

EDUCATION, EXPERIENCE, AND GENDER: EVIDENCE ON THE DETERMINANTS OF EARNINGS IN THE BRAZILIAN FORMAL LABOR MARKET

EDUCAÇÃO, EXPERIÊNCIA E GÊNERO: EVIDÊNCIAS SOBRE OS DETERMINANTES DOS GANHOS NO MERCADO DE TRABALHO FORMAL BRASILEIRO

EDUCACIÓN, EXPERIENCIA Y GÉNERO: EVIDENCIA SOBRE LOS FACTORES DETERMINANTES DE LOS INGRESOS EN EL MERCADO LABORAL FORMAL BRASILEÑO

Max Richard Coelho Verginio¹, Liara Darabas Ronzani², Caroline da Graça Jacques³, Dimas de Oliveira Estevam⁴

DOI: 10.54899/dcs.v22i82.3387

Recibido: 16/09/2025 | Aceptado: 17/09/2025 | Publicación en línea: 29/09/2025.

ABSTRACT

This article estimates the determinants of earnings and the gender wage gap in Brazil's formal labor market, combining insights from human capital theory and feminist economics. Using 2022 microdata from the Relação Anual de Informações Sociais (RAIS), we estimate a semilog Mincer wage equation with controls for schooling, tenure, age, gender, and region. A stratified sample of 100,000 formal employment records was drawn to preserve geographic representativeness, and educational levels were standardized. Quadratic terms for age and tenure capture diminishing returns, and a progressive modeling strategy assessed individual and joint effects. Inference relies on heteroskedasticity-consistent standard errors. Results indicate a persistent wage penalty for women, with hourly earnings about 20 percent lower than those of men with similar profiles after controls. Higher education is associated with higher pay, although a negative differential appears for master's degree holders, suggesting institutional specificities. Regional disparities are notable, with lower wages in the North and Northeast. Overall, the evidence points to entrenched gender and regional inequalities not fully explained by productivity, reinforcing the need for policies that promote gender equity and the valuation of human capital across Brazil.

Keywords: Gender Inequality. Human Capital. Education. Labor Market. Wage Discrimination.

¹ Doctor of Socioeconomic Development, Universidade do Extremo Sul Catarinense (UNESC), Criciúma, Santa Catarina, Brazil. E-mail: verginio@unesc.net Orcid: <https://orcid.org/0000-0002-5235-2687>

² Ph.D. candidate in Economics, Universidade Federal do Rio Grande do Sul (UFRGS), Araranguá, Santa Catarina, Brazil. E-mail: liadarabas@hotmail.com Orcid: <https://orcid.org/0000-0001-5748-0736>

³ Doctor of Political Sociology, Universidade Federal de Santa Catarina (UFSC), Criciúma, Santa Catarina, Brazil. E-mail: carolinejacques@unesc.net Orcid: <https://orcid.org/0000-0001-7377-3782>

⁴ Doctor of Political Sociology, Universidade Federal de Santa Catarina (UFSC), Criciúma, Santa Catarina, Brazil. E-mail: doe@unesc.net Orcid: <https://orcid.org/0000-0002-8116-2209>

RESUMO

Este artigo estima os determinantes dos rendimentos e a disparidade salarial entre homens e mulheres no mercado de trabalho formal brasileiro, articulando teoria do capital humano e economia feminista. Com microdados da RAIS 2022, estima-se uma equação de salários à la Mincer em semilog com controles para escolaridade, tempo de emprego, idade, gênero e região. Foi extraída amostra estratificada de 100.000 vínculos para preservar a representatividade geográfica, com padronização dos níveis educacionais. Termos quadráticos para idade e tempo de emprego capturam retornos decrescentes, e a estratégia progressiva de modelagem avaliou efeitos individuais e combinados. A inferência utiliza erros-padrão consistentes à heterocedasticidade. Os resultados mostram penalidade salarial persistente para as mulheres, com remuneração horária cerca de 20 por cento inferior à dos homens de perfil semelhante após controles. A educação superior eleva rendimentos, embora surja diferencial negativo para portadores de mestrado, sugerindo especificidades institucionais. Observam-se ainda disparidades regionais, com salários menores no Norte e Nordeste. Em conjunto, as evidências indicam desigualdades de gênero e regionais enraizadas, não explicadas apenas por produtividade, reforçando a necessidade de políticas de equidade de gênero e de valorização do capital humano no país.

Palavras-chave: Desigualdade de Gênero. Capital Humano. Educação. Mercado de Trabalho. Discriminação Salarial.

