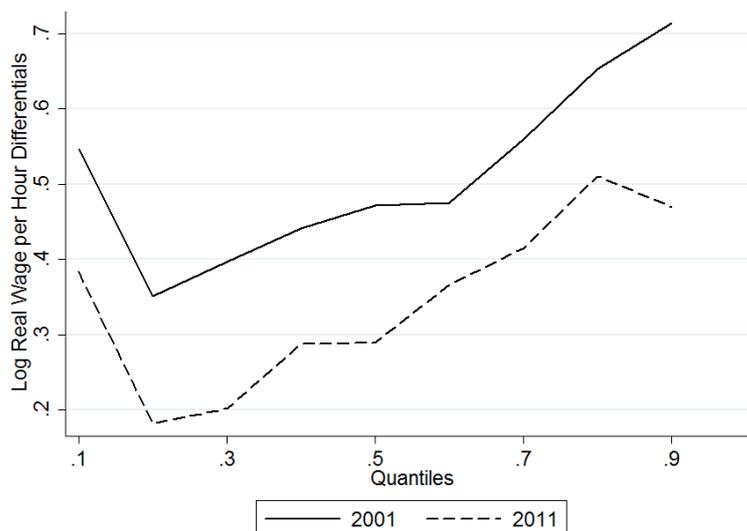
Source: Author's calculations from PNAD 2001 & 2011

The purpose of this paper is thus to analyse the effect the outstanding increase in workers skills has had in the heterogeneous evolution of racial wage gaps for women in urban Brazil between

2001 and 2011. Skills in this context are understood in terms of human capital as a combination of schooling and work experience following Mincer (1958). Our focus on the skill effects on racial wage inequality answers both to the theoretical and empirical underpinnings expressed above, together with the aim to further understand previous findings by Giles (2012). This last paper established that changes in education levels and racial wage gaps were the predominant factors in driving the reduction in poverty and inequality for metropolitan Brazil between 2001 and 2009.

By focusing on the issue of racial wage gaps amongst women during this time, we are also contributing to the existing literature in three ways. Firstly we are extending existing evidence on racial wage gaps which only covers up to 2006 in Salardi (2013). Instead, we are using the latest *Pesquisa Nacional por Amostra dos Domicílios* (PNAD) data covering from 2001 to 2011; coinciding with the outstanding economic and educational improvements in the country. Secondly, our focus on women answers to previous shortfalls in the literature as discussed by Soares (2000). We believe a pure male-based study of racial discrimination fails to understand how white and non-white women are experiencing policy changes in the labour market. This has profound effects for current and future generations in Brazil, as much evidence reveals black women in particular to be the most vulnerable sector in the Brazilian labour market (Fredman, 2012). Their strong discrimination also has effects on the vulnerability of future generations, as greater education levels and equal occupational access for women have strong externalities on child survival and education (Hill & King, 2010). Finally, we make use of some of the latest quantile regression methods available today. By means of this contribution are reviewing previous assumptions of decomposition frameworks, to obtain new insights into the role that skills and their remuneration are playing in the racial wage gap across the whole wage distribution. In particular, to our present knowledge the application of Melly (2005) has never been carried out in Brazil. This will allow us to extend the basic Melly (2006) framework based on Machado & Mata (2005) and assess racial wag gaps on the basis of a tripartite effect decomposition.

By means of these contributions we therefore aim to answer two questions: in what way have racial wage gaps amongst women reduced across the earnings distribution between 2001 and 2011? And how has a greater skill composition of the workforce contributed to this phenomenon?

The methodology used to answer these questions is based on Melly (2005, 2006) quantile wage decomposition methods. Melly (2006) portrays the effects of skills in a multidimensional framework, differentiating between the effect of increased skills per se and wage structures.

Consequently we decompose the racial wage gap by year into the effects of characteristics and coefficients in each quantile (Salardi, 2013). Whereas Melly (2005) takes the analysis a step further in the essence of the Juhn et al. (1993) by decomposing wage gaps into the effect of characteristics, median coefficients and residuals within the quantile framework and taking into account heteroscedasticity in the error term[2].

On a final note, we are also aware of the limitations present in this study. Firstly we are only focusing on urban areas. Urban areas provide meaningful economic units, less based on political decisions and portraying a greater concentration of highly educated individuals (Bacolod, Blum & Strange 2007). Moreover urban areas are now the central setting for formal labour markets in the country, as they home 85% of the population (World Bank Data, 2011). Secondly we only cover the period between 2001 and 2011. In so doing we feel this can portray in our model the effects of economic growth and educational expansion more accurately, without capturing the effects of the strong political and economic changes which took place in the 1980s and 1990s. Our period of interest coincides with the outstanding increase in education presented above and is, therefore, more relevant to the question at hand. Finally, due to the aggregate nature of our model, we cannot include other institutional functions likely to have affected our results (Melly, 2005). However, we will discuss this issue in more detail throughout the course of this paper.

Part II outlines previous research on racial wage gaps in Brazil. Part III reviews the theoretical approach behind the evolution of racial wage gaps across the wage distribution. Part IV analyses the data and variables used in the model. Part V discusses the methodology and Part VI comments on results which are concluded together with their policy implications in Part VII.

Part II. Previous Research

Despite the shortfall of specific studies on racial discrimination amongst women in Brazil, the past 50 years have seen a rise in microeconomic level studies worldwide due to the availability of reliable micro-level data and the advancement in econometric techniques (Dickson & Harmon, 2011). These methodologies were mostly developed in the US and were later applied to Brazil due to similarities in ethnic composition and geographical size. Stemming from Becker

---

[2] A large sector of the literature has worked at developing decomposition methods for quantiles which apply residual analysis; however most of it has been applied to explaining earning inequality changes across time. Instead, we are following Salardi (2013) to analyse racial wage gaps separately across two years.

(1957) who treated investment in education as a capital investment[3], these studies have allowed for the analysis of the effect investment in skills has had at an individual level. In particular following Mincer (1958) by proxying skills acquisition with schooling and experience, a large strand of the literature has worked towards shedding more light onto the understanding how higher skill level and its return impact the distribution of income subject to differences in class, gender or race (Dickson & Harmon, 2011). Some of the most notable examples in the literature were developed for the US; however due to the very vast literature on racial inequality in Brazil, we will limit the literature review to work carried out in this latter country. Moreover this section will be organized based on the conceptual and methodological development of these types of analysis in Brazil.

The study of racial discrimination in the Brazilian literature can be divided across three generations of scholars who exemplify how recent this phenomenon has been acknowledged in the country: The *First Generation* ignored the existence of a racial gap in the labour market and actually initiated their studies to exemplify the case of Brazil as a "racial democracy" as compared to other ethnically diverse countries around the world (highlighting Gilberto Freire in the 1930s) (Osório, 2008). The *Second Generation* however, recognized the existence of strong discriminating structures in the Brazilian labour market, attempting to highlight the national structures behind the prevalence of racial discrimination. Fernandes (1978) Telles (2003) or Garcia (2005) are notable examples of this group of academics. Finally it has not been until the emergence of the so called *Third Generation* of scholars who have taken on a more quantitative approach to racial discrimination by analyzing not only differing wage structures, but also historically reinforced differences in education or experience. This strand in the literature has thus gone beyond observed wage differences to embrace disparities across skills and occupation structures in order to determine why, despite political commitment to reduce wage disparities, racial discrimination persists in Brazil (Osório, 2008). Langoni (1973) was the pioneer in these types studies. By means of a dummy variable based regression, he underscored differing access to property as a major contributor towards racial wage gaps. Following his work, Hasenbalg (2005) and Silva (1980) developed the existing conceptual framework regarding wage differentials by carrying out studies which separated pure wage differential effects from other skill composition differences across the population in the country (Bailey et al. 2012). Hasenblag highlighted the importance of historical discriminating structures, not just in wages, but also in the composition factors which determine those wages. The latter study by Silva, N. (1980) revealed that coloured workers in Brazil (both black and mixed) systematically earned

---

[3] Becker (1957) carried out the first formal study of wage discrimination. He analysed why workers with the same level of skills, schooling, experience and productivity had different salaries according to their skin colour, sex or religion.

16% less than their white co-workers due to a pure discrimination effect[4]. Therefore by means of these studies, the understanding of discrimination was broadened to encompass social and historical differences across races.

The application of economic techniques for analysing racial discrimination in Brazil has thus developed from this *Third Generation* of authors. Most of these studies have built on econometric models based on Oaxaca (1973) & Blinder (1973) wage decomposition technique and its extensions; portraying a complex structure behind racial wage gaps by breaking down observed differences in the labour market into the effect of skill composition and discrimination to its returns (Bailey et al. 2012). Noteworthy examples in the literature include Reis & Paes de Barros (1989) who determined that black vs. white wage differentials were not due to pure discrimination; as differences in educational attainment explained a significant part of the racial differentials in wages. They quantified that almost 50% of wage inequality in Brazil was due to educational differentials. Indeed education has been repeatedly highlighted as the greatest factor behind wage differentials, accounting for between one-third and one-half of the total observed inequality in the labour market across time (Barros & Mendonca, 1995 and Ramos & Vieira, 2001). This evidence thus provides support for the use of a multifaceted framework, based on our arguments outlined in previous paragraphs. Soares (2000) is also good example of the development of multifaceted approaches to discrimination. This study decomposed racial wage gaps in Brazil by distinguishing discrimination in terms of pure human capital, different opportunities in the labour market and differential wage structures. The conclusion was that black males suffered a mixture of education differential and wage discrimination, whereas black females suffered these two consequences and also an opportunity differential in the labour market. Indeed Soares (2000) is one of the few studies in the literature of racial discrimination which places special emphasis on the differentials between black and white women. For the purpose of our study, it is especially noteworthy for the purpose of our article to note that despite white women suffering a pure discrimination effect, black women on the other hand suffer all forms of discrimination which are envisioned in the literature (Soares, 2000). Whereas Oliveira & Gonçalves (2006) extended this line of work, focused on women, highlight the important effect of both lower educational attainment and lower remuneration of this education for black women in response to white women; making this latter sector the worst off in Brazilian society.

More recent work in this field has built on this past evidence and attempted to explain the changes and implications of wage gaps during the New Millennium, when the country has

---

[4] Silva, N. (1980) reached this result controlling for a set of economic factors such as human capital and labour market characteristics.

performed outstandingly well in economic terms, while racial gaps persist (Giles, 2012). Garcia, Ñopo & Salardi (2009) found that wage gaps have been decreasing hand in hand with the recent economic improvement of the country, though racial wage gaps have decreased considerably less than gender wage gaps. Arias, et al (2004) embraced more recent techniques by taking into account the whole wage distribution rather than just the mean. Their findings distinguish different effects across deciles, where black males earn 24% less than their white counterparts in the bottom decile; a gap which increases to 56% in the top decile. These widening racial wage differentials would thus suggest the existence of glass ceiling effects for racial wage gap in Brazil and reinforce the need for the use of techniques which go beyond the mean (Soares, 2000).

Finally, an interesting fact revealed by most studies is that the unexplained change in wage inequality has decreased over time; becoming less significant for wage trends from 2001 onwards. This line of evidence stems from advancements in methodology techniques previously framed and is of special interest for the purpose of our paper. Nevertheless, this result is controversial and subject to methodology. Arias, Yamada, & Tejerina (2004) find that the unexplained component of wage distribution is greater for higher than lower quantiles of the distribution, where there are productivity enhancing factors which are not captured in survey data, and which could explain a dragging effect of the residual on improvements in wage distribution in Brazil. Cunningham & Jacobsen (2008), by means of a combination of the Oaxaca & Blinder and Bourguignon, Ferreira & Lustig (1998) decompositions look at wage inequalities across and within gender and race in Bolivia, Brazil, Guatemala and Guyana. They find that residual inequality is the major contributor to overall inequality. Foguel & Acevedo (2006) perform one of most recent updates to the literature by expanding their analysis between 1984 and 2005 following Juhn et al. (1993) methodology. They determine Price, Composition and Residual Effects played significant roles in racial wage gaps. Non-observables however played the greatest role in racial wage gaps during this time, except during 2001-2005; when the Price Effect also increased its role considerably. However these results most likely capture the effect of the drastic economic and political reforms with the country underwent throughout this time: trade liberalizations, hyperinflation, the 1988 Constitution or the IMF intervention for example. Finally, Salardi (2013) embraced the most recent decomposition methods in the literature by carrying out a comparative study of each of them for Brazil between 1996 and 2006. To our knowledge this is the most recent update to racial wage differential in the country. The author confirmed the persistence of glass ceiling effects for racial wage gaps amongst men

mostly due to a Composition Effect; whilst gender wage gaps presented both sticky floor and glass ceiling effects, largely due to a Price Effect.[5]

We conclude this section by highlighting that the development of studies on racial discrimination in the country has developed complex multifaceted frameworks around this issue. The latest strands in the literature have thus attempted to explain wage differentials across race in a quantitative fashion, distinguishing between skill differentials and their returns across the wage spectrum. Accordingly it has been highlighted that it is lower education attainment, access to inferior jobs and lower market remunerations to skills which shape and have maintained the racial wage gap in Brazil. However a large part of this gap remains unexplained in the form of the residual. Despite a modest increase in studies on wage distributions across female workers, the focus on men has been prevalent. Therefore, understanding how these observed and unobserved mechanisms affect racial wage gaps across the wage distribution in the Brazilian female workforce remains a gap in the literature today.

Part III Theory: The Drivers of Labour Income Inequality

The consequences of discrimination in the labour market on development and growth are severe. As, stemming from Becker (1957), discrimination leads to inefficiencies in the labour market which originate from the reallocation of resources from more to less productive individuals in the economy (Lang & Lehmann, 2011). Consequently much research in applied economics works towards understanding what brings about and maintains wages gaps has been carried out in applied economics by building on the Mincer (1958) equation framework:

$$logY = \alpha + \beta S + \gamma X + \gamma X^2 + \varepsilon \qquad (1)$$

Wage per hour(Y) is portrayed as a function years of schooling attained (S) and experienced (X),[6] both of which are remunerated by β and γ respectively. The reason for proxying skill with years of schooling and experience is that Mincer originally formulated earnings outcome from a process where individuals invest in these two types of human capital (Lemieux, 2002). Any other factor which affects wage, together with unobserved skills of workers are portrayed as a residual component (ε) which has been determined to play a varying role in wage structures across the literature (Rauch (1993) Glaeser, Scheinkman & Shleifer (1995) Armington & Acs, (2004) and Moretti, (2004)).

---

[5] Both Composition and Price Effect will be defined in detail in Section III.
[6] Note that introducing experience squared addresses the non linear pattern of these variables in the data.

Consequently based on the Juhn et al. (1993) decomposition, from this framework wage differentials are thought to be composed of three elements:

- A quantity effect, determined by differences in the observable characteristics of workers.
- A price effect, determined by differing returns to observable characteristics.
- A residual effect, determined by the differing non-observable factors which determine workers earnings.

(Foguel & Acevedo, 2006)

Contrasting these ideas, traditional definitions of racial discrimination only encompass differing returns to workers with the same observable characteristics based on race;[7] i.e. a price effect based definition: The expression *money whitens* is very common in Brazil, as observations of the labour market reveal that the higher the end of the wage spectrum the whiter the population seems to become (Bailey et al. 2012). However, despite the traditional idea of the importance of the Price Effect on wage gaps, growing evidence and the application of decomposition frameworks for racial wage gaps, have determined that quantity and residual components of the wage equation are also determinants of wage gaps (Foguel & Acevedo, 2006). The former has been tackled in the literature predominantly as a historical lower level of education amongst coloured people; whereas the latter refers to non-observed factors which affect wages and are not modelled in the explanatory covariates of the Mincer Equation, such as family background (proxied by parent's education) or economic structures (proxied by unions, trade openness or minimum wage) (Juhn et al. (1991) and Bourguignon, Ferreira & Menendez (2004)). Nevertheless, the conclusions as to the role played by each component in racial wage gaps have been subject to much controversy depending on the method, time period and even geographical location employed. Likewise it must be noted that the Price and Composition effects are not mutually exclusive: it is the combination of both factors in observed and unobserved skills which seems to drive wage differentials in the labour market (Lemieux, 2006).[8] We therefore now proceed to analyse these three effects in detail and present stylized facts.

---

[7] In terms of the labour market, this is perceived by the differing labour returns across individuals who are equally productive and employed in the same sector and/or region.

[8] Most importantly for our second decomposition, under the assumption of stationary economic environment, the analyst can identify both the effect of both the observed and unobserved level of skills and composition of the workforce (the Composition Effect) and the rate of return to that composition (the Price Effect) in the evolution of the wage gap across time (Lemieux, 2006).

3.1 Changes in the Composition of the Workforce

It is often argued that differences in wages stem from difference in skills across the labour market (Henriques, 2001). Juhn et al. (1993) also acknowledged that changing composition can affect overall earnings dispersion by increasing or reducing heterogeneity of observed skills. From this perspective, racial discrimination is thus not only understood as the consistent choice of hiring white worker instead of hiring an observationally equivalent black worker at the same wage (Lang & Lehmann, 2011).

Many authors have highlighted changes in the education and experience endowment of workers as the major factor behind the rise in wages in Brazil (Salardi, 2013). Its importance is also deemed to be increasing, especially in urban areas due to a labour market that demands high skilled workers in greater amounts. These factors and not pure discrimination per se are most likely driving racial differentials (Acemoglu, 2003). As such, traditional studies of the composition of the workforce include the analysis of the age composition, educational and experience levels at a given time.

Firstly, higher age is perceived in conjunction with higher earnings, as older workers are generally more experienced (Johnson, 1993). The age structure of Brazil corresponds to a country which is undergoing demographic changes at a fast pace: with an aging population, decreasing fertility rates and higher quality of life (Lee, 2003). This is reflected in the workforce, which partly due to the effect of longer compulsory education and partly due to population aging has resulted in a more mature active female labour force. Figure 2 presents the age composition of the urban workers for both black and white females which were employed the week of the PNAD survey in 2001 and 2011. It can clearly be seen that the columns of the population below 24 years of age have decreased significantly, most probably due to the greater affluence of women obtaining higher levels of education and thus joining the workforce at a later age. However the workforce 39 years old and above has spread out more evenly across age groups, increasing the amount of workers between 39 and 65 of age.

