WAGE DIFFERENTIAL BETWEEN MOTHERS AND NON-MOTHERS IN BRAZIL: A DECOMPOSITION STUDY FOR SELECTED AGE GROUPS

1. Introduction

The term "motherhood penalty" refers to the disadvantages suffered by women who have children in the labor market. The constraints faced by mothers in the labor sphere can be manifested in different indicators, being the wage gap the main one (WALDFOGEL, 1997; ANDERSON ET AL, 2003; BUDIG AND ENGLAND, 2001; GRIMSHAW AND RUBERY, 2015). Other measures include labor participation rate (MISRA, BUDIG, AND MOLLER, 2007), occupational status (prestige) (ABENDROTH ET AL, 2014) and career progression time (MCINTOSH ET AL, 2012). Most studies measure these penalties by comparing women mothers and childless women.

From a wage perspective, motherhood is often penalized due to various factors. Some of them, listed by Budig and England (2001), are: i) the limitation on the accumulation of experience due to temporary absence from the paid market to look after children; ii) the reduction in productivity due to the difficulty of reconciling the responsibilities of paid and unpaid work; iii) the choice of lower paid jobs to allow a better reconciliation between motherhood and work, such as jobs with more flexible hours, which generally have lower salaries; and iv) discrimination against mothers by employers.

Although there are several studies on female participation and the wage gap in the Brazilian labor market (some examples are LEME AND WAJNMAN, 2000; SOARES, 2000; GIUBERTI AND MENEZES FILHO, 2005), most of these studies focus on comparisons between men and women. Few studies have analyzed the differences in salaries between mothers and childless women. The aim of this paper is therefore to decompose the wage differential between women with and without children, analyzing different age groups , in order to separate the portion explained by the productive characteristics of these women from that not explained by the observed attributes.

Decomposition analysis has been widely used in the labor market literature to study wage differentials between certain groups, with sex and race/color being the most explored categories (FORTIN, LEMIEUX AND FIRPO, 2011). This widespread use stems from the fact that the wage differential between these groups is partly explained by differences in their productive characteristics (such as levels of education and experience) and partly by differences in the return on attributes, caused by the structure of the market and the way it remunerates individuals.

Since mothers divide their time between caring for their children and the labor market, they probably face greater obstacles to investing in education or acquiring work experience, which would contribute to them having lower incomes on average than non-mothers who are at the same stage of the life course. However, differences related to attributes are not expected to explain the entire income gap between the two groups under analysis, since the structure of the labor market can cause differences in the returns of mothers and non-mothers, despite their characteristics. The former, due to their double shift, may be perceived as less productive and therefore have lower average earnings than the latter.

The productive attributes of women with and without children may vary according to the time of life they are in, just as the structure of the market may present different dynamics according to the age of these women (ABENDROTH ET AL, 2014; KHAN ET AL, 2014; LANDIVAR,

2018, NEUMEIER, SØRENSEN AND WEBBER, 2018). For this reason, the decomposition analysis will be carried out for different age groups, since it will thus be possible to observe the evolution of both parts of the wage differential between the different moments of the life course.

In addition to the decomposition analysis of the average wage of mothers and non-mothers, it is important to carry out the same exercise across the distribution of salaries of these two groups of women. In a country characterized by marked income inequality, such as Brazil, it is expected that the motherhood penalty will not be homogeneous across income quantiles. Although some national studies have already carried out a similar analysis, it should be noted that the relationship between the motherhood wage penalty and income level is not conclusive, even in the international literature: there are studies that show greater differentials in the lower portion of the wage distribution (BUDIG AND HODGES, 2010), in the middle of the distribution (KILLEWALD AND BEARAK, 2014) and at the top (ANDERSON ET AL, 2002; ENGLAND, 2016). Thus, the decomposition along the wage distribution intended here will serve as an additional subsidy, with the difference of being carried out for different age groups.

In order to meet the abovementioned objectives, this paper is structured in four sections, in addition to this introduction. The first contains a brief literature review of the motherhood penalty, with a focus on the wage perspective and the age-related aspects; the second describes the database and the methods used for the analysis; the third presents and discusses the results obtained, and the last brings the final considerations.

2. A brief literature review on the mothehood wage penalty

The literature on the maternity penalty shows a consensus on the fact that mothers receive, on average, lower wages than non-mothers. Some examples are Waldfogel (1997), Budig and England (2001) and Anderson, Binder and Krause (2002) for the United States; Kühhirt and Ludwig (2012) for Germany; Livermore, Rodgers and Siminski (2011) for Australia; Jia and Dong (2012) for China; Gamboa and Zuluaga (2013) for Colombia; Molina and Montuenga (2009) for Spain, and Gafni and Siniver (2015) for Israel. Although there is no disagreement in the literature about the existence of the penalty itself, several points are still up for debate. For example, what factors contribute most to the wage gap between mothers and non-mothers? Which mothers are penalized more, those with the worst socioeconomic status or those with the best? How does the penalty tend to behave over the life course?

With regard to socioeconomic status (SES), some studies have looked at the issue from the point of view of the level of income, education or occupations of women with and without children (ANDERSON ET AL, 2002; BUDIG AND HODGES, 2010; KILLEWALD AND BEARAK, 2014; ENGLAND, 2016, GLAUBER, 2018). An obvious inference is that women with lower SES would suffer a greater penalty, given their condition of greater social vulnerability. However, this conclusion is not unanimous, either conceptually or empirically.

From a theoretical point of view, Budig and Hodges (2010) present two lines of reasoning: on the one hand, women at the top of the income distribution live in households with better conditions for outsourcing care tasks, thus reducing the conflict between work and family. On the other hand, the greater availability of resources may allow high-earning mothers to reduce or even interrupt their work activity for longer, leading to a greater wage penalty for them due to the loss of market experience. Anderson et al (2002) and England (2016) also point out that, due to the lower level of human capital of women with low SES, these workers would be less vulnerable to a deterioration in wages due to motherhood. Thus, interruptions in labor market experience would not have a major impact on their wages, and the penalty would be greater for the more qualified, which was confirmed in their results.

The findings of Anderson et al (2002) and England (2016) differ from those found by Budig and Hodges (2010), which indicated greater penalties in the lower quantiles of the income distribution, and Killewald and Bearak (2014), in which the penalty was concentrated in the middle of the wage distribution. The last two studies used the same database, but different methodologies - conditional quantile regressions, in the case of Budig and Hodges, and unconditional quantile regressions, used by Killewald and Bearak[1] . The discrepancy in the results showed that the question of the motherhood penalty by socio-economic group seems to be sensitive to the method used, raising the debate as to which one is the most appropriate.

With regard to the motherhood penalty at different ages, studies show that, on the one hand, greater experience in balancing motherhood and work and the growth and independence of children can contribute to reducing the penalty at older ages (NEUMEIER, SØRENSEN AND WEBBER, 2018). However, the discontinuity of mothers' experience can make it difficult for them to enter higher-paying occupations or delay their occupational progression, causing a growing disadvantage compared to non-mothers (ABENDROTH ET AL, 2014; KHAN ET AL, 2014).

When analyzing occupational prestige over the life course for European countries, Abendroth et al (2014) showed that motherhood has short- and long-term costs. The passage of time has not compensated for the negative consequences of a first birth, and losses in occupational status accumulate as the first child grows up. For the second birth, occupational penalties did not worsen over time, but occupational status did not recover. Khan et al (2014) carried out similar research, considering participation, wages and prestige for the United States. The authors observed that the difference in participation between women with and without children became almost non-existent from the age of forty; for wages, there is an increase in the penalty around the age of thirty, but the difference narrows at forty. The prestige penalty is reduced between the ages of thirty and forty. Neumeier, Sørensen and Webber (2018) also analyzed the American context, showing that the wage gap between mothers and non-mothers increases over the life course, but at a decreasing rate as age advances.

A key aspect of the penalty over the life course is the *timing* of motherhood, i.e. the moment in life when the child arrives. For example, women who have postponed motherhood will probably have children at a time when they already have a large amount of accumulated human capital. Thus, the penalty suffered by such women would be different from that experienced by women who had their children at a young age. Landivar (2018) showed that in the United States, mothers in occupations categorized as "professional" – which offers higher wages - experienced a greater wage penalty from early motherhood. Delaying the first birth reduced the penalty in these high-wage occupations, while women in low-wage occupations showed little economic benefit from late motherhood.

In Brazil, the analysis of the labor participation and wages of women with children in the Brazilian labor market generally appears within broader contexts of analyses that relate labor variables to types of household and family arrangements. The findings generally show that men's participation in the labor market is not significantly affected by the family structure of which they are a part, but for women there is a wide variation, depending on marital status or

[1] The differences between the two methods will be explained later in this paper.

the presence of children (LAVINAS AND NICOLL, 2006; SORJ ET AL, 2007; WAJNMAN, 2012; MADALOZZO AND BLOFIELD, 2017).

The aforementioned studies provide unanimous evidence that the presence of children in the household tends to depress participation in labor market or the wages of Brazilian women. However, the national literature still has few studies that delve into this issue. Table 1 provides a summary of the studies that have looked more specifically at the motherhood penalty in Brazil.

Table 1. Summary of studies on the maternity penalty in the Brazilian labor market

Author	Description of work
Pazello and Fernandes (2004) Pazello (2006) Souza, Rios-Neto and Queiroz (2011)	These studies have in common the attempt to control the preferences of the women analyzed. To do so, they compare mothers and women who have given birth to stillbirths, assuming that both have preferences oriented towards motherhood, even though the latter had never experienced it. Another comparison is that of women who have had twins with women who have not, which again assumes a preference for children for both groups of analysis, even though the former has experienced an exogenous fertility shock. Generally speaking, the results of these studies showed a negative relationship between motherhood and labor market participation or hourly wages, but the effects diminish in a "long-term" perspective (here demarcated by the age of the mother or child).
Paulo (2012)	Compared mothers and childless women at two points in time, 1992 and 2009, and showed that the negative impact of motherhood on earnings increased between the two years. This result is attributed to the greater demands of the market, making it difficult for mothers to reconcile work and domestic activities.
Paulo (2013)	In addition to analyzing the motherhood penalty for average wages, the author studied the phenomenon along the conditional distribution of wages. The initial hypothesis of a greater penalty in the higher income quantiles was not confirmed, since the quantile with the greatest difference between mothers and non-mothers varied between the years analyzed. The author also investigated the impact of the timing of the birth of the first child, showing that there is a premium for postponing motherhood only for wealthier women.
Guiginski (2015)	They analyzed four indicators of access to work for men and women: participation, quality of the job held, earnings and risk of unemployment. The results, for 2013, showed that the presence of children in the household had a negative impact on all female indicators; for males, however, the results were mostly not statistically significant.
Souza (2016)	Used the Oaxaca-Blinder method to decompose the hourly wages of mothers and non-mothers, finding a wage penalty for the former and an increase in this according to the number of children. The author also analyzed the distribution of conditional and unconditional wages, and in both cases found greater penalties at the top of the distribution, justifying the results by the opportunity cost, in terms of potential earnings, of high-income women .

(continued)

(conclusion)

Muniz and Veneroso (2019)	Observed the impact of the presence of young children in the household for men and women in two main dimensions: participation, using a sample selection model, and wages, using conditional quantile regressions. They showed that at the top of the income distribution (the richest 5%) there is no difference in wages between the sexes. At the bottom, on the other hand, there was a wage gap, but the authors consider that the impact of motherhood is small, and other factors predominate, such as schooling, type of occupation and discrimination. The penalty, according to the authors, is more significant in relation to market participation than in terms of wages, and varies according to the order of childbirth.

Source: Own elaboration.

Although some of the studies shown in Table 1 present analyses in which age is considered as an explanatory variable for the differences between women with and without children, none of them emphasize the differences between these women resulting from the stage of the life course they are at. This is a gap that this paper seeks to fill, since decomposition analyses will make it possible to verify possible changes in the composition of the wage differential between different age groups. The next section describes the data source for this study and the methods used.

3. Data and methods

In order to achieve the objectives set out in the introductory section, we used microdata from the National Household Sample Survey (*Pesquisa Nacional por Amostra de Domicílios*, PNAD), carried out by the Brazilian Institute of Geography and Statistics (*Instituto Nacional de Geografia e Estatística*, IBGE) for 2015. The period chosen took into account the most recent survey in which the fertility questionnaire was carried out. Women living in urban areas were selected for the sample, divided into mothers and non-mothers, and the division of the two groups was made using the variable V1101, "Had any children born alive up to the reference date". The sample was also divided into two age groups: 25 to 34 years and 35 to 44 years. It was decided to use ten-year age groups to avoid the sample becoming too small.

With regard to the empirical models, the Oaxaca-Blinder decomposition was used for the average hourly wages and the quantile decomposition according to the approach of Firpo, Fortin and Lemieux (2009, 2018) for the hourly wages quantiles. The central idea of these models is to identify the effect of the personal and productive attributes of two groups of interest on the wage gap between them, as well as to separate the effect in question from that not explained by observable characteristics. Sections 3.1 and 3.2 explain these methods in detail.