RESUMEN

Este artículo estima los determinantes de los ingresos y la brecha salarial de género en el mercado laboral formal de Brasil, integrando la teoría del capital humano y la economía feminista. Con microdatos de la RAIS 2022, se estima una ecuación salarial de Mincer en semilog con controles por escolaridad, antigüedad en el empleo, edad, género y región. Se construyó una muestra estratificada de 100.000 vínculos para preservar la representatividad geográfica y se estandarizaron los niveles educativos. Términos cuadráticos para edad y antigüedad capturan retornos decrecientes, y una estrategia progresiva de modelado evaluó efectos individuales y combinados. La inferencia se basa en errores estándar consistentes frente a heterocedasticidad. Los resultados indican una penalidad salarial persistente para las mujeres, con remuneraciones horarias alrededor de 20 por ciento inferiores a las de los hombres con perfiles similares después de los controles. La educación superior aumenta los ingresos, aunque se observa un diferencial negativo para quienes poseen maestría, lo que sugiere especificidades institucionales. También se evidencian disparidades regionales, con salarios más bajos en el Norte y el Nordeste. En conjunto, la evidencia señala desigualdades de género y regionales arraigadas, no explicadas solo por productividad, y subraya la necesidad de políticas de equidad de género y de valorización del capital humano en todo Brasil.

Palabras clave: Desigualdad de Género. Capital Humano. Educación. Mercado Laboral. Discriminación Salarial.



Esta obra está bajo una [Licencia Creative Commons Atribución- NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/)

INTRODUCTION

In recent decades, women's participation in the Brazilian formal labor market has expanded, yet wage gaps persist even among workers with similar schooling and occupations, indicating structural and institutional barriers beyond individual productivity. Feminist scholarship argues that such inequalities are not fully explained by observable characteristics: the sexual division of labor structures sex-based social relations, separating and hierarchizing "masculine" and "feminine" activities as a historically constructed valuation regime that privileges male occupations (Hirata & Kergoat, 2007). Gender thus operates as a constitutive element of social relations and a primary lens through which power is organized, situating wage differences within a broader system of inequality that permeates even formal, regulated markets (Scott, 1995). Feminist economics further exposes the androcentric bias of mainstream theory—its neglect of unpaid work and of gendered access to resources—calling for models that embed social and normative constraints shaping choices and trajectories (Coelho, 2009). Evidence for Brazil echoes this pattern: women's lower pay reflects not only qualifications but also discrimination, sectoral segmentation, and the devaluation of female-dominated occupations (Teixeira, 2008).

However, even so, women with higher education have increased their relative participation in the composition of the labor force and in the economy as a whole. In many countries, there has been a significant increase in their participation in the labor market, especially since the 1990s. However, women's struggle for inclusion in the labor market and economic activities has been a field of dispute, guided by prejudice and inequalities in access compared to men. The historical bonds involving women have generally been and are patriarchal, their functions have always been well defined: domestic responsibilities, children's education, dedication to the husband, taking care of the sick, among other functions considered "feminine" (Sucupira, 2016).

In view of this scenario, this article aims to analyze the determinants of remuneration in the Brazilian formal labor market, with emphasis on gender wage disparities, based on the application of Mincer's equation (1984) to the microdata of RAIS 2022. An econometric approach with log-linear specification is used, incorporating human capital variables (years of education and experience), as well as sociodemographic (gender, age) and regional controls. The proposal is to evaluate to what extent the observed wage differences can be explained by

productive factors or if they persist after controlling for such variables, indicating the presence of structural inequalities in the formal labor market.

GENDER INEQUALITIES, FEMINIST ECONOMICS AND RETURNS TO HUMAN CAPITAL

Gender inequalities in the labor market are not exhausted by differences in schooling, experience, or other observables: they are anchored in a sexual division of labor that separates and hierarchizes “masculine” and “feminine” activities, privileging male-associated roles (Hirata & Kergoat, 2007). In this sense, gender is a constitutive principle of social relations and a key to interpreting power asymmetries; wage disparities are thus part of a broader, persistent system of inequality that operates even in formal, regulated markets (Scott, 1995).

Feminist economics critiques the androcentric core of mainstream theory for overlooking unpaid domestic and care work and for ignoring gendered barriers to resources and opportunities. Drawing on diverse traditions, it calls for models that embed social and normative constraints shaping choices and trajectories; the historical classification of domestic/care activities as “unproductive” helps explain their under-valuation in analysis and policy (Coelho, 2009). In Brazil, empirical evidence shows that women’s lower pay stems not only from qualifications but also from discrimination, occupational segregation, and the devaluation of feminized professions, even as women—especially those with higher education—expanded their participation under unequal terms (Sucupira, 2016; Teixeira, 2008).

This pattern intersects with labor-market segmentation: long-standing debates—revived during World War I as women replaced men in various jobs—juxtaposed “productivity” and “subsistence” notions of wages, often naturalizing lower female pay and the idea of a “secondary” female labor force (Pujol, 1992; Teixeira, 2008). Discrete but pervasive practices still penalize women, from motherhood-related hiring filters to role stereotyping (Sina, 2005). As a result, women remain concentrated in occupations aligned with socially feminized attributes—care, services, and clerical work—whose market value mirrors their association with reproductive activities (Bruschini & Lombardi, 2002; Salvaro et al., 2014). Contemporary productive restructuring has compounded this with precarization—informality, instability, and weaker protections—which affects both sexes but more intensely women, feeding the

“feminization of poverty” and the rise of female-headed households (Brasil, 2017; Hirata, 2005, 2018), 2005; Hirata, 2018; Brasil, 2017).

