

Unpacking gender and race segregation along occupational skills and socio-economic status in Brazil

Ben-Hur Francisco Cardoso¹ [0000-0003-1969-5669],

Laís Souza¹,

Flávio L. Pinheiro^{2*} [0000-0002-0561-9641] fpinheiro@novaims.unl.pt,

Liana Bohn¹ [0000-0003-3443-5048],

Dominik Hartmann^{1*} [0000-0002-4292-8654] dominik.hartmann@ufsc.br

¹ Departamento de Economia e Relações Internacionais Universidade Federal de Santa Catarina,
Florianópolis, SC, Brazil.

² NOVA Information Management School (NOVA IMS), Universidade NOVA de Lisboa, Campus
de Campolide, Lisboa 1070-312, Portugal

* Corresponding author

This is the Author Peer Reviewed version of the following article published by Springer Nature:

Cardoso, B-H. F., Souza, L., Pinheiro, F. L., Bohn, L., & Hartmann, D. (2025). Unpacking gender and race segregation along occupational skills and socio-economic status in Brazil. *Nature human behaviour*. <https://doi.org/10.1038/s41562-025-02272-9>

This version of the article has been accepted for publication, after peer review and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections.

Unpacking Gender and Race Segregation along Occupational Skills and Socioeconomic Status in Brazil

Ben-Hur Francisco Cardoso¹, Laís Souza¹, Flávio, L. Pinheiro²,
Liana Bohn¹, and Dominik Hartmann¹

¹*Departamento de Economia e Relações Internacionais Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil.*

²*NOVA Information Management School (NOVA IMS) – Universidade Nova de Lisboa, Portugal*

Abstract

The occupational specialization of social groups is closely tied to gender, racial, and class identities, segmenting the labor market into perceived White/Black and male/female roles and skills sets. Using data from 100 million formal workers in Brazil (2003–2019), we examine patterns of occupational segmentation across 426 occupations, identifying distinct skill demands and socioeconomic statuses linked to race/skin color and gender. Classifications of “male” or “female” occupations are shaped by required skills, while distinctions between “White” and “Black” occupations reflect socioeconomic status and historical inequalities. Women and men are segmented by gender-associated skill sets, such as engineering versus caregiving skills. Within these skill sets, strong hierarchical segregation persists, with Black individuals disproportionately concentrated in lower socioeconomic status positions. Despite recent socioeconomic changes, occupational specialization patterns have remained stable. Our findings highlight that the strong association between race and lower-status occupations must be addressed for a more inclusive society

Keywords: *skills, segregation, intersectionality*

Corresponding Authors:

- Flavio L. Pinheiro (fpinheiro@novaims.unl.pt)
- Dominik Hartmann (dominik.hartmann@ufsc.br)

INTRODUCTION

In the contemporary world, a distinct division of labor necessitates complementary occupations with specialized skills for the overall functioning of societies¹. To provide essential care, healthcare professionals require specialized medical knowledge, diagnostic skills, and a compassionate bedside manner. Information technology specialists navigate the digital realm, employing coding and problem-solving skills. Skilled artisans and designers contribute to the aesthetic fabric of society, utilizing creativity and a keen eye for detail in fields such as architecture or graphic design. While it is true that different occupations require specialization in different skills², such specialization does not translate to silos of disconnected occupations. Rather, the contrary. Recent methods from network and data science have helped to show that the relatedness between occupations in terms of skills relatedness forms a complex network structure that impacts the functioning of labor markets^{3,4} and their evolution^{5,6}. However, little is known in an empirically comprehensive manner, using data on the entire (formal) labor market of a large and heterogenous country, how skills and social status assign and segregate different social groups, such as White and Black persons, women and men, to different roles in these complex labor market structures.

This study leverages a dataset of 100 million Brazilian workers and employs econometric and network science methods to reveal that each gender tends to concentrate in distinct skill clusters. Moreover, our findings reveal a clear pattern within these gender-specific skill clusters: Black workers consistently occupy lower socioeconomic positions compared to their white peers.