3.1.Average wages: Oaxaca-Blinder decomposition

The methodology developed by Oaxaca (1973) and Blinder (1973) for decomposing the wage gap between two groups of interest is considered pioneering in studies related to the labor market, and has become a widely used tool in the analysis of wage differentials by race/color, sex and other factors (FORTIN, LEMIEUX AND FIRPO, 2011). Thus, in view of the scope of this paper, the Oaxaca-Blinder decomposition model applied to the income differential between mothers and non-mothers is briefly presented. The starting point of the model is the estimation of linear regressions for each group of interest, whose dependent variable is the logarithm of the hourly wages, and the regressor variables are attributes that influence the behavior of earnings, as shown in (1).

$$\ln W_{gi} = \beta_{g0} + \sum_{k=1}^{K} X_{ik}\,\beta_{gk} + v_{gi}, \qquad g = M, NM \tag{1}$$

In (5),$\ln W_{gi}$ is the logarithm of hourly wages,β_{g0} is the angular coefficient,$\sum_{k=1}^{K} X_{ik}$ is the vector of explanatory variables,$\sum_{k=1}^{K} \beta_{gk}$ is the vector of linear coefficients,v_{gi} is the error term, which is expected to equal zero, and M and NM correspond, respectively, to the groups of mothers and non-mothers. The difference between the average hourly wages of the two can be represented as in equation (2):

$$\hat{\Delta}_0^{\mu} = \ln \overline{W}_M - \ln \overline{W}_{NM} \tag{2}$$

Substituting equation (1) into (2) and rearranging the terms, we have:

$$\hat{\Delta}_0^{\mu} = \underbrace{\sum_{k=1}^{K} (\overline{X}_{Mk} - \overline{X}_{NMk})\,\hat{\beta}_{NMk}}_{\hat{\Delta}_X^{\mu}\text{ (composition effect)}} + \underbrace{(\hat{\beta}_{M0} - \hat{\beta}_{NM0}) + \sum_{k=1}^{K} \overline{X}_{Mk}\,(\hat{\beta}_{Mk} - \hat{\beta}_{NMk})}_{\hat{\Delta}_S^{\mu}\text{ (wage structure effect)}} \tag{3}$$

Equation (3) expresses the structure of the decomposition. The first term, the composition effect ($\hat{\Delta}_X^{\mu}$), represents the portion of the wage gap arising from differences in the attributes of women with and without children. The second, the wage structure effect ($\hat{\Delta}_S^{\mu}$) shows the effect of the coefficients of the predictor variables, i.e. the returns to the characteristics of these women. The literature usually attributes the effect of discrimination in the market to this part of the decomposition, but with reservations, since the part not explained by characteristics can include other problems, such as the non-inclusion of other variables that can explain the wage differential between the groups analyzed.

As well as distinguishing between the effects of wage composition and structure, the Oaxaca-Blinder decomposition makes it possible to estimate the contribution of each covariate to these two effects. Thus, the aggregate decomposition is the separation of $\hat{\Delta}_0^{\mu}$ into $\hat{\Delta}_X^{\mu}$ and $\hat{\Delta}_S^{\mu}$, while the detailed decomposition consists of the subdivision of $\hat{\Delta}_X^{\mu}$ into $\hat{\Delta}_{X,k}^{\mu}$, as well as $\hat{\Delta}_S^{\mu}$ into $\hat{\Delta}_{S,k}^{\mu}$, for which k represents each covariate in the model (FORTIN, LEMIEUX AND FIRPO, 2011).

3.2.Wage distribution: Firpo, Fortin and Lemieux decomposition

Quantile regression analysis was introduced by Koenker and Bassett (1978). Since then, various studies have sought to combine this modeling with decomposition methods in line with Oaxaca-Blinder (FORTIN, LEMIEUX AND FIRPO, 2011). While some of these were successful in running quantile decomposition models in aggregate terms (composition and wage structure effect), the detailed decomposition - which presents the contribution of each variable in the vector of predictors to the composition and return effects - proved to be a more complex challenge to solve.

Fortin, Lemieux and Firpo (2011) explain that, in an OLS model, the relationship between the vector of predictors (X) and the predicted variable (Y) is defined by a function of conditional averages, which show how the averages of Y vary according to the variables in the vector of X.

$$E(Y \mid X) = X\beta \tag{4}$$

Alternatively, from the perspective of the Law of Iterated Expectations, the average value of Y can be obtained unconditionally from the vector of explanatory variables, as shown in (5):

$$\mathrm{E(Y)} = \mathrm{E_X(Y \mid X)} = \mathrm{E(X)\beta} \quad (5)$$

According to equation (10), β could be interpreted as the effect of increasing the mean of X on the unconditional mean of Y.

When there is interest in the distribution of the dependent variable, reasoning analogous to that represented by equation (4) can be applied to quantiles; thus, each conditional quantile $\mathrm{Q_\tau(X)}$ can be expressed as a function of the vector of predictors:

$$\mathrm{Q_\tau(X)} = \mathrm{X\beta_\tau} \quad (6)$$

Where β_τ is the effect of X on the τ-th conditional quantile of Y . The same analogy, however, cannot be applied to the unconditional mean of the dependente variable, since $Q_\tau \neq E_X[Q_\tau(X)] = E(X)\beta_\tau$, for which Q_τ is the unconditional quantile. This makes it impossible to interpret β_τ as the effect of increasing the average of X on Q_τ.

The infeasibility of applying the Law of Iterated Expectations in the context of quantile regression has led to the development of some alternatives. The most widespread of these is the proposal by Machado and Mata (2005), which consists of estimating regressions for all quantiles of the distribution. Once the entire conditional distribution of Y given X has been estimated, it would be possible to carry out a counterfactual exercise to analyze differences between two groups, such as women with and without children, for example. Another widely known proposal is that of Melly (2006), based on the work of Machado and Mata (2005), but with some differences regarding the estimation of this counterfactual exercise. In Machado and Mata (2005), the counterfactual unconditional wage distribution is constructed from the generation of a random sample; in turn, the estimation of the unconditional distribution by Melly (2006) is carried out by integrating the conditional distribution over a range of covariates (SALARDI, 2013).

Firpo, Fortin and Lemieux (2009, 2018) argue that the Machado-Mata-Melly models have the limitation of only providing an aggregate decomposition, and that it is not possible to estimate the detailed contribution by components of the vector X . Therefore, the authors created an alternative formulation, using the *Recentered Influence Function* (RIF), which makes it possible to estimate the detailed decomposition by covariates, as well as being path independent*,* i.e. the order of the predictors does not influence the model's results. For these reasons, we chose to use the model proposed by Firpo, Fortin and Lemieux (2009, 2018) for the quantile decomposition exercises in this study.

In order to understand the use of RIF in regression (FIRPO, FORTIN AND LEMIEUX, 2009) and decomposition (FIRPO, FORTIN AND LEMIEUX, 2018) models, Appendix A provides an introduction to the concept of Influence Function - IF - which, in general terms, represents the influence of an individual observation on the distributional statistic of interest. Considering an influence function $IF(y; v)$, where y is an observation of the dependent variable Y, F_Y is the density of the cumulative distribution of Y and $v(\cdot)$ is a functional that estimates the statistic of interest, Firpo, Fortin and Lemieux (2009) define the RIF as shown in (7):

$$RIF(y; v) = v(F_Y) + IF(y; v) \quad (7)$$

In the context of this work, the distributional statistic is given by the quantile Q_τ; therefore, equation (7) could be rewritten as (8):

$$RIF(y; Q_\tau) = Q_\tau + IF(Q_\tau) \tag{8}$$

The RIF regression model resembles a standard linear regression, except that the dependent variable is replaced by the RIF of the statistic of interest, in this case,$RIF(y; Q_\tau)$. Thus, the income differential between women with and without children in a quantileQ_τ can be defined according to equation (9),

$$\Delta Q_\tau = E\left[RIF\{y, Q_{\tau_M}\}\right] - E\left[RIF\{y, Q_{\tau_{NM}}\}\right] \tag{9}$$

where $E\left[RIF\{y, Q_{\tau_M}\}\right] = \bar{X}^{M'}\hat{\beta}^M$ and $E\left[RIF\{y, Q_{\tau_{NM}}\}\right] = \bar{X}^{NM'}\hat{\beta}^{NM}$. In order to identify the composition and wage structure effects on ΔQ_τ, it is necessary to create a counterfactual scenario. The insertion of the counterfactual quantile Q_{τ_C} to calculate the wage differential is done as shown in (10),

$$\Delta Q_\tau = \underbrace{Q_{\tau_M} - Q_{\tau_C}}_{\Delta Q_{\tau_S}} + \underbrace{Q_{\tau_C} - Q_{\tau_{NM}}}_{\Delta Q_{\tau_X}} \tag{10}$$

in which ΔQ_{τ_X} and ΔQ_{τ_S} are the composition and wage structure effects, in that order. In a traditional extension of the Oaxaca-Blinder decomposition to quantiles, one would have $Q_{\tau_C} = \bar{X}^{M'}\hat{\beta}^{NM}$, i.e. the counterfactual quantile would reflect the characteristics of mothers, who are paid according to the wage structure of non-mothers. However, Firpo, Fortin and Lemieux (2018) point out that this definition of Q_{τ_C} may not be appropriate in cases of poor model specification or if the RIF does not reflect a good approximation of Q_τ. The authors therefore propose a strategy based on estimating a reweighting factor to obtain Q_{τ_C}, according to the methodology originally formulated by DiNardo, Fortin and Lemieux (1996). Equation (11) shows the definition of this reweighting factor, denoted by $\psi(X)$.

$$\psi(X) = \frac{1-p}{p}\frac{P(T = 1|\, X)}{1 - P(T = 1|\, X)} \tag{11}$$

In (11), p represents the proportion of individuals belonging to the group $T = 1$ – in the context of this work, the group of mothers –, and $P(T = 1|\, X)$ is the conditional probability of an individual with atributes X constituting the group $T = 1$. In practice, $\psi(X)$ can be measured by estimating binary choice models.

After obtaining the factors $\psi(X)$, the counterfactual quantile can be obtained by a regression model, so that $E\left[RIF\{y, Q_{\tau_C}\}\right] = \bar{X}^{c'}\hat{\beta}^c$. Once this regression has been estimated, the reweighted decomposition of ΔQ_τ takes the following form:

$$\Delta Q_\tau = \underbrace{\underbrace{\bar{X}^{M'}(\hat{\beta}_M - \hat{\beta}_c)}_{\Delta Q_{\tau_S}^p} + \underbrace{(\bar{X}^M - \bar{X}^c)'\hat{\beta}_c}_{\Delta Q_{\tau_S}^e}}_{\Delta Q_{\tau_S}} + \underbrace{\underbrace{(\bar{X}^c - \bar{X}^{NM})'\hat{\beta}_{NM}}_{\Delta Q_{\tau_X}^p} + \underbrace{\bar{X}^{c'}(\hat{\beta}_c - \hat{\beta}_{NM})}_{\Delta Q_{\tau_X}^e}}_{\Delta Q_{\tau_X}} \tag{12}$$

It can be seen from equation (12) that the reweighting procedure implies a "double decomposition". In the first stage, the differential ΔQ_τ is decomposed into ΔQ_{τ_X} and ΔQ_{τ_S}, which represent the aggregate composition effect and the aggregate wage structure effect, respectively. Next, each of these components is broken down again, generating two "pure" terms, $\Delta Q_{\tau_X}^{p}$ and $\Delta Q_{\tau_S}^{p}$, and two error terms, $\Delta Q_{\tau_X}^{e}$ and $\Delta Q_{\tau_S}^{e}$. The portion $\Delta Q_{\tau_X}^{e}$ is called "specification error" and is used to measure the quality of the model specification and the approximation made by the RIF. The portion $\Delta Q_{\tau_S}^{e}$ is called "reweighting error" and measures the quality of the specification of the binary choice model used to estimate the reweighting factors.

Based on the literature review carried out in section 2, there has been an discussion about the most appropriate method for estimating the motherhood penalty in the distribution of wages, since the results have been shown to be sensitive to the modeling chosen (BUDIG AND HODGES, 2010; KILLEWALD AND BEARAK, 2014; BUDIG AND HODGES, 2014). Due to the possibility of estimating the detailed decomposition by covariates, we chose to use the model proposed by Firpo, Fortin and Lemieux (2009, 2018). Another advantage of the method developed by the authors is that it is *path independent*, i.e. the order of the predictors does not influence the results of the model. Table 2 shows the explanatory variables used to apply the decomposition.

The RIF models were estimated using STATA 14 statistical software, which includes the *rif* package developed by Rios-Avila (2020). The RIF regression and decomposition models were estimated using the *rifhdreg* and *oaxaca_rif* commands available in this package. Logistic regressions were used to calculate the reweighting factors, implemented directly in the aforementioned commands using the *rwlogit* argument. Firpo, Fortin and Lemieux (2009) recommend estimating *bootstrap* standard errors for unconditional quantiles, so all models were estimated using *bootstrap* standard errors with 1000 replications.

Table 2. Explanatory variables in the decomposition models for the income differential between mothers and non-mothers

Variables	Description
Age	Continuous variable, defined by the number of years lived.
Age squared[2]	Continuous variable, defined by the square of the number of years lived.
Years of schooling	Multinomial variable for the woman's level of education: no schooling; primary education (incomplete or complete); secondary education (incomplete or complete); higher education (incomplete or complete).
Marital status	Binary variable for the woman's marital status. If she has a spouse in the household, it takes on a value of 1; otherwise, it takes on a value of 0.
Race/Color	Binary variable for the woman's race/color. If black or brown, it takes on a value of 1; if white, it takes on a value of 0.

(continued)

[2] In the main regression model, the age squared variable was constructed by centering it on the mean; thus, for a given age a_i , we have $a_i^2 = (a_i - a^{\mu})^2$. For the binary choice model used to calculate the reweighting factors, we simply calculated .$a_i^2 = a_i \cdot a_i$

(conclusion)

Variables	Description
Hours dedicated to household chores	Continuous variable, defined by the total number of hours worked in the week.
Region of the country	Multinomial variable for the main Brazilian regions: Southeast, South, Midwest, Northeast and North.
Occupational position	Multinomial variable for position in occupation: employee with signed work contract; employee without signed work contract; military or statutory employee; domestic worker with signed work contract; domestic worker without signed work contract; self-employed worker; employer.

Source: Own elaboration.

3.3.Methodological limitations

One issue that is relevant in studies on the labor market is sample selectivity, which has a significant impact on the results of the models and arises from different mechanisms. The estimation of wage equations implies the selection of women who participate in the labor market and earn positive incomes (i.e. those who report an income of zero are not considered). In this case, sample selection bias is corrected by estimating a *probit* model, whose dependent variable is whether or not the woman participates in the labor market. Next, the variable responsible for the correction is obtained, the Inverse Mills Ratio (λ)[3] , which is then inserted as a covariate in the income equation.