Against this backdrop, the next subsection turns to the sexual division of labor within higher-education teaching in Brazil, connecting these structural mechanisms to a specific, knowledge-intensive segment of the labor market.

Segmentation Theory and Returns to Schooling

This study adopts the human capital framework formalized by Mincer (1984), in which earnings reflect investments in schooling and experience. The baseline specification is the semilog model

$$\ln(y_i) = \beta_0 + \beta_1 \cdot Education_i + \beta_2 \cdot Experience_i + \beta_3 \cdot Experience_i^2 + \epsilon_i \quad (1)$$

Where y_i is the hourly wage of individual i , $Education_i$ represents years of formal schooling, $Experience_i$ denotes potential labor market experience, measured in years, and ϵ_i is the error term. The model includes both linear and quadratic terms for experience to capture potential non-linearities in returns to experience, typically interpreted as diminishing marginal returns (Moretto, 2000).

Recent Brazilian evidence motivates these extensions. Moura (2008) documents non-linearities in schooling and experience—suggesting dummies for levels and direct tenure measures can outperform linear proxies—while Cangussu et al. (2010) show that, despite alternative growth-type formulations, the Mincerian approach retains superior explanatory power for income determination. Addressing identification, Teixeira and Menezes-Filho (2012) use instruments tied to school availability and the 1971 education law and find that OLS overstates returns to education, highlighting endogeneity concerns. Beyond averages, Silva et al. (2022) uncover heterogeneous and non-monotonic returns (linear–cubic in schooling) with interactions by informality, urban status, and race, and Faustino et al. (2023) combine Mincer with Blinder–Oaxaca decompositions to separate explained and unexplained components of gender-regional wage gaps. Complementing the formal urban focus, Reis and Moreira (2017) confirm positive—though smaller—schooling effects in family farming incomes. Together with broader reviews

(Fernandes, 2015), this literature supports a parsimonious but context-aware Mincer model that is transparent about scope and limitations.

In line with this evidence, the present paper estimates a cross-sectional Mincer equation using RAIS-2022 microdata for Santa Catarina, emphasizing clarity and compatibility with the dataset. We include core controls (gender, age, experience/tenure, and schooling—via years or level dummies) and report robust inference, while refraining from techniques such as Heckman selection or Oaxaca–Blinder that would require additional assumptions not directly supported by the data structure. The interpretation is thus firmly descriptive of returns within a segmented and unequal formal labor market, informed by the identification cautions and heterogeneity patterns documented in prior Brazilian studies.

METHODOLOGY

This is an empirical, quantitative, cross-sectional study that estimates the determinants of formal labor earnings using the Mincer (1984) framework. We adopt a log-linear specification in which the logarithm of hourly remuneration depends on years of schooling, tenure in the current job with a quadratic term, age with a quadratic term, gender, and geographic location. These covariates allow the estimation of conditional marginal effects associated with human capital and sociodemographic characteristics in the Brazilian institutional context. The analysis uses large-scale secondary microdata to identify structural patterns in returns to education and experience.

Education is operationalized as years of schooling, mapped from RAIS educational levels to the minimum years in the Brazilian system, yielding a continuous measure of formal human capital. To capture individual and institutional heterogeneity, we add controls for gender, age, higher-education attainment, and regional location. The extended specification is presented in equation (2).

$$\begin{aligned} \ln(y_i) = & \beta_0 + \beta_1 \cdot Education_i + \beta_2 \cdot Tenure_i + \beta_3 \\ & \cdot Tenure_i^2 + \beta_4 \cdot Age_i + \beta_5 \cdot Age_i^2 + \beta_6 \\ & \cdot Female_i + \beta_7 \cdot College_i + \sum_r \gamma_r \\ & \cdot Region_i + \epsilon_i \end{aligned} \quad (2)$$

The equation above models the logarithm of hourly remuneration, $\ln(y_i)$, as a function of years of formal education ($Education_i$), tenure in the current formal job ($Tenure_i$) and its quadratic term, age and age squared (Age_i , Age_i^2), a binary variable for sex ($Female_i$, equal to 1 for women), a binary variable for higher education attainment ($College_i$, equal to 1 for those with a completed college degree), and a set of regional dummies ($Region_i$), with the reference region omitted to avoid multicollinearity. The coefficients represent the marginal effects of the explanatory variables on the log of wages, while ϵ_i denotes the random error term. This specification follows the Mincerian framework and allows for the estimation of the marginal returns to education and experience, while controlling demographic and geographic characteristics.

As the RAIS 2022 microdata does not provide a continuous variable to represent formal schooling, it was necessary to construct the variable years of schooling from the categorical coding of the field "schooling after 2005". To this end, the minimum number of years of study required by the Brazilian educational system to complete the respective level was associated with each category, as presented in the Table 1. In cases in which the category included ranges of years (such as incomplete schooling), the arithmetic mean of the interval was adopted. The new resulting variable, expressed in continuous values, allows the measurement of formal human capital in a way that is more appropriate to the structure of Mincer's equation. Observations with unknown schooling were excluded from the sample.