Figure 2. Age Composition of the Female Workforce

Age Structure of Female Workforce 2001   **Age Structure of Female Workforce 2011**
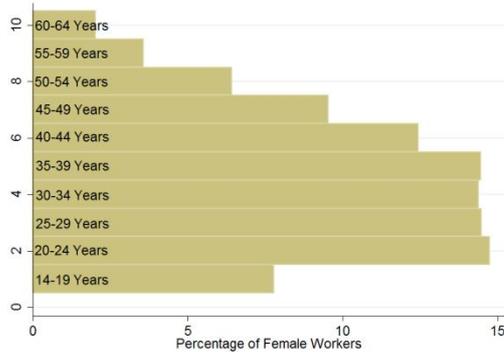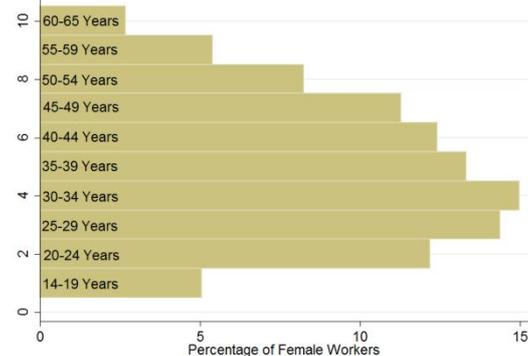
Source: Author's calculations from PNAD 2001 & 2011

Moreover as perceived in Table 8 in Appendix 1, the average age by quantile increased as we progress to higher quantiles; which is consistent with the overall positive relation between skill level and wages per quantile in Salardi (2013). The average age per quantile also generally remained circa 3% higher for white than for non-white sectors of the population as well as higher in 2011 than in 2001. It must be noted that there has been a homogenization of the workforce across races as the 2011 white female workforce is on average 2-3 years older than in 2001 and the 2011 non-white workforce is 2-5 years older than in 2001. This factor could point out at an increased maturity of the black workforce due to greater education and longer life expectancy.

Secondly, average education amongst women has increased from 9.24 average years of education in 2001 to 10.74 in 2011[9]. White women experienced a 15% mean increase in education across the workforce; whilst non-white women increased education 26% throughout the whole wage spectrum (PNAD 2001, 2011)[10]. These improvements have been determined to exert an important effect on wage structures (O'Neill, 1990). Previously it was determined that the reason behind institutionalized differentials according to race in Brazil was due to the historical slavery based roots of this sector of the population (Loureiro, 2003). As a Portuguese colony, Brazil received almost four million African slaves during three centuries, which

---

[9]These figures are higher than men's figures; confirming overall higher education levels amongst women in urban Brazil than amongst men.

[10] Figures obtained from PNAD 2001 & 2011 data.

together with powerful European colonizers set racial disparities as the base for the socioeconomic composition of the country (Heringer, 2002). A good example is the fact that the South of the country, composed primarily of European descendants is considered the rich sector of the country, whilst the Northeast, where the slave ports were historically located is the poorest sector of the country. Despite the end of the colonial period in the 1820s and the abolition of slavery in, Brazil has never carried out an active policy for racial equality as the USA did (The Economist, 2012). Instead, racial differentials have traditionally instead been disregarded in the hopes that time would bring equality. However, studies have pointed towards family and individual surroundings as major contributors towards education acquisition (Hanushek & Woessmann, 2010). This would therefore suggest that historical low quality and quantity of education for coloured population would have trickling effects for coloured individuals today in the Brazilian community (Ferreira, 2000).

Despite the above, more years of education overall provides little information as to how this improvement is important in terms of the level of education actually obtained: as noted by Holanda-Filho & Pessõa (2008) and Moura (2008), years of education do not mean completed years of education nor completed educational levels; a distinction which is noteworthy in a country such as Brazil where levels and threshold levels exist. Therefore it is not important how many years of education one has per se, but how many education cycles they have successfully completed. As such, the table below has broken down the increase in the education level for White and Non-white women according to each education cycle. Years of study are grouped into 5 groups by educational level: primary, secondary, tertiary and post graduate. Primary education comprises the first seven years of primary education, known as *Ensino Fundamental[11]*; secondary school, *Ensino Medio*, corresponds to 8-10 years of education; a Bachelor Degree comprises 11-14 years of study. Finally those carrying out post-graduate studies are present 15 or more years of education. Table 9 in Appendix 1 reveals increasing years of education the higher the quantile, which are around 1-2 years less for non-white workers with respect to white workers and also around 2 years more of education per person in 2001 with respect to 2001 for all races and quantiles. Education differentials by quantile are lower for non-white than for white female workers throughout the whole wage distribution. However wage differentials have been smallest at the tails of the wage distribution and wage gaps have reduced considerably over time.

Contrasting education composition across quantiles, Table 1 which presents the population structure by year of study, reveals a decrease in the number of women who only coursed *Ensino*

---

[11] However due to ocular inspection of the data, we have divided primary education category into two separate groups, from 0 to 3 years of study and from 4 to 7 years of study.

*Basico o*r less and an overall increase in women coursing Bachelor Degrees or higher. Our results show the considerable increase (from 12% to 22%) in of the number of black women with years of study equivalent to a degree (12-15 years of study) ;a figure which is truly outstanding. s. This confirms the facts published in the latest SIS (2012) report: that a greater number of women are achieving higher levels of education today, especially so for postgraduate studies[12]. Moreover, despite fewer white women completing only *Ensino Medio* in 2011 than 2001, more non-white women achieved this level of education.

Table 1: Composition of the Workforce

Summary Statistics for Women				
	2001		2011	
Sample Size	49228		53898	
	White	Non-White	White	Non-White
	27115	22113	26730	27168
	55%	*45%*	*50%*	*50%*
	100%		*100%*	
By Years of Schooling				
Less than 4	2619	4595	1548	3051
	5%	*9%*	*3%*	*6%*
4-8	8702	8736	5760	7508
	18%	*18%*	*11%*	*14%*
9-11	1909	1691	1528	1956
	4%	*3%*	*3%*	*4%*
12-15	9709	6020	11901	11773
	20%	*12%*	*22%*	*22%*
16 or more	4176	1071	5993	2880
	8%	*2%*	*11%*	*5%*
	100%		*100%*	

Note: Brasilia is exlcluded from the sample.
Source: Authors calculations from PNAD 2001 & 2011

Finally, within the Mincerian equation (Equation 1), skills are not only measured in terms of education achieved, but also by the level of on the job training (Lemieux, 2002). It is assumed that returns to experience follow a concave pattern where people with greater experience in a field will have higher skills and thus greater remuneration but only to a certain level from when

---

[12] The SIS (2012) report determined that non-white workers increased their access to higher education by 30%. However, although our number is more modest, (SIS) 2012 refers to the whole country, and our sample is only restricted to women and to those who completed the equivalent of 4 years in this category of education.

this pattern reverses (Lemieux, 2002).[13] Years of experience for the female working population between 2001 and 2011 remained at a stable average of 19.50 years. White women had an average of 18.9 years of experience in 2001, a percentage which slightly increased to 19.5 in 2011. However Non-white women had a higher average experience of 20.22 years in 2001, which increased to 20.33 years in 2011. The slightly higher level of experience for non-white women is expected and coincides with the lower level of education of this latter sector of the population: Non-whites, on average, seem to be entering the workforce at a younger age. Similar patterns across quantiles are found in Table 3 of Appendix 1. This, together with similar results for Salardi (2013) on education and age patterns confirms the fact that age, experience and education all present consistent positive relationship between wages and human capital endowments (Salardi, 2013).

To conclude this section, beyond observed characteristics of the workforce, it must be noted however that the Composition Effect also plays a role in the unobserved characteristics of the population. On the one hand it is known that individuals have different levels of skill, both in terms of education and experience which will lead them to obtain differing wage levels in the labour market. However authors such as Acemoglu (2003) extended this analysis to unobserved characteristics. The arguments posed by this article establishes that new technologies are complementary to certain skills which are specific to each education group, where certain skills are rewarded more highly in society today, and where such skills are imperfectly correlated to years of education. Returns to education are thus higher at the upper quantiles; a fact which is present throughout the tables in Appendix 1. This part of the literature therefore reinforces the need to analyse both the observable and unobservable factors in the traditional Mincer equation. Modern econometric techniques and detailed data have allowed for a more thorough understanding of these issues (Autor, Katz & Kearney, 2005).

3.2 The Price Effect

It is not only the effect of the composition of workers' skills which determines wages in the labour market, but also how those skills are paid for (Juhn et al. 1993). Soares (2000) for example found that pure discrimination is the major factor behind wage differentials. This is known as the Price Effect of individuals in the labour market and just as the Composition Effect, can be measured both for observed and unobserved skills (Lemieux, 2002).

---

[13] It must be noted that this assumption excludes people who gained non transferable skills in a particular environment and which later on were not applied to a new job market.

The mechanism behind racial discrimination in the case of Brazil is clearly one where the market uses race to infer information about productivity (Lang, & Lehmann, 2011). Race and social class come hand in hand in this country: the idea that social mobility may lead to racial mobility is a popular idea captured in the colloquial phrase *money whitens* (Bailey et al. 2012). Moreover the traditional definition of racial discrimination in the labour market refers to discrimination in terms of the Price Effect: differing labour returns across individuals who are equally productive and employed in the same sector and/or region (Loureiro, 2003).

In Table 2, racial discrimination amongst women is clearly visible within all education levels; where non-white women with the same level of education consistently earn less than their white counterparts. However these gaps vary according to the level of education: wage gaps have decreased less for women below 4 and above 16 years of education. Moreover we can perceive higher wages per year of education across time, which have increased radically for lower education levels; particularly so for non-white female workers. Together with this evidence, Table 2 also plots average log wage by year of experience. These trends follow a concave pattern with higher returns to experience for white compared to non-white workers; returns which have reduced across time.

Table 2. Wage Gaps by Year of Schooling and Experience

Average Log Real Wage per hour by Education Level (R$)					
	White	**Non-white**	**White**	**Non-white**	**2001**	**2011**
Average Hourly Wage by Years of Schooling					Racial Gap	
Less than 4	2.35	1.86	3.88	3.15	21%	19%
4-7	3.09	2.29	4.46	3.56	26%	20%
8-10	3.3	2.65	4.25	3.82	20%	10%
11-15	5.85	4.16	6.63	5.58	29%	16%
16 or more	15.15	12.14	15.55	12.82	20%	18%
Average Hourly Wage by Years of Experience					Racial Gap	
Less than 5	4.06	2.51	5.52	4.42	38%	20%
5-9	5.28	2.93	7.57	5.48	45%	28%
10-14	6.58	3.22	8.37	5.55	51%	34%
15-19	6.89	3.72	8.73	5.77	46%	34%
20-24	7.15	3.66	8.85	6.08	49%	31%
25-29	7.02	3.56	9.15	6.09	49%	33%
30-34	6.21	3.35	8.97	5.64	46%	37%
35-39	5.22	3.6	8.02	5.30	31%	34%
40-44	4.73	2.78	6.38	4.20	41%	34%
45-49	3.59	2.56	5.05	4.07	29%	19%
50-54	2.89	2.21	4.80	3.71	24%	23%

Note: Brasilia is exlcluded from the sample.

Base year for real wages: 2005

1R$=$0,5

Source: Authors calcuations from PNAD 2001 & 2011

Graphing wages per level of education and also by years of experience can help clear the patterns in wage differentials in Brazil. Figure 3 reveals that average salaries per year of study have increased overall during the period of analysis. Patterns of returns to education for both races in 2001 and 2011 are convex shaped with a greater growth at higher levels than in lower levels of education, thus revealing higher demand for skilled labour in cities. Moreover, confirming threshold effects analysed by Holanda-Filho & Pessõa (2008) and Moura (2008), a clear peak is discernible at the 9th year of study, coinciding with the end of *Ensino Medio* cycle. However despite most wage structures following similar patterns, it must be noted that white wage remuneration structures for 2011 seem to have changed slightly, still following fluctuations until the 9th year of study. However, for this group, returns fall before the first year of study and increase more notably around the 5th year of study. This change thus means that by 2011, real wage differentials have increased for the primary education levels for the white sector of the workforce.

Figure 3. Returns to Education

Log Real Wage per Hour by year of study 2001　　　　**Log Real Wage per Hour by year of study 2011**
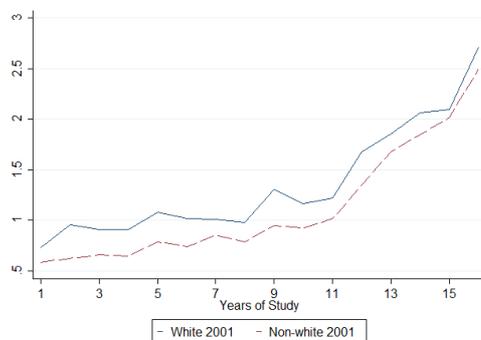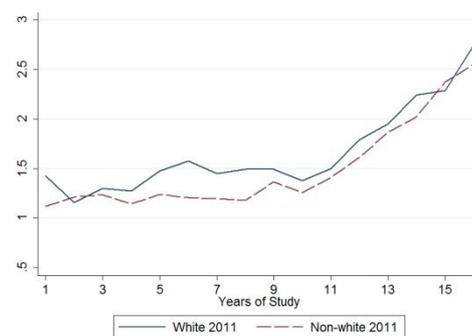
Source: Author's calculations from PNAD 2001 & 2011

Trends in log real hourly wages per year of experience, depicted in Figure 4, also follow patterns highlighted in Table 2. These patterns are concave (where returns to experience fall from 30 years of experience onwards), higher for whites than for non-whites and also higher in 2011 than in 2001. This result coincides with previous findings in urban Brazilian labour market by Ferreira & Barros (1999). The literature explains this pattern is the result of a decrease in productivity for workers above a particular age, especially in terms of health and physical condition (Johnson, 1993). Contrasting with the returns to education, wage gaps are larger per year of experience than by year of study and just as with education.

Figure 4. Returns to Experience

Log Real Wage per Hour by experience 2001

Log Real Wage per Hour by experience 2011
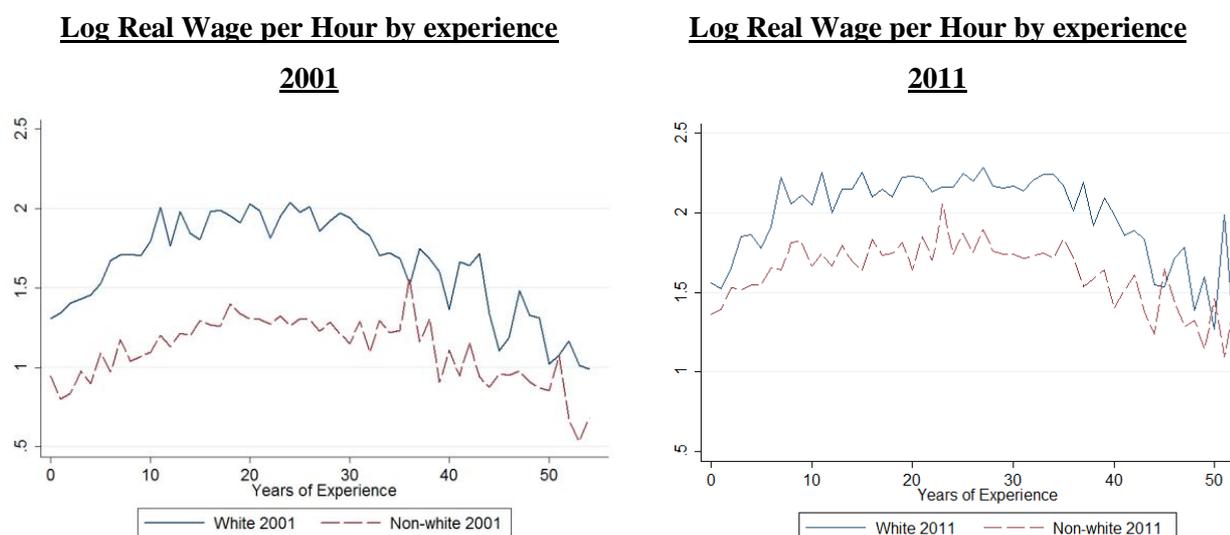
Source: Author's calculations from PNAD 2001 & 2011

3.3 The Residual Effect

Finally, there are also unobserved components of the Mincer Equation (Equation 1) which affect observed wage structures. Although their study in the literature is still largely contested and subject to assumptions, previous studies have highlighted three major factors behind changes in the residual components of the original Mincer Equation (Equation 1). Namely omitted variables, institutional factors and measurement errors (Lemieux, 2006).

Firstly, it is believed that education and experience alone are an imperfect proxy for skills in the labour market. This means that there are unobserved factors which also determine the skill level of a worker and are not captured by the original Mincer formulation (Equation 1), such as innate abilities of schooling quality. The literature has developed a large range of studies trying to control for this effect by introducing instrumental variables in wage equations for personal background (parent's educations) or quality of schooling (Bourguignon et al. (2004) and Ferreira, (2000)). Moreover, recent work by Florida has also made use of new datasets which measure skills in terms of demand and not of supply, although this part of the literature is restricted to the US for lack of data in other countries (Florida, Mellander, Stolaricl & Ross, 2012).