While crucial for development⁷⁻⁹, the division of labor often leads to labor market segmentation, a historical process driven by political and economic forces dividing the labor market into submarkets¹⁰. In this process, occupational disparities in social valuation restrict certain social groups' access to high-value roles as the division of labor assigns different roles to different social groups¹¹. Those limits upon occupational opportunities, akin to any restriction on access to social spaces, constitute a form of segregation¹². Segregation within occupational settings can appear in two forms: vertical and horizontal¹³. In vertical segregation, social and economic status variations between occupations are evident, with certain roles perceived as better or worse than others, primarily based on their wages. For instance, comparing a bricklayer to a civil engineer highlights this vertical disparity. In the horizontal segregation, distinctions are based on

the nature of the work rather than perceived status, negating the idea of one occupation being significantly superior or inferior to another. For example, comparing a doctor to a civil engineer illustrates this horizontal segregation, emphasizing differences in job roles without implying a strong hierarchical ranking.

The recent availability of large-scale data sources has revealed novel patterns of segregation¹⁴. Notably, recent research highlights pronounced divisions among racial and income groups in various domains, including workplaces, neighborhoods, and transit systems^{15–17}. In this study, we empirically explore the dynamics of vertical and horizontal occupational segregation and its contribution to the disparities in occupational opportunities among different social groups defined by gender and race^{18–22}. Using a dataset of 100 million workers in Brazil, we employ econometric and network science techniques to reveal that each gender tends to specialize within specific skill clusters. More importantly, we show that within these gender-tagged skill clusters, there exists a discernible pattern: black workers are found to be occupying lower socioeconomic positions compared to their white counterparts. Hence, in the Brazilian labor market, workers experience vertical segregation based on their race/skin color and horizontal segregation based on their gender. **Figure 1** illustrates this pattern that is discernible along the socioeconomic status of occupations and their levels of racial and gender segregation.

For several reasons, Brazil is an exceptional case for understanding racial and gender segregation^{23–25}. As a nation marked by significant gender and racial disparities, Brazil provides a clear lens through which occupational specialization can be studied. Furthermore, the Brazilian RAIS dataset provides an extensive resource for a detailed analysis of gender, race/skin color, and 426 occupations (at the 4-digit CBO classification), encompassing over 100 million formal labor market workers from 2003 to 2019. Robustness tests can be conducted using Census Data (see the Methods for details), providing insights into the informal labor market.

In exploring the occupational distribution among social groups, the literature on intersectionality underscores discrimination experiences, particularly among women of different races, with a notable emphasis on black women in the United States^{19,26} and Latin America^{18,27}. Conversely, the literature on racial discrimination reveals more similarities among women of different races, a pattern less evident among men^{20–22}. This intricate interplay of gender and race in longstanding and stable occupational differences shows how structural intersectionality shapes the distribution of social groups in the labor market and how these attributes influence the allocation of workers to specific roles.

The current body of literature delves deeply into the segmentation of occupations based on gender and race, exposing the specific occupations favored by each group^{28–33}. However, a notable gap in the existing research lies in interpreting racial and gender specialization as vertical or horizontal segregation, a void our study seeks to address. Economic, political, and social processes significantly influence the segregation by race and gender. During the period analyzed in this article, Brazil underwent various cycles, including periods of accelerated growth and economic and political crises. Our results suggest that the economic expansion in the early 2000s could have been a “segregated expansion,” where women and Black individuals remained in occupations distinct from those of white men. Furthermore, although the labor supply became more qualified in educational terms and more diverse in terms of race and gender, this expansion did not automatically translate into an egalitarian labor market with increased opportunities for all social groups. Our research significantly contributes to large-scale empirical investigations by employing robust statistical methods. We specifically explore how skill and occupational status assignments play a quantifiable role in shaping the segregation patterns of social groups, focusing on gender and race/skin color within diverse occupations.

RESULTS

Figure 2A shows the Occupation Space, a network where pairs of occupations are connected based on the overlap of skills necessary to perform both³⁴, with nodes (*i.e.*, occupations) colored according to their ISCO 08 Major Group. Thus, occupations requiring similar skills tend to be closer in the network, forming skill-related clusters (see Section 9 of the **Supplementary Information** for more details).^{35,36}

Figure 2B-E shows each social group’s *Relative Specialization* (RS) (men/women and black/white) in the occupation space. RS gauges how a social group engages in an occupation more frequently than “expected” based on age structure (see Methods for details). In that sense, an RS greater than one implies that a social group is overrepresented in a particular occupation, indicating more individuals working in it than expected from a random distribution.