Table 1

Conversion of RAIS schooling codes into years of study

Code	RAIS	Description	Years of study
	1	Illiterate	0
	2	Up to 5th incomplete	1–4 $\Rightarrow$ 3 (average)
	3	5th complete, fundamental inc.	5–8 $\Rightarrow$ 6.5 (average)
	4	6th to 9th complete elementary school	until 9 $\Rightarrow$ 9
	5	Complete elementary school	9
	6	Incomplete high school	10–11 $\Rightarrow$ 10.5
	7	Complete high school	12
	8	Incomplete Superior	13.5
	9	Complete Superior	16
	10	Masters	18
	11	Doctorate	21
	-1	Ignored	NA (remove or impute)

Source: Prepared by the authors based on the RAIS 2022 schooling codes.

The data come from the *Relação Anual de Informações Sociais (RAIS)* for 2022, provided by the Ministry of Labor and Employment. Processing was carried out in R (2025). We excluded records with inconsistencies or atypical characteristics: contracts inactive on December 31, weekly hours below 10, zero or negative average pay, intermittent or part-time contracts, zero tenure in the job, and unknown or invalid schooling.

To enable the estimation with adequate computational performance, a sample of 100,000 employment relationships was extracted through stratified sampling proportional to the Federation Units. This technique preserves the geographic representativeness of the formal workforce. The decision for the sample size considered the statistical robustness of the model, based on two complementary approaches: Green's heuristic (1991), which recommends observations for multiple regressions, and Cohen's statistical power analysis $n \geq 50 + 8k(1988)$, which relates the number of predictors to the expected effect size ($f^2 = 0,15$) and level of significance. Both suggest minimum sizes much smaller than the sample actually used, which ensures stability of the coefficients and inferential reliability, even considering the stratification design effect.

Or Table 2 It presents the description of the variables used in the regression, accompanied by the theoretical hypothesis as to their expected sign, according to the literature on labor economics and human capital.

Table 2

Description of the variables used in the regression model and their expected signals

Variable	Description	Expected signal
Logarithm of hourly remuneration	Continuous dependent variable, obtained by the ratio between the average monthly remuneration and the total hours contracted, transformed via natural logarithm.	(dependent variable)
Female	Dichotomous variable that assumes a value of 1 for women and 0 for men. It captures potential gender inequalities.	Negative (–)
Years of study (centralized)	Minimum number of years of formal schooling, estimated on the basis of declared education. It reflects the human capital accumulated through formal education.	Positive (+)
Time on link (centered)	Duration of the current employment relationship, in years. It is related to accumulated experience.	Positive (+)
Time in Link Squared	Square of time in the link, used to capture diminishing returns from experience.	Negative (–)
Age (centered)	Worker's age in years, representing professional maturity and seniority.	Positive (+)
Age squared	Age square, to capture nonlinear effects (e.g., decline after a certain point).	Negative (–)
Complete higher education	Binary indicator for workers with complete higher education or more (master's/doctorate).	Positive (+)

Source: Prepared by the authors based on RAIS microdata (2022).

Mincer's equation was used as a basis for the specification of the empirical model, considering the logarithm of hourly remuneration as a dependent variable. Different combinations of explanatory variables related to experience (length of employment and age, with quadratic terms), education (years of study and complete degree), sex and geographic location were tested. The selection of the most parsimonious and informative model was carried out through an exhaustive search (exhaustive method) using the Leaps (2024) in the R environment, which evaluated all possible combinations of regressors. The final choice considered the lowest values of the informative criteria (AIC and BIC), the highest adjusted R^2 and the statistical significance of the coefficients. This strategy ensures greater robustness in the specification and minimizes the risk of bias in the selection of explanatory variables.

The residuals of the final model were subjected to formal normality tests, including Shapiro-Wilk (1965), Anderson-Darling (1952) and Jarque-Bera (1980). The presence of heteroscedasticity was investigated using the Breusch-Pagan test (1979), while autocorrelation was assessed using the Durbin-Watson test (1950). To correct for potential heteroscedasticity effects identified in the residuals, the standard errors were robustly estimated using the consistent variance-covariance matrix proposed by Huber (1967) and formalized by White (1980), known as an estimator sandwich or rugged HC1 type.

RESULTS

The Table 3 presents the descriptive statistics of the variables used in the empirical analysis. Continuous variables centered around the mean (years of schooling, experience, age, and their quadratic terms) were included, in addition to the dependent variable, which represents the logarithm of hourly remuneration. Dichotomous variables that capture individual characteristics (such as gender and education) and regional location of employment relationships are also displayed.