Secondly, institutional changes are also determined to influence unexplained wage gap patterns in urban Brazil (Foguel & Acevedo, 2006). The existing literature in particular highlights the effect of 6 variables: antidiscrimination institutions, minimum wage, unions, open trade, skill bias technological change and economic growth (Lemieux, 2006). Antidiscrimination institutions have greatly developed in Brazil during our period of analysis. Based on Articles 3, 4 and 7 of the 1988 Constitution, which condemns racial discrimination in the labour market; Lula and Rousseff's governments have set specific political objectives and national institutions towards tackling this problem. Today, the government has two institutions which work solely towards lobbying and promoting measures towards racial equality: the *Secretaria Especial de Políticas de Promoçao da Igualdade Racial da Presidência da República*, set in 2003 (SEPPIR), and the *National Council for Combating Discrimination*, which became the *National Council for the Promotion of Racial Equality* in 2004. Alongside these organizations, the Government has also set various legislative measures towards reducing racial disparities: A *National Affirmative Action Program* together with a *National Human Rights Programmee* both created in 2002, which condemn exclusionary practices and set quotas by race for work posts in the public administration and for contracting services by government agencies. Moreover, universities have established quotas for non-white students, together with scholarships for diplomatic studies for Afro-descendant Brazilians (Fredman, 2012). With reference to effect of minimum wages, we confirm it has strongly increased from monthly 246,79R$ to 406,09R$ between 2001 and 2011[14]. According to the ILO (2012), minimum wage is one of the most effective policy tools to reduce wage disparities along the bottom end of the wage distribution. However as pointed out in Table 11 in Appendix 2, our data reveal that the bottom decile of white workers and the two bottom deciles of non-white workers earn below minimum wage level and they do not seem to have been affected by this measure. Adding to the fact that only around 50% of urban workers are currently formally employed in Brazil (Salardi, 2013), we conclude that the positive effect of minimum wage is most likely a strong determinant of wage gaps across race, which leaves out the bottom end of the wage distributions. Just as with minimum wage, trade unions are also believed to be a strong political institution which affects wage structures, particularly so for workers earning around the median of the wage distribution (Lemieux, 2006). In Brazil, although trade unions historically did not have much of a role, their affiliation expanded during our period of study (Arbache, 2002). Nevertheless, from Figure 7 in Appendix 2 we can confirm that the majority of female workers are still not part of trade unions, which are also known to marginalise the least paid and lowest educated workers (Arbache, 2002). Furthermore the literature has reiterated the important role of skill-biased technological

---

[14] Both values are expressed at 2005 values in Brazilian Reais, where 1R$=$0.5 (IPEA, 2001, 2011)

change and trade liberalization are in determining wages ((Katz & Murphy, 1992), (Krugman, 1995), (Acemoglu, 2003) and (Di Nardo, Fortin & Lemieux, 1996)). On the one hand, skill-biased technological change explains how a greater demand of skilled workers in the labour market due to higher technological advancements, leaves those with access to lower skill levels worse off (Lemieux, 2006). However although Figure 3 depicts higher returns to education for the top end of the wage spectrum, the increase in returns to education across time was stronger for the lower level of wages. This fact thus contradicts the original skill-based technological change hypothesis and confirms that this factor is not likely to be having an effect in our case study. On the other hand the economic theory behind the trade liberalization hypothesis is based on Heckscher-Ohlin theory of international commerce. According to this view, trade liberalization has led to the specialization in a range of products, mostly high skilled, which increased their price relative to low skilled products, mostly analogous to non-black sectors and thus induced greater wage differentials in the country (Krugman, 1995).[15] Trade liberalization did have a strong impact on economic structures between 1988 and 1995, when Brazil carried out radical economic reforms with the instalment of trade liberalisation and control of inflation. However during the past decade, Brazil has not undergone particularly strong tariff cuts in comparison to the 80s and 90s and therefore also renders this factor marginal in our case study (Castilho, Menéndez & Sztulman, 2012). Ultimately, economic growth has also been known to affect wage structures. This argument is based on evidence surrounding the business cycle, particularly recessions, which affect non-whites more acutely (Juhn, Murphy & Pierce, 1991). Between 2001 and 2011 Brazil experienced positive but limited economic growth rates compared to other BRICS countries. GDP per capita ranged between 1.31% and 6.09% during this period and no economic recessions (World Bank Data, 2001-2011). Moreover Arbache (2001) found that the Brazilian wage structure was quite unresponsive to short-run economic changes.

Finally, the literature has also highlighted the effect of measurement errors in the Residual Effect (Lemieux, 2006). Although this factor is likely to be present in Brazilian economic models, we will discuss further the conclusion and methodology its possible implication in our decomposition.

3.4 Combining the Price, Composition & Residual Effects

Based on the theoretical framework above, decomposition methods following the traditional Blinder Oaxaca (1973) decomposition technique thus divide wage inequality between Price

---

[15]Acemoglu (2003) in particular theorized that emerging markets have undergone a joint process of technological change with product specialization through trade liberalization which could indeed portray this theory in the case of Brazil; which opened its borders in the 1980s.

Effects (changes in the wage structure) and Composition Effects (changes in the wage dispersion attributed to explained components in the wage regression). However, the introduction of quantile regressions and techniques now allow assessing the whole wage distribution instead of the mean. This has led to deeper analysis of relative changes not only in specific points of the wage distribution, but also in relative changes across the whole distribution (Lemieux, 2006). From Table 3 we can perceive wage inequality was higher for white than for non-white female workers and reduced across time for both races (as measured by 90-10 and 50-10 percentile ratios), a trend which reverses at the bottom end of the distribution.

Under the original Mincer framework and as developed above, the distribution of wages (W) is determined by the distribution of human capital (proxied by schooling and experience) and its price (determined to be the wage structure in an economy) (Lemieux, 2002). Therefore the variance of log wages across the wage spectrum is thus assumed to be the product of the squared return to human capital (r) times the variance of human capital (H):

$$Var(LnW) = r^2 x Var(H) \qquad (2)$$

Together with the variance of human capital and of wage structure, Mincer differentiates between the Price Effects of observed variables and other unobserved economic factors by means of the third component to the analysis of the wage distribution: the conditional distribution of the error term from Equation 1, (ε) (Lemieux, 2002). By means of an example whereby individuals with the same level of schooling invest in different quantities of work experience, Mincer points towards the fact that there are still wage variances within each skill category which are not captured in the variance of observed skills alone. This implies that variance of observed skills in the equation does not fully explain wage variance and thus renders the variance of the error term (ε) important[16]. In the particular case of Brazil between 2001 and 2011, this would explain why despite such a radical increase in education, this has not translated in the same measure to a reduction in wage gaps. Therefore, beyond possible measurement errors, the fact that residuals vary across skill levels would imply that there are unmeasured skills in the original Mincer equation and other structures in the economy which are varying across these skill levels and affecting overall wage inequality (Lemieux, 2002). Past decomposition methods base d on the original Blinder (1973) & Oaxaca (1973) methodology

---

[16]Lemieux (2006) also points towards the fact that error variance is not only important, but also varied across the wage distribution, thus rendering the variance of ε heteroscedastic. Heteroscedasticity implies that the variance of the error term is not constant, but varies with levels of skill: residual variance is thought to increase the higher the level of education for example.

would thus be only framing part of the picture by not taking into account the Residual Effect on log wage variances. This idea was further developed by Juhn et al. (1991) who provided the evidence in the literature that USA racial wage gap varied partly due to an increase in residual inequality.[17] Therefore under this light, racial wage inequality can thus formally divide wage variances into two components: "Between Group Inequality" and "Within Group Inequality" (Reis & Paes de Barros, 1989):

On the one hand, Between Group Inequality (BGI) is the difference in wages between workers who have the same level of education and experience but are differed by another factor: race or gender for example (Lemieux, 2002). Research on racial discrimination is thus normally studied from this perspective; where BGI is represented by the independent variables of the Mincer Equation (observed factors). In Table 3 we perceive that ,comparatively, BGI decreased much more for non-white workers than for their white counterparts throughout time, except at the bottom end of the distribution, where observed skill dispersion reduced considerably more for white than for non-white workers. However its effect was much smaller than the former and also decreasing with time. Moreover, negative values for 50-10 percentile ratio in 2011 point towards the fact that inequality has actually increased for the 10th centile white workers in 2011 in comparison to 2001. Non-white women in particular had lower between group inequality than white women, confirming findings by Salardi (2013) that differentials based on skills are higher amongst white than amongst non-white sectors of the population.

On the other hand, Within Group Inequality (WGI) is defined as the inequality in mean wage which would be observed if mean wage across all observed education and experience groups were the same (Lemieux, 2002). WGI is portrayed by unobserved factors and analysed by means of the variance of the residual in equation (1)[18]. The study of WGI has recently taken on a greater role in the literature due to the fact that it has been widely proven that the uncorrelation of the residual with measured variables in the Mincer Equation is far from true (Lemieux, 2006). From Table 3 we can infer that WGI followed the same overall trends than BGI, however to a larger extent. This means that residual wage inequality has actually been the greatest factor contributing to overall inequality during the period covered in our model specification[19].

---

[17] However it must be noted that the Juhn et al. (1993) method for residual imputation within their decomposition is based on the assumption of conditional rank preservation (Fortin et al. 2011).

[18] McCall (2000) went as far as including the residuals as independent variables in a two stage regression.

[19] It must be noted that, according to the number of explanatory variables in the model, the relative importance of BGI and WGI will change.

Finally together with changes in overall observed wage inequality, BGI and WGI; Table 3 also depicts the difference across races for all three measures in 2001 and 2011[20]. Differences in wage inequality measures across races allow us to conclude that overall whites have had higher wage, coefficients and residual dispersions than non-white workers across the whole distribution. However non-white workers have had a greater dispersion at the lower end of the wage distribution. Moreover while overall inequality and BGI have reduced with time, residual inequality has actually increased. We can thus confirm that this is due to a greater reduction in wage dispersion of the non-white workforce compared to the white workforce.

Table 3. Observed and Residual Inequality

	Wage Inequality					
	Overall Wage Inequality		Within Inequality		Between Inequality	
White Women						
	2001	2011	2001	2011	2001	2011
90-10	2.24	1.88	1.71	1.54	0.53	0.34
90-50	1.37	1.28	0.89	0.89	0.48	0.39
50-10	0.86	0.60	0.81	0.65	0.05	-0.05
Non-white Women						
	2001	2011	2001	2011	2001	2011
90-10	2.09	1.77	1.71	1.49	0.38	0.28
90-50	1.11	1.05	0.86	0.78	0.25	0.27
50-10	0.99	0.71	0.85	0.70	0.14	0.01
Racial Inequality Gaps						
	2001	2011	2001	2011	2001	2011
90-10	14.64	11.34	-0.34	5.77	14.99	5.57
90-50	27.03	22.77	3.71	10.72	23.31	12.05
50-10	-13.31	-11.42	-4.06	-4.94	-9.25	-6.48

Note: Racial Inequality Gaps were computed by the formula (white_ineq-nonwhite_ineq)*100

Source: Authors calculations from PNAD 2001 & 2011

To conclude this section, theory has determined that it has been a mixture of Composition, Price and Residual Effects which have determined the pattern of female wage gaps throughout the past decade (Juhn et al. 1993). Assessing the wage gaps in a framework which takes into account the whole wage distribution instead of just the mean allows determining how the reductions in racial wage differentials have changed relatively across different points of the wage distribution (Lemieux, 2002). In this light we have analysed how patterns in WGI and BGI have affected racial wage gaps across different wage percentiles. All measures of dispersion are higher for whites than for non-whites at the top end of the spectrum; yet they are

---

[20] Residual wage inequality was computed following Autor, Katz & Kearney (2008). Trends in residual inequalities were obtained by fitting the following model by sample and year:

$$lnw_i = \alpha_i + s_i + x_i + (s_i \times x_i) + \varepsilon_i$$

w_i is real hourly wages, s_i is a vector of five schooling completion categories, x_i is a vector of thirteen experience categories ranging between 0 and 58 years in 5 year increments, ε_i is an error term. Finally the inclusion of $(s_i \times x_i)$ provides a very flexible wage equation (Autor et al. 2008).

also higher for non-whites than for whites at the bottom end. In determining overall inequality, WGI has played a greater role than BGI. Despite the fact that racial differentials of overall wage inequality have reduced across time, racial differentials for residual inequality have actually increased. This means that unobserved factors are thus promoting wage dispersion for sectors of the population with lower wage inequalities, inequalities that are not being targeted appropriately.

The next sections will allow us to disentangle the contribution of the Composition, Price and Residual Effects to the reduction in racial wage gaps amongst women in Brazil between 2001 and 2011.

Part IV. Data

The data used for the model is composed of two cross section data bases provided by PNAD for the years 2001 and 2011. Due to the nature of the data, it is impossible to follow individuals across time; therefore quantile regressions will allow us to follow quantiles and their evolution instead. Moreover, we have restricted the sample to working women between the ages of 14 and 65.

The surveys cover 100,000 households per year in urban areas from 26 Federal States, where the city of Brasília is left out due to the strong bias arising from public sector wages (Salardi, 2013). PNAD is a multipurpose survey which is greatly used in the literature for its high quality data and ample coverage of variables including education, migration, fertility, income, earnings, housing and household access to services and facilities, making special emphasis on the country's participation in the labour market. The survey was first carried out in 1967 and the latest covers 2011, conducted yearly except for census years and 1994. Most importantly, as it has kept its main structure since 1976, it allows for long term comparisons with minimum comparability issues (Ravallion, Ferreira & Leite, 2010). In our case, the only variables which required recoding were education and work occupation, due to a slight modification in the grouping of both variables. Otherwise all variables were coded in the same way throughout both surveys.

However as in most surveys, there are drawbacks to our data. PNAD reports race and socioeconomic status by self-reporting, thus possibly suffering from endogeneity bias and self-selection (Bailey et al. 2012). It is also known to underreport income, which is self-reported, both at high and low income levels. Moreover, due to the lack of panel surveys in Brazil, the data presented is cross-sectional micro-data, fielded every September and carried out in three

sampling stages (Da Mata, Deichmann, Henderson, Lall & Wang, 2007). Other national datasets such as Pesquisa de Orçamentos-Familiares (POF) and the PPV survey are available and are widely used for Brazilian household-level analysis. Nevertheless PPV seems incomplete for the purpose of this study and POF carried out sampling at wider time intervals than was required; therefore PNAD seemed more appropriate for the purpose of this paper (Ferreira, 2000). Misinterpretations are also quite frequent amongst participants due to the format of the questionnaire and also due to the low reading skills of some sectors of the population (Ferreira, 2000). Finally, PNAD does not count with any measure of post-tax earnings, so all studies carry out income based analysis using gross earnings before taxes and without equivalence scales (Deaton, 1997).

4.1 Variables to be used

The total number of observations used in this study added up to 103126 throughout the two years. An analysis of outliers was carried out and 4 variables dropped due to unlikely level of wages per hour[21]. Ultimately the variables used for the purpose of this paper comprise socioeconomic characteristics of individuals above 14 years of age who have reported to be earning a salary in PNAD 2001 or PNAD 2011 including:

Wage per Hour

For the purpose of this paper, a variable for hourly wages was created from both data sets, with information on monthly wages and monthly hours worked for the individual's main job[22]. Wages comprise monetary payment of each individual aged 14 and over in every job performed[23]. The reason for using 14 year olds instead of 16 (which is the legal age to start working in the country) is due to the existence of apprenticeship periods in the country as from the age of 14 (ILO, 2009). Hours worked was transformed from a reported measure of weekly hours worked in the week of the survey to monthly, by multiplying by 4. For comparability across years, all quantities were set based using 2005 consumer price index measures from OECD.

---

[21] These observations declared log hourly earnings over 800R$ and did not carry out job descriptions which fitted these high wages, namely administrators and school teachers. Their omission does not alter results much but does indeed smooth the upper section of the wage distribution.
[22] PNAD also includes variables for earnings from all jobs and household earnings.
[23] Wages are only reported in monthly form; we are thus aware of possible measurement errors arising from possible errors in reporting hours worked.

Education

Education is measured in years of education[24] from 1 (less than one year of education) to 16 (equal to 15 or more years of education); as provided by the PNAD database. Education categories were built based on completion of education levels as described in section 2.

Experience

Due to the fact that there is no direct measure of work experience in our sample, this variable is built by the following expression following Ferreira & Barros (1999):

$$Exp = age - years\ of\ education - 6 \qquad (3)$$

The reason for subtracting 6 is that children start compulsory education at this age; thus it is assumed children below this age cannot be part of the workforce. Experience has been modelled in single and quadratic forms, to capture changing returns to scale across the wage distribution, as perceived in Figure 4.

Race

The sample is divided by race, selecting only individuals who are either reported to be white, black or mixed. A dummy variable for white was generated with value 1 if white and 0 if not white; whereas due to the model specifications requiring the selecting variable to be bivariate dummy, black and mixed were grouped together as "non white". For the purpose of this analysis we will thus not include those who are Asian nor indigenous. Asians have been grouped with white citizens in previous studies such as in Soares (2000); however this sector of the population only represents 1% of the total Brazilian population (Census, 2010). Indigenous population has previously been included in "non-white" categories. However, we did not deem this to be appropriate for our analysis for two main reasons: the indigenous population has very different historical roots to the black and mixed population[25], and in urban areas this racial group is a very small minority of the population (less than 0.5% of the population) (Census, 2010).

---

[24] We are aware other measures of education such as – level of education completed – are more appropriate measures of education. The fact that you coursed a year of education provides no guarantee that the year was completed successfully. However quality of data for our variable used was far superior and thus chosen above other measures.

[25] Due to the fact that quantile analysis covers not only a pure ethnicity, but also social dimensions which have historical roots, we feel that including indigenous as "non –white" could potentially distort results.

Occupation

In order to address possible issues arising from omitted variable bias and sample selection bias, Appendix 5 presents the Melly (2005, 2006) decomposition framework including occupation categories. This variable was built on the four occupational categories reported in PNAD: Employee, domestic worker, self employed and employer. Each of the four categories was then presented as an independent dummy with value 1 if the individual belonged to the particular occupation and 0 if otherwise. We note that all individuals who reported wages also reported belonging to one of these four occupation categories.

Part V. Methodology

Traditional studies of wage inequality have focused on the mean by building the Oaxaca (1973) and Blinder (1973) model. However, it is known that different measures of inequality yield varying results depending on the part of the wage distribution on which more weight is put. Accordingly, a growing number of methods have extended the traditional Oaxaca (1973) and Blinder (1973) model in order to account for the whole wage distribution as well as the distribution of the residuals (Melly, 2005).

The methodology used for this paper will be based on Melly (2005,2006); who based on Juhn, et al (1993), developed a semiparametric quantile regression technique[26] where the conditional distribution is integrated to obtain the unconditional distribution. The objective is thus to study the changes in the whole distribution of wages[27]. The reason for extending the analysis beyond the mean in the quantile regression model is twofold. Firstly it offers a more flexible model for heterogeneous data at various points of the conditional earnings distribution. It is useful to investigate whether the effect of race differs according to an individual's position across the distribution of income after controlling for the segmentation of the labour market and other individual characteristics (Bailey et al. 2012). Furthermore as the characteristics are allowed to influence the whole conditional distribution of Y. Any measures of inequality can be used to decompose the statistics (Melly, 2005). Secondly quantile regressions are also less sensitive to outliers, providing more robust estimators when facing non-normality and also have better properties than OLS in the presence of heteroscedasticity (Deaton, 1997). Accounting for heteroscedasticity is precisely the other main contribution of Melly's methodology which will be discussed at the end of this section. Not assuming for the independence of the error term thus

---

[26]The method is considered semiparametric in that it is assumed the conditional quantiles satisfy a parametric restriction without the need of distributional assumptions; i.e. the covariates are assumed to influence the whole conditional distribution (Melly, 2005)

[27] It is assumed the conditional quantiles satisfy a parametric restriction without the need to fulfil a distributional assumption (Salardi, 2013).

allows us to account for the effect of the residual in the decomposition, improving the quality of our results[28] and adding depth to the analysis of unobservable skills (Melly, 2005).