However, the inclusion of λ in models that use the Refocused Influence Function has not yet been widely applied in the literature. Studies such as Lemieux (2008), Firpo, Fortin and Lemieux (2009, 2018), Firpo and Pinto (2016) and Rios-Avila (2020) use RIF regressions to analyze income inequality, but do not correct for selection bias related to participation in the labor market. Since the inclusion of the Inverse Mills Ratio in RIF models would have unclear implications for the models, we opted not to correct for this source of bias.

Another source of sample selection bias, more specific to the context of this study, is related to motherhood status itself, which varies according to the moment in the life course as childless women migrate to the group of mothers. The latter becomes larger and larger as they get older, while the group of non-mothers becomes smaller and has specific characteristics. In order to overcome this selection bias, the reweighting procedure described in section 3.2 was used. Thus, decomposition models were estimated for the average and distribution of salaries with and without the reweighting procedure, in order to verify differences in both strategies.

In addition to the observations on selection bias, it is necessary to clarify the use of hourly wages as the dependent variable in the models. By using this measure instead of the level of eanrnings, the effect of the amount of work is suppressed, despite its importance in explaining the differences in the entry of mothers and non-mothers into the paid market. However, the use

[3] According to Greene (2000), selection bias correction models can be understood as special cases of censored models. In these, part of the distribution of a variable x is above or below some value a. Assuming that x has a normal distribution with mean μ and deviation σ , then the density function of a truncated distribution is given by$f(x \mid x > a) = \left[\frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\right]/[1 - \Phi(\alpha)]$, where$\alpha = (a - \mu)/\sigma$ and $\phi(.)$ and $\Phi(.)$ represent, respectively, the standard normal density function and the cumulative normal distribution function. The Inverse Mills Ratio, the variable responsible for correcting for selectivity bias, is determined by $\lambda(\alpha)$, where $\lambda(\alpha) = \phi(\alpha)/[1 - \Phi(\alpha)]$ if$x > \alpha$ and $\lambda(\alpha) = -\phi(\alpha)/\Phi(\alpha)$, if $x < \alpha$.

of hourly wages allows us to verify the possibility of mothers being confined to lower-paid occupations in search of greater flexibility, justifying its adoption in this study.

It is also important to comment on the potential endogeneity between having children and hourly wages, due to a possible bidirectional causality between the explanatory variable and the explained variable. In other words, while motherhood would contribute to lower wages for women with children compared to those without, women in the worst economic situation - and with the lowest incomes - would be the most likely to become mothers. In this regard, it should be noted that this study is limited to exploring the motherhood penalty in the Brazilian labor market and how it is constituted, so no hypotheses about the direction of causality were tested.

The problem of endogeneity can also be observed between hourly wages and the number of hours spent on household chores. The opportunity cost of time spent on housework tends to be higher for workers who earn high wages. They would therefore be more likely to outsource household production, reducing the time spent on chores. Therefore, the hours spent on housework would be endogenously determined by wages (HERSCH AND STRATTON, 1997; HERSCH, 2009).

Another situation concerns unobservable individual characteristics, such as an innate high work-oriented productivity. Since this is not an observable attribute, it can be interpreted as an error term in the wage equation. If workers with higher levels of innate market-oriented productivity specialize more in paid work to the detriment of domestic work, then there will be a negative correlation between the error term and the hours spent on chores (HERSCH AND STRATTON, 1997; HERSCH, 2009).

In both situations described above, the estimate of the parameter related to hours spent on household chores would be biased downwards, i.e. housework would have a greater negative impact on wages than it actually does. Despite these considerations, the empirical studies by Hersch and Stratton (1997) and Hersch (2009) attested to the exogenous relationship between hours spent on chores and hourly wages, so that the estimators were not biased.

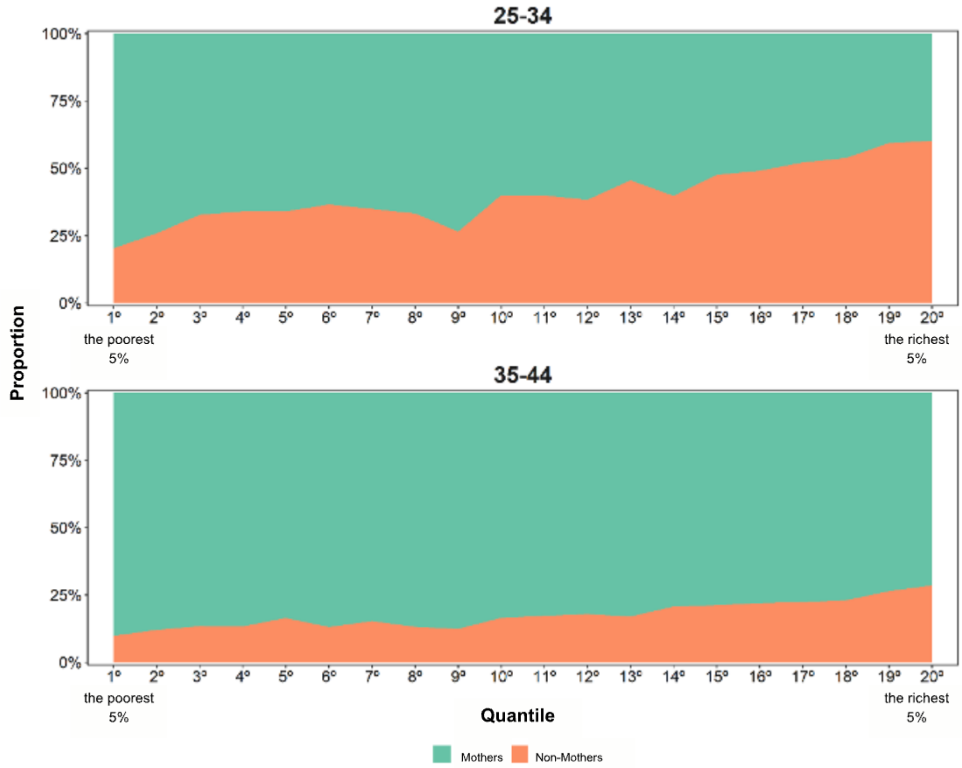
4. Results and Discussion

This section is divided into three parts. The first presents the profile of the women under analysis according to some variables of interest, which is related to the composition effects on the wages differential. The second subsection shows the behavior of hourly wages throughout the wage distribution by these variables, representing the "price effect" of the wage gap. Finally, the decomposition models estimated using RIFs (*Recentered Influence Functions*) are presented, in the light of the work by Firpo, Fortin and Lemieux (2018).

4.1. Composition of the profile of the women under analysis

This subsection presents the composition of the sample established according to some variables of interest, for their average values and in selected quantiles of hourly wages. But first, a brief general description of the sample of women with and without children along the distribution of wages is presented (Figure 1). There is a clear relationship between motherhood and hourly wages in both age groups, since there is a higher concentration of women with children in the first quantiles of the distribution and, similarly, a higher proportion of women without children in the higher quantiles. This configuration is consistent with studies on the phenomenon of *childlessness* in Brazil, which point to a higher prevalence of childlessness in the higher income strata (CAVENAGHI E ALVES, 2013; FUJIWARA, 2018; LEOCÁDIO, 2019; SIMÃO ET AL., 2020).

Figure 1. Composition of the sample of mothers and non-mothers aged between 25 and 34 and between 35 and 44 along the wage distribution - Brazil, 2015



Source: PNAD 2015 microdata (IBGE).

Figure 1also shows a predominance of mothers over non-mothers, which intensifies in the older age group, which was expected since the mother/non-mother ratio naturally tends to rise with increasing age. In addition, phenomena related to the dynamics of fertility in Brazil, such as the postponement of fertility and consequent increase in the average age at the birth of the first child (MIRANDA-RIBEIRO, GARCIA E FARIA, 2019) and the increase in *childlessness* mentioned above, contribute to a higher proportion of non-mothers among women in younger cohorts.

Table ***1*** shows the composition of the sample of mothers and non-mothers aged between 25 and 34, according to the variables of interest, andTable2 shows similar information for women aged between 35 and 44. With regard to participation in the labor market, non-mothers are more active in labor market than mothers in both age groups, but participation rates converge in the older age group. This may reflect the fact that mothers in the older age group probably have children at an age when it is no longer necessary to spend a large number of hours caring for them, making these women available for the paid market. On the other hand, mothers in the younger age groups usually have young children, who demand more of their time, often making paid work impossible. In the case of women without children, there is the question of the selectivity effect, since the number of women in the 35 to 44 age group who do not have children is relatively small, so they have very specific characteristics.

The distribution of women among the categories of education shows that mothers have a lower average level of education than non-mothers. In both age groups, there is a higher proportion of mothers with only primary education than non-mothers. On the other hand, the opposite is true for higher education, with a higher proportion of non-mothers compared to mothers. This also applies to the quantiles of the wage distribution, because even among the richest women, women with children have less schooling than those without children. Thus, there is a motherhood penalty in terms of schooling, regardless of income level.

In addition, it can be seen that women belonging to the oldest age group are more present in the categories with fewer years of schooling, regardless of whether or not they have children. This is a reflection of the expansion of access to higher education in the country in the mid-2000s, which led younger cohorts to increase their level of schooling (OLIVEIRA AND COLOMBI, 2014). Thus, women in the 35 to 44 age group did not benefit from the educational transition in the same way as women aged 25 to 34, since they were able to enter university at a time when access to higher education was increasing.

The race/color variable shows, on average, a higher proportion of black mothers than non-mothers, especially in the younger age group. Black women are doubly penalized compared to white women in the Brazilian job market, both because of the inequality of opportunities, which produces a gap between the attributes of both, and because of racial discrimination itself (PASSOS AND WAJNMAN, 2021). Thus, the higher proportion of black women among mothers may contribute to explaining the wage differential between mothers and non-mothers. This situation also occurs for the quantiles, although the proportion of white women rises at the top of the distribution for mothers and non-mothers.

The proportion of women with and without a spouse in the household reveals a greater presence of a spouse among mothers - which is to be expected, since the transition in marital status usually precedes the transition to motherhood (OLIVEIRA, RIOS-NETO AND OLIVEIRA, 2006). In addition, there is a direct relationship between the proportion of mothers with a spouse in the household and the level of income, since the higher the quantile of the wage distribution, the higher the proportion of women with a spouse in both age groups. This behavior reflects the vulnerability of the family arrangement composed of women without a spouse and with children, which is usually more recurrent among the lower income strata (SORJ ET AL, 2007; MADALOZZO AND BLOFIELD, 2017). For women without children, there is no such relationship between the presence of a spouse in the household and income.

With regard to the hours dedicated to household chores, there is a high number of hours dedicated by mothers, and little change throughout the wage distribution. For non-mothers, on the other hand, the inverse relationship between income level and hours spent on chores is clearer. This is because women in the higher income strata are better able to outsource household chores, but caring for other members of the household is not fully outsourceable. Analyzing this variable from an age perspective, there is a convergence in the number of hours dedicated by women with and without children in the older age group, with a reduction in the case of the former and an increase in the case of the latter.

Table 1. Composition of the sample of mothers and non-mothers aged between 25 and 34, according to variables of interest - Brazil, 2015

Variable	Mothers						Non-mothers					
	Mean	0,1	0,25	0,5	0,75	0,9	Average	0,1	0,25	0,5	0,75	0,9
Economically active	69,4%						83,0%					
Economically inactive	30,6%						17,0%					
No schooling	2,04%	2,81%	2,53%	1,29%	1,91%	1,34%	2,02%	1,18%	0,10%	0,40%	0,74%	0,94%
Primary education	29,95%	47,34%	30,09%	26,84%	17,43%	5,25%	8,20%	10,37%	8,54%	4,08%	0,58%	0,89%
Secondary education	51,16%	44,97%	59,57%	62,28%	54,76%	31,01%	42,81%	73,00%	70,46%	51,38%	19,11%	10,71%
Higher education	16,85%	4,88%	7,81%	9,58%	25,89%	62,40%	46,96%	15,45%	20,89%	44,14%	79,56%	87,46%
Black and Brown	56,99%	69,09%	65,65%	51,55%	50,84%	42,22%	45,03%	57,98%	58,39%	48,19%	30,58%	25,33%
White	43,01%	30,91%	34,35%	48,45%	49,16%	57,78%	54,97%	42,02%	41,61%	51,81%	69,42%	74,67%
With spouse in the household	53,23%	38,57%	47,38%	47,98%	53,62%	52,86%	26,75%	27,76%	24,87%	22,66%	24,40%	33,39%
No spouse in the household	46,77%	61,43%	52,62%	52,02%	46,38%	47,14%	73,25%	72,24%	75,13%	77,34%	75,60%	66,61%
Weekly hours to household chores	23,19	23,32	20,55	21,15	21,85	21,63	14,92	14,84	15,06	14,84	13,58	12,98
Employee with work contract	48,3%	9,0%	56,5%	58,7%	47,3%	42,3%	60,0%	66,7%	79,2%	63,8%	57,8%	48,6%
Employee without work contract	12,8%	22,9%	15,5%	10,6%	14,7%	10,1%	12,3%	13,0%	8,9%	13,1%	12,7%	10,9%
Military and statutory	7,4%	0,0%	8,3%	3,1%	8,2%	19,9%	10,0%	0,0%	3,5%	6,4%	15,7%	23,9%
Domestic workers with contract	3,4%	1,7%	8,1%	4,0%	3,0%	0,0%	2,2%	8,1%	2,5%	2,1%	0,0%	1,2%
Domestic workers without contract	10,2%	31,0%	6,9%	10,2%	8,9%	5,0%	3,1%	4,2%	2,4%	4,3%	1,4%	0,2%
Self-employed workers	16,0%	33,9%	4,5%	12,5%	16,2%	17,0%	10,5%	7,2%	3,1%	8,9%	8,6%	13,1%
Employers	1,9%	1,6%	0,2%	0,9%	1,8%	5,7%	1,9%	0,8%	0,4%	1,4%	3,9%	2,1%
Southeast	42,8%	24,4%	37,5%	49,9%	45,8%	45,7%	49,4%	33,7%	52,8%	44,7%	54,4%	60,6%
South	13,7%	9,9%	8,8%	18,5%	20,6%	18,4%	13,8%	7,0%	12,1%	18,8%	18,9%	16,0%
Northeast	25,8%	44,2%	35,2%	16,0%	17,4%	17,4%	22,5%	46,2%	25,4%	19,8%	12,2%	9,9%
North	9,2%	15,0%	9,9%	5,8%	7,0%	6,4%	5,9%	7,7%	4,5%	5,3%	5,3%	4,7%
Midwest	8,5%	6,5%	8,6%	9,9%	9,3%	12,0%	8,3%	5,5%	5,3%	11,5%	9,2%	8,8%

Source: PNAD 2015 microdata (IBGE).