Table 3

Descriptive statistics of the sample (centralized continuous variables)

Variable	N	Average	Standard Deviation	Median	Minimum	Maximum
Years of schooling (centered)	100,000	0.00	2.74	-0.32	-12.32	8.68

Tenure in current job (centered)	100,000	0.00	7.26	-3.29	-5.36	44.62
Age (centered)	100,000	0.00	11.96	-1.10	-24.10	57.90
Tenure squared (centered)	100,000	52.75	138.89	19.82	0.00	1,991.26
Age squared (centered)	100,000	143.15	172.30	82.82	0.01	3,352.36
Log hourly wage	100,000	4.19	0.79	3.97	1.97	10.16
Female (1 = woman)	100,000	0.45	0.50	0.00	0.00	1.00
Middle education (1 = yes)	100,000	0.57	0.50	1.00	0.00	1.00
Higher education (1 = yes)	100,000	0.23	0.42	0.00	0.00	1.00
Midwest region (1 = yes)	100,000	0.14	0.35	0.00	0.00	1.00
Northeast region (1 = yes)	100,000	0.14	0.35	0.00	0.00	1.00
North region (1 = yes)	100,000	0.02	0.15	0.00	0.00	1.00
South region (1 = yes)	100,000	0.22	0.41	0.00	0.00	1.00

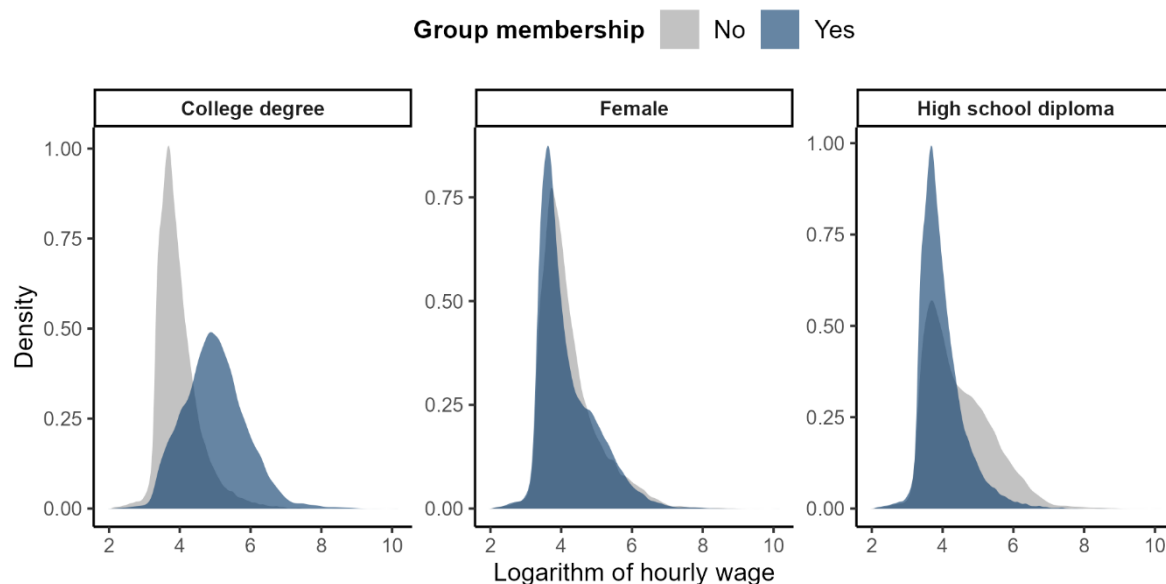
Source: Prepared by the authors based on RAIS microdata (2022), based on a sample stratified by federation unit.

It is observed that the average logarithm of the hourly remuneration ($\ln y$) is 4.19, with a standard deviation of 0.79, which corresponds to an approximate hourly average salary of R\$ 66.40 in nominal values (based on $\exp(4.19)$). The variable indicates that 45% of the sample is composed of women, while 23% of the workers have completed education. The upper variable and the other regional dummies reflect the proportional sample composition of the federative units. The continuous variables were centralized, therefore, they have a null mean per construction. On the other hand, the standard, minimum and maximum deviation values show the heterogeneity of the sample in terms of experience, age and education, suggesting wide variation in the occupational profiles analyzed.

To examine the distribution of hourly remuneration in the formal labor market, the Figure 1 based on individual variables indicative of structural inequalities: female gender, complete high school and higher education. The logarithm of hourly remuneration was used as a measure of comparison between groups, which allows us to visualize relative differences on a more symmetrical scale. The densities were plotted separately according to whether they belonged to each group, making it possible to identify displacements in distribution and possible asymmetries in access to remuneration.

Figure 1

Distribution of hourly remuneration according to individual characteristics in the formal labor market (Brazil, 2022)



Source: Prepared by the authors based on RAIS 2022 (data from the Ministry of Labor and Employment).

It is observed that, in the three panels, there is a shift to the right in the curves of the individuals belonging to the groups analyzed ("Yes"), indicating higher average remunerations. The difference is particularly marked among those with higher education, whose distribution is more dispersed and concentrated in higher income brackets. In the case of females, although there is overlap, the corresponding curve shows a slight concentration at lower levels, suggesting the persistence of gender wage inequalities. As for complete secondary education, the distribution of those with this education also shifts positively, but with a smaller amplitude compared to the group with higher education, reflecting lower relative gains.