However Melly (2005, 2006) is not the only quantile regression based wage decomposition. Machado & Mata (2005) also developed important methods applied in the literature, but Melly extends these works by solving the problem of crossing of different quantile curves and by determining the asymptotic distribution of the estimator (Melly, 2005). As we are using a large number of quantile regressions performed (as is the case in our model: one for each percentile), Melly's methodology is also considered more efficient than the original Machado & Mata (2005). Fortin (2007) also developed a detailed quantile regression decomposition method which is said to improve various shortcomings of Melly (2006) and Machado & Mata (2005) decompositions. However it has been applied to provide a greater understanding of individual observable factors in the Mincer equation, not to analyze the residuals. We thus feel that for the purpose of this paper, where we focus on observed and unobserved human capital skills, Melly (2006, 2005) is more appropriate.

Despite the advantages which Melly's methodology provides compared to previous work, we are aware these methods are not without fault. It must firstly be noted that through the semiparametric wage decomposition, it is not the purpose of this paper to identify causal effects from the factors computed in our model (Salardi, 2013).[29] We are instead interested in analysing the joint contribution of experience and years of schooling on the wage distribution as our analysis focuses on the demand side of labour force inequalities. Therefore exogeneity does not produce biases per se but instead reduces effectiveness of the estimators. Secondly, strong assumptions behind quantile regressions are also maintained throughout our analysis, summarized in Appendix 3. Scepticisms exist as to whether the reality decompositions attempt to simulate actually fits all these assumptions (Fortin, Firpo & Lemieux, 2011). For this reason we will deal with some of the more problematic assumptions regarding our model at the end of this section. Thirdly we can only carry out an aggregate decomposition, however due to the nature of our question this is not much of a drawback. Lastly the method used is

---

[28] Assuming independent error terms is correct if the error term is indeed independent and normally distributed. However if the location model is unsuitable, a decomposition assuming homoscedasticy produces misleading results (Melly, 2005).

[29] It is not possible to identify causality by means of these techniques because of lack of exogeneity of the independent variables in these models (which are also determined by the wage structure) and also due to lack of choice in the binary treatment which defines our groups of interest (in this case race) (Fortin et al. 2011). This last point is not so relevant in the case of Brazil: although it is true one cannot choose their race, it has been proven that due to the self reporting nature of the questionnaire used, people had a greater propensity to describe themselves "whiter" than they really are.

computationally demanding, especially when carrying out Melly (2006) due having to carry out bootstrapping in our large dataset. For Melly (2005) we instead estimate a large number of quantile regressions (one for each percentile 1 to 99) and draw random variables of τ at random (Melly, 2005).

Taking all advantages and shortfalls into account, this paper carries out two types of wage decomposition. The first decomposition follows Melly (2006) and extends the original Machado & Mata (2005) model, decomposing wage differentials between white and non-white female workers into the effects of characteristics and coefficients at different quantiles of the wage distribution. This is carried separately for 2001 and 2011. The second decomposition extends Machado & Mata (2005) and is based on Juhn et al. (1993) by taking into account for heteroscedasticity. This is carried out following Melly (2005), which will add insights on how the residual has affected racial gaps across time. These decompositions propose estimators of distribution functions in the presence of covariates to simulate what the wage structure would be if non-white workers were rewarded according to the wage structure of white workers (Salardi, 2013).

5.1 The Model

Following Koenker & Bassett, (1978) by means of a quantile regression framework, where $\{y_i, x_i\}_{i=1}^N$ is an independent sample from a population where x_i is a *K* x *1* vector of regressions, the estimated regression for each quantile is defined as:

$$F_{Y|x}^{-1}(\tau|x) = x_i\beta(\tau), for\forall\tau\epsilon(0,1) \qquad (4)$$

The $F_{y|x}^{-1}(\tau|x)$ is the τth quantile of log wages per hour (*Y*) conditionally on human capital characteristics (x_i). Results for these equations are presented in section 6.1. In our model we have only taken into account the effect of variables which proxy workers' skills, where $x_i = \{Educ, Educ^2, Exp, Exp^2\}$ [30]. Following Ferreira & Barros (1999) we thus assume that the labour market is not segmented by region, firm size, or any other attribute other than race. The reason for restricting the quantile regression to the most austere Mincer specification is because Melly's decomposition can only account for the aggregate covariate's effect on wages. Therefore, as the focus of our paper is on worker's human capital alone, adding more factors to

---

[30]Different forms of this equation were tested, including proxying education with 5 dummy variables for education categories, or proxying experience with 13 experience categories. All specifications yielded similar results, however for the purpose of comparison with previous studies and due to the shape of earning return functions, we settled for the quadratic specification of the Mincer Equation.

the explanatory variables would simply add effects onto the returns of these variables that are not directly associated with skills.[31]

Under the assumption of linear relationship[32], it is thus understood that $\beta_t(\tau)$ are rates of return to characteristics as specified at the quantile of the conditional distribution (Melly, 2005). $\beta(\tau)$ is estimated for each τ by:

$$\hat{\beta}(\tau) = \arg\min_{b \in R^K} \frac{1}{N} \sum_{i=1}^{N} (y_i - x_i b)(\tau - 1(y_i \leq x_i b)) \qquad (5)$$

Where 1(.) is the indicator function.

Asymptotically, we can estimate an infinite number of quantile regressions. In a finite sample, the number of numerically different quantile regressions is $O(nlog(n))$, where each coefficient vector prevails on an interval; being $(\tau_0 = 0, \tau_1, \ldots, \tau_J = 1)$ the points where the solution changes. Letting $\hat{\beta}$ be the vector of all different quantile regression coefficients: $\hat{\beta} = \left(\hat{\beta}(\tau_1), \ldots, \hat{\beta}(\tau_j), \ldots, \hat{\beta}(\tau_J)\right); \hat{\beta}(\tau_j)$ prevails for τ_{j-1} to τ_j for j=1,...,J. (Melly, 2005). This is the model for the conditional quantiles of y.

Therefore to apply the Melly (2006) distribution, the conditional quantile function (Equation 4) is calculated both for whites and non-whites in each year:

$$\hat{q}(\hat{\beta}^W, x^W) = F_{Y|x}^{-1}(\tau_W|x_W) = x_{W,i}\beta_W(\tau), for \forall \tau \epsilon (0,1) \qquad (6)$$

$$\hat{q}(\hat{\beta}^W, x^N) = F_{Y|x}^{-1}(\tau_N|x_N) = x_{N,i}\beta_N(\tau), for \forall \tau \epsilon (0,1) \qquad (7)$$

With the corresponding counterfactual regression at quantile level[33]:

$$\hat{q}(\hat{\beta}^W, x^N) = F_{Y|x}^{-1}(\tau_W|x_N) = x_{N,i}\beta_W(\tau), for \forall \tau \epsilon (0,1) \qquad (8)$$

---

[31] However, we are aware of possible shortcomings of this specification which will be dealt with later on in the paper.

[32] Please refer to Appendix 3 for further explanation.

[33]This procedure is thus used to simulate the counterfactual distributions for white and non-white women in 2001 and 2011 (Melly, 2005). Following Salardi (2013) the choice of the reference group has been arbitrary, making white women the base for the first two decompositions and 2001 the base for the latter two decompositions. The reason for choosing 2001 as the base year is because we are interested in isolating the effect over time of the decomposition effects, answering "what if" questions: what would have happened to wages if only the return to characteristics was set as in 2011, ceteris paribus?

The counterfactual represents the average wages if non-whites were paid according to the wage structure of whites (Salardi, 2013).

The unconditional distribution is then obtained by integrating the conditional distribution unction over a range of covariates. Inverting the unconditional distribution function thus gives us the unconditional quantiles of interest (Salardi, 2013):

$$\theta = F_Y^{-1}(\theta) = E[F_{Y|x}^{-1}(\tau|x)(Q_\theta(y|x)] = \int F_{Y|x}^{-1}(\tau|x)((Q_\theta(y|x))dF_x(x) \quad (9)$$

$F_Y^{-1}(\theta)$ is the conditional cumulative distribution of wages and $F_{Y|x}^{-1}(\tau|x)$ is the quantile function.

The wage gaps of the unconditional quantile functions between each racial group are thus denoted:

$$\hat{q}(\hat{\beta}^W, x^W) - \hat{q}(\hat{\beta}^N, x^N) = \left(\hat{q}(\hat{\beta}^N, x^N) - \hat{q}(\hat{\beta}^W, x^N)\right) + \left(\hat{q}(\hat{\beta}^W, x^N) - \hat{q}(\hat{\beta}^W, x^W)\right) \quad (10)$$

This decomposition is then performed both for 2001 and 2011, the results of which are presented in section 6.2.

In addition, performing the Melly (2005) decomposition follows the same initial procedure as Melly (2006). However the residual effect in Melly (2006) asymptotically disappears. This new decomposition adds the effect of the residual to analyzing racial wage gaps,[34] which is achieved by separating the median coefficient effect to the residual effect following Juhn et al. (1993) (Salardi, 2013).

Firstly the counterfactual unconditional wage distribution is constructed using estimates from the conditional quantile regressions based on Juhn et al. (1993) (Equation 4)[35]. As such, by taking the median as a measure of the central tendency of a distribution, the simple wage equation for each race:

$$y_i^r = x_i^r \beta^r(0.5) + u_i^r, r = White, Non - white \quad (11)$$

---

[34] This is similar to the Machado & Mata (2005) method implemented by Albrecht, Bjorklund & Vroman (2003).
[35] Taking into account that the law of iterated expectations does not apply in the case of quantile regressions.

Where $\beta^t(0.5)$ is the coefficient vector of the median regression in race r (Melly, 2005). The effects of changes in characteristics x, coefficients $\beta(0.5)$ and residual u are isolated when building each counterfactual as in equation (9). Thus the θth quantile of the counterfactual distribution of wages:

$$\hat{q}(\widehat{\beta^N}, X^W) = \inf\left\{q: \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{J}(\tau_j - \tau_{j-1})1\left(x_i^{11}\widehat{\beta^N}(\tau_j) \leq q\right) \geq \theta\right\} \quad (12)$$

Finally, the Melly (2005) decomposition including residuals becomes can be expressed as:

$$\hat{q}(\hat{\beta}^W, x^W) - \hat{q}(\hat{\beta}^N, x^N)$$
$$= \left(\hat{q}(\hat{\beta}^W, x^W) - \hat{q}(\hat{\beta}^{mW,rN}, x^W)\right) + \left(\hat{q}(\hat{\beta}^{mW,rN}, x^{11}) - \hat{q}(\hat{\beta}^N, x^W)\right)$$
$$+ \left(\hat{q}(\hat{\beta}^N, x^W) - \hat{q}(\hat{\beta}^N, x^N)\right) \quad (13)$$

The first bracket represents the effect of changes in the residuals[36]; the second the effects of changes in the median coefficients and the third the effects of changes in the distribution of covariates (Melly, B: 2005). This decomposition is thus carried out both for white and non-white female workers are also performed for 2001 and 2011; the results are presented in section 6.3.

5.2 Reassessing Assumptions

As stated above, a drawback of quantile decomposition methods in general is that they are subject to strong assumptions which do not always hold in reality (Fortin et al. 2011). In our case, out of the assumptions explained in Appendix 3, we feel four issues in our model require clarifications: selectivity, monotonicity, omitted variables and heteroscedasticity

---

[36]In order to separate the effects of the coefficients from the effects of residuals, the τth quantile of the residual distribution conditionally on x is consistently estimated by $x\left(\hat{\beta}(\tau) - \hat{\beta}(0.5)\right)$. We thus calculate the distribution that would have prevailed if the median return to characteristics had been the same as in 2011 but the residual had been distributed as in 2001.

Firstly a major concern in the decomposition literature, which ought to be taken into account in our analysis, is that of selection bias. The problem with the exogeneity assumption which prevails in our method is that it can fail; which would lead to three possible self-selection and general endogeneity problems (Fortin et al. 2011):

- White and non-white women present differing decision processes when entering the labour market. Self-selection can be based on observables, on unobservables and on bounds (Machado, 2011). This is a prevalent issue in the analysis of female participation rates, especially taking into account that only around 60% (World Bank Data, 2001-2011) of the female population participates in the workforce in Brazil. In labour models, black women have been found to be positively selected in work occupation, while white women appear to be negatively selected (Neal, 2004). We believe this source of selectivity could be problematic, yet controlling for sample selection in Melly's methodology is still being developed to date in Melly (2011).

- Selection bias can also take place if the individual can decide whether to belong to group A or B. This, in theory, does not affect results, as technically speaking nobody can choose their race. However, the self-reporting nature of PNAD can lead to selection bias because of the often ambiguous distinction between being "mixed" and being "black" or "white"[37]. Nevertheless, due to our binary classification of white and non-white, we believe that our specification is less likely to suffer from this source of bias.

- General endogeneity bias could take place due to the correlation of the covariates to the error term. This can be corrected by means of instrumental variables and has been done so for the case of Brazil in recent studies (Bourgignon et al. 2004 for example). However in the case of quantile regressions, the validity of instruments is still under question. Consequently controlling for instrumental variables is beyond the scope of this paper and instead propose different solution: On the one hand we exclude the principle of exogeneity and instead adhere to the assumption of weak exogeneity, which does allow a degree of correlation between observables and unobservables, as long as this is the same for both groups (Firpo et al, 2011). On the other hand we acknowledge the possible effect of self-selection and endogeneity in our model. One of the most cited factors which affects wage inequality and the levels of skills an individual will have, not included in our model, is occupation (Salardi (2013) or Soares (2000)). Despite the fact that the nature of our question is not based on occupation structures, we have re-run the

---

[37] Due to the social stigma of being coloured, it is believed a greater number of people consider themselves "white" when in reality they are "mixed", for example (SIS, 2012).

model controlling for occupation in Appendix 6 to determine if there indeed is a bias in our results[38]. Comparing the results in Appendix 6 with Figures 5 and 6 thus reveals the robustness of our results, as the role of all effects does not vary much by adding occupation controls. However, and as expected, the explanatory role of the Composition Effect increases slightly, particularly for the lower end of the earnings spectrum. This means that our results present a slight upward bias in the role of the Residual Effect and a slight downward bias for the Composition Effect if we do not control for occupation. Furthermore, sample selection within the quantile framework is still problematic. Albrecth, van Vuuren & Vroman (2009) developed an extension to Machado & Mata (2005) which controls for sample selection. However in view of their similar results for the effect of education and experience with and without controls, together with previous mean decompositions for Brazil which are almost unchanged with sample selection; we are confident of the uncorrected findings at the quantile level. Moreover we assume instead weak exogeneity

Secondly, the *Monotonicity Assumption* can break with measurement error in a model; meaning that the residual is the result of both this measurement error and unobserved skills (Firpo et al. 2011). Consequently it is theoretically impossible to invert the conditional quantile function and reach Equation (9) (Salardi, 2013). We are aware of possible measurement errors in our data; however we follow Melly (2005), who also commented on the possible lack of monotonicity in these models: Considering the following property of q_0, the population's θ^{th} quantile of Y:

$$q_0 = F_Y^{-1}(\theta) \iff \int 1(y \le q_0)dF_y(y) = \theta \iff \int \left(\int 1(y \le q_0)f_{Y|X}(y|x)dy \right)dF_X(x)$$

$$= \theta \iff \int \left(\int_0^1 \left(F_{Y|X}^{-1}(\tau|x) \le q_0 \right)d_\tau \right)dF_X = \theta \qquad (14)$$

This problem is overcome by changing the variable of integration, noting that $f_\tau(\tau_j) = 1$, $\forall_{\tau_j} \epsilon(0,1)$. Replacing $F_{Y|X}^{-1}(\tau_j|x_j)$ by its consistent estimate $x_i\hat{\beta}(\tau_j)$ and following the convention of taking the infimum of the set if the finite sample solution is not unique. The sample analogue of q_0 is therefore:

---

[38] As stated in the data section, four controls for occupation have been introduced in our model based on occupational category data provided in PNAD. These variables stand for employees, self-employed, domestic workers and employer. Summary statistics for the occupation structure of whites and non-whites for both years are also presented in Appendix 6.

$$\hat{q}(\widehat{\beta}, x) = \inf\left\{q: \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{J}(\tau_j - \tau_{j-1})1(x_i\hat{\beta}(\tau_j) \le q) \ge \theta\right\} \qquad (15)$$

Assuming traditional restrictions of the quantile regression model, one can thus prove that $\hat{q}$ is a consistent and asymptotically normally distributed estimator of q_0 and overcome the problem of lack of monotonicity (Melly, 2005).

Thirdly, following O'Neill (1990) the prevalent study of males in racial wage decompositions is due to the statistical bias which arises from strong changes in female labour market participation rates. However we argue this specification was true for the US in the 1990s, when this fact did indeed bias results, but it need not be so clear for Brazil today. On the one hand, the strong prevalence of the informal market means that only between 40-50% of the population is represented in formal labour market studies in Brazil (Salardi, 2013). This means that whilst female participation rate has indeed increased more than male participation rate, when analysing our data both sexes have varied composition in similar terms. We have confirmed that men's participation rates overall and in our four occupation categories have varied as much as women's participation rates across races. Therefore while we have dealt with possible biases arising from workers' occupation composition, there is no evidence to suggest that these same biases would not be present for men as well.

Finally by means of Melly (2005), we find that homoscedasticity is not fulfilled in our case (as is probably the case in many other country studies). We can prove this by running quantile regressions for the difference between the 90th and 10th percentiles which have been plotted for Table 4 below[39]. This table provides evidence of the effect on the dispersion of wages of the explanatory covariates in the quantile regressions. If the error term is independent of a characteristic then the coefficient of its fitted value should not be significantly different from zero. Moreover we can also infer the effect each of the characteristics modelled has on WGI: if the 90th-10th centile difference is positive (negative), then a higher value increases (decreases) WGI (Melly, 2005).