Moreover, we use the Internacional Socio-Economic Index (ISEI)^{35,36} scores to rank the social status of occupations in **Figure 2B-E**. Status is not randomly distributed in the occupation space (see **Supplementary Figure 7**). Indeed, occupations with a higher

socioeconomic status, such as Doctors and Judges, tend to cluster in the center towards the right side of the network where more socio-cognitive skills-based occupations are located. (**Supplementary Table 23** provides a complete list of occupations ranked by ISEI). More importantly, **Figure 2B-E** allows us to see that, for instance, White Women exhibit RS typically in occupations of high socioeconomic prestige (high ISEI score), while Black Men reveal the opposite pattern.

Also, clear specialization and segregation patterns can be observed here, summarized in the scatter **Figure 3**, which plots occupations by their ISEI and a one-dimensional PCA-based projection of skills. First, women tend to focus rather on occupations, such as psychologists and domestic housekeepers, requiring socio-cognitive skills (on the right side of the occupation space). In contrast, on average, men are more biased towards occupations, such as civil engineers and bricklayers that require more physical skills. However, there is also a strong racial divide. Black women have relative specializations in more peripheric parts of the occupation space with lower status values, exemplified by domestic housekeepers. Black men also have significantly fewer relative specializations than white men, and these specializations are strongly biased towards lower ISEI occupations on the right side of the occupation space, exemplified by bricklayers. **Supplementary Table 5** provides the top ten occupations by gender and race/skin color, including information on ISEI values and key skills associated with these occupations.

DECODING OCCUPATIONAL SEGREGATION

Next, we make use of regressions (1) to analyze to which extent different types of skills and levels of status are linked with the occupational specialization of four social groups, (2) to show if indeed a significant horizontal skills segregation by gender, and vertical status segregation by race/skin color can be observed, and (3) to analyze to which extent these results are stable between 2003 and 2019; noteworthy a period with substantial economic up- and downswings, and institutional changes and crises. In summary, for workers in social group g working in occupation o in year t , we make the following OLS regression:

$$RS_{g,o}^t = \alpha_{g,t} + \beta_{g,s,t} \cdot \widehat{ISEI}_o + \sum_h \beta_{g,h,t} \cdot \theta_{o,h} + \epsilon_{g,o}^t,$$

where $\widehat{ISEI}_o$ (adjusted ISEI) is our metric for social status of occupation o , $\theta_{o,h} = 1$ indicates that occupation o is specialized in skill h and $\theta_{o,h} = 0$ otherwise. See **Methods**

for more details. The coefficient of adjusted ISEI ($\beta_{g,s,t}$) serves as an indicator of vertical segregation and the coefficients of skill specialization ($\beta_{g,h,t}$) measures the extent of horizontal segregation for a social group.

Table 1 provides an illustrative example showing that, in 2010, different types of skills exhibited substantially different associations with the four social groups. This example also allows for robustness checks using census data from the same year. The core qualitative results remain consistent across different years (see **Supplementary Tables 5 to 21**) and when using census data (see **Supplementary Table 22**). For instance, while “communication, collaboration and creativity” shows a strong, positive, and significant effect with women (in particular white women), it has a negative association with men (in particular Black men). “Working with machinery and equipment” instead is negatively associated with women but strongly positive with men. Management skills do not show a significant association, but the coefficient for white persons is notably positive and Black ones negative. As discussed below, the temporal evolution of the Adjusted ISEI (see **Figure 4 and 5**) shows that these patterns are remarkably stable over time.

Table 1 also includes the “adjusted ISEI” indicator to gauge socioeconomic status, capturing the status component independent of the average human capital required in the occupation (see **Methods** and **Supplementary Information** for details). This approach is crucial to mitigate multicollinearity issues, particularly given that certain skills are associated with occupations of higher socioeconomic status. The regressions for each social group and year are conducted individually. So, implicitly, all standard deviations of the estimators are based on a heteroscedasticity-robust covariance matrix regarding social groups and years. See **Supplementary Table 4** for multicollinearity tests. Additionally, results remain qualitatively robust when accounting for various potential confounders, including age, education level, and region (see **Supplementary Information Section 10**).

GENDER AND HORIZONTAL SEGMENTATION

The regression results indicate that skills segregate people – primarily – based on gender. **Figures 4A and 4B** show the temporal evolution of the coefficient – for different social groups – of skills positively related to female specialization. The results are remarkably stable over time, with a small exception of Black women, whose positive association with caring and assisting increased in the early 2000s, and whose negative coefficient of working with machinery has become slightly less negative over the period 2003-2019.