Table 4 shows the results of the estimates according to the equation proposed by Mincer, the natural logarithm of remuneration explained by experience at the level of (Moretto, 2000), so it is possible to interpret the values of the betas as percentage variations of the explained variable, based on unit variations of the explanatory variables. Age in years, and length of employment in months, were used as a proxy for experience. Age squared and length of employment squared were added to these to capture the effect of reduced labor productivity at the end of the professional career. In addition, the effects on education are calculated and the hypothesis of differences statistically determined by the sex of the worker are tested a. com

except for the variation in gender in model number two, all other variables were statistically significant.

Table 4

Estimates of the logarithm of hourly remuneration with individual variables: gender, education, age and experience

	Dependent variable: log of hourly remuneration				
	Log hourly wage				
	M1 (1)	M2 (2)	M3 (3)	M4 (4)	M5 (5)
Female (1 = woman)	-0.059*** (0.005)	-0.204*** (0.005)	-0.206*** (0.004)	-0.206*** (0.004)	-0.204*** (0.004)
Years of schooling (centered)		0.137*** (0.001)	0.126*** (0.001)	0.124*** (0.001)	
Higher education (1 = yes)					0.917*** (0.006)
Tenure in current job (centered)			0.038*** (0.0003)	0.038*** (0.001)	0.034*** (0.001)
Age (centered)				0.011*** (0.0002)	0.006*** (0.0002)
Tenure squared (centered)				-0.0004*** (0.00003)	-0.0003*** (0.00003)
Age squared (centered)				-0.0005*** (0.00001)	-0.0004*** (0.00001)
Constant	4.215*** (0.003)	4.281*** (0.003)	4.281*** (0.003)	4.373*** (0.003)	4.153*** (0.003)
Observations	100,000	100,000	100,000	100,000	100,000
R ²	0.001	0.217	0.334	0.363	0.405
Adjusted R ²	0.001	0.217	0.334	0.363	0.405

Note: Estimates obtained by linear regression with robust standard errors (HC1). The dependent variable is the logarithm of hourly remuneration. The models progressively incorporate centered explanatory variables: gender, years of schooling, potential experience (and its square), age (and its square), and higher education dummy. $p < 0.1$; * $p < 0.05$; ** $p < 0.01$.

Source: Prepared by the authors based on RAIS 2022 (Ministry of Labor and Employment).

The results consistently reveal the presence of a wage penalty associated with the female sex. In all models, the coefficient of the variable remains negative and statistically significant, indicating that, even after controlling for education, age and experience, women have lower

hourly earnings than men — around 20.4% less in the complete model (M5). The increase in years of schooling is strongly associated with an increase in remuneration, but this effect is outweighed by the magnitude of the coefficient of the variable, which indicates that having a higher education degree is associated with a remuneration about 91.7% higher. The variables of experience and age have a positive relationship with earnings, while their quadratic terms have negative coefficients, showing diminishing returns. These findings are in line with the human capital theoretical framework, which holds that education and accumulated experience contribute to individual productivity and, therefore, to remuneration. The persistence of the gender penalty, even with the controls, also dialogues with the critical literature on labor market segmentation and structural inequality, suggesting that discriminatory factors still influence wage returns.

The results of the Table 5 show marked regional inequalities in the formal labor market. Even after the introduction of individual variables (model R5), workers in the regions North and, above all, Northeast continue to have significantly lower wages than those in the Southeast: approximately 12% and 24.3% lower, respectively. These differences are not explained by education, age, or experience, which points to regional structural constraints — such as the lower presence of high-productivity sectors or historically unequal public policies. The inclusion of the dummy variable for higher education shows a robust and positive impact (coefficient of 0.913 in the model R5), reaffirming the high salary return associated with the university degree. The gender penalty remains significant and stable in magnitude, even with regional controls, reinforcing the literature that highlights gender discrimination as a persistent component in determining earnings (Blau & Kahn, 2017). Negative quadratic effects for age and experience are still present, reflecting diminishing returns compatible with Mincer's model. In summary, the analysis shows that regional disparities are added to individual inequalities, structuring a deeply segmented labor market, in which not even qualification eliminates distortions of territorial origin or gender.

Table 5

Estimates of the logarithm of hourly remuneration with the inclusion of regional dummies and progressive individual variables

Dependent variable: log of hourly remuneration				
Log hourly wage				
R1	R2	R3	R4	R5
(1)	(2)	(3)	(4)	(5)