---

[39] Note that we have only reported education instead of education and education squared. This has been done so as not to capture the negative effect of low levels of education portrayed in Table 5.

Table 4. Differential Regression Results

	Differential Quantile Regression Results			
	White 2001	Non-white 2001	White 2011	Non-white 2011
Education	0.0412***	0.0121***	0.0576***	0.0101***
	(0.00378)	(0.00308)	(0.00345)	(0.00268)
Experience	0.0203***	0.00945***	0.0365***	0.0131***
	(0.00260)	(0.00305)	(0.00317)	(0.00336)
Experience Sq	-5.63e-05	8.49e-06	-0.000308***	2.42e-05
	(5.39e-05)	(5.43e-05)	(7.11e-05)	(7.22e-05)
Constant	0.982***	1.417***	0.412***	1.147***
	(0.0636)	(0.0461)	(0.0543)	(0.0543)
Observations	27,059	21,998	26,709	27,114

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculations from PNAD 2001 & 2011

Consequently we show that characteristics do affect WGI, except for Experience Squared in three out of four cases and Education for White 2001. We can thus infer that the error term is not independent from the covariates and actually changes according to most explanatory variables in our framework. Due to this composition effect on the residual, methods which do not take into account the dependence between characteristics and residuals underestimate the effect of the former and overestimate the effect of the latter (Melly, 2005). Therefore the higher the level of education, the higher the variances of wages and thus the greater WGI. Experience follows a similar pattern until a certain value level, after which the effect becomes insignificant or even negative.

Part VI. Results

The results for the Melly (2005, 2006) decomposition are divided into three subsections below: Section 64.1 presents results for the median coefficients of the quantile regressions. Section 6.2 summarizes the results for the Melly (2006) racial wage gap decomposition both in 2001 and 2011. Finally section 6.3 discloses the results for the Melly (2005) decomposition. These results are presented on tables as well as graphs, plotting the decomposition results at 99 different quantiles (θ =0.01, 0.02, 0.03…0.99):

6.1 Quantile Regressions

We have firstly plotted the results for the Koenker & Bassett (1978) quantile regressions (please refer to Figures 8 & 9 in Appendix 4). Furthermore we have presented below the results for the 50th centile in Table 5[40]. Stars indicate that results are significantly different from zero[41]:

Table 5. Quantile Regression Results

	Median Quantile Regression Results			
	White 2001	**Non-white 2001**	**White 2011**	**Non-white 2011**
Education	-0.0379***	-0.0177***	-0.115***	-0.0753***
	(0.00645)	(0.00494)	(0.00501)	(0.00388)
Education Sq	0.00982***	0.00845***	0.0118***	0.00901***
	(0.000325)	(0.000285)	(0.000242)	(0.000220)
Experience	0.0411***	0.0415***	0.0275***	0.0233***
	(0.00110)	(0.00136)	(0.000948)	(0.000916)
Experience Sq	-0.000521***	-0.000507***	-0.000337***	-0.000273***
	(2.33e-05)	(2.67e-05)	(2.08e-05)	(2.01e-05)
Constant	-0.0842**	-0.384***	0.751***	0.552***
	(0.0337)	(0.0264)	(0.0282)	(0.0228)
Observations	27,115	22,113	26,73	27,168

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculations from PNAD 2001 & 2011

The coefficients of these median regressions provide evidence of how the level of wage depends on the explanatory variables in the quantile regressions (Melly, 2005). Otherwise the results for the median regressions reveal generally intuitive results taking in mind the stylized facts presented in previous sections.

Years of study presents non-linear returns to education, negative at the smallest number of years of study, which then become positive the higher the years of education. Looking back at Figure 3, for those with less than *Ensino Fundamental,* returns to years of study were low, quite stable

---

[40]Melly (2005) using Chernozhukov & Hong (2002) 3-step censored quantile regressions due to income top-coding on US data. However, Brazilian data is not top coded and we thus feel unnecessary when using PNAD data. Instead we decided to run Koenker & Basser (1978) quantile regressions as stated in the methodology (Velez, Barros & Ferreira, 2008).

[41] Standard errors are estimated by bootstrapping 100 times, which estimates the distribution of $\hat{\beta}(\theta)$ consistently (Melly, 2005).

and fluctuating around a given value. This phenomenon has been previously encountered in wage studies for Brazil, especially in urban labour markets, where there is a strong demand for qualified labour. Those who do not possess at least *Ensino Fundamental* have thus very limited formal employment options and are most likely exempt from the formal labour remuneration structures presented in this model (Stevens & Weale, 2003). Moreover the positive squared value for education captures increasing returns to education for those with higher levels of education. Finally unsurprisingly, returns to education have increased in 2011 with respect to 2001 and were also higher for white than for non-white sectors of the population throughout both years.

Experience also has positive albeit diminishing returns across the two years of study for the 50th centile. The variable Experience Square provides very low returns compared to Experience for all sectors of the population in both years for example. These are smaller in 2011 than in 2001 and also lower for White than for Non-white female workers. Furthermore higher years of experience, as proxied by Experience Squared, yield negative returns for all races across both years, which is explained by the inverted U shaped returns to education graphed in previous sections. This result also coincides with previous findings in the literature by Ferreira, F. & Barros, R. (1999), who explain the fact that from a certain age, irrespective of your experience, salaries no longer rise in the same proportion as when a worker is young[42].

Finally, results for all quantiles across the Quantile Regressions are graphically presented in Figures 8 & 9 of Appendix 4. These results capture differing trends across races and years which are not depicted in the median regression.

On the one hand, Education presents and L-shaped decreasing returns to education pattern across quantiles for white workers in 2011 and non-white workers in 2001 and 2011. Indeed white workers in 2011 follow a smooth, downward sloping curve across quantiles which flattens out from the 70th centile onwards. However non-white workers experience a common peak in returns to education around the 10th centile, followed by falling wage trends between the 20th and 40th centiles which then smooth out until around the 90th centile. From the 90th centile onwards, non-white workers earning above the 90th centile in 2001 then decrease returns dramatically; whilst those non-white workers located in the same wage percentile during 2011 slightly increase. Contrasting these results, white workers in 2001 experience a different pattern to all other sectors, with U-shaped returns to education. For this sector of the population, the lowest returns to education are experienced by the 20th centile, whilst the highest returns are

---

[42] It must be taken into account that Experience was built on age, so we are thus also capturing effects of age on its returns to experience.

obtained at the bottom and top of the earnings distribution. Moreover, just as in the median quantile regression results, returns to education are higher in 2011 than in 2001.

On the other hand, returns to experience follow quite different patterns for white and non-white workers. White workers experience increasing returns to experience from the 20th centile onwards and fluctuating returns beforehand. Whereas non-white workers present U-shaped returns to experience. From Figure 9 in Appendix 4 this pattern is more clearly discernible in 2011: with the bottom peak located at the 20th centile. However in 2001, the U-shape is smoother across the entire wage distribution: starting at with a peak at the 20th centile and followed by a W shaped pattern with a slight rise around the 40th centile. Finally just as in the median results, returns are lower for 2011 than for 2001.

6.2 Decomposition of Racial Wage Gap by Year

In order to understand how and why racial wage gaps have decreased over time across the whole earnings distribution we have carried out the decomposition proposed by Melly (2006), developed in section V. This decomposition, which is very similar to Machado & Mata (2005), allows disentangling the aggregate contribution of differences in characteristics (the explained component) and differences in returns to those characteristics (the unexplained component or wage structure) on racial wage gaps by quantile (Salardi, 2013). Note that the residual component asymptotically disappears, whereas it is still present when we implement the decomposition of the unconditional quantile wage gap using the Melly (2005) (Salardi, 2013). Figure 5 graphically depicts the results presented in Table 1 of Appendix 5 for the Melly (2006) decomposition, whilst Table 6 reports the decomposition of indices of inequality for racial wage gaps.

Figure 5. Decomposition of Differences in the Distribution
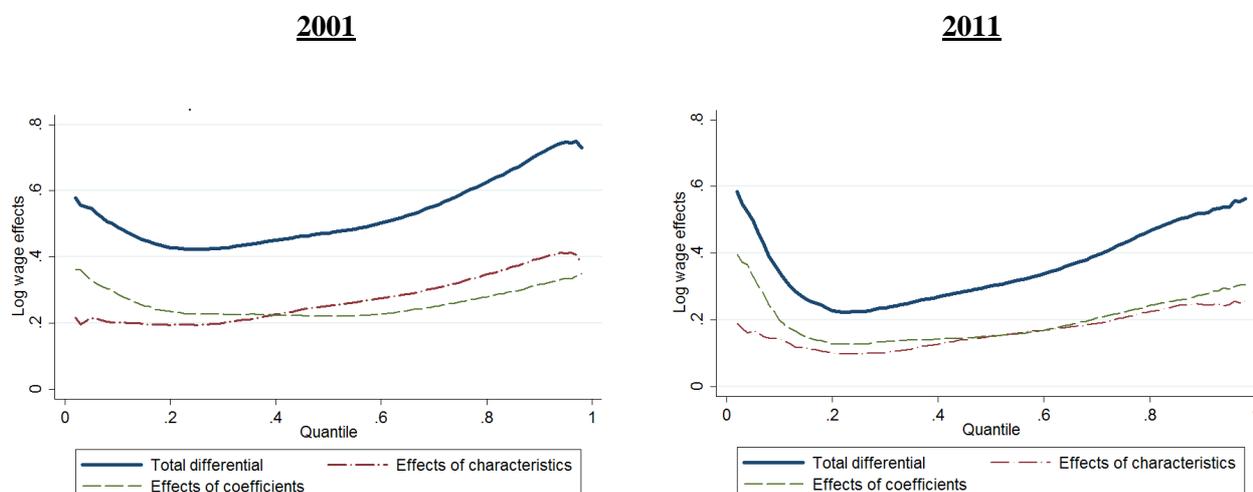
2001   **2011**

Source: Author's calculations from PNAD 2001 & 2011

Table 6 Decomposition of Inequality Measures

	2001						2011					
Statistic	Total Change		Effects of Coefficients		Charactertitics		Total Change		Effects of Coefficients		Charactertitics	
90-10	21,00	(0,018)	2,00	(0,019)	19,00	(0,013)	17,70	(0,059)	5,50	(0,044)	12,20	(0,043)
	100%		10%		90%		100%		31%		69%	
50-10	-2,00	(0,12)	-7,00	(0,012)	5,00	(0,006)	-4,00	(0,043)	-5,00	(0,036)	1,00	(0,018)
	100%		350%		-250%		100%		125%		-25%	
90-50	24,00	(0,013)	10,00	(0,0167)	14,00	(0,011)	21,50	(0,041)	10,50	(0,037)	11,00	(0,04)
	100%		42%		58%		100%		49%		51%	

Note: All numbers have been multiplied by 100. Negative Values in Percentage Change Reveal a Decrease in inequality measures whereas positive results reveal an increase in inequality measures

The percentage weight of each effect is presented each variable.

Bootstrap standard errors with 100 replications in parenthesis.

Source: Authors calculations from PNAD 2001 & 2011

The results in Figure 5 present the observed and simulated racial wage gaps across the wage distribution in 2001 on the left graph and in 2011 on the right graph. The most striking common result across both years is a U-shaped total racial differential log wage effect across earnings. This takes place as a result of total wage gaps being smallest at the centre compared to the ends of the wage distribution; these findings confirm previous findings in the literature for urban Brazil by Salardi (2013), Salardi (2009) and Ñopo(2012). Numerically this can be perceived in Table 12, Appendix 5: in 2001 white women workers for the 10th centile on average earned 50% more than non-white female workers, 48% more in the 50th centile and 72% more in the 90th centile of the wage distribution. These effects diminish in 2011, where white women earned

34% more in the 10th centile, 30% more in the 50th centile and 52% more in the 90th centile. Moreover, despite the fact that wage differentials overall did decrease across time (Table 6 reveals a fall in 90-10 percentile ratio from 21% to 17.70%); Figure 5 also reveals an apprehensively deeper U-shaped racial wage gap distribution in 2011 compared to 2001. The reason for this effect across time results from the greatest change in racial wage differentials between 2001 and 2011 being experienced across the middle of the wage distribution: as seen in Table 12 in Appendix 5, wage gaps for the 50th centile decreased from 48% differential to 30% in a decade. Conversely the smallest change occurs at the top end, where wage racial gaps decrease from 72% to 52%, emphasizing the maintenance of glass ceiling effects during our period of study. Prevalence of glass ceiling effects are also present in Table 6 by diminishing, albeit positive 90-50 percentile ratios: 24% and 21.50% for 2001 and 2011 respectively. It is the maintenance of wage differentials within the lowest 10th centile which is worrying, as it underscores the appearance of sticky floor effects amongst non-white working women. . As confirmed by Figure 5 and our decomposition results, wage inequality by race for the first three centiles actually does not vary across years and remains around 57% racial wage differentials. Likewise, Table 6 reveals the rise of sticky floor effects by means of negative 50-10 percentile ratios, which increase from 0.02 to 0.04 between 2001 and 2011. Following Salardi (2013), who discusses the existence of sticky floor effects for gender wage gaps, we believe this result is likely captured in our results too. We can thus infer that, despite the fact that overall inequality has fallen, mostly due to the reduction in the median of the distribution, black women, particularly at the lowest wage levels, are not benefiting as much from these improvements: a marginalisation of the lowest sectors, as described by Ferreira & Barros (1999).

The combination of Composition and Price Effects which compile the variation of observed wages also follow similar previous findings in the literature. Salardi (2013) finds a strong Composition Effect on the total wage differential across races and a strong Price Effect on total wage gaps across gender. However, due to the fact that we are capturing both a gender and racial discrimination effect in our model, we find a combination of both these results for racial gaps amongst women: as seen in Table 12 of Appendix 5 characteristics in 2001 play a smaller role amongst lower earnings, which then becomes predominant at the top end of the wage distribution: Their role composes 42% of observed wage gaps for the 10th centile and 56% for the 90th centile in 2001. Moreover from Table 6 we can confirm that it has been the skill composition of workers which plays the greatest role in wage gap inequalities across the top half and whole wage distribution as seen in 90-50 and 90-10 percentile ratios; whilst as perceived from 50-10 percentile ratios, coefficients play a greater role in the lowest paid quantiles. Quantity and quality of education are thus important in the labour market amongst women, particularly at the top 50% of the distribution (Lemieux, 2006). Results which are also

reached by Salardi (2013) using PNAD 2006 data and Bartalotti & Leme (2007) using PNAD 2004 data. Notwithstanding by 2011, the value of workers' skill composition diminish below the incidence of the Price Effect for both the top and bottom end of the wage spectrum. From Table 12 in Appendix 5, characteristics comprise 41% of total wage gaps at the 10th centile and 48% at the 90th centile. However by 2011, coefficients become predominant at both the ends of the wage distribution. In Table 1 of Appendix 5 and Table 6 we can confirm the greater weight of coefficients in determining wage gaps and inequality measures. Consequently, the wage structure has been important in the development of higher sticky floor effects: In Table 6, we can perceive that coefficients play a much larger role in determining 50-10 percentile ratio than characteristics, particularly so for 2011. Taking into account that wages have a positive relation with education and experience in our data, we can thus confirm that discrimination in the labour market is indeed significant for determining wage structures across those with the lowest level of skills, and increasingly so for those with the highest skill level.

6.3 Extended Decomposition Results

In contrast with the results in section 6.2, figure 6 and Table 1 of Appendix 5 reveal the results for Melly (2005) decomposition. Table 7 presents the decomposed percentile ratios for wage gaps across the earnings distribution just as in Table 6 in the previous section. In comparison to Melly (2006), this method now takes into account the effects of residuals as a component to total wage variation. Consequently, we distinguish between the effect of the wage structure and returns to observed skills, as suggested by Juhn et al. (1993), Lemieux (2002) and Autor et al. (2005).

Figure 6. Results of the Melly (2005) Racial Gap Decomposition
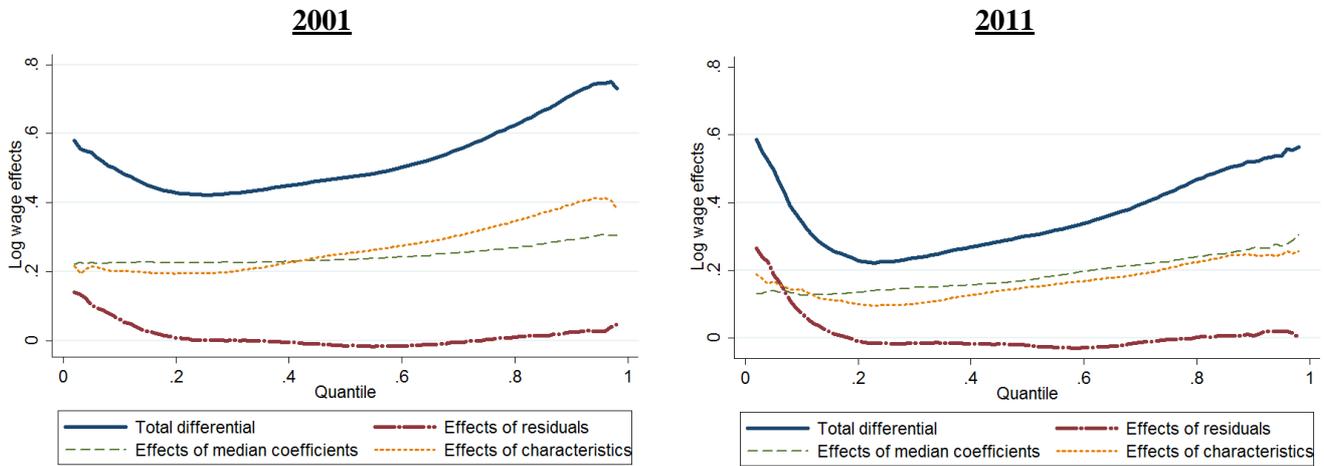
2001 **2011**

Source: Author's calculations from PNAD 2001 & 2011

Table 7. Decomposed Inequality Measures

	Decomposition of Inequality Indexes															
	2001								2011							
	Total		Residual		Coefficient		Characteristics		Total		Residual		Coefficient		Characteristics	
90/10	0,22	(0,017)	-0,04	(0,021)	0,07	(0,016)	0,19	(0,015)	0,18	(0,018)	-0,06	(0,015)	0,14	(0,021)	0,10	(0,016)
	100%		-17%		30%		87%		100%		-33%		77%		56%	
50/10	-0,02	(0,009)	-0,08	(0,011)	0,01	(0,008)	0,05	(0,006)	-0,04	(0,010)	-0,09	(0,006)	0,04	(0,011)	0,01	(0,008)
	100%		434%		-54%		-280%		100%		225%		-100%		-25%	
90/50	0,24	(0,012)	0,04	(0,015)	0,06	(0,011)	0,14	(0,011)	0,22	(0,012)	0,03	(0,011)	0,10	(0,015)	0,09	(0,011)
	100%		17%		24%		60%		100%		13%		43%		43%	

Note: Negative Values in Percentage Change Reveal a Decrease in inequality measures whereas positive results reveal an increase in inequality measures

The percentage weight of each effect is presented each variable

Source: Authors calculations from PNAD 2001 & 2011

From Figure 6 and Table 12 in Appendix 5 we can discern that the Composition Effect across the distribution of wages remains virtually the same as in the previous section: This effect is prevalent in explaining wage gaps across the entire wage distribution in 2001. The greater incidence of the Composition Effect on the overall wage distribution inequality (90-10 percentile ratio) was also found by Melly (2005) and Salardi (2013). However, the Price Effect overtakes the Composition Effect and, by 2011, it had become prevalent in racial wage gaps for the top end of the earnings distribution. Residuals remain marginal during the whole period. From Table 12 in Appendix 5, we reveal that characteristics decrease from 55% to 47% of total wage gaps at the 90th centile between 2001 and 2011; whereas coefficients increase their role from 41% to 51%. However at the bottom end of the distribution, the role of the Composition and Price Effects changes with respect to Melly (2006) due to the inclusion of the Residual Effect. In comparison to the previous results, whilst the Price Effect was predominant at the lower end of the distribution in 2001 this effect reduces considerably by 2011. At the bottom

10th centile of the wage distribution, the Price Effect evolved from composing 47% of total wage gaps in 2001 to 37% in 2011. It has thus been a combination of the maintenance of the Composition Effect and rise of the Residual which have become the main effects in the growth of sticky floors: From Table 12 in Appendix 5 we confirm that the Composition Effect remains explaining 41% of total wage gaps throughout our period of study. Conversely the Residual Effect increases from 12% to 22%. These results contrast with Foguel & Acevedo (2006), who found that residuals decreased across time, for the period 2001-2005.