As shown in **Figure 4A**, regardless of race, an occupation requiring *communication, collaboration, and creativity* increases the relative specialization of women in that occupation by ~39% and decreases the relative specialization of men in that occupation by ~28%. Similarly, as shown in **Figure 4B**, almost regardless of race, an occupation requiring *assisting and caring* increases the relative specialization of women in that occupation by ~27% and decreases the relative specialization of men in that occupation by ~14%. Despite the same direction of the effect, the impact is marginally more pronounced for white men.

Figures 4C and 4D show the temporal evolution of the coefficient – for different social groups – of skills positively related to male specialization. Thus, as shown in **Figure 4C**, regardless of race, an occupation requiring *constructing* increases the relative specialization of men in that occupation by ~28% and decreases the relative specialization of women in that occupation by ~37%. Despite the same direction of the effect, the impact is marginally more pronounced for Black men. Similarly, as shown in **Figure 4D**, almost regardless of race, an occupation requiring *working with machinery and specialized equipment* increases the relative specialization of men by ~21% and decreases the relative specialization of women in that occupation by ~33%.

We conducted the same regression for each age cohort to test the stability of our results across generations, see **Supplementary Information Section 11**. We found that, for every age group and year, the coefficients associated with “female”-related skills in this analysis are consistently higher for women than for men in all cases, and vice versa. The other skills analyzed in the regression proved neutral, not showing a clear relationship with the social groups. In other words, gender horizontal segregation does not rely on all skills; rather, it predominantly manifests in skills culturally more associated with each gender. The temporal evolution of the coefficients for these other skills can be found in **Supplementary Figures 2 to 4**.

RACE AND VERTICAL SEGREGATION

In **Figure 5**, we show the temporal evolution of the coefficient – relative to the adjusted socioeconomic status – for different social groups. There is almost no statistical difference between the values for Black women and Black men, and white women and white men. Thus, occupational segregation based on socioeconomic status is only expressed regarding race/skin color. Therefore, regardless of gender, an occupation having one standard deviation higher in adjusted socioeconomic status increases the

relative specialization of white individuals in that occupation by $\sim 7\%$ and decreases the relative specialization of Black individuals by $\sim 10\%$. These measures appear relatively stable over time. We conducted the same regression for each age cohort to test the stability of our results across generations, see **Supplementary Information Section 11**. We found that, for every age group and year, the coefficient associated with status is higher for white individuals than for Black individuals in 99,41% of cases. Notably, stronger differences in the ISEI coefficient between gender and race/skin color are evident in younger age cohorts. White women, initially in higher-status jobs, experience a relative decline in their ISEI coefficient during childbearing years (ages 20-40), while Black men show a relative increase over time. The asymmetric effects of parenting likely contribute to this pattern, but further research is needed to proxy these effects, as the data lacks direct information on childbearing and parenting.

DISCUSSION

In this work, we investigate the unequal occupational opportunities encountered by distinct social groups, delineated by gender and race/skin color. This inquiry is rooted in the recognition that certain occupations, shaped by their socioeconomic status and requisite skills, become linked to specific social characteristics. In other words, occupational segregation by gender and race can, in its aggregate, be explained by pre-market discriminations – such as historical factors and various socialization processes –, as well as by exclusionary employment practices^{37–39}. This connection adds to the intricate network of social and economic divisions within the labor market.

Previous studies have identified specific occupations in which different genders and races tend to concentrate, such as women in caregiving roles or low-paid jobs in the textile industry. Here, we apply network science and econometric methods to map the relative specialization of social groups (men/women, Black/White) within the occupation space—a network based on the (dis)similarity of skills across hundreds of occupations. This approach allows us to provide empirical evidence that social groups specialize in distinct parts of the occupation space, and this specialization is not random but driven by common factors like skill requirements and socioeconomic status.

Our analysis demonstrates that, in terms of socioeconomic status, Black workers face a longstanding disadvantage compared to their White counterparts, irrespective of gender. This holds true when considering specific occupations, with Black individuals more likely

to occupy lower socioeconomic positions. Interestingly, the study highlights that gender-based occupational segregation operates independently of race/skin color, with no strong statistical difference observed between Black women and men or White women and men. Examining the temporal evolution of skills related to specialization, the results underscore the persistence of gender-based occupational patterns. Occupations requiring skills emphasizing communication, collaboration, and creativity exhibit increased relative specialization of female participation, while those associated with construction and working with machinery attract higher male relative specialization.