Table2. Composition of the sample of mothers and non-mothers aged between 35 and 44, according to variables of interest - Brazil, 2015

Variable	Mothers						Non-mothers					
	Mean	0,1	0,25	0,5	0,75	0,9	Average	0,1	0,25	0,5	0,75	0,9
Economically active	73,0%						78,4%					
Economically inactive	27,0%						21,6%					
No schooling	3,89%	5,76%	3,38%	3,47%	1,83%	1,08%	5,17%	1,72%	0,63%	0,00%	0,34%	0,00%
Primary education	37,72%	57,43%	42,97%	35,01%	21,83%	6,77%	16,80%	31,83%	22,24%	12,12%	0,45%	3,88%
Secondary education	38,65%	32,90%	45,12%	50,16%	42,91%	23,45%	39,35%	52,55%	63,30%	44,92%	20,90%	9,93%
Higher education	19,75%	3,91%	8,52%	11,37%	33,43%	68,70%	38,67%	13,90%	13,83%	42,96%	78,30%	86,20%
Black and Brown	53,75%	67,33%	65,95%	51,65%	48,68%	36,37%	46,39%	70,34%	53,28%	47,37%	37,90%	36,23%
White	46,25%	32,67%	34,05%	48,35%	51,32%	63,63%	53,61%	29,66%	46,72%	52,63%	62,10%	63,77%
With spouse in the household	54,71%	45,05%	47,02%	51,32%	53,86%	58,31%	29,10%	25,55%	28,28%	36,92%	20,95%	41,26%
No spouse in the household	45,29%	54,95%	52,98%	48,68%	46,14%	41,69%	70,90%	74,45%	71,72%	63,08%	79,05%	58,74%
Weekly hours to household chores	22,39	23,52	19,70	20,82	21,24	20,09	16,32	16,11	17,00	15,41	15,02	13,79
Employee with work contract	37,6%	10,7%	44,2%	55,6%	34,2%	28,9%	47,0%	45,2%	62,1%	65,4%	41,4%	42,7%
Employee without work contract	10,2%	19,2%	16,5%	9,5%	9,3%	9,4%	8,8%	13,1%	9,0%	4,0%	6,2%	6,6%
Military and statutory	12,6%	0,5%	11,7%	9,7%	16,2%	32,6%	15,7%	0,0%	8,5%	11,5%	27,1%	25,5%
Domestic workers with contract	6,6%	5,6%	15,8%	8,9%	3,4%	0,8%	3,8%	9,9%	6,2%	3,3%	0,6%	0,9%
Domestic workers without contract	12,0%	27,8%	6,8%	6,3%	14,2%	3,3%	7,0%	19,0%	4,4%	4,0%	1,9%	1,4%
Self-employed workers	17,7%	34,4%	4,6%	9,8%	20,0%	14,2%	15,2%	10,3%	9,4%	10,5%	17,0%	18,5%
Employers	3,2%	1,9%	0,6%	0,3%	2,7%	10,7%	2,5%	2,4%	0,4%	1,4%	5,7%	4,5%
Southeast	45,6%	31,6%	32,9%	51,4%	48,4%	46,7%	50,5%	44,8%	49,5%	52,2%	56,0%	56,1%
South	13,9%	8,6%	7,5%	18,7%	19,1%	17,9%	12,9%	4,0%	10,6%	19,4%	14,4%	13,2%
Northeast	24,4%	41,6%	38,1%	13,5%	16,4%	16,8%	22,3%	36,6%	25,6%	12,3%	15,2%	15,8%
North	7,7%	11,7%	10,0%	6,6%	5,2%	7,8%	5,6%	6,3%	4,8%	3,3%	5,4%	6,3%
Midwest	8,4%	6,4%	11,6%	9,7%	10,9%	10,8%	8,7%	8,4%	9,5%	12,8%	8,9%	8,7%

Source: PNAD 2015 microdata (IBGE).

The occupational profile of the women under analysis shows a higher proportion of mothers working as self-employed and domestic workers without a formal contract compared to non-mothers. The latter, in turn, are more likely to work in positions that offer stability and protection, such as jobs with a formal contract and the civil service. The high proportion of mothers in self-employment and domestic service reflects the search for more flexible working hours, which allow for a better balance between work and motherhood (BUDIG AND ENGLAND, 2001). At the same time, the search for flexibility, especially among lower-income mothers, is often related to the precariousness of work and informality.

With regard to self-employment, it should be pointed out that not every activity classified as such is related to informality, since there are professionals who are self-employed and contribute to the social security system, characterizing them as formal. It is reasonable to infer, however, that this is not the situation for mothers at the lower end of the wage distribution. Women with children have high rates of informality, which in turn tend to be more prevalent among the low-income population (VILLANUEVA AND LIN, 2015). Domestic service, in turn, is traditionally characterized by a high degree of informality. In order to enter the market, workers in this sector tend to sell their working hours for competitive prices, which most of the time involves involuntarily giving up their labor rights (WAJNMAN, 2016).

Finally, there are marked regional differences between the quantiles of the wage distribution and between the different motherhood statuses. From an income perspective, in the lower part of the distribution there is a high proportion of women living in the North and Northeast regions, while in the upper part there is a predominance of women living in the Southeast, regardless of age group and number of children. On the other hand, when analyzing from the perspective of motherhood, there is a higher percentage of women without children in the Southeast, while mothers are more prevalent in the North and Northeast. Since the latter are economically less developed regions with higher levels of fertility (CAMARANO E CARNEIRO, 2016), the regional perspective may also be an explanatory factor for the motherhood penalty in Brazil.

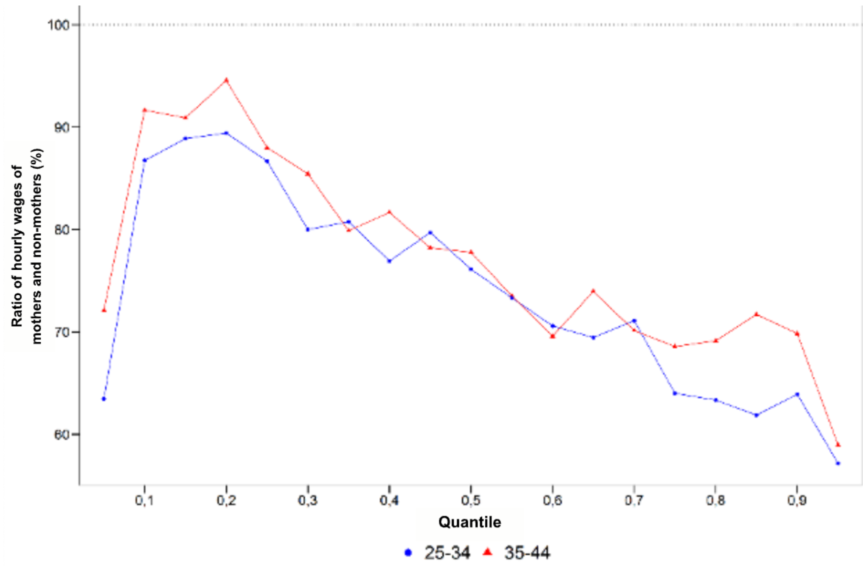
4.2.Wage structure of women with and without children according to different variables

This subsection shows the wage structure of women with and without children throughout the distribution, by age group. In order to make it easier the visualization, it was decided to show only the ratios of hourly wages between mothers and non-mothers for the two age groups under analysis. Thus, values less than unity (100%) denote that the hourly wages for mothers are lower than for non-mothers.

Figure ***2*** shows that the biggest difference in income occurs in the higher quantiles, so that the women most penalized by motherhood would be the richest. This result is consistent with studies in the literature (ANDERSON ET AL, 2002; ENGLAND, 2016), which justify the result by the higher opportunity cost that motherhood brings to such women. England (2016) argues that mothers belonging to the lower income strata would be less susceptible to the wage losses resulting from motherhood, since they would not show major differences in their productive attributes compared to low-income childless women.

Another factor that could explain the greater income gap in the higher quantiles is wage rigidity in the lower portion of the distribution, which occurs through a number of mechanisms. The presence of the minimum wage is one of them, since it contributes to greater income homogeneity in the lower quantiles. There is also the effect of the wage structure of occupations. Technological changes, for example, tend to have little effect on the wages of occupations typical of the lower income strata, making the heterogeneities in this part of the distribution smaller (LEMIEUX, 2008, AUTHOR ET AL, 2014).

Figure 2. Ratio of income/hour of mothers and non-mothers along the wage distribution, by age group - Brazil, 2015



Source: PNAD 2015 microdata (IBGE).

The greater penalty of motherhood among wealthier women also finds support in the national literature from the study by Souza (2016), which, like international studies, justified its results using the argument of the high opportunity cost among such women . Other studies carried out for the Brazilian context, such as Paulo (2013) and Muniz and Venerozo (2019), presented different conclusions: in the first study, several years were analyzed, and the quantile with the greatest difference between women with and without children varied. The second study, on the other hand, pointed to the existence of a maternity wage penalty only in the lower part of the wage distribution; it should be noted, however, that this study compared men and women rather than women with and without children.

Another relevant aspect observed inFigure ***2*** concerns the age perspective. It can be seen at the extremes of the distribution that the income ratio of women in the 35 to 44 age group is higher than the ratio of those aged between 25 and 34. Although women in both age groups do not belong to the same cohort, this may be an indication that mothers are unable to fully recover the effects of motherhood on their income at older ages.

Figure ***3*** shows the hourly wages ratio of women with and without children along the wage distribution, by age group, according to the variables of interest. In general, there is a similar pattern to that seen inFigure ***2*** , in which the motherhood penalty is more pronounced at the top of the wage distribution. From an age perspective, the differences between the two age groups are not very prominent for most of the variables under analysis.

The hourly wages ratios according to the different education categories show a discrepant behavior in the higher quantiles of women aged between 25 and 34 and those aged between 35 and 44 who have no formal education, but it should be borne in mind that this group of women composes a very small proportion of the sample. Despite this category of education, there is a motherhood penalty in the lower portion of the wage distribution for women with only primary education, especially for the younger curve; the same occurs for higher quantiles in the higher education category. This configuration probably reflects the fact that the typical occupations of those at the top of the distribution are related to a higher level of education; therefore, heterogeneities will be more visible at higher quantiles. At the lower end of the distribution, occupations may not require higher education, but primary education is usually necessary, which would explain the greater heterogeneities at lower quantiles for this category of education.

When analyzing the hourly wages ratios according to race/color, for most quantiles there is a greater penalty among white women - the exceptions are the 0.05 and 0.1 quantiles, in both age groups, and from the 0.8 quantile onwards, for women aged between 35 and 44. Some studies in the literature (WALDFOGEL, 1997; GLAUBER, 2007) have already pointed to a greater penalty for white women, although no clear explanation has been given for this result. It is inferred that the other penalties suffered by black women, arising from discrimination based on race and unequal opportunities, make the effect of motherhood on earnings lower for these women compared to white women.

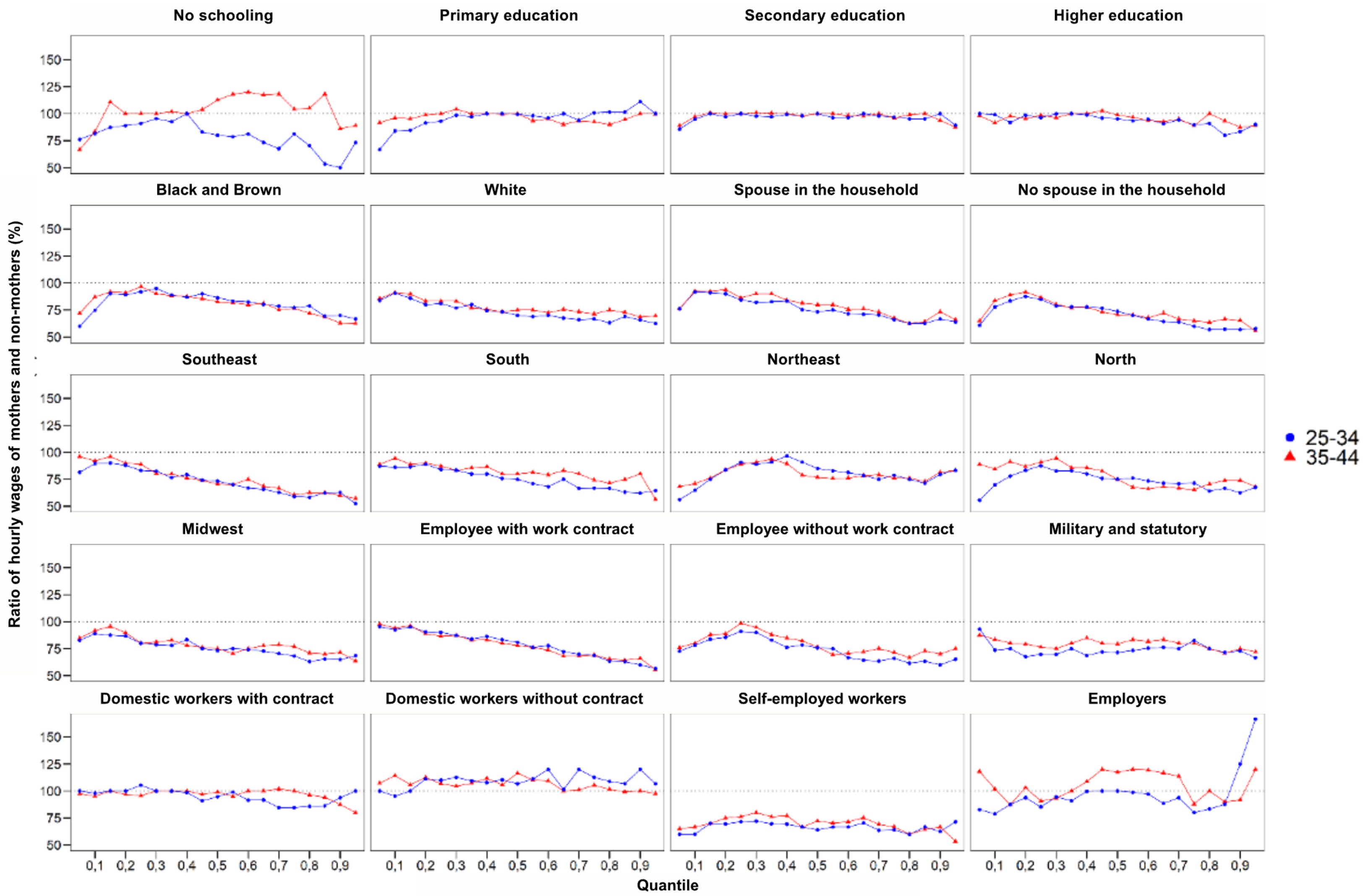
With regard to the presence of a spouse in the household, for almost the entire wage distribution there is a higher motherhood penalty among women who do not have a spouse. This result is in line with some studies in the literature (BUDIG AND ENGLAND, 2001; GLAUBER, 2007), which point to a higher motherhood penalty for married women compared to those who are not in a union. This is because married women spend more time on household chores, which has a negative impact on their productivity and, consequently, their wages. However, this mechanism is not necessarily valid for the Brazilian context, mainly due to the worse economic situation of women with children but no spouse.