Residuals in particular, although exerting a very modest dimension compared to the Price and Composition Effect, presents varying trends throughout the distribution. The Residual Effect is positive at the bottom and top end of the wage distribution but became negative around the centre of the wage distribution. In Table 12, Appendix 5, residuals present an increasing role at the bottom 10th centile of the wage distribution, contributing to the rise of sticky floor effects: from 12% to 22% of observed wage gaps as seen in Table 12 in Appendix 5. However the role of the Residual Effect at the median opposes that of both the Composition and Price Effects: -2% of racial wage gaps at the 50th centile in 2001 and -8% in 2011. These results seem to indicate that the Residual Effect is a component of the disproportionate fall in racial wage gaps during this time in the median. The strong effect of the residual on the bottom 10% of wage earners is also present in Table 7, where residuals amount to 434% of the total 50-10 percentile ratio in 2001, and to 225% in 2011. Moreover, negative and increasing values for the Residual Effect in 90-10 percentile ratio decomposition (from -0.04 to -0.06) would imply that this effect is higher across time for the $10^{th,}$ comparatively to the 90th centiles, across earnings. Therefore, in comparison to Melly (2006), it seems that it is the Residual Effect and not the Price Effect which accounts for a greater part of the more pronounced U-shaped pattern across racial wage gaps. Differentiating between returns to observed prices and wage structure effects is indeed important and ought to be taken into account in future research (Melly, 2005).

6.4 Robustness Checks

Finally, following Melly (2005) we have carried out robustness checks for our model. Firstly, we have run the model with 20, 40 and 100 bootstrap replications, all of which have reached similar results to those reported in previous sections. Moreover, as the decomposition framework is sensitive to the order of explanatory variables in our model, we have run each decomposition again, but rearranging Education and Experience. Results specified with Experience first are presented in Appendix 7 and yield identical findings to our original specification. On a final note, despite the fact that our decomposition results are built on fitted values for log hourly wages; the robustness of our results is also perceived when comparing our

fitted values in Figures 5 and 6 to raw wage differentials present in the data in Figure 1. The bottom quantile and the top 3 deciles experienced much racial wage differences both in 2001 and 2011 (around 40% compared to an average 23% for the middle quantiles). These middle quantiles also experienced the greatest decline in wage differentials compared to the ends (the 5th quantile decreased from 37% to 25% compared much smaller changes in the tails of the wage distribution). Therefore, the checks performed make us confident of the results presented in this paper.

VII Conclusion & Policy Implications

Throughout this article we have analysed the reduction in women's racial wage gaps across the earnings distribution, and how changes in the skill composition of the female workforce in urban Brazil contributed to this phenomenon between 2001 and 2011. In so doing, we have provided a truly original contribution by focusing on women, a gap in the literature which is likely to impact present and future generations (Hill & King, 2010) (Gradín, 2007)). In addition, we have extended wage decomposition frameworks applied to the country by using Melly (2005), discussing the impact of heteroscedasticy and updating wage gaps to 2011.

Our results thus confirm past trends in the literature but also reveal concerning new wage gap patterns and their potential sources. In the course of our analysis, we have confirmed falling U-shaped racial wage gap distributions mostly due to the fact that improvements in racial disparities have been greater around the median. Glass ceiling effects have fallen, but still prevailed between 2001 and 2011, together with the development of sticky floor effects (Please refer to Figures 5 & 6). Our findings coincide with previous research in the country carried out by Salardi (2013) and Santos & Ribeiro (2006), relative to the persistence, albeit lowering, of glass ceiling effects amongst women. However, the original contribution of this paper is the finding of growing sticky floor effects across time. By means of Melly (2005) we have determined them to be related to increased skills in the form of the Composition Effect and to unobserved economic wage structures captured in the Residual Effect. Therefore as predicted in the introduction, higher skill composition has indeed played strong and varying role in determining racial wage gaps along the distribution of female wages:

On the one hand, as seen in Figure 5, Melly´s (2006) decomposition reveals that the Composition Effect has played a greater role at the top end of the wage distribution while the Price Effect has become predominant at the bottom end, in 2001. However, by 2011, the Price Effect became the driving force behind both glass ceilings and sticky floors (please refer to Table 12 in Appendix 5). We thus confirm that the soaring increase in tertiary education has

reduced the Composition Effect of glass ceilings. Instead, by 2011 it was the greater returns to education which became the major factor at the top and bottom levels of the wage distribution.

On the other hand, extending our analysis to Melly (2005) reveals a different and more complete insight to the reasons behind changes in racial wage gaps amongst working women in urban Brazil (please refer to Figure 6). Whilst the Composition Effect still played a predominant role in glass ceiling effects in 2001; it diminished with time and became the predominant factor in sticky floors by 2011. Consequently, positive externalities of higher levels of education and experience have benefited the upper wage sectors and, conjunctionally, bypassed the lowest wage earners; this trend seems to be a common occurrence in Brazilian public policy, as perceived by Ferreira & Barros (1999). Furthermore, the delineation of returns to observed skills and wage structure effects in our model reveals the Residual Effect and not the Price Effect to have had an increasing role in sticky floor effects during our period of analysis. This is clear in Table 12 in Appendix 5, where the Residual Effect presents higher negative 50-10 percentile ratios for both years: -0.08 for 2001 and -0.09 for 2011; whereas the Price Effect shows positive values: 0.01 in 2001n and 0.04 in 2011. In contrast to this finding, we have also found that the Residual Effect has also contributed to lowering wage gaps around the median of the wage distribution in 2011. As opposed to the Composition and Price Effects, the residuals present negative values around the 50th centile, accounting for -2% and -8% of the total variation in racial wage gaps for 2001 and 2011, respectively (please refer to Table 1 Appendix 5). These results thus point towards the fact that certain institutional factors and unobserved skills are strongly benefitting female workers in the median earnings distribution, whilst at the same time marginalising those below the lowest earnings decile. Moreover, they also reinforce findings from Section III that WGI is also an important factor which determines overall wage inequality and should not be omitted from these types of analysis.

Beyond these results, as discussed in the introduction it is not possible to directly discern what unobserved factors compose the Residual Effect due to the aggregate nature of our model and because of the heteroscedasticity of the residual inequality[43]. We do know however that changes in the skill composition of workers positively affected WGI. Therefore as revealed in Table 4, we can determine that greater education and experience alone made residual inequality higher; which contributed to greater overall racial wage inequality during this time. This partly explains the positive Residual Effect on sticky floors; however other factors are clearly counteracting the residual's outcome across the rest of the wage distribution. Therefore by means of the

---

[43] Autor et al. (2005), Lemieux (2002) and Autor et al. (2008) all propose residual decomposition methods based on a combination of Di Nardo et al. and Juhn et al . (1993) decomposition frameworks. However they rely on the unrealistic assumption of residual inequality homoscedasticity.

theoretical framework and empirical evidence presented in Section III, we infer possible explanations behind these changes:

Firstly, the literature has highlighted the role of unobserved variables affecting wage dispersion in our model. On the one hand, by running an extended model including occupation categories, we confirmed that the omission of this variable is slightly downward biasing the Composition Effect. However, we conclude that this difference is not very significant, as our results have remained almost identical across both models (see Figure 3 in Appendix 6, and Figures 5 & 6). On the other hand, unobserved variables in terms of family background or school quality could also be affecting our results. Bourguignon et al. (2004) and Ferreira, (2000) have highlighted the role of parent's education and quality of schooling as major explanatory factors for wage inequality at the bottom of the wage distribution. Following this evidence, we assume that these unobserved factors are very likely to be impacting on our results for the Residual Effect.

Secondly, the literature has highlighted the important effect of institutional factors in the Residual Effect, particularly concerning six issues: antidiscrimination institutions, minimum wage, trade unions, open trade, skill biased technological change and economic growth (Lemieux, 2006). It has previously been outlined that the very low trade union membership, the limited trade policy variations, patterns of returns to education in Figure 3 and the unresponsiveness of wage structures to short-run economic changes (Pavcnik et al (2002) and Arbache (2001)); render trade unions, trade liberalization, skill-biased technological change and economic growth marginal explanations for the patterns of the Residual Effect. However, the important institutional changes undertaken to tackle discrimination in the Brazilian labour market, with the creation of organizations such as SEPPIR or the instalment of university quotas, are most likely having a strong effect on wage gaps and skill differentials. This is particularly true at the top end of the earnings distribution, as university students have been targeted by these policies (Fredman, 2012). Moreover, the increase in minimum wage during our period of study has surely had a powerful impact across the middle and bottom of the wage distribution(ILO, 2012). In Table 11, Appendix 2, we reveal that the first decile of white workers and the first two deciles of non-white workers earn below the minimum wage and are thus marginalised from its effect. Minimum wage increases can thus explain why the bottom 10% of workers are being left out of wage gap improvements, but those located in the centre of the distribution have disproportionately improved.

Thirdly there may also be some measurement error in our data. As noted in section V, PNAD is not without fault because of the self-reporting nature of many questions, the lack of panel data and the sample selection bias due to 50% informality (Da Mata et al. 2001). However, to date

there is no other database as complete as PNAD, and these same shortcomings are faced by all micro-data based studies (Deaton, 1997).

All things considered, in answer to our original question we confirm that the increase in workers' skills has indeed positively contributed to lower wage gaps at the top and middle of the wage distribution between 2001 and 2011. However, it has also had a negative effect amongst the lowest wage earners in the female workforce with the rise of sticky floor effects. Contrastingly, changing returns to education have also played an important role in the lowering of racial wage gaps around the median of the wage distribution. Although they have had a predominant effect in the maintenance of glass ceiling effects by 2011. Together with greater education, we have highlighted unobserved skills, minimum wage and antidiscrimination measures as the most likely unobserved effects prevailing in sticky floors. Therefore although wage gaps have overall decreased, the prevalence of glass ceilings and the growth of sticky floors render the racial differential distribution more heterogeneous. Policy implications for these results highlight first and foremost the need to target the workforce located at the lowest levels of earnings. Moreover there must be further work towards reducing the large informal market in the country which marginalizes its workers from any positive labour policy developments. Finally, glass ceiling effects must continue to be reduced by means not only of higher skill levels but also by changes in the remuneration structure for those skills. Government controls of these returns to skill could promote a more equal payment structure both for white and non-white workers. Future research ought to engage with the factors that we have suggested to be behind these changes in racial wage gaps, and also further control for sample selection and omitted variable bias. This will provide a more accurate and complete understanding of what determines wage discrimination for women in the workforce.

References

Acemoglu, D. (2003). Patterns of Skill Premia, *The Review of Economic Studies*, Wiley Blackwell, Vol. 70, Issue 2, pp.199-230.

Albrecht, J., A. Bjorklund, and S. Vroman (2003, January). Is there a glass ceiling in Sweden? Journal of Labor Economics Vol.21, Issue 1, pp.145-177.

Arbache, J. (2001). Wage Differentials in Brazil: Theory and Evidence. Journal *of Development Studies* Vol.38 Issue 2.

Arbache, J. (2002). Unions and Wage Inequality in Brazil in Unions *and Labor Market in Brazil*. IPEA/The World Bank Brazil Jobs Report.

Arias, O. Yamada, G. & Tejerina, L. (2004). Education, Family Background and Racial Earnings Inequality in Brazil, *International Journal of Manpower*, Vol.25, Issue 3, pp. 355 – 374.

Armington, C. & Acs, Z. (2004). Job Creation and Persistence in Services and Manufacturing, *Journal of Evolutionary Economics*, Springer, vol. Vol.14, Issue 3, pp.309-325.

Arulamplam, W. Booth, A. & Bryan, M. (2006). Is There a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution. *Industrial and Labor Relations Review* Vol. 62, No.2, pp.163-186.

Atal, J. Ñopo, H. & Winder, N. (2009). New Century, Old Disparities: Gender and Ethnic Wage Gaps in Latin America, RES Working Paper 4640, Inter-American Development Bank.

Autor D. Katz L. & Kearney M. (2008). Trends in U.S. Wage Inequality: Revising the Revisionists, *Review of Economics and Statistics,* Vol.90,No.2, pp. 300–323.

Autor, D. Katz, L. & Kearney, M. (2005). Rising Wage Inequality: the Role of Composition and Prices. NBER Working Paper No. 11628.

Bacolod, M. Blum, B. S. & Strange, W. (2009). Skills in the City. *Journal of Urban Economics* Vol.65, pp.136-153.

Bailey, S. Loveman, M. & Muniz, J. (2012). Measures of ''Race'' and the Analysis of Racial Inequality in Brazil. *Social Science Research* Vol.42, Issue 1, pp.106–119.

Barros, R. & Mendonca, R. (1995). Os Determinantes da Desigualdade no Brasil. Texto para Discussao, no 377, Rio de Janeiro, Instituto de Pesquisa Economica Aplicada.

Bartalotti, O. & Leme, M. (2007). Discriminação Salarial por Cor e Gênero Revisitada: Uma abordagem de decomposição contrafactual utilizando regressões quantílicas, Mimeo.

Becker, G. (1957). The Economics of Discrimination. Chicago: University of Chicago Press.

Blinder, A. (1973). Wage Discrimination: Reduced Form and Structural Variables. *Journal of Human Resources*, Vol.8.

Bourguignon, F. Ferreira, F. & Lustig, N. (1998). The Microeconomics of Income Distribution Dynamics in East Asia and Latin America. World Bank-IADB Research Proposal

Bourguignon, F. Ferreira, F. & Menendez, M. (2004). Inequality of Outcomes, Inequality of Opportunities, and Intergenerational Education Mobility in Brazil, in World Bank, ed. *Inequality and Economic Development in Brazil*. Washington, D.C.

Campante, F. Crespo, A. &Leite, P. (2004). Desigualdade Salarial Entre Raças No Mercado de Trabalho Urbano Brasileiro: Aspectos Regionais, *Revista Brasileira de Economia*, Vol.58, No.2, pp.185-210.

Chernozhukov, V. & Hong, H. (2002). Three-step Censored Quantile Regressions and Extramarital Affairs, *Journal of the American Statistical Association* Vol. 97, pp.872-783.

Castilho, M. Menéndez, M. & Sztulman, A. (2012). Trade Liberalization, Inequality and Poverty in Brazilian States. World Development, Vol.40, Issue 4, pp.821–835

Census (2010). http://censo2010.ibge.gov.br/en/

Cunningham, W. & Jacobsen, J. (2008). Earnings Inequality Within and Across Gender, Racial, and Ethnic Groups in Latin America, Policy Research Working Paper No. 4591 World Bank.

Da Mata, D. Deichmann, U. Henderson, J. Lall, S. & Wang, H. (2007). Determinants of city growth in Brazil, *Journal of Urban Economics* Vol.62, pp.252-272.

Deaton, A. (1997). The Analysis of Household Surveys: A Microeconometric Approach to Development Policy. World Bank Publications.

Di Nardo, J. Fortin, N. & Lemieux, T. (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semi-parametric Approach. *Econometrica*, Vol.64 Issue 5. pp. 1001-44.

Dickson, M. & Harmon, C. (2011). Economic returns to education: What We Know, What We Don't Know, and Where We Are Going—Some brief pointers. *Economics of Education Review* Vol. 30, pp.1118– 1122.

Fernandes, F. (1978). A Integraçao do Negro na Sociedade de Clases. Sao Paulo, Ática.

Ferreira, F. & Barros, R. (1999). Slippery the Slope: Explaining the Increase in Extreme Poverty in Urban Brazil: 1976-1996. *Texto para Discussão* No.404, Departamento de Economia, PontifíciaUniversidade Católica do Rio de Janeiro.

Ferreira, F. (2000). Education for the masses? The interaction between wealth, educational and political inequalities, PUC. Processed.

Ferreira, F. Lanjouw, P. &Neri, M. (2003). A Robust Poverty Profile for Brazil Using Multiple Data Sources. *RevistaBrasileria de Economia* Vol.57, Issue 1, pp.59-92.