Female (1 = woman)	-0.059*** (0.005)	-0.212*** (0.004)	-0.204*** (0.004)	-0.206*** (0.004)
Higher education (1 = yes)		1.065*** (0.006)	0.943*** (0.006)	0.913*** (0.006)
Tenure in current job (centered)			0.028*** (0.0004)	0.035*** (0.001)
Age (centered)			0.005*** (0.0002)	0.006*** (0.0002)
Tenure squared (centered)				-0.0003*** (0.00003)
Age squared (centered)				-0.0005*** (0.00001)
North region (1 = yes)	-0.068*** (0.017)	-0.069*** (0.017)	-0.093*** (0.014)	-0.119*** (0.012)
Northeast region (1 = yes)	-0.207*** (0.008)	-0.207*** (0.008)	-0.199*** (0.006)	-0.238*** (0.006)
Midwest region (1 = yes)	0.012 (0.008)	0.012 (0.008)	-0.008 (0.007)	-0.011* (0.006)
South region (1 = yes)	-0.015** (0.006)	-0.014** (0.006)	0.004 (0.005)	0.019*** (0.005)
Constant	4.221*** (0.004)	4.248*** (0.004)	4.071*** (0.004)	4.099*** (0.003)
Observations	100,000	100,000	100,000	100,000
R ²	0.008	0.009	0.319	0.405
Adjusted R ²	0.008	0.009	0.319	0.405

Note: The dependent variable is the logarithm of the hourly remuneration. All models include regional dummies (with reference to the Southeast region). Individual variables are added progressively, covering gender, education, experience (and its square) and age (and its square). Estimates were obtained by linear regression with robust standard errors (HC1). Statistical significance is indicated by: p<0.1; *p<0.05; **p<0.01.

Source: Prepared by the authors based on RAIS 2022 microdata, made available by the Ministry of Labor and Employment.

The predicted values of the logarithm of hourly remuneration, controlling for education, age, experience, region and educational level, reinforce the existence of a wage penalty associated with the female sex. The adjusted estimate for males (female = 0) was 4.32, with a confidence interval of [4,32; 4,33], while for females (female = 1), the prediction was 4.10, with an interval of [4,10; 4,11]. The highest confidence interval of the female group does not even reach the lower limit of the male group, which shows that the observed difference is not only statistically significant, but also substantial. This is a robust and persistent wage penalty, which is not eliminated even after controlling for the main observable productivity variables. This result closes the section by reaffirming the central findings of this study: gender inequality in earnings

remains in the formal labor market, even among individuals with similar levels of education, age, and experience.

CONCLUSION

This study aimed to analyze the determinants of remuneration in the Brazilian formal labor market, focusing on gender wage disparities, in the light of human capital theory and feminist economics. Using microdata from RAIS 2022 and econometric estimates based on Mincer's equation (1984), the results showed a robust wage penalty associated with the female sex, not explained by differences in education, experience or regional location.

The mobilized theoretical literature — especially the concepts of sexual division of labor (Hirata & Kergoat, 2007), market segmentation (Teixeira, 2008) and critiques of feminist economics (Coelho, 2009) — offers support for interpreting these findings as an expression of structural and historically constructed inequalities. The insertion of women in the labor market has expanded in quantitative terms, but qualitatively remains marked by symbolic and institutional barriers that limit the full recognition and appreciation of their workforce.

The presence of a significant wage differential, even among women with a higher education degree, reinforces the thesis that the Brazilian formal market continues to operate under logics of gender segmentation and hierarchization. This reality imposes urgent challenges for public policies, which must not only expand access to education and professional qualification, but also promote cultural and institutional transformations capable of addressing the bases of gender inequality in the world of work.

It is concluded that the observed wage differentials are not only due to observable differences in human capital, but reflect mechanisms of structural and institutional discrimination that perpetuate the asymmetry between men and women. Future research can deepen this analysis by incorporating unobservable factors, such as type of contract, effective working hours, and institutional selection processes, in addition to expanding the scope to specific sectors such as higher education, which presented ambiguous results regarding the effect of *stricto sensu* degrees. The struggle for gender equality in formal work thus remains a central field of dispute and social transformation.

REFERENCES

- Anderson, T. W., & Darling, D. A. (1952). Asymptotic Theory of Certain «Goodness of Fit» Criteria Based on Stochastic Processes. *The Annals of Mathematical Statistics*, 23(2), 193-212. <http://www.jstor.org/stable/2236446>
- Blau, F. D., & Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3), 789-865. <https://doi.org/10.1257/jel.20160995>
- Brasil. (2017). *Relação Anual de Informações Sociais*. Programa Nacional de Disseminação das Estatísticas do Trabalho. <http://pdet.mte.gov.br/>
- Breusch, T. S., & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, 47(5), 1287-1294. <https://doi.org/10.2307/1911963>
- Bruschini, C., & Lombardi, M. R. (2002). Instruídas e trabalhadeiras trabalho feminino no final do século XX. *Cadernos Pagu*, 17-18, 157-196. <https://doi.org/10.1590/S0104-83332002000100007>
- Cangussu, R. C., Salvato, M. A., & Nakabashi, L. (2010). Uma análise do capital humano sobre o nível de renda dos estados brasileiros: MRW versus Mincer. *Estudos Econômicos (São Paulo)*, 40(1), 153-183. <https://doi.org/10.1590/S0101-41612010000100006>
- Coelho, L. (2009). Economia Feminista. En *Dicionário Internacional da Outra Economia* (pp. 128-133). Almedina.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates. <https://utstat.utoronto.ca/brunner/oldclass/378f16/readings/CohenPower.pdf>
- Durbin, J., & Watson, G. S. (1950). Testing for Serial Correlation in Least Squares Regression: I. *Biometrika*, 37(3/4), 409. <https://doi.org/10.2307/2332391>
- Faustino, I., Maia, K., Gomes, M. R., Mourao, P., & Araujo, E. (2023). Decomposing wage differences in Brazilian regions: A revised insight about traditional discrimination. *International Journal of Social Economics*, 50(1), 1-19. <https://doi.org/10.1108/IJSE-09-2021-0569>
- Fernandes, G. A. D. A. L. (2015). BRAZILIAN FEMALE LABOR MARKET: RACIAL-SKIN COLOR DISCRIMINATION AND INEFFICIENCY. *Economia Aplicada*, 19(2), 241-259. <https://doi.org/10.1590/1413-8050/ea85456>
- Green, S. B. (1991). How Many Subjects Does It Take To Do A Regression Analysis. *Multivariate Behavioral Research*, 26(3), 499-510. https://doi.org/10.1207/s15327906mbr2603_7
- Hirata, H. (2005). Globalização, trabalho e gênero. *Revista de Políticas Públicas*, 9(1), 111-128. <http://www.periodicoseletronicos.ufma.br/index.php/rppublica/article/view/3770/1848>