The categories of the variable referring to the country's regions show that the motherhood penalty is higher in those regions where higher wages are earned - the Southeast, South and Center-West. On the other hand, the North and Northeast regions have hourly wages ratios closer to 100%, even at the top of the wage distribution, where the motherhood penalty has been more pronounced.

Finally, the position in occupation variable shows a downward movement in the ratios of hourly wages throughout the distribution for female employees with and without a formal contract, as well as for military and statutory employees. This shows that, in these occupations, the motherhood penalty is more pronounced at higher hourly wages levels. For domestic workers with and without a formal contract, the ratios are always close to 100%, which is explained by the wage rigidity typical of the occupation[4] , contributing to the homogeneity of the incomes of women with and without children. In relation to self-employed women, the ratios also show a relatively constant trajectory, but unlike domestic service, they indicate the existence of a motherhood penalty, which may reflect different types of activities carried out by mothers and non-mothers. In relation to the category of employers, the higher income of mothers at the top of the wage distribution stands out in the case of women aged between 25 and 34, and in almost the entire distribution for those in the 35 to 44 age group, producing hourly wages ratios of over 100%.

[4] In addition, the number of female workers belonging to this occupation is very small in the upper quantiles of the distribution.

Figure 3. Ratio of earnings/hour of mothers and non-mothers along the wage distribution, by age group, according to variables of interest - Brazil, 2015



Source: PNAD 2015 microdata (IBGE).

In short, the analysis of the hourly wages ratios of women with and without children for different variables indicated a more accentuated motherhood penalty for the more socially privileged groups: women with higher education, white, with a spouse in the household, in occupations offering better working conditions and living in more developed regions of the country. These findings reveal greater homogeneity among the poorest and most socially vulnerable women, so that motherhood has little impact on their income.

The following sections explore these issues in more depth, as will be shown from the estimated decomposition models, which will make it possible to verify the magnitude of the composition and wage structure effects on the wage gap, as well as assessing which variables contribute to or attenuate each of these effects.

4.3.Quantile decomposition models via RIF

The results of the RIF decompositions at the midpoint[5] of the wage distribution are shown below. Table ***3*** shows the following indicators: the natural logarithm of hourly wages for women with and without children; the counterfactual hourly wage of non-mothers - i.e. the wage they would receive if they had the same characteristics as mothers - and the wage gap between the two groups. It is worth noting that the estimation of the counterfactual hourly wage refers only to the model with reweighting, according to the procedure explained in section 3.2 of this paper.

Table 3. Average hourly earnings differential between mothers and non-mothers aged between 25 and 34 and between 35 and 44 - Brazil, 2015

	25 to 34 years		35 to 44 years	
	Coefficient	$exp\ (coef)$	Coefficient	$exp\ (coef)$
Mothers	2.121***	R$ 8.34	2.241***	R$ 9.40
	(0.00846)		(0.00772)	
Non-mothers	2.427***	R$ 11.32	2.520***	R$ 12.43
	(0.0117)		(0.0184)	
Counterfactual (with reweighting)	2.159***	R$ 8.66	2.172***	R$ 8.78
	(0.0403)		(0.0489)	
Total difference	-0.306***	0.74	-0.279***	0.76
	(0.0142)		(0.0203)	

Source: PNAD 2015 microdata (IBGE).
Note: Statistical significance levels:*** $p \leq 0.01$,** $p \leq 0.05$,* $p \leq 0.1$. Bootstrap standard errors were used throughout the procedure (1000 replications) to calculate the p-value.

The results ofTable ***3*** show that motherhood is penalized in both age groups, since the hourly wages of mothers are lower than those of non-mothers. In addition, there is little change in the hourly wages differential between the two age groups. Another aspect of interest is the counterfactual income estimated by the reweighted model. For women aged between 25 and 34, its value is R$8.66, indicating a significant reduction in relation to the R$11.32 previously

[5] It should be noted that estimating a RIF regression without reweighting to the mean is equivalent to estimating a simple linear regression model, since in this case the RIF of any observation y_i will be the observation itself. Similarly, estimating a RIF decomposition without reweighting to the mean provides similar results to those produced by the Oaxaca-Blinder decomposition. In order to make it feasible to estimate the reweighted models and compare them with the models without reweighting, it was decided to also use RIF regression for the decompositions of average hourly wages.

found for women without children in this age group. For childless women aged between 35 and 44, the counterfactual income/hour is R$8.78, which is lower than that already estimated for childless women in this age group (R$12.43) and also lower than the mothers' own hourly wage of R$9.40. This result may once again be related to the specificities of the group of older childless women caused by the selectivity effect.

Table ***4*** shows the average hourly earnings into two parts: the portion explained by the attributes of the women under analysis (composition effect) and the portion not explained by these attributes (wage structure effect). The unweighted model showed that both the composition effect and the wage structure effect were statistically significant for both age groups; however, the contribution of the former was higher, at 77.9% for women aged between 25 and 34 and 78.5% for those aged between 35 and 44. In other words, this means that mothers have attributes that put them in a worse position in the labor market compared to non-mothers, causing them to receive lower wages. This result is in line with one of the hypotheses put forward by Budig and England (2001) to explain the motherhood wage penalty, which concerns the differences in productive characteristics between women with and without children.

The application of reweighting led to changes in the contribution of each group to the income gap. There was a greater share of the composition effect in both age groups - 85.5% for women aged between 25 and 34 and 124.9% for those aged between 35 and 44 - even with the equalization of characteristics in the reweighted model. The wage structure effect, in turn, was no longer statistically significant for the two age groups. The results therefore suggest that the difference in average hourly wage between women with and without children is only due to the composition effect, i.e. the differences in attributes between the women under analysis.

Table 4. Average hourly earnings of mothers and non-mothers aged between 25 and 34 and between 35 and 44, with and without reweighting, by effect type - Brazil, 2015

Type of model	Share of differential	25 to 34 years		35 to 44 years	
		Coefficient	%	Coefficient	%
Without reweighting	Composition effect	-0.238*** (0.0142)	77.9	-0.219*** (0.0184)	78.5
	Salary structure effect	-0.0676*** (0.0159)	22.1	-0.0600*** (0.0209)	21.5
	Total difference	-0.306*** (0.0142)	100	-0.279*** (0.0203)	100
With reweighting	Composition effect	-0.267*** (0.0398)	87.5	-0.348*** (0.0486)	124.9
	Wage structure effect	-0.0382 (0.0407)	12.5	0.0694 (0.0491)	-24.9
	Total difference	-0.306*** (0.0142)	100	-0.279*** (0.0203)	100

Source: PNAD 2015 microdata (IBGE).

Note: Statistical significance levels:*** $p \leq 0.01$,** $p \leq 0.05$,* $p \leq 0.1$. Bootstrap standard errors were used throughout the procedure (1000 replications) to calculate the p-value.

A detailed analysis of the composition effect (Table ***5***) shows that, regardless of the reweighting procedure, the main contributions to the difference in average income/hour between women with and without children aged between 25 and 34 come from the variables of schooling, weekly hours spent on household chores and position in the occupation. In other words, if mothers had the same level of education, dedicated the same number of hours to household chores and held similar positions to non-mothers, their hourly wage would be higher. For women aged between 35 and 44, education and position in the occupation remained the main contributors, along with the region variable, which was less important in the younger group. Hours spent on household chores, on the other hand, were no longer statistically significant for the older age group.

The differences between mothers and non-mothers for the attributes mentioned above - schooling, hours devoted to chores and position in the occupation - had already been shown in section 4.1 of this paper. In terms of level of education, there was a high percentage of mothers (especially younger mothers) who only had primary education, while non-mothers were more likely to have secondary and higher education. Thus, it can be inferred that the motherhood implies an interruption, often permanent, in investment in qualifications.

With regard to hours spent on household chores, the impact on mothers in the younger age group probably reflects the age of their children, since the proportion of young children among them is higher than among older women. Given that caring for young children takes many hours, especially when they have not yet reached school age, it is to be expected that the younger age group will be more penalized. Finally, the contribution of position in the occupation is justified by the high proportion of mothers working as self-employed and domestic workers without a formal contract, occupations characterized by low wages and greater flexibility.

Table 5. Composition effect for the average hourly earnings of mothers and non-mothers aged between 25 and 34 and between 35 and 44, with and without reweighting - Brazil, 2015

	Without reweighting				With reweighting			
	25 to 34 years		35 to 44 years old		25 to 34 years old		35 to 44 years	
	Coefficient	%	Coefficient	%	Coefficient	%	Coefficient	%
Age	0.0390***	-16%	0.00154	-1%	0.0468***	-17%	0.00128	0%
	(0.00475)		(0.00345)		(0.00679)		(0.00405)	
Schooling	-0.192***	81%	-0.182***	83%	-0.216***	81%	-0.235***	68%
	(0.00903)		(0.0132)		(0.0131)		(0.0202)	
Spouse	0.0175***	-7%	0.0123*	-6%	0.0175***	-7%	0.0188*	-5%
	(0.00421)		(0.00714)		(0.00538)		(0.0107)	
Race / Color	-0.0133***	6%	-0.00592**	3%	-0.0182***	7%	-0.00498	1%
	(0.00232)		(0.00250)		(0.00489)		(0.00486)	
Hours household chores	-0.0495***	21%	-0.0172*	8%	-0.121***	45%	-0.0514	15%
	(0.00879)		(0.00942)		(0.0350)		(0.0314)	
Brazilian Region	-0.0116***	5%	-0.00815**	4%	0.00116	-0.4%	-0.0375***	11%
	(0.00280)		(0.00393)		(0.00843)		(0.0128)	
Position in occupation	-0.0281***	12%	-0.0191***	9%	-0.0364***	14%	-0.0378***	11%
	(0.00551)		(0.00602)		(0.00956)		(0.0124)	
Specification error	-	-	-	-	0.0583**	-22%	-0.00124	0%
					(0.0279)		(0.0190)	
Composition effect	-0,238***	100%	-0,219***	100%	-0.267***	100%	-0.348***	100%
	(0.0142)		(0.0184)		(0.0398)		(0.0486)	

Source: PNAD 2015 microdata (IBGE).
Note: Statistical significance levels:*** $p \leq 0.01$,** $p \leq 0.05$,* $p \leq 0.1$. *Bootstrap* standard errors were used throughout the procedure (1000 replicates) to calculate the p-value.

Still on the detailed composition effect (Table ***5***), age and the presence of a spouse in the household are the main factors mitigating the wage differential between women with and without children aged between 25 and 34. Thus, if mothers in the younger age group were similar to non-mothers in terms of age and marital status, their wages would be lower. With regard to the last variable, the result reflects the positive effect of the presence of a spouse for mothers already discussed in the descriptive analysis, in which the greater vulnerability of the family arrangement made up of women without a spouse and with children was also found. For women aged between 35 and 44, both variables were not statistically significant.

In the specific context of the reweighted model, the attenuating impact of the specification error on the wage gap between mothers and non-mothers aged between 25 and 34 is noteworthy. As presented in section 3.2, the specification error is a measure of the accuracy of the estimated model. When its coefficient is statistically significant, it means that other factors not included in the model have an influence on the wages gap. Thus, it can be said that other variables not specified in the model estimated for women aged between 25 and 34 would act to reduce the wage gap between mothers and non-mothers.

With regard to the effect of the wage structure (Table 6), it can be seen for the model without reweighting that education and hours spent on chores were the only statistically significant variables for women aged between 25 and 34. In other words, if mothers and non-mothers had the same returns in terms of schooling and hours spent on chores, the income gap between the two would be smaller. For women aged between 35 and 44, no statistically significant variables were observed. Among the coefficients of the wage structure effect for the reweighted models, the only ones that were statistically significant were those relating to the race/color variable and the reweighting error, both for the 35-44 age group. With regard to race/color, the results show that the wage gap between mothers and non-mothers would be smaller if they received similar returns for this attribute. The reweighting error, in turn, is a measure of the quality of the probit model used in the reweighting procedure. Firpo, Fortin and Lemieux (2018) point out that, in large samples, the reweighting error tends to be statistically equal to zero; thus, its relevance for the 35-44 age group may be related to the small sample size, especially with regard to non-mothers.