Our results are derived from an extensive dataset encompassing 100 million Brazilian workers. However, our findings on gender and race/skin color have broader implications that extend beyond the borders of Brazil, emphasizing challenges faced by historically marginalized groups, although these differences in opportunities are more pronounced in some countries than in others^{40–42}. The negative effects of slavery, a pervasive historical injustice, have echoed across the globe^{43–45}. Consequently, we anticipate that the challenging positioning of Black individuals, particularly Black men, is not confined to Brazil but resonates globally^{46,47}. Similarly, considerations about the roles of women in the workforce transcend national boundaries. The universality of such challenges underscores the need for a comprehensive understanding of the structural factors influencing race and gender dynamics in the labor market. By acknowledging the historical context, our study contributes to a broader dialogue on social inequalities.

The findings reveal a nuanced landscape where workers experience distinct forms of segregation. Despite some studies indicating a decrease in occupational segregation, particularly gender-based^{48,49}, our study reveals persistent segregation in the last 20 years within occupations distinguished by both gender and race/skin color. This persistent phenomenon stands as a hallmark of socioeconomic structure of modern society, underscoring the need for continued exploration and intervention to address these deeply ingrained disparities.

The study recognizes its limitations. First, we did not examine how generational differences, defined by age cohorts⁵⁰, or ethnical differences influence labor market segmentation in the Brazilian context. There is also scope for further exploration of regional disparities and the effects of varying racial compositions on local labor market outcomes, particularly given Brazil's vast territorial and demographic diversity. Findings on regional variation may also inform comparative analyses in other large, diverse countries such as the United States and China. Moreover, by focusing exclusively on

formal labor market participants, the analysis excludes a significant share of the working-age population—especially women—who are outside the formal labor force due to caregiving responsibilities or limited local employment opportunities. The reliance on a single indicator of occupational status, the International Socio-Economic Index (ISEI), may further limit the analysis by inadequately reflecting the complexity of Brazil’s sociocultural and labor market structures. Finally, greater attention to institutional variation across countries and regions could improve our understanding of global patterns of both vertical and horizontal occupational segregation. Despite these limitations, the study draws on a large dataset and presents robust evidence of horizontal gender-based occupational segregation by skill, as well as pronounced vertical stratification by race/skin color.

METHODS

We use microdata sourced from the *Relação Anual de Informações Sociais* (RAIS), an annual administrative record provided by all companies, designed to meet the information needs of government agencies in Brazil. The dataset encompasses all individuals with formal employment, representing approximately 70% of the Brazilian labor market. We focus on adult workers (at least 18 years old) employed on December 31, between 2003 and 2019.

Understanding social race in Brazil involves recognizing that visible physical traits such as skin color predominantly determine racial classification, more than family ancestry⁵¹ — a contrast to contexts where ancestral/ethnic lineage fundamentally defines racial categorization. Unlike classifications in other contexts, Brazil acknowledges a spectrum of racial identities through its largest census categories: *branco* (light-skinned), *pardo* (brown-skinned), and *preto* (dark-skinned)⁵². In this context, the term *“negro”* serves as an inclusive category that encompasses the diverse racial composition of the population by intersecting both *“pardo”* and *“preto”* classifications, making it the most relevant racial category for affirmative action policies in the country⁵². Here, when we refer to a “Black person,” we align with the Brazilian understanding of *“negro.”* Therefore, the Asian and indigenous categories, representing less than 1% of the sample, have been excluded.

Gender markers follow a binary categorization – women and men – facilitated by RAIS. Gender and race/skin color are analyzed jointly, meaning the results are considered

for Black women, white women, Black men, and white men. In **Table 2**, we observe the large number of workers comprising these data for each social group and year. The substantial number of observations, approximately 30 million workers per year, is the primary advantage of RAIS microdata.