- Hirata, H. (2018). Gênero, patriarcado, trabalho e classe. *Trabalho Necessário*, 16(29), 14-27. <http://periodicos.uff.br/trabalhonecessario/article/view/4552/4195>
- Hirata, H., & Kergoat, D. (2007). Novas configurações da divisão sexual do trabalho. *Cadernos de Pesquisa*, 37(132), 595-609. <https://doi.org/10.1590/S0100-15742007000300005>
- Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. En *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics: Vol. 5.1* (pp. 221-234). University of California Press.
- Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3), 255-259. [https://doi.org/10.1016/0165-1765\(80\)90024-5](https://doi.org/10.1016/0165-1765(80)90024-5)
- Miller, T. L. based on F. code by A. (2024). *leaps: Regression Subset Selection*. <https://doi.org/10.32614/CRAN.package.leaps>
- Mincer, J. (1984). *Schooling, experience, and earnings*. Columbia University Press.
- Moretto, C. F. (2000, novembro). Função minceriana de determinação dos rendimentos individuais: Uma aplicação do método de variáveis instrumentais. *Revista Teoria e Evidência Econômica*, 8(15), 47-65.
- Moura, R. L. D. (2008). Testando as hipóteses do modelo de mincer para o Brasil. *Revista Brasileira de Economia*, 62(4), 407-449. <https://doi.org/10.1590/S0034-71402008000400003>
- Pujol, M. A. (1992). *FEMINISM AND ANTI-FEMINISM IN EARLY ECONOMIC THOUGHT*. Edward Elgar Publishing. <https://econpapers.repec.org/RePEc:elg:eebook:362>
- R Core Team. (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Reis, C. V. S., Moreira, T. B. S., & Cunha, G. H. de M. (2017). *O Efeito Marginal do Capital Humano na Agricultura Familiar*.
- Salvaro, G. I. J., Estevam, D. de O., & Felipe, D. F. (2014). Mulheres em cooperativas rurais virtuais: Reflexões sobre gênero e subjetividade. *Psicologia: Ciência e Profissão*, 34(2). <https://doi.org/10.1590/1982-3703000262013>
- Scott, J. (1995). Gênero: Uma categoria útil de análise histórica. *Educação e Realidade*, 20(2), 71-99. <https://seer.ufrgs.br/educacaoerealidade/article/view/71721/40667>
- Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4), 591. <https://doi.org/10.2307/2333709>
- Silva, T. B. D., Bondezan, K. D. L., & Lucas, M. D. S. (2022). Retornos para educação no Brasil: Enfoque nas regiões Sul e Nordeste para o ano de 2017. *Interações (Campo Grande)*, 67-83. <https://doi.org/10.20435/inter.v23i1.3052>

- Sina, A. (2005). *Mulher e trabalho: O desafio de conciliar diferentes papéis na sociedade*. 2005.
- Sucupira, F. (2016). Dupla presença feminina: Efeitos da divisão sexual do trabalho sobre o tempo cotidiano das mulheres urbanas de baixa renda. *Idéias*, 7(1), 15-40.
- Teixeira, M. O. (2008). Desigualdades salariais entre homens e mulheres a partir de uma abordagem de economistas feministas. *Desigualdades Salariais Entre Homens e Mulheres a Partir de Uma Abordagem de Economistas Feministas*, 31-45.
<https://periodicos.uff.br/revistagenero/article/view/30952/18041>
- Teixeira, W. M., & Menezes-Filho, N. A. (2012). Estimando o retorno à educação do Brasil considerando a legislação educacional brasileira como um instrumento. *Revista de Economia Política*, 32(3), 479-496. <https://doi.org/10.1590/S0101-31572012000300008>
- White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48(4), 817. <https://doi.org/10.2307/1912934>