Table 6. Effect of salary structure on the average hourly earnings of mothers and non-mothers aged between 25 and 34 and between 35 and 44, with and without reweighting - Brazil, 2015

	Without reweighting				With reweighting			
	25 to 34 years		35 to 44 years old		25 to 34 years old		35 to 44 years	
	Coefficient	%	Coefficient	%	Coefficient	%	Coefficient	%
Age	0.348	-520%	0.886	-1475%	-0.0541	145%	0.954	1375%
	(0.698)		(0.683)		(1.206)		(0.979)	
Schooling	0.310**	-463%	0.0510	-85%	0.113	-302%	0.172	248%
	(0.129)		(0.218)		(0.171)		(0.271)	
Spouse	-0.0144	22%	-0.0148	25%	-0.0299	80%	0.0163	23%
	(0.0130)		(0.0190)		(0.0245)		(0.0251)	
Race / Color	0.0293*	-44%	-0.0157	26%	0.0336	-90%	-0.0803***	-116%
	(0.0154)		(0.0218)		(0.0323)		(0.0264)	
Hours household chores	0.0933***	-139%	0.00702	-12%	-0.0170	45%	0.00391	6%
	(0.0273)		(0.0375)		(0.0548)		(0.0362)	
Brazilian Region	0.0294	-44%	0.0215	-36%	-0.0223	60%	0.0446	64%
	(0.0181)		(0.0240)		(0.0384)		(0.0329)	
Occupational position	-0.000485	1%	0.0109	-18%	0.0169	-45%	0.00201	3%
	(0.0134)		(0.0221)		(0.0273)		(0.0265)	
Constant	-0.862	1289%	-1.006	1674%	-0.117	313%	-1.182	-1703%
	(0.699)		(0.695)		(1.217)		(0.977)	
Re-weighting error	-	-	-	-	0.0394	-105%	0.139***	200%
					(0.0327)		(0.0428)	
Salary structure effect	-0.0676***	100%	-0.0600***	100%	-0.0382	100%	0.0694	100%
	(0.0159)		(0.0209)		(0.0407)		(0.0491)	

Source: PNAD 2015 microdata (IBGE).
Note: Statistical significance levels:*** $p \leq 0.01$,** $p \leq 0.05$,* $p \leq 0.1$. *Bootstrap* standard errors were used throughout the procedure (1000 replicates) to calculate the p-value.

The results of the estimated decompositions for quantiles 0.05 to 0.95 are presented below. The coefficients of these models are shown in Figure 4 to Figure 7[6] . Analyzing the difference in hourly wage between women with and without children, it can be seen that the motherhood penalty is increasing throughout the wage distribution, since the coefficients are statistically significant in all quantiles and denote a negative effect which becomes greater as one moves towards the top of the distribution (Figure 4).

As pointed out in previous sections of this paper, the greater penalty at the top of the distribution is supported by literature (ANDERSON ET AL, 2002; ENGLAND, 2016, SOUZA, 2016). The lower accumulation of human capital among women belonging to the lower social strata makes them less susceptible to wage erosions resulting from motherhood compared to those located in higher social strata, who would face a higher opportunity cost. In addition, the effect of wage rigidity in the lower quantiles caused by the minimum wage must be taken into account, reducing the income gap in the lower quantiles of the distribution.

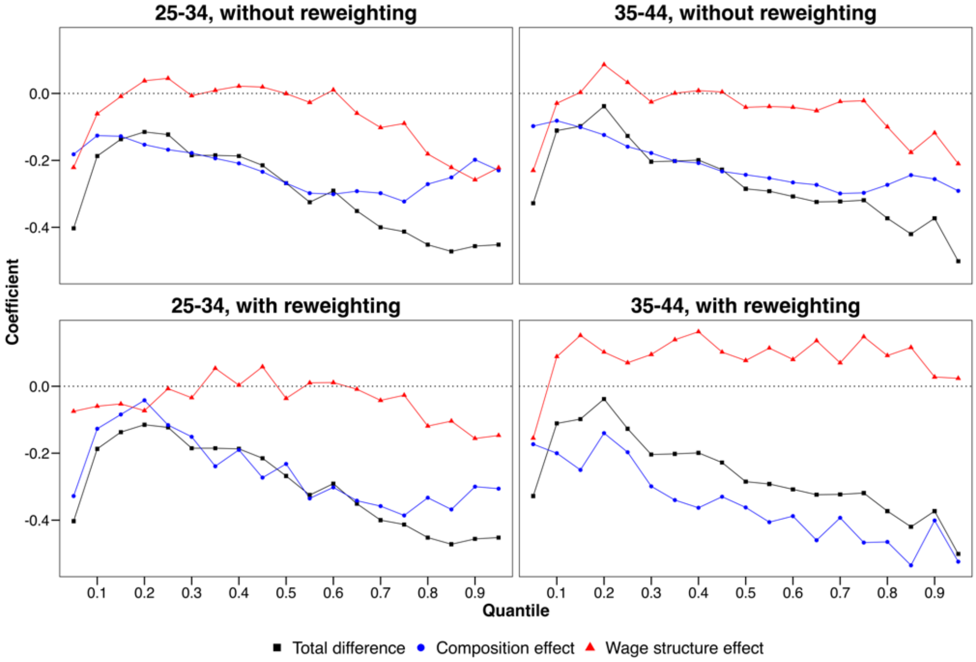
The finding of a greater motherhood penalty at higher quantiles of the distribution applies to both women aged between 25 and 34 and those aged between 35 and 44. However, for most quantiles, the wage gap is higher in the younger age group. Specifically, this situation occurs between the 0.05 and 0.20 quantiles, as well as between the 0.65 and 0.90 quantiles, although the differences between the two age groups are generally small.

The composition effect was statistically significant throughout the wage distribution - with the exception of the 0.02 quantile of the reweighted models - and contributed the largest share of the wage gap in both age groups, although it was smaller in magnitude for the older age group in most quantiles. Considering the models without reweighting, it can be seen that the negative effect of the difference related to characteristics is linearly increasing up to the 0.75 quantile for women aged between 25 and 34, with a consequent attenuation of the differential in higher order quantiles. For women aged between 35 and 44, this attenuating effect at the top of the distribution becomes more discrete. With the application of the reweighting procedure, the negative impact of the portion related to characteristics became greater, and for the oldest age group, it began to exhibit a decreasing behavior throughout the distribution.

The effect of the wage structure - the portion not explained by attributes - accounted for a smaller share of the differential for both age groups. For women aged between 25 and 34, the pattern shown by the wage structure effect over the quantiles changed little when comparing the models with and without reweighting. However, it should be noted that for the reweighted models, only the 0.15 and 0.20 quantiles were statistically significant, while in the models without reweighting, the 0.05; 0.10; 0.20; 0.25 and 0.65 and above quantiles were statistically significant. The opposite situation occurred among women in the 35-44 age group, for whom a greater number of statistically significant quantiles were observed after applying the reweighting procedure. Thus, controlling for the selectivity related to motherhood *status* meant that the salary structure effect appeared more strongly for the older age group, for which there was a greater incidence of selection bias.

[6] The tables with the complete estimates have not been added to this work, but can be made available on request to the authors.

Figure 4. Quantiles of hourly wages of mothers and non-mothers aged between 25 and 34 and between 35 and 44 - Brazil, 2015



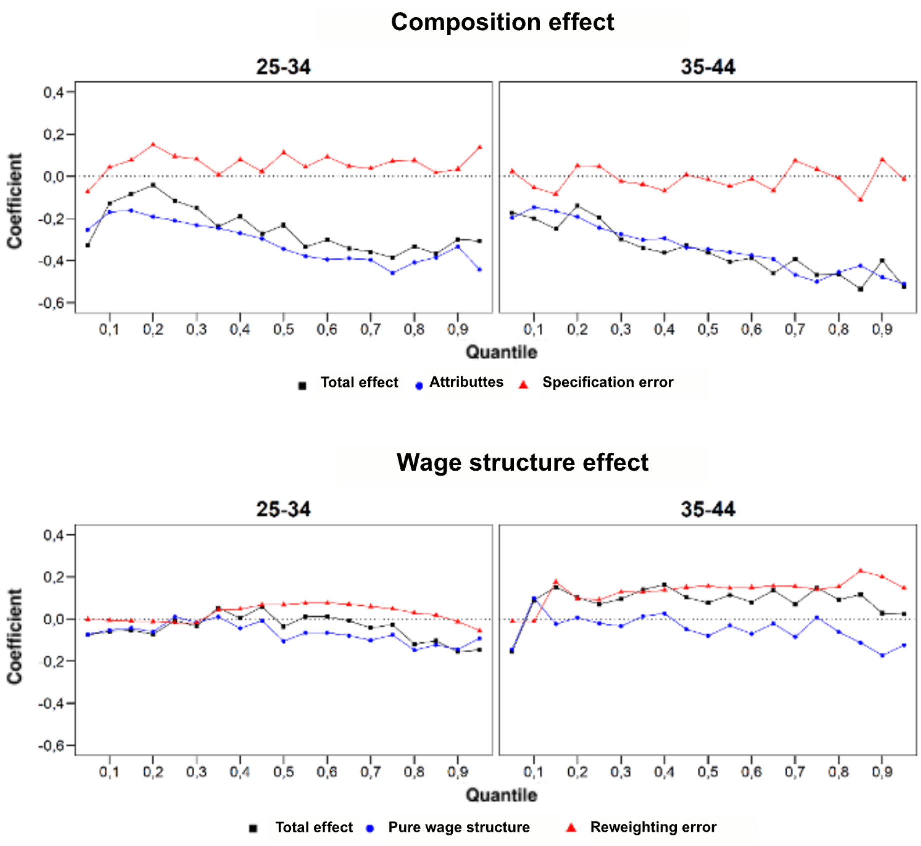
Source: PNAD 2015 microdata (IBGE).

Figure 5 shows the composition and wage structure effects for the reweighted estimates. As shown in section 3.2 of this paper, the application of the reweighting procedure involves a two-stage decomposition, with each of the two main parts - the composition effect and the wage structure effect - being broken down into a "pure" effect and an error term. Thus, the composition effect is split between the portion of the differential related to attributes and the specification error, which is a measure of the quality of the estimated decomposition model. For the wage structure effect, the contribution of returns to attributes is separated from the reweighting error, which measures the quality of the *probit* model used to equalize the characteristics of the two groups of interest.

Analysis of the composition effect shows that it is mostly made up of the differential related to attributes, which was statistically significant in all quantiles of the wage distribution for the two age groups of interest. As can be seen inFigure **5** , the contribution of the attributes over the course of the quantiles resembles the behavior of the composition effect itself. The specification error, in turn, proved to be small in magnitude and did not show statistical significance for most of the quantiles. In particular, the residual portion of the composition effect was statistically significant between the 0.15 and 0.30 quantiles and in the 0.50 and 0.60 quantiles for women aged between 25 and 34; for those in the 35 to 44 age group, only the 0.85 quantile coefficient was statistically significant. The greater number of coefficients statistically different from zero among younger women indicates that other variables not specified in the model are important in explaining the income gap between mothers and non-mothers in this age group, which had already been verified when analyzing the decomposition of average hourly wages.

As far as the aggregate wage structure effect is concerned, it can be seen that the largest part of this is the differentials in the returns to attributes, which was statistically significant in all quantiles for both age groups. The reweighting error, on the other hand, was not statistically significant for women aged between 25 and 34; however, for those aged between 35 and 44, it was statistically significant at the 0.15 quantile and between the 0.30 and 0.95 quantiles. As explained above, the reweighting error is expected to tend towards zero in large samples. The fact that this error is different from zero in many quantiles for the older age group may be associated with the small sample size, especially for women without children.

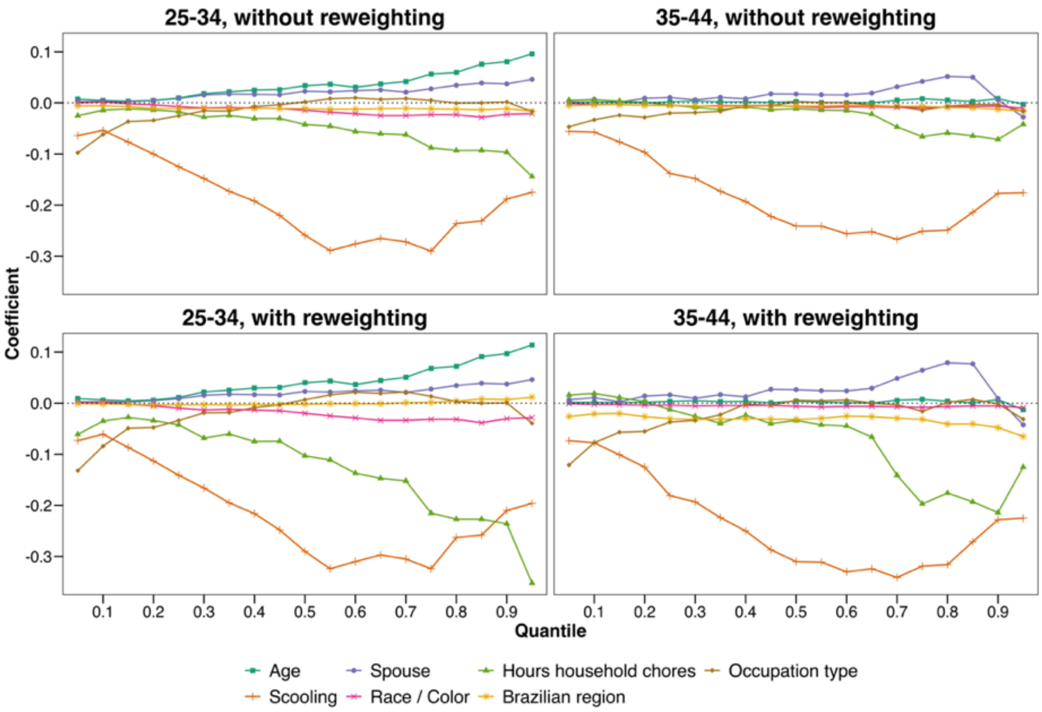
Figure 5. Components of the composition and wage structure effect of the decomposition with reweighting of the hourly wages quantiles of mothers and non-mothers aged between 25 and 34 and between 35 and 44 - Brazil, 2015



Source: PNAD 2015 microdata (IBGE).