We use the 4-digit ISCO-08 to classify occupations for Brazilian workers. The *International Socioeconomic Index* (ISEI)³⁵ assigns a socioeconomic status value to each ISCO-08 occupation. Ganzeboom³⁶ defines ISEI as a measure that captures the ability of an occupation to “maximize the conversion of education into income.” Judges have the highest ISEI (88.96), and Hunters have the lowest (11.01). Moreover, we utilize the *European Skills, Competences, Qualifications, and Occupations* (ESCO), a database mapping the set of skills required by each occupation following the ISCO-08. We use version 1.1 of ESCO, available from <https://esco.ec.europa.eu/en/publication/skills-occupations-matrix-tables>. ESCO tables present various levels of detail. The first level has eight broad skills: “communication, collaboration, and creativity”, “information skills”, “assisting and caring”, “management skills”, “working with computers”, “handling and moving”, “constructing”, and “working with machinery and specialized equipment”.

Regarding the classification of workers' occupations, RAIS follows the *Classificação Brasileira de Ocupações* (CBO). As described below, this work aims to identify how horizontal segregation is based on the skills of occupations and how vertical segregation is based on socioeconomic status segment workers in the Brazilian labor market. However, information about the set of demanded skills and the socioeconomic status of occupations is not available for CBO but for the *International Standard Classification of Occupations* 2008 (ISCO-08). Currently, there is the Quadro Brasileiro de Qualificações, which maps the skills of each CBO occupation, but is still in a testing phase, so we opted to use already consolidated data to ensure the robustness of our analysis.

A correspondence table for the 4-digit ISCO-08 is feasible with the CBO-2002, which is available after 2003. The Brazilian Census of 2010 classifies the jobs of interviewed Brazilian workers into two classifications: (i) 4-digit CBO and (ii) *Classificação de Ocupações para Pesquisas Domiciliares* (COPD), the Brazilian version of ISCO-08. No correspondence table between COPD and ISCO-08 is necessary, except for a small number of trivial cases. With this census data, we have an “empirical” correspondence table between 4-digit CBO 2002 and ISCO-08. As given, for each 6-digit CBO 2002, we manually chose the most suitable 4-digit ISCO-08. The *Brazilian Census* of 2010

employs the *Classificação de Ocupações para Pesquisas Domiciliares* (COPD), the Brazilian version of ISCO-08 for occupation coding, eliminating the need for a correspondence table between COPD and ISCO-08.

All data were analyzed using the Python programming language. The original datasets were read and processed using the Pandas library (version 1.4.0), and basic mathematical operations were performed with NumPy (version 1.20.3). Statistical models and regressions were estimated using statsmodels (version 0.13.2), the PCA analysis was done in Wolfram Mathematica (version 13.0). Figures and visualizations were produced with Matplotlib (version 3.2.0) and Wolfram Mathematica (version 13.0).

RELATIVE SPECIALIZATION

Let be $N_{g,a,o}^t$ the number of workers from social group g and age group a working in occupation o in year t . Consequently, the Relative Specialization of social group g working in occupation o in year t is:

$$RS_{g,o}^t = \frac{N_{g,o}^t}{Z_{g,o}^t}, \quad (1)$$

where the omission of a subscript indicates the summation over the respective set. $Z_{g,o}^t$ provides a baseline expectation of the relative specialization of such occupation from each social group. Therefore, since certain occupations are generally only accessed at an older age or the beginning of careers, and different social groups have different age compositions, we could calculate the expected value for each social group by the average of each age group in the occupation:

$$Z_{g,o}^t = \sum_a N_{g,a}^t \frac{N_{a,o}^t}{N_a^t}. \quad (2)$$

The age groups considered in this study are intervals of 5 years (see Supplementary Information for the detailed breakdown of observations per age group and social group). If $RS_{g,o}^t > 1$, we say that social group g is overrepresented in occupation o in year t . Similarly, if $RS_{g,o}^t < 1$, we say that social group g is underrepresented in occupation o in year t . Thus, this indicator will serve as the basis for identifying how occupations are segmented in terms of the different social groups studied here.

OCCUPATIONS, SKILLS, AND SOCIOECONOMIC STATUS

Having mapped the occupations associated with each social group, it is now important to understand how this segmentation occurs. In other words, to determine whether the

occupational specialization of different social groups is random or if it is possible to identify the common denominators among the occupations in which each social group specializes. To that end, we use the analysis based on the eight skills required by each occupation (ESCO).