Florida, R. Mellander, C. Stolaricl, K. & Ross, A. (2012). Cities, skills and wages. Journal of *Economic Geography* Vol.12, pp.355–377.

Foguel, M. & Acevedo, J. (2006). Uma Decomposiçao da Desigualdade de Rendimentos do Trabalho no Brasil: 1995-2005. in Barrs, R.P. de. Foguel, M. N. Ulyssea, G. (Orgs). *Desigualdade de Renda no Brasil: Uma Análise da Queda Recente*, Rio de Janeiro.

Fortin, N. M., Lemieux, T., and Firpo, S. (2010). Decomposition Methods in Economics, in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Vol. 4 A, Elsevier B.V., pp. 1-102.

Fredman, (2012). Anti-Discrimination Laws and Work in the Developing World: A Thematic Overview. World Bank, Washington, DC.

Garcia, L. Ñopo, H. &Salardi, P. (2009). Gender and Racial Wage Gaps in Brazil 1996-2006: Evidence Using a Matching Comparisons Approach, Research Department Publications No.4626, Inter-American Development Bank, Research Department.

Garcia, V. (2005). Questoes de Raça e GênerpnaDesigualdade Social Brasileira Recente. Dissertaçao de Mestrado- Universidade Estadual de Campinas. Instituto de Economia

Giles, L. (2012). Metropolitan Urban Poverty and Inequality: Explaining Brazil's 21st Century Economic Miracle. Unpublished, Masters Thesis Universidad Carlos III Madrid, History of Economics.

Glaeser, E. Scheinkman, J. & Shleifer, A. (1995). Economic Growth in a Cross-Section of Cities, WorkingPapers No.e-95-4, Hoover Institution, Stanford University.

Gradín, C. (2007). Why Is Poverty So High Among Afro-Brazilians? A Decomposition Analysis of the Racial Poverty Gap. IZA Discussion Paper No. 2809.

Hansen, B. (2000). Sample Splitting and Threshold Estimation. *Econometrica* Vol. 68, No.3, pp.575-603.

Hanushek, E. & Woessmann, L. (2010). The Economics of International Differences in Educational Achievement. IZA Discussion Paper No. 4925.

Hasenbalg, C. (1979).Discriminaçao e Desigualdades Raciais no Brasil. Ediçoes Graal, Rio de Janeiro.

Henriques, R. (2001). Desigualdade Racial no Brasil. Evoluçao das Conióes de Vida na Década de 90. IPEA Texto para Discussao, No. 807.

Heringer, R. (2002). Racial inequalities in Brazil: a Synthesis of Social Indicators and Challenges for Public Policies. *Cadernos de Saude Publica* No18, S57–S65.

Hill, A. & King, E. (2010). Women's education and economic well-being. *Feminist Economics,* Vol.1, Issue 2.

Holanda-Filho, F. & Pessõa, S. (2008). Retorno na Educaçao no Brasil, *Pesquisa e Planejamento Econômico*. Instituto de Pesquisa Econômica Aplicada - IPEA.

Instituto Pesquisa de Economía Aplicada (IPEA) (2001, 2011). http://www.ipeadata.gov.br

International Labor Organization (2009). Decent Work Country Profile Brazil. Geneva and Brasilia: Vol. 15, pp.57.

International Labor Organization (ILO) (2012). Justicia social y crecimiento: el papel del salario mínimo. *Boletín Internacional de Investigación Sindical*. Vol.4, No.1.

Johnson, P. (1993). Aging and European economic demography, in: Johnson, P. and K.F. Zimmermann (eds.) *Labor markets in an ageing Europe*, Cambridge University Press, Cambridge.

Juhn, C. Murphy, K. & Pierce, B. (1991). Accounting for the Slowdown in Black-White Wage Convergence in M. Kosters (eds): *Workers and Their Wages: Changing Patterns in the United States*. AEI Studies, no. 520 Washington, D.C.: AEI Press, pp.107-43.

Juhn, C. Murphy, K. & Pierce, B. (1993): Wage Inequality and the Rise in Returns to Skill. *Journal of Political Economy*, Vol.101, Issue 3, pp.410-42.

Katz, L. & Murphy, K. (1992). Changes in Relative Wages, 1963-1987: Supply and Demand Factors, *The Quarterly Journal of Economics*, MIT Press, Vol.107, Issue 1, pp.35-78.

Koenker, R. & Bassett, G. (1978). Regression Quantiles, *Econometrica,* Vol.46, Issue 1, pp.33-50

Krugman, P. (1995). Technology, Trade, and Factor Prices, NBER Working Papers No. 5355,

Lang, K. & Lehmann, J. (2011). Racial Discrimination in the Labor Market: Theory and Empirics. *Journal of Economic Literature*, Forthcoming.

Langoni, C. (1973). Distribuçao da Renda e Desenvolvimento Econômico no Brasil. Rio de Janeiro: Expressao Cultural.

Lee, R. (2003). The Demographic Transition: Three Centuries of Fundamental Change. *Journal of Economic Perspectives*, Vol.17, No.4, pp.167–190.

Leite, P. (2005). Race Discrimination or Inequality of Opportunities: the Brazilian Case, Ibero America Institute for Economic Research Discussion papers, No. 118.

Lemieux, T. (2002). Decomposing Changes in Wage Distributions: A Unified Approach. *Canadian Journal of Economics*, Vol.35, Issue4, pp. 646-88.

Lemieux, T. (2006). Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill? *The American Economic Review*, Vol.96, No.3. pp.461-498.

Loureiro, P. (2003). Uma Resenha Teórica e Empírica Sobre Economia da Discriminaçao. *Revista Brasileira de Economia*, Vol.57, No1, pp.125-157.

Loureiro, P. Carneiro, F. & Sachsida, A. (2004). Race and Gender Discrimination in the Labor Market: An Urban and Rural Sector Analysis for Brazil. Journal of Economic Studies, Vol.31, Issue 2, pp.129-143.

Machado J. & Mata J. (2005). Counterfactual Decompositions of Changes in Wage Distributions Using Quantile Regression, *Journal of Applied Econometrics*, Vol.20, Issue 4, pp.445-65.

Machado, C. (2011). Selection, heterogeneity and the gender wage gap. Mimeo, University of

Marques, L. Ñopo, H. &Salardi, P. (2009). Gender and Racial Wage Gaps in Brazil 1996-2006: Evidence Using a Matching Comparisons Approach, Inter-American Development Bank Working paper No68.

McCall, L. (2000). Explaining Levels of Within-Group Wage Inequality in U.S. Labor Markets. *Demography* Vol.37, pp.415–430.

Melly, B. (2004). Decomposition Differences in Distribution Using Quantile Regression, Mimeo.

Melly, B. (2005). Decomposition of Differences in Distribution Using Quantile Regression, *Labour Economics*, Vol.12, No.4, pp.577-590.

Melly, B. (2006). Estimation of counterfactual distributions using quantile regression. mimeo, Swiss Institute for International Economics and Applied Economic Research (SIAW), University of St. Gallen.

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution, *Journal of Political Economy* Vol.66, No.4, pp. 281-302

Moretti, E. (2004). Estimating the Social Return to Higher Education Evidence from Longitudinal and Repeated Cross Sectional Data, *Journal of Econometrics* Vol.121, pp.175-212.

Morrison, A. White, E. & the Centre of Creative Leadership (1987). Breaking the Glass Ceiling, New York: Addison-Wesley.

Moura, R. (2008). Testando as Hipóteses do Modelo de Mincer para o Brasil, *Revista Brasileira de Economia*, Vol.62, pp.407-449.

Neal, D. The Measured Black-White Wage Gap Among Women is Too Small, *Journal of Political Economy*, No.112, S1-28.

Neal, D. & Johnson, W. (1996). The Role of Premarket Factors in Black-White Wage Differences, *Journal of Political Economy* Vol. 104, No. 5 (Oct., 1996), pp. 869-895.

O'Neill, J. (1990). The Role of Human Capital in Earnings Differences Between Black and White Men. *Journal of Economic Perspectives,* Vol.4, Issue 4, pp.25-45.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets, *International Economic Review*, Vol.14.

Oliveira, A. Gonçalves, E. (2006). Tendências da Desigualdade Salarial para Coortes de Mulheres Brancas e Negras no Brasil, *Estudios Eeconometricos*, São Paulo, Vol. 36, No.2, pp.205-236.

Osório, R. (2008). Desigualdade Racial e Mobilidade Social no Brasil: Um Balanço das Teorias. In: *As Políticas Públicas de Desigualdade Racial no Brasil: 120 Anos Após a Abolição*. Mário Theodoro (org.), Brasília: Ipea.

Pavcnik, N. Blom, A. Goldberg, P. & Schady, N. (2002). Trade Liberalization and Labor Market Adjustment in Brazil. World Bank Policy Research Working Paper No. 2982.

Ramos, L. & Vieira, M. (2001). Desigualdade de Rendimentos no Brasil na Década de 90: Evoluçao e Principais Determinantes. IPEA, Rio de Janeiro.

Rauch, J. (1993). Productivity Gains from Geographic Concentrationof Human Capital: Evidence From the Cities, *Journal of Urban Economics*, Vol. 34, Issue 3, pp.380-400.

Ravallion, M. Ferreira, F. & Leite, P. (2010): Poverty reduction without economic growth? Explaining Brazil's poverty dynamics, 1985–2004. *Journal of Development Economics,* Vol.93, pp.20–36.

Reis, J. &Paes de Barros, R. (1989). Um Estudo sa Evolução das Diferenças Regionais da Desigualdade no Brasil. Rio de Janeiro: IPEA. Processed.

Salardi, P. (2013). An Analysis of Pay and Occupational Differences by Gender and Race in Brazil - 1987 to 2006, PhD Thesis, University of Sussex, UK

Santos, R. & Ribeiro, E. (2006). Diferenciais de rendimentos entre homens e mulheres no brasil revisitado: explorando o "teto de vidro". Centro Universitario Newton Paiva/MG e PPGE/UFRGS.

Silva, N. (1980). O Preço da Cor: Diferenciais Raciais na Distribuçao da Renda no Brasil. Pesquisa e Planejamento Econômico, Rio de Janeiro, Vol. 10, No.2.

Síntesis de Indicadoes Sociais (2012). http://saladeimprensa.ibge.gov.br/es/noticias?view=noticia&id=1&busca=1&idnoticia=2268

Soares, S. (2000). O Perfil da Discriminação no Mercado de Trabalho - Homens Negros, Mulheres Brancas e Mulheres Negras, Texto para discussão No769 , IPEA.

Telles, E. (2003). Racismo à Brasileira: Uma Nova Perspective Sociológica – Traduçao de Nadjeda Rodrigues Marques e Camila Olsen. Rio de Janeiro: Relume Dumará: Fundaçao Ford.

The Economist (2012). "Race in Brazil: Affirming a divide" The Economist 28th January.

Velez, C. Barros, R. & Ferreira, F. (2008). Inequality and Economic Development in Brazil, World Bank Publication, USA.

World Bank Data (2011). http://data.worldbank.org/

Appendixes

Appendix 1. Summary Statistics by Quantile

Table 8. Average Age per Quantile

Average Age per Quantile						
Quantile	White 2001	Nonwhite2001	White 2011	Non White 2011	Racial Difference 2001	Racial Difference 2011
1	33.09	30.89	36.83	35.28	7%	4%
2	33.15	32.31	35.12	35.99	3%	-2%
3	33.50	32.92	34.78	34.49	2%	1%
4	33.41	33.15	34.93	35.68	1%	-2%
5	33.94	34.37	35.47	34.7	-1%	2%
6	34.38	34.39	36.32	35.50	0%	2%
7	35.29	34.86	37.12	36.04	1%	3%
8	36.15	35.56	38126	37194	2%	2%
9	38.22	36.65	39.68	38.81	4%	2%

Note: Brasilia is exlcluded from the sample.

Base year for real wages: 2005

1R$=$0,5

Source: Authors calculations from PNAD 2001 & 2011

Table 9. Average Years of Study per Quantile

Average Years of study per Quantile						
Quantile	White 2001	Nonwhite2001	White 2011	Non White 2011	Racial Difference 2001	Racial Difference 2011
1	6.68	5.44	8.41	7.19	19%	15%
2	7.84	6.13	9.69	8.33	22%	14%
3	8.29	6.64	10.01	9.09	20%	9%
4	8.86	7.31	10.42	9.59	17%	8%
5	9.26	7.59	10.69	9.58	18%	10%
6	10.09	8.00	11.42	9.96	21%	13%
7	10.87	8.57	12.09	10.32	21%	15%
8	12.06	9.16	13.32	11.12	24%	17%
9	13.38	10.2	14.22	12.21	24%	14%

Note: Brasilia is exlcluded from the sample.

Base year for real wages: 2005

1R$=$0,5

Source: Authors calculations from PNAD 2001 & 2011

Table 10. Average Years of Experience per Quantile

Average Experience per Quantile						
Quantile	White 2001	Nonwhite2001	White 2011	Non White 2011	Racial Difference 2001	Racial Difference 2011
1	20.46	19.46	22.41	22.09	-5%	-1%
2	18.96	20.17	19.43	21.69	6%	12%
3	19.21	20.28	18.77	19.4	6%	3%
4	18.55	19.84	18.52	20.09	7%	8%
5	18.67	20.77	18.78	19.12	11%	2%
6	18.30	20.39	18.91	19.54	12%	3%
7	18.42	20.29	19.03	19.71	10%	4%
8	18.09	20.41	18.81	20.07	13%	7%
9	18.85	20.46	19.46	20.60	9%	6%

Note: Brasilia is exlcluded from the sample.

Base year for real wages: 2005

1R$=$0,5

Source: Authors calculations from PNAD 2001 & 2011

Appendix 2 . Summary Statistics for Institutional Factors

Table 11. Average Real Monthly Wages by Quantile

Quantile	White 2001	Non-white2001	White 2011	Non-white 2011
		Average Real Monthly Wage by Quantile		
1	171.43	100.00	298.50	149.25
2	257.14	142.86	406.71	298.51
3	314.29	257.14	447.76	406.72
4	394.29	257.14	522.39	406.72
5	471.43	285.71	597.01	447.76
6	571.43	357.14	746.27	522.39
7	714.29	428.57	895.52	597.01
8	1071.42	542.86	1208.96	746.27
9	1714.29	800	2089.55	1119.40

Note: Red lines delimit where the minimum wage earners are located.

Brasilia is exlcluded from the sample.

Base year for real wages: 2005.

1R$=$0,5

Source: Authors calculations from PNAD 2001 & 2011.

Figure 7. Union Membership by Racial Group & Year
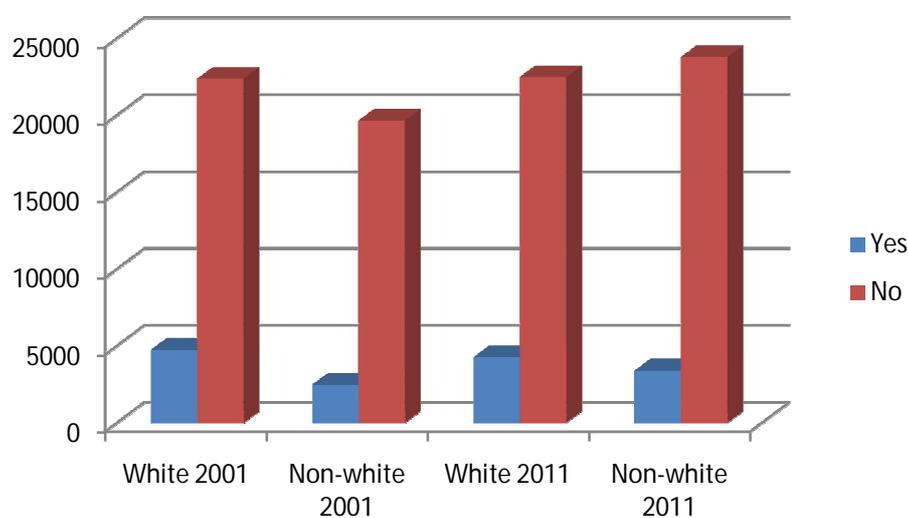
Appendix 3. Assumptions of the Quantile Regression Based Decompositions

1. *Mutually exclusive groups*: The population under study must be divided into two mutually exclusive groups: A and B, where for an agent i $D_{Ai} + D_{Bi} = 1$ and $D_{gi} = 1\{i \text{ is in } g\}, g = A, B$ and $1\{\cdot\}$

2. *Partial Equilibrium:* prices and quantities can be treated as independent

3. *Structural Form*: each worker i belonging to either group A or B is paid according to the wge structure of each group: m_A or m_B. These wage structures are functions of workers' observable characteristics (x) and unobservable characteristics (e):
$$Y_{Ai} = m_A(X_i, e_i) \text{ and } Y_{Bi} = m_B(X_i, e_i)$$
This means that there are only three possible reasons for this assumption to hold: differences in the wage functions of A and B, differences in the distribution of observable characteristics (x) across each group and differences in the distribution of unobservable characteristics (e) across each group.

4. *Simple Counterfactual Treatment*: A counterfactual wage structure (m_c) is assumed to be a simple counterfactual when $m_c(..) \equiv m_A(..)$ for workers in group B, or $m_c(..) \equiv m_B(..)$ for workers in group A.

5. *Overlapping Support*: the support for all wage setting factors $[X', \varepsilon']'$ is $X \times \varepsilon$. For all $[x', e']'$ in $X \times \varepsilon, 0 < \Pr[D_B = 1 | X = x, \varepsilon = e] < 1$. Therefore, no value of $X = x, \varepsilon = e$ can identify workers into one group or another.

6. *Conditional Ignorability/Indepdence/Exogeneity:* The distribution of wages, conditional on X depends on the conditional distribution of ε, and the wage structure $m_g(\cdot)$. This is to say for $g = A, B, having \left(D_g, X, \varepsilon\right) a\ joint\ distribution.$ For all x in X, ε is independent on D_g given $X = x$. This the same as assuming: $D_g \perp \varepsilon | X$. As we do not fulfill this assmption, we instead assume *weak ignorabilit/exogeneity* where unoservables are not independent of the mean *X,* however their conditional distribution on *X* is assumed to be the same for both groups.

7. *Invariance of the Conditional Distributions:* The conditional wage distribution $F_{Y_A|X,D_A}(y|X = x)$ can be extrapolated for $x \epsilon X$, therefore it remains valid when the marginal distribution of $F_{X|D_B}$ replaces $F_{X|D_A}$.