Analyzing the composition effect detailed by the covariates along the wage distribution (Figure **6**), it can be seen that schooling, hours dedicated to chores and position in the occupation represented the greatest contributions to the wage gap between mothers and non-mothers, in a similar way to what happens with the models estimated for average hourly wages. The differences in schooling were statistically significant in all quantiles of the distribution in both age groups. With regard to hours dedicated to chores, there was statistical significance in all quantiles of the wage distribution for women aged between 25 and 34, while for those aged between 35 and 44, the variable was statistically significant only between quantiles 0.7 and 0.9. The position in occupation variable was statistically different from zero up to the 0.35 quantile for both age groups. These results apply to the models with and without reweighting.

Figure 6. Composition effect detailed by the variables of the decomposition of the quantiles of the hourly wages of mothers and non-mothers aged between 25 and 34 and between 35 and 44 - Brazil, 2015



Source: PNAD 2015 microdata (IBGE).

The coefficients of the education variable along the quantiles show a similar pattern for both age groups, with an increasing impact up to approximately the 0.7 quantile, attenuating at the top of the distribution. This behavior can be explained by the smaller relevance of qualification in occupations typical of the lower income quantiles, a scenario that gradually changes as the individuals move up the wage distribution. In the higher quantiles, many mothers are able to access higher levels of education, contributing to the attenuation of educational differentials. In terms of the level of the coefficients, the negative effect of education differentials is slightly higher for mothers aged between 25 and 34, for the model without reweighting. In the case of the reweighted models, the differences between the two age groups become less noticeable, since the negative effect of the education differentials became greater for women aged between 35 and 44.

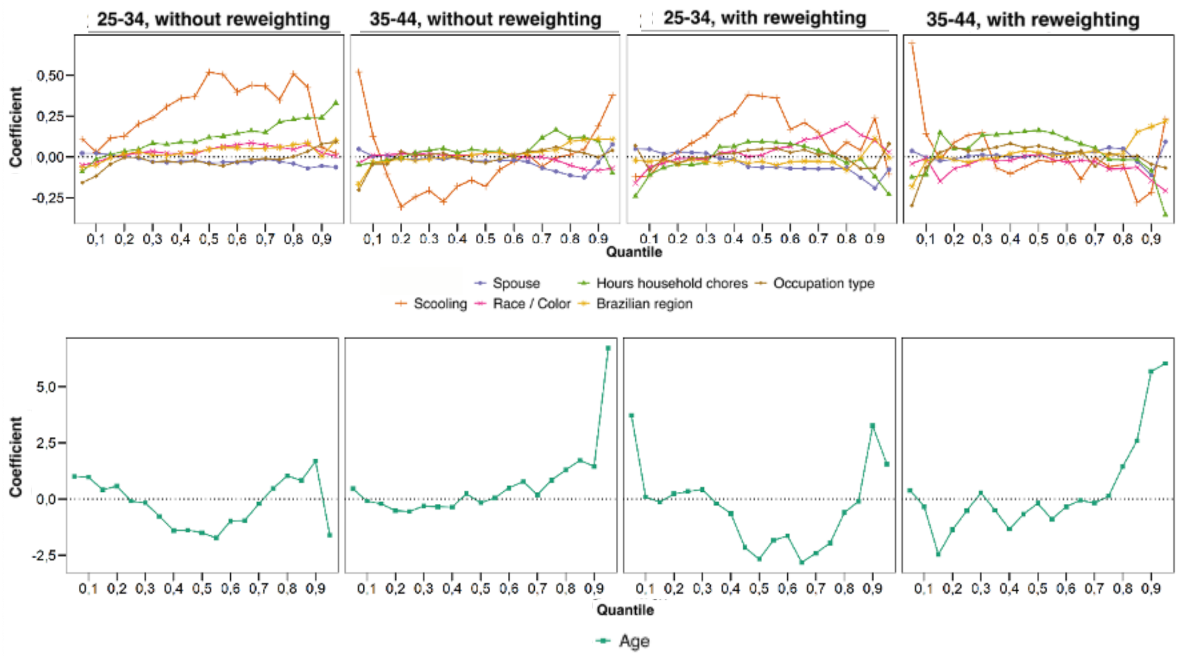
With regard to the chores variable, section 4.1 showed a negative relationship between the number of hours spent by women without children and their level of income; however, similar behavior was not observed for mothers, since even in the higher quantiles of the distribution, they spent a high number of hours a week on household chores. Although wealthier mothers have the economic means to outsource domestic and care tasks, it should be borne in mind that many of these tasks may not be easily attributed to third parties, especially those involving a direct dependent, such as children (FOLBRE AND NELSON, 2000). This may be one reason why the hours devoted by mothers to chores are high even in the higher quantiles, causing a greater differential in this part of the wage distribution. With regard to the age perspective, we assume the same hypothesis as for the decomposition of average hourly wages, that the presence of young children makes the penalty related to hours spent on chores more relevant for younger women.

The behavior observed for the occupational position variable can be explained by the difference in the occupational composition of the group of mothers and non-mothers in the lower quantiles, since for mothers there was a high proportion of self-employed women and domestic workers without a formal contract, differing from the occupational composition found for non-mothers in the same quantiles. In the upper part of the distribution, the differences between women with and without children in terms of occupational position categories are smaller, thus explaining why the variable was not statistically significant for quantiles above 0.35.

Finally, for the models without reweighting, the variables of age and the presence of a spouse in the household attenuated the wage differential between mothers and non-mothers. Among women in the 25-34 age group, this effect is greater as they move up the wage distribution, in both variables. In the older age group, age was no longer statistically significant, while the presence of a spouse in the household remained relevant in most quantiles. When the reweighting procedure was applied, both variables lost statistical relevance for women aged between 25 and 34. For those in the oldest age group, the age variable remained statistically insignificant, while the presence of a spouse in the household proved to be different from zero between quantiles 0.45 and 0.55 and between 0.65 and 0.85.

Figure **7** shows the wage structure effect detailed by the model's covariates. Regardless of the reweighting procedure, this effect was not statistically significant for most quantiles, in all variables and in both age groups. It is therefore not possible to make inferences from its coefficients.

Figure 7. Wage structure effect detailed by the variables in the decomposition of the quantiles of the hourly earnings of mothers and non-mothers aged between 25 and 34 and between 35 and 44 - Brazil, 2015



Source: PNAD 2015 microdata (IBGE).
Note: Coefficients of the age variable shown separately due to the difference in scale in relation to the other variables.

5. Conclusion

The results of the decomposition models showed the great importance of the attributes of mothers and non-mothers in explaining the wage differential between these two groups, regardless of age group. The greater importance of the composition effect was observed both for average hourly wages and for the quantiles of the wage distribution. With regard to the quantile decompositions, it was found that the motherhood penalty is greater at the top of the distribution, which is in line with previous studies on the subject. The situation of vulnerability faced by women with and without children at the bottom of the distribution makes the two groups more homogeneous in terms of their attributes, reducing the impact of motherhood on earnings.

From an age perspective, there was a slight reduction in the wage differential between mothers and non-mothers in the older age group compared to the younger group. However, it cannot be said that this result is due to an improvement in the characteristics of mothers, since the composition effect was greater in the 35-44 age group. This occurred in the models estimated for average hourly wages with and without the reweighting procedure, as well as for the models reweighted for the quantiles of the wage distribution.

The attributes that made the greatest contribution to widening the wage gap between mothers and non-mothers were schooling, hours devoted to household chores and position in the occupation, in both age groups analyzed. The quantile decompositions showed that the differentials in schooling increased up to approximately the 0.7 quantile, and then narrowed from that point onwards. On the other hand, the differentials in hours dedicated to chores became increasingly larger as the wage distribution progressed. This result is particularly interesting, as wealthier women can afford to outsource chores, but still suffer a greater penalty in this variable. On the other hand, the differences in occupation were only significant up to the 0.35 quantile, due to the high proportion of mothers working as self-employed workers and in domestic service without a formal contract.

On the other hand, the variables of age and the presence of a spouse in the household attenuated the income differential between women with and without children. With regard to the age variable, this effect was only relevant among younger women. The importance of the spouse reveals the vulnerability faced by women in a single-parent arrangement with children, since they generally have an inelastic labor supply, making them more willing to accept lower wages. Thus, the results show that these women would have higher wages if they had a spouse in the household.

The estimation of the detailed decomposition models by covariates showed the need to promote mothers' access to qualifications, given the large impact of the schooling variable on the income differential between mothers and non-mothers. To this end, it is necessary for mothers to receive support in carrying out care tasks, which still fall mostly to women. Consideration should be given to the role of public policies, such as the increased provision of child care services, which are of great importance to mothers who cannot afford to outsource care tasks. The expansion of policies related to domestic workers could also act collaterally in reducing the penalty among low-income women, given the high percentage of mothers with this occupational profile.

It should also be borne in mind that the results of this study showed that the penalty for hours spent on household chores is greater for women at the top of the wage distribution, who are precisely those for whom outsourcing household chores is economically viable. This scenario shows that not all tasks are fully "outsourceable", especially those related to caring for children and other household members. Although public sector action in this regard is more limited, there are some policies that can mitigate inequality in the distribution of tasks within the home, and therefore reduce the burden of care tasks on mothers. One of them, for example, concerns maternity leave, which in many countries has already been replaced by a shared leave, allowing the couple to jointly decide who will take time off work and for how long. Above all, awareness needs to be raised about the importance of a more equal division of tasks in the private sphere.

REFERENCES

ABENDROTH, Anja-Kristin; HUFFMAN, Matt L.; TREAS, Judith. The parity penalty in life course perspective: Motherhood and occupational status in 13 European countries. *American Sociological Review*, v. 79, n. 5, p. 993-1014, 2014.

ANDERSON, Deborah J.; BINDER, Melissa; KRAUSE, Kate. The motherhood wage penalty: Which mothers pay it and why? American economic review, v. 92, n. 2, p. 354-358, 2002.

___________. The motherhood wage penalty revisited: Experience, heterogeneity, work effort, and work-schedule flexibility. *ILR Review,* v. 56, n. 2, p. 273-294, 2003.

AUTOR, David H. et al. Trade adjustment: Worker-level evidence. *The Quarterly Journal of Economics*, v. 129, n. 4, p. 1799-1860, 2014.

BUDIG, Michelle J.; ENGLAND, Paula. The wage penalty for motherhood. *American sociological review,* p. 204-225, 2001.

___________; HODGES, Melissa J. Differences in disadvantage: Variation in the motherhood penalty across white women's earnings distribution. *American Sociological Review*, 2010, 75.5: 705-728.

CAMARANO, Ana Amélia; CARNEIRO, Isabella Gomes. Padrões de formação de família por regiões brasileiras e grupos sociais: diferenças ou semelhanças? *Anais do XI Encontro Nacional de Estudos Populacionais*, p. 31-50, 2016.

CAVENAGHI, Suzana; ALVES, José Eustáquio Diniz. Childlessness in Brazil: socioeconomic and regional diversity. In: *Ponencia presentada en el XXVII IUSSP International Population Conference*, Busan Korea. 2013.

DINARDO, John; FORTIN, Nicole M.; LEMIEUX, Thomas. Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica* 64, no. 5, p. 1001–44, 1996.

ENGLAND, Paula et al. Do highly paid, highly skilled women experience the largest motherhood penalty? *American Sociological Review*, v. 81, n. 6, p. 1161-1189, 2016.

FIRPO, Sergio; FORTIN, Nicole M.; LEMIEUX, Thomas. Unconditional quantile regressions. *Econometrica*, v. 77, n. 3, p. 953-973, 2009.

___________. Decomposing wage distributions using recentered influence function regressions. *Econometrics*, v. 6, n. 2, p. 28, 2018.

FIRPO, Sergio; PINTO, Cristine. Identification and estimation of distributional impacts of interventions using changes in inequality measures. *Journal of Applied Econometrics*, v. 31, n. 3, p. 457-486, 2016.

FOLBRE, Nancy; NELSON, Julie. For love or money – or both? *Journal of Economic Perspectives*, v. 14, n. 4, p. 123-140, 2000.

FORTIN, Nicole; LEMIEUX, Thomas; FIRPO, Sergio. Decomposition methods in economics. In: *Handbook of labor economics*. Elsevier, 2011. p. 1-102.

FUJIWARA, Joyce Caroline Alecci Meneghim. *Mulheres sem filhos no Brasil: uma análise de características socioeconômicas e demográficas, razões e repercussões*. Tese de Doutorado – Univerisidade de Campinas, 2018.

GIUBERTI, A. C.; MENEZES-FILHO, N.. Discriminação de rendimentos por gênero: uma comparação entre o Brasil e os Estados Unidos. *Economia Aplicada*, v. 9, n. 3, p. 369-384, 2005.

GLAUBER, Rebecca. Marriage and the motherhood wage penalty among African Americans, Hispanics, and Whites. *Journal of Marriage and Family*, v. 69, n. 4, p. 951-961, 2007.

___________. Trends in the motherhood wage penalty and fatherhood wage premium for low, middle, and high earners. *Demography*, v. 55, n. 5, p. 1663-1680, 2018.

GREENE, William H. *Econometric analysis 4th edition*. International edition, New Jersey: Prentice Hall, p. 201-215, 2000.

GRIMSHAW, Damian; RUBERY, Jill. The motherhood pay gap: A review of the issues, theory and international evidence. *Conditions of Work and Employment Series,* 2015.

GAFNI, Dalit; SINIVER, Erez. Is There a Motherhood Wage Penalty for Highly Skilled Women? *The BE Journal of Economic Analysis & Policy*, v. 15, n. 3, p. 1353-1380, 2015.

GAMBOA, Luis Fernando; ZULUAGA, Blanca. Is there a motherhood penalty? Decomposing the family wage gap in Colombia. *Journal of family and economic issues*, v. 34, n. 4, p. 421-434, 2013.

GLAUBER, Rebecca. Marriage and the motherhood wage penalty among African Americans, Hispanics, and Whites. Journal of Marriage and Family, v. 69, n. 4, p. 951-961, 2007.