We denote $0 \leq \varphi_{o,h} \leq 1$ as the fraction that occupation o requires skill h . By construction,

$$\sum_h \varphi_{o,h} = 1 \quad \text{for all occupation } o. \quad (3)$$

where $\theta_{o,h} = I(\varphi_{o,h} > \langle \varphi_h \rangle)$ and represents the indicator of skill specialization in each occupation. In other words, if an occupation o demands a skill h above the average, we say that occupation o specializes in skill h ($\theta_{o,h} = 1$). On the other hand, if this demand is below or equal to the average, we say that occupation o is not specialized in skill h ($\theta_{o,h} = 0$).

Certain skills are associated with occupations of higher socioeconomic status; however, we may face multicollinearity issues in econometric regressions that consider both skills and socioeconomic status together. To avoid this, we perform the regression:

$$\text{ISEI}_o = \mu + \sum_{h=1}^7 \rho_h \theta_{o,h} + \varepsilon_o \quad (4)$$

and define the adjusted ISEI of occupation o ($\widehat{\text{ISEI}}_o$) as the z-transformation of the regression residual:

$$\widehat{\text{ISEI}}_o = \frac{\hat{\varepsilon}_o - \langle \hat{\varepsilon} \rangle}{\text{std}(\hat{\varepsilon})}. \quad (5)$$

In the regression, we use up to the seventh skill to avoid issues of strict linear dependence, given that the sum of φ s for each occupation, which always adds up to 1. Thus, using the adjusted ISEI in regressions, we can measure the isolated effect of socioeconomic status beyond what can be explained by skills. As we will demonstrate in the next section, the adjusted ISEI is highly correlated with the ISEI, facilitating the interpretation of results.

SKILLS PROJECTION

To obtain a one-dimensional projection of the eight ESCO Skills, we perform a correlation-based *Principal Component Analysis* (PCA) to Φ , a matrix that informs on the fraction that occupation o requires skill h ($\varphi_{o,h}$). The projection obtained corresponds

to the linear component of the original eight-dimensional space of ESCO Skills that captures the highest variance (32%).

The obtained projection is validated by matching its ability to discriminate skills according to gender and between skills groups. **Figure S6** of the **SI** shows that positive values of the projection tend to represent occupations mostly dominated by women and dominated by skills associated with “communication collaboration and creativity” and “management skills” while negative values are linked with men and with skills on “handling and moving” and “working with machinery and specialised equipment”.

ECONOMETRIC STRATEGY

Both skills and the socioeconomic status of occupations can explain occupational segregation regarding gender and race/skin color. We formalize this hypothesis for workers in social group g working in occupation o in year t through the regression:

$$RS_{g,o}^t = \alpha_{g,t} + \beta_{g,s,t} \cdot \widehat{ISEI}_o + \sum_{h=1}^7 \beta_{g,h,t} \cdot \theta_{o,h} + \epsilon_{g,o}^t. \quad (6)$$

The coefficient of adjusted ISEI ($\beta_{g,s,t}$) serves as an indicator of the vertical segregation of a social group. A positive value indicates that the social group is vertically segregated in occupations with higher status. On the other hand, the coefficients of skill specialization ($\beta_{g,h,t}$) measure the extent of horizontal segregation for a social group. A positive value suggests that the social group is horizontally segregated in occupations that specifically demand those skills. The residuals will be robust to heteroscedasticity by social group and year (that is, conducting a regression for each year and social group). Once again, we use up to the seventh skill to avoid problems of strict linear dependence, given that the sum of φ s for each occupation is always 1.

Here, we measure socioeconomic status using the *International Socioeconomic Index* (ISEI) rather than the occupational average wage. This decision was motivated by the potential issues of inverse causality and endogeneity bias associated with the occupational average wage. Simple regressions using the average wage might not distinguish whether social groups are inherently in low-wage occupations or if the occupational average wage is lower due to being predominantly composed of social minorities. By opting for ISEI, an independent measure derived from international data, we circumvent this concern, allowing for a more accurate and unbiased assessment of the relationship between social groups and occupational status.

DATA AVAILABILITY

RAIS data is available from the Brazilian Ministry of Labor and Employment portal (<https://www.gov.br/trabalho-e-emprego/pt-br/assuntos/estatisticas-trabalho/microdados-rais-e-caged>).

Census data is available from the Brazilian Institute of Geography and Statistics (IBGE) portal (<https://www.ibge.gov.br/estatisticas/multidominio/cultura-recreacao-e-esporte/9662-censo-demografico-2010.html?=&t=microdados>).