8. *Strict Monotoniciy in the Randome Scalar ε*: For g=A,B and for all values of x in X, ε is a scalar random variable and $m_g(X, \varepsilon)$ is strictly increasing in ε.

### Assumption for Functional form restrictions applied to more general decompositions

9. *Constant Returns to Unobservables:* For g=A,B, $v_g = \sigma_g \varepsilon$.

10. *Homoskedasticity:* For g=A,B, $\text{Var}[\varepsilon | X, D_g = 1] = 1$.

11. *Conditional Rank Preservation:* The rank for workers in both groups are distributed equally conditionally on explanatory variables. For all indivual i, there is $\tau_{Ai}(x_i) = \tau_{Bi}(x_i)$, where $\tau_{Ai}(x_i) = F_{v_{A|X}}(v_{Ai}|X = x_i)$ and $\tau_{Bi}(x_i) = F_{v_{B|X}}(v_{Bi}|X = x_i)$ are the rankings of v_{Ai} and v_{Bi} in their conditional distributions.

12. *Heterogeneous Returns to Observables:* For g=A,B, $Y_{gi} = X_i\beta_{g,\tau} + h_{g,\tau}(\varepsilon_i)$.

13. *Linear Relationship:* As in the OLS framework, we assume a linear relationship between the quantiles of y and x.

14. *Complete Collection of Linear Conditional Quantiles:* For $g=A,B$, and $\forall \in (0,1)\tau = \Pr\left(Y_g \leq x\beta_{g,\tau}|X = x, D_g = 1\right)$.

(Fortin, et al. 2011 ch2)

Appendix 4. Quantile Regression Results

<u>**Figure 8. Quantile Regression Results for Education**</u>
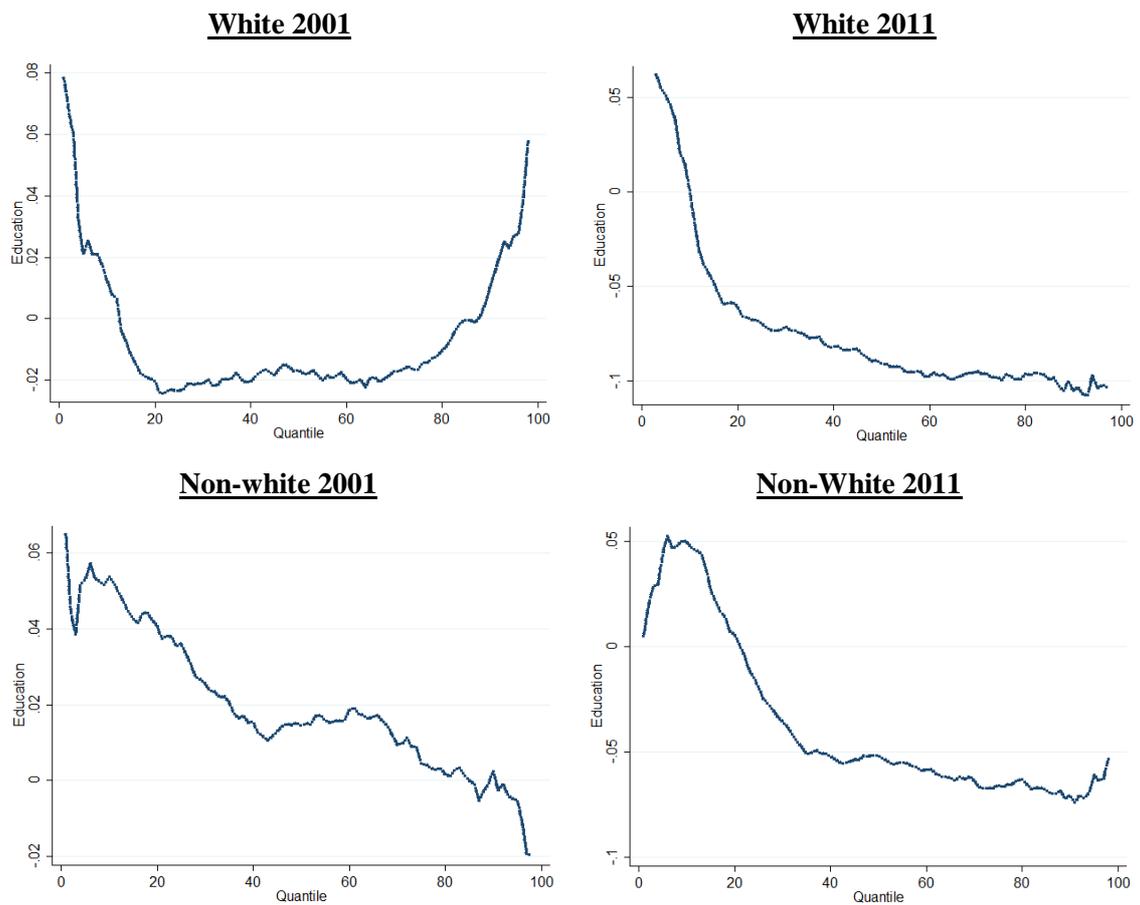
Source: Author's calculations from PNAD 2001 & 2011

Figure 9. Quantile Regression Results for Experience

White 2001

White 2011

Non-white 2001

Non-white 2011
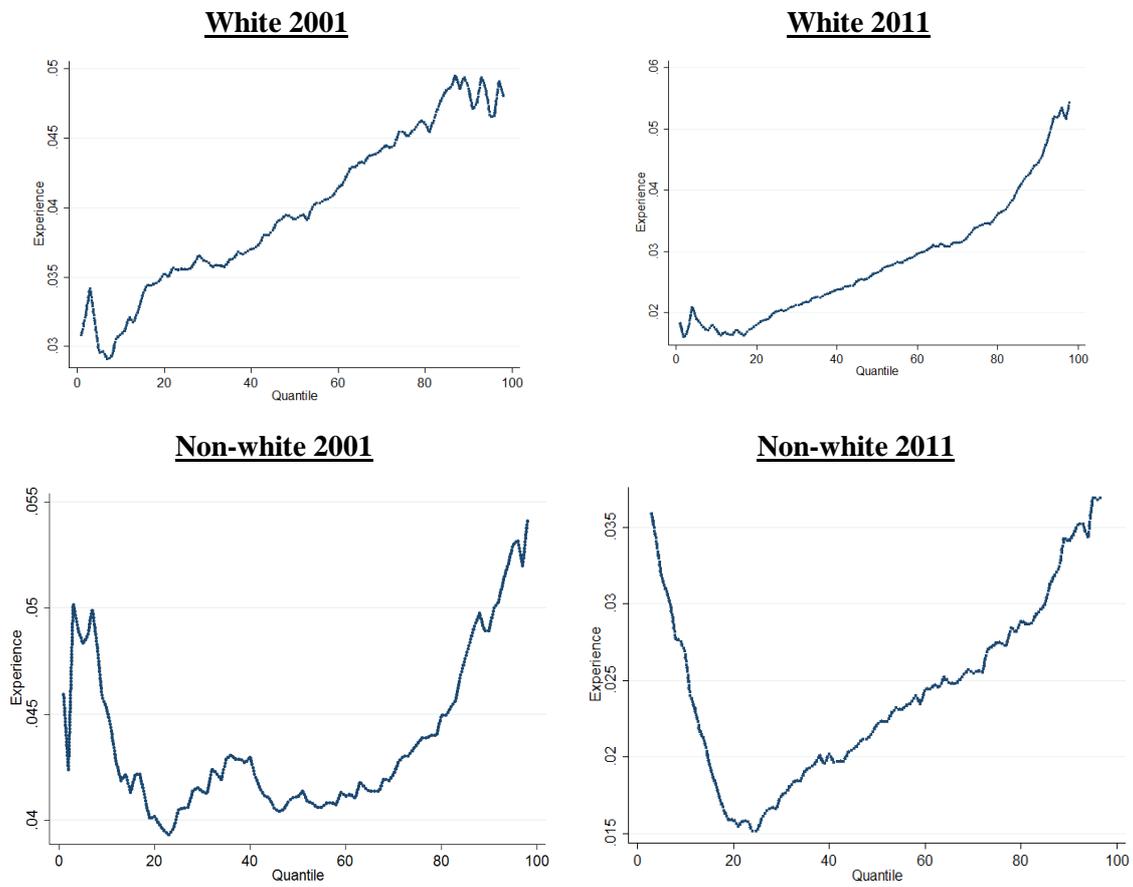
Source: Author's calculations from PNAD 2001 & 2011

Appendix 5. Decomposition Results

Table 12. Decomposition Results

	Racial Wage Differentials By Quantile											
	2001						**2011**					
	0,10		**0,50**		**0,90**		**0,10**		**0,50**		**0,90**	
	Melly (2006) Decomposition											
Total Variation	0,50	(0 ,007)	0,48	(0,005)	0,72	(0,012)	0,34	(0,007)	0,30	(0,005)	0,52	(0,012)
	100%		100%		100%		100%		100%		100%	
Characteristics	0,21	(0,009)	0,26	(0,007)	0,40	(0,018)	0,14	(0,008)	0,15	(0,006)	0,25	(0,012)
	42%		54%		56%		41%		50%		48%	
Coefficients	0,29	(0,011)	0,22	(0,007)	0,32	(0,013)	0,20	(0,011)	0,15	(0,004)	0,27	(0,011)
	58%		46%		44%		59%		50%		52%	
	Melly (2005) Decomposition											
Total Variation	0,50	(0,015)	0,48	(0,009)	0,72	(0,017)	0,34	(0,013)	0,30	(0,007)	0,52	(0,017)
	100%		100%		100%		100%		100%		100%	
Characteristics	0,20	(0,006)	0,25	(0,007)	0,39	(0,012)	0,14	(0,007)	0,16	(0,005)	0,25	(0,008)
	41%		53%		55%		41%		52%		47%	
Coefficient	0,23	(0,008)	0,23	(0,008)	0,29	(0,015)	0,12	(0,007)	0,17	(0,006)	0,27	(0,013)
	47%		49%		41%		37%		56%		51%	
Residual	0,06	(0,010)	-0,01	(0,005)	0,02	(0,011)	0,08	(0,009)	-0,02	(0,005)	0,01	(0,011)
	12%		-2%		3%		22%		-8%		2%	

Note: Bootstrap standard errors with 100 replications in parenthesis.

The percentage weight of every decomposition is presented below each value.

Source: Authors calculations from PNAD 2001 & 2011

## Appendix 6. Model Results Including Occupation

6.1 Summary Statistics for Occupational Categories for Women

Figure 10. Occupation Composition for White Female Workers
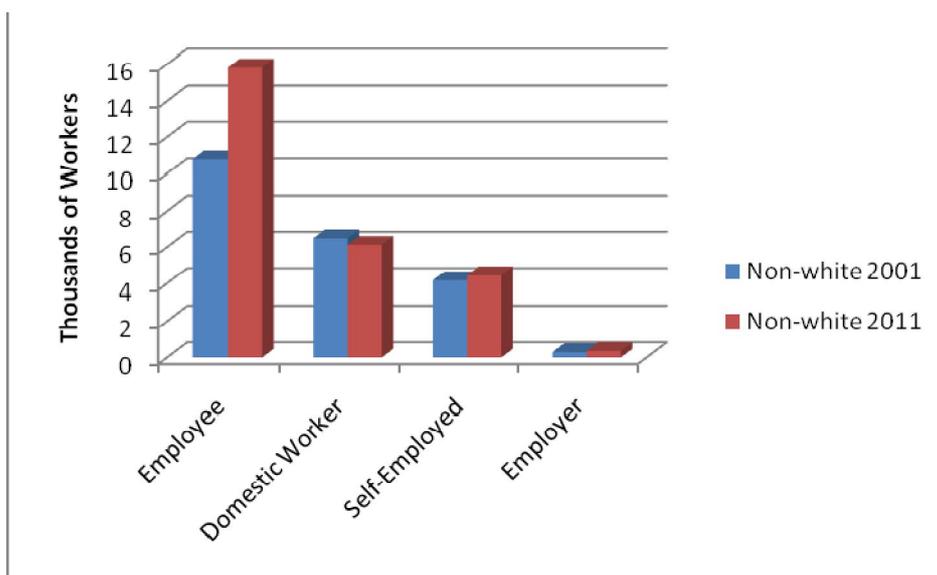
*Source: Author's calculations from PNAD 2001 & 2011*

Figure 11. Occupation Composition for Non-white Female Workers
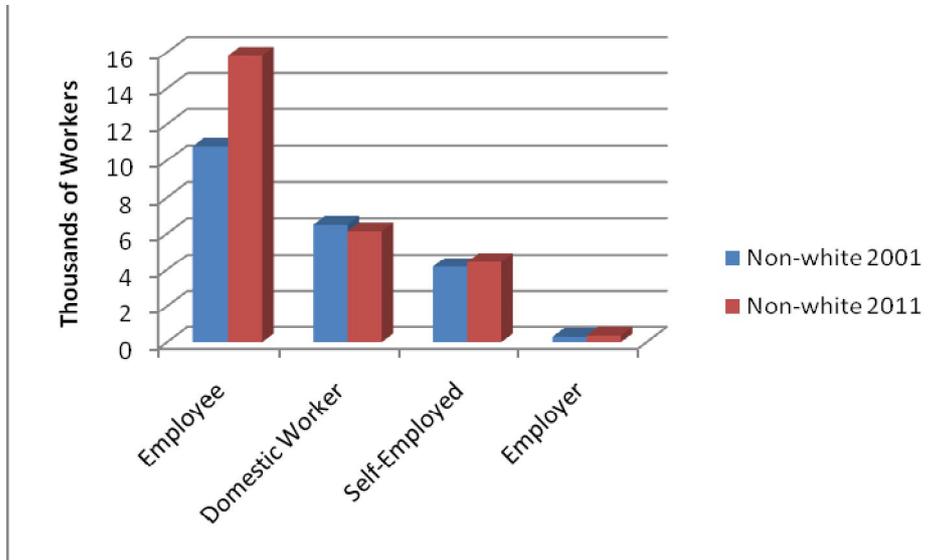
*Source: Author's calculations from PNAD 2001 & 2011*

6.2 Decomposition Results

Figure 12. Decompositions Results Including Occupational Categories

Melly 2006

2001

2011

Melly 2005
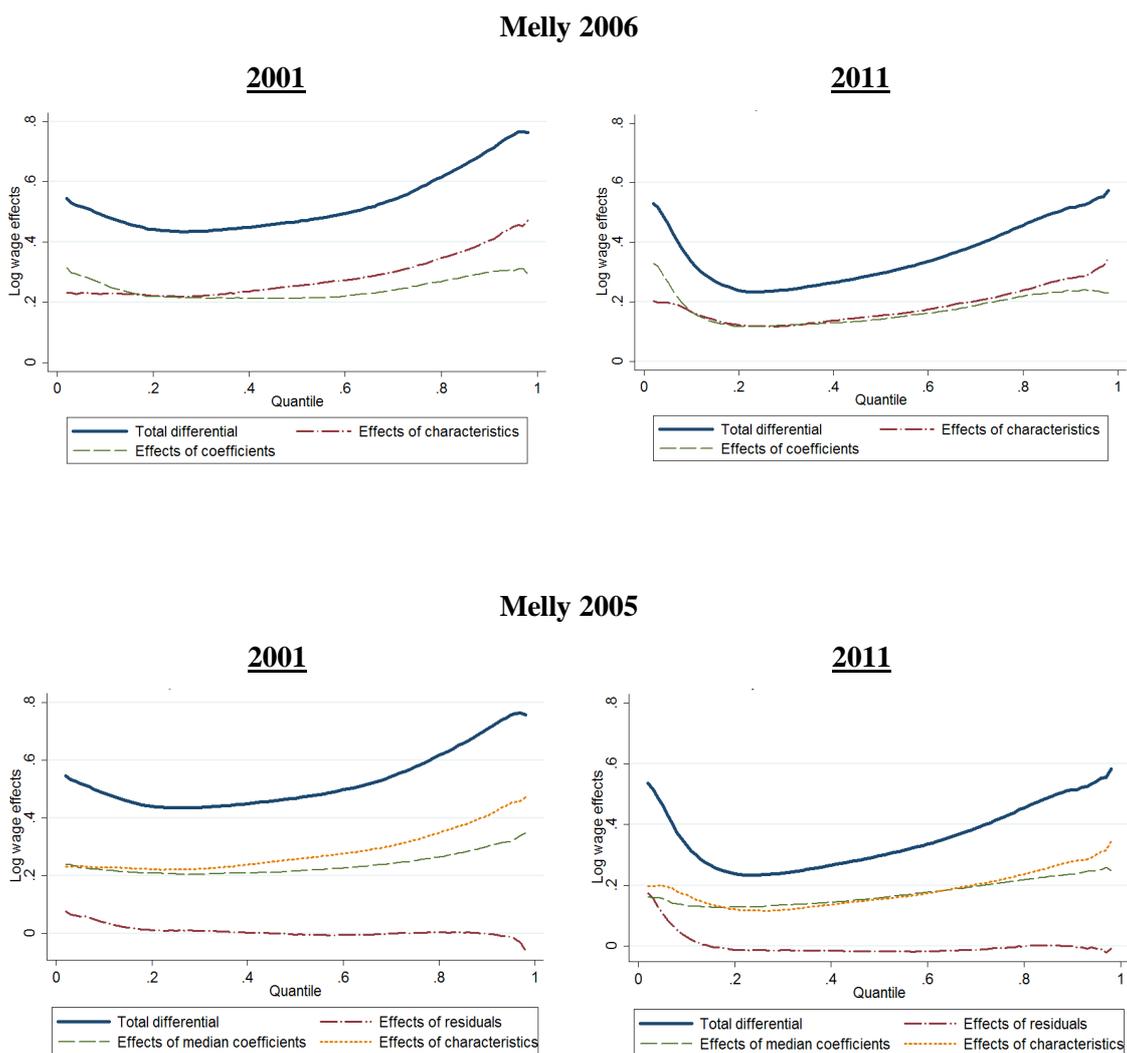
2001

2011

Source: Author's calculations from PNAD 2001 & 2011

Appendix 7. Robustness Checks

Figure 13. Decompositions with Variable Experience First

Melly 2006

2001   **2011**

Melly 2005

2001   **2011**

Source: Author's calculations from PNAD 2001 & 2011