___________. Trends in the motherhood wage penalty and fatherhood wage premium for low, middle, and high earners. Demography, v. 55, n. 5, p. 1663-1680, 2018.

GUIGINSKI, Janaína. *Mercado de trabalho e relações de gênero: associação entre a presença de filhos e as condições de acesso ao trabalho das mulheres*. Dissertação de Mestrado – Universidade Federal de Minas Gerais, Belo Horizonte, 2015.

HERSCH, J. Home production and wages: evidence from the American Time Use Survey. *Review of Economics of the Household*, v. 7, n. 2, p. 159–178, 2009.

HERSCH, J.; STRATTON, L. S. Housework, Fixed Effects, and Wages of Married Workers. *The Journal of Human Resources*, v. 32, n. 2, p. 285, 1997

INSTITUTO BRASILEIRO DE GEOGRAFIA DE ESTATÍSTICA (IBGE). Censo Demográfico 2010. Pesquisa Nacional por Amostra de Domicílios. Disponível em: <https://ww2.ibge.gov.br/home/estatistica/pesquisas/pesquisa_resultados.php?id_pesquisp=40>. Acesso em 14 nov. 2018.

JIA, Nan; DONG, Xiao-yuan. Economic transition and the motherhood wage penalty in urban China: investigation using panel data. *Cambridge Journal of Economics*, v. 37, n. 4, p. 819-843, 2012.

KAHN, Joan R.; GARCÍA-MANGLANO, Javier; BIANCHI, Suzanne M. The motherhood penalty at midlife: Long-term effects of children on women's careers. *Journal of Marriage and Family*, v. 76, n. 1, p. 56-72, 2014.

KILLEWALD, Alexandra; BEARAK, Jonathan. Is the motherhood penalty larger for low-wage women? A comment on quantile regression. *American Sociological Review*, v. 79, n. 2, p. 350-357, 2014.

KOENKER, Roger; BASSETT JR, Gilbert. Regression quantiles. *Econometrica: journal of the Econometric Society*, p. 33-50, 1978.

KÜHHIRT, M., & LUDWIG, V. Domestic Work and the Wage Penalty for Motherhood in West Germany. Journal of Marriage and Family, 74(1), 186–200, 2012

LANDIVAR, Liana Christin. How the Timing of Children Affects Earnings in 20 Occupations. SocArXiv Papers, apr. 2018.

LAVINAS, Lena; NICOLL, Marcelo. Atividade e vulnerabilidade: quais os arranjos familiares em risco? *Revista de Ciências Sociais, Rio de Janeiro*, v.. 49, n. 1, 2006.

LEME, M. C.; WAJNMAN, S. Tendências de coorte nos diferenciais de rendimentos por sexo.In: *Desigualdade e Pobreza no Brasil*, coordenado por Ricardo Henriques. Rio de Janeiro: IPEA, p. 251-270, 2000.

LEOCÁDIO, Victor Antunes. *Childlessness no Brasil: a contribuição das mudanças sociodemográficas para a tendência de zero filho*. Dissertação de Mestrado – Universidade Federal de Minas Gerais, Belo Horizonte, 2018.

LEMIEUX, Thomas. The changing nature of wage inequality. *Journal of population Economics*, v. 21, n. 1, p. 21-48, 2008.

LIVERMORE, Tanya; RODGERS, Joan; SIMINSKI, Peter. The effect of motherhood on wages and wage growth: evidence for Australia. *Economic Record*, v. 87, p. 80-91, 2011.

MACHADO, José AF; MATA, José. Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of applied Econometrics*, v. 20, n. 4, p. 445-465, 2005.

MADALOZZO, Regina; BLOFIELD, Merike. Como famílias de baixa renda em São Paulo conciliam trabalho e família? *Estudos Feministas*, v. 25, n. 1, p. 215-240, 2017.

MCINTOSH, Bryan; MCQUAID, Ronald; MUNRO, Anne; DABIR-ALAI, Parviz. Motherhood and its impact on career progression. *Gender in Management: An International Journal*, Vol. 27 Iss 5 pp. 346 - 364, 2012).

MELLY, B. Estimation of Counterfactual Distributions Using Quantile Regression. *University of St.Gallen*, 2006, 68.

MIRANDA-RIBEIRO, Adriana; GARCIA, Ricardo Alexandrino; FARIA, Tereza Cristina de Azevedo Bernardes. Baixa fecundidade e adiamento do primeiro filho no Brasil. *Revista Brasileira de Estudos de População*, v. 36, 2019.

MISRA, Joya; BUDIG, Michelle; MOLLER, Stephanie. Reconciliation policies and the effects of motherhood on employment, earnings and poverty. *Journal of Comparative Policy Analysis: Research and Practice*, 9(2): 135–155, 2007.

MOLINA, José Alberto; MONTUENGA, Víctor M. The motherhood wage penalty in Spain. *Journal of Family and Economic Issues*, v. 30, n. 3, p. 237-251, 2009.

MUNIZ, Jerônimo Oliveira; VENEROSO, Carmelita Zilah. Diferenciais de participação laboral e rendimento por gênero e classes de renda: Uma investigação sobre o ônus da maternidade no Brasil. *Dados*, v. 62, 2019.

NEUMEIER, Christian; SØRENSEN, Todd; WEBBER, Douglas. The Implicit Costs of Motherhood over the Lifecycle: Cross-Cohort Evidence from Administrative Longitudinal Data. *Southern Economic Journal*, v. 84, n. 3, p. 716-733, 2018.

OLIVEIRA, A. L. M.; COLOMBI, A. P. F. Avanços e Contradições do Mercado de Trabalho Brasileiro de 2003 a 2012: uma análise em perspectiva de gênero. *Anais do V Seminário de Trabalho e Gênero e III Seminário Internacional do PPGCS: Teorias, pesquisas e práticas sociais*, Uberlândia (MG), 2014.

OLIVEIRA, Elzira; RIOS-NETO, Eduardo; OLIVEIRA, Ana Maria Hermeto Camilo de. Transições dos jovens para o mercado de trabalho, primeiro filho e saída da escola: o caso brasileiro. *Revista Brasileira de Estudos Populacionais*, v. 23, n. 1, 2006.

OAXACA, R. Male-female wage differentials in urban labor markets. *International economic review*, p. 693-709, 1973.

PASSOS, Luana; WAJNMAN, Simone. Diferencial salarial entre as mulheres: a dupla penalidade das negras. *Migração, trabalho e gênero: textos selecionados*. Belo Horizonte: ABEP, p. 192-206, 2021.

PAULO, Maira Andrade. Uma análise do diferencial de salários entre as mulheres que são mães e aquelas que não são mães, e entre aquelas que postergam e não postergam a fecundidade no período entre 1984 a 2009. *Anais do XVIII Encontro Nacional de Estudos Populacionais*, p. 1-19, 2012.

——————. *A maternidade, o momento do nascimento do primeiro filho e a situação do mercado de trabalho no Brasil: uma análise do salário e da adequação escolaridade/ocupação por grupos socioeconômicos*. Tese de Doutorado – Universidade Federal de Minas Gerais, Belo Horizonte, 2013.

PAZELLO, Elaine Toldo. A maternidade afeta o engajamento da mulher no mercado de trabalho: um estudo utilizando o nascimento de gêmeos como um experimento natural. *Estudos Econômicos* (São Paulo), v. 36, n. 3, p. 507-538, 2006.

——————; FERNANDES, Reynaldo. A maternidade e a mulher no mercado de trabalho: diferença de comportamento entre mulheres que têm e que não têm filhos. *Anais do 32º Encontro Nacional de Economia (Anpec)*, 2004.

RIOS-AVILA, Fernando. Recentered influence functions (RIFs) in Stata: RIF regression and RIF decomposition. *The Stata Journal*, v. 20, n. 1, p. 51-94, 2020.

SALARDI, Paola. *An analysis of pay and occupational differences by gender and race in Brazil, 1987 to 2006*. Tese de Doutorado. University of Sussex, 2013.

SIMÃO, Andréa Branco; COUTINHO, Raquel Zanatta; GUEDES, Gilvan Ramalho. Desejo por filhos entre mulheres de alta escolaridade: conflitos, mudanças e permanências. *Revista Brasileira de Estudos de População*, v. 37, 2020.

SORJ, Bila; FONTES, Adriana; MACHADO, Danielle Carusi. Políticas e práticas de conciliação entre família e trabalho no Brasil. *Cadernos de pesquisa*, v. 37, n. 132, p. 573-594, 2007.

SOUZA, Laetícia Rodrigues de; RIOS-NETO, Eduardo Luiz Gonçalves; QUEIROZ, Bernardo Lanza. A relação entre parturição e trabalho feminino no Brasil. Revista. *Brasileira de estudos populacionais*. São Paulo, v. 28, n. 1, p. 57-79, June 2011.

SOUZA, Paola Faria Lucas de. *Efeitos da Maternidade e do Casamento sobre o Diferencial de Salários entre Gêneros no Brasil para o Ano de 2014*. Tese de Doutorado – Universidade Federal do Ceará, Fortaleza, 2016.

VILLANUEVA, Aida; LIN, Ken-Hou. Motherhood Wage Penalties and the Informal Sector in Latin America. *Annual Meeting of the Population Association of America*. Disponível em: https://paa.confex.com/paa/2016/meetingapp.cgi/Paper/7930. Washington, DC, 2016, Acesso em 12 nov. 2018.

WAJNMAN, Simone. Relações familiares e diferenciais de rendimentos por sexo no Brasil. *População e desenvolvimento em debate: contribuições da Associação Brasileira de Estudos Populacionais*. Belo Horizonte: ABEP, p. 187-192, 2012.

——————. "Quantidade" e "qualidade" da participação das mulheres na força de trabalho brasileira. *Até onde caminhou a revolução de gênero no Brasil?* Belo Horizonte: ABEP, p. 45-58, 2016.

WALDFOGEL, Jane. The effect of children on women's wages. *American sociological review,* p. 209-217, 1997.

APPENDIX A - INFLUENCE FUNCTION (IF) AND REFOCUSED INFLUENCE FUNCTION (RIF)

The following is the concept of Influence Function (IF), as presented by Rios-Avila (2020)[7] . The IF can be understood as a measure of robustness, by representing the influence of an individual observation on the distributional statistic of interest. Considering a regressed variableY - in the context of this work, income/hour - and its Cumulative Distribution Function,$F_Y(\cdot)$, the information needed to analyze the distribution can be summarized at $F_Y = [\{y, F_Y(y)\} | y \in \mathbb{R}]$, where y is a given observation from Y. Any distributional statistic, such asQ_τ , can be obtained from the vectorsY and F_Y.

In order to measure the impact of a change in the distributionF_Y to a counterfactual distribution G_Y on the distributional statistic of interest, one must additionally consider a functional$v(\cdot)$, which estimates this statistic from the information contained in Y and F_Y . Thus, we have

$$\Delta v = v(G_Y) - v(F_Y) \quad (A.1)$$

where Δv represents the change in the statistic of interest. This, in turn, can be rewritten as a standardized rate for the change fromF_Y toG_Y , as shown in (A.2).

$$\Delta^s v = \frac{\Delta v}{\Delta(G_Y - F_Y)} = \frac{v(G_Y) - v(F_Y)}{\Delta(G_Y - F_Y)} \quad (A.2)$$

The measurement of the standardized rate$\Delta^s v$ can be extrapolated to infinitesimal changes from F_Y to G_Y , which summarizes the very concept of IF. Considering, for example, the addition of an observation to the sample whose income/hour value is y_i , its Cumulative Distribution Function $H_{y_i}(y)$ will have only one element. This means that $H_{y_i}(y) = 0 \; \forall \, y < y_i$ and

[7] Rios-Avila (2020) presents a package developed for the *STATA* statistical software that allows easy implementation of RIF regressions and decompositions. The author gives a detailed and intuitive exposition of the concept of IF, as well as the models of Firpo, Fortin and Lemieux (2009,2018). For this reason, it was decided to replicate the demonstrations developed in his work.

$H_{y_i}(y) = 1\ \forall\, y \geq y_i$; in other words, the distribution H_{y_i} places mass only at y_i. With this, the distribution G_Y can be rewritten in terms of H_{y_i}, according to equation (A.3).

$$G_Y = (1-\varepsilon)F_Y + \varepsilon H_{y_i} \tag{A.3}$$

The term ε plays a similar role to $[\Delta(G_Y - F_Y)]$, measuring the change from F_Y to G_Y . From (14) it is finally possible to obtain the FI, according to (A.4):

$$IF\{y_i, v(F_Y)\} = \lim_{\varepsilon \to 0} \frac{v\{(1-\varepsilon)F_Y + \varepsilon H_{y_i}\} - v(F_Y)}{\varepsilon} = \frac{\partial v(F_Y \to H_{y_i})}{\partial \varepsilon} \tag{A.4}$$

Once the FI has been defined, Firpo, Fortin and Lemieux (2009) introduce the concept of the Recentred Influence Function (RIF), which consists of adding the FI to the distributional statistic of interest:

$$RIF\{y_i, v(F_Y)\} = v(F_Y) + IF\{y_i, v(F_Y)\} \tag{A.5}$$

In the context of quantiles, the IF and RIF are defined, respectively, by equations A.6 and A.7,

$$IF\{Y, Q_\tau\} = \frac{(\tau - \mathbb{1}\{Y \leq Q_\tau\})}{f_y(Q_\tau)} \tag{A.6}$$

$$RIF\{y_i, v(F_Y)\} = Q_\tau + \frac{(\tau - \mathbb{1}\{y_i \leq Q_\tau\})}{f_y(Q_\tau)} = c_{1,\tau} \cdot \mathbb{1}\{y_i > Q_\tau\} + c_{2,\tau} \tag{A.7}$$

where $\mathbb{1}\{\cdot\}$ is a function indicating the relationship between the observationy_i and the quantile; $Q_\tau f_y(\cdot)$ is the marginal density of Y ; and $c_{1,\tau}$ and $c_{2,\tau}$ are constants, so that $c_{1,\tau} = 1/f_y(Q_\tau)$ and $c_{2,\tau} = Q_\tau - c_{1,\tau} \cdot (1-\tau)$.