Both RAIS and Census data are freely available for unrestricted use without licensing, patenting, or control mechanisms, as explicitly stated on the government's open data policy page at <https://www.gov.br/trabalho-e-emprego/pt-br/aceso-a-informacao/dados-abertos>.

ESCO Skills-Occupation Matrix Tables are available from <https://esco.ec.europa.eu/en/publication/skills-occupations-matrix-tables>.

The official titles and hierarchical structure of ISCO-08 occupation groups are available from the International Labour Organization's ILOSTAT portal at <https://ilostat.ilo.org/methods/concepts-and-definitions/classification-occupation/>.

The datasets necessary for reproducing the findings presented in this manuscript and supplementary information are available at <https://doi.org/10.6084/m9.figshare.27184176>.

CODE AVAILABILITY

All the code necessary to reproduce the results presented in this manuscript and supplementary information can be downloaded at <https://doi.org/10.6084/m9.figshare.27184176>.

ACKNOWLEDGMENTS

BFC and LS would like to express their gratitude for the financial support of the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Finance Code 001, and DH for the support of CNPq (406943/ 2021-4 and 315441/2021-6). BFC and DH would like to thank the Ministry of Economics of Brazil for access to the 2003-2019 RAIS database Processo SEI n.º 19965.106679/2021-1. FLP acknowledges the funding

support from FCT (Fundação para a Ciência e a Tecnologia), under the project - UIDB/04152 - Centro de Investigação em Gestão de Informação (MagIC)/NOVA IMS. DH would like to thank Cristian Jara-Figueroa, Mary Kaltenberg and Cesar Hidalgo for the insights on occupational networks in Brazil and the many fruitful discussions held at the MIT Media Lab in 2014-2016. All authors are grateful to Celi Scalón, Eva Catela, Guilherme Moura, Karine de Souza Silva, participants at the SASE Rio conference 2023, and members of Nocode-UFSC group meetings for valuable comments.

The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

AUTHOR CONTRIBUTION STATEMENT

B.F.C., D.H., L.S., L.B. and F.L.P. conceived the study and designed the experiments. B.F.C. performed the experiments and analyzed the data. All Authors participated in the writing and revision of the manuscript.

COMPETING INTERESTS STATEMENT

The Authors declare no competing interests.

TABLES

Table 1 - OLS estimates of the marginal values of relative occupation specialization for each social group in 2010. Results for the remaining years of analysis can be found in the SI. Asterisks indicate the level of significance at p-values 0.001 (***), 0.01 (**), and 0.05 (*).

Variable	Occupational Relative Specialization											
	Black Women			White Women			Black Men			White Men		
	Coef	95% CI	P	Coef	95% CI	P	Coef	95% CI	P	Coef	95% CI	P
Intercept	0.709***	0.478, 0.940	<0.001	0.823***	0.610, 1.036	<0.001	0.998***	0.862, 1.134	0.000	1.213***	1.081, 1.344	<0.001
θ: communication, collaboration, and creativity	0.308**	0.124, 0.492	<0.001	0.498***	0.328, 0.667	<0.001	-0.352***	-0.460, -0.243	<0.001	-0.199***	-0.304, -0.095	<0.001
θ: information skills	-0.070	-0.226, 0.085	0.374	-0.005	-0.148, 0.138	0.949	-0.069	-0.160, 0.023	0.141	0.065	-0.023, 0.154	0.147
θ: assisting and caring	0.340***	0.184, 0.497	<0.001	0.256***	0.111, 0.400	0.001	-0.022	-0.115, 0.070	0.635	-0.267***	-0.356, -0.177	<0.001
θ: management skills	-0.106	-0.273, 0.061	0.213	0.001	-0.153, 0.155	0.988	-0.025	-0.124, 0.073	0.615	0.044	-0.051, 0.139	0.366
θ: handling and moving	0.067	-0.118, 0.253	0.476	-0.049	-0.220, 0.121	0.571	0.124*	0.015, 0.233	0.026	-0.061	-0.166, 0.044	0.254
θ: constructing	-0.372**	-0.625, -0.119	0.004	-0.419***	-0.652, -0.186	<0.001	0.456***	0.307, 0.605	<0.001	0.099	-0.045, 0.243	0.175
θ: working with machinery and specialized equipment	-0.317***	-0.502, -0.132	<0.001	-0.372***	-0.542, -0.202	<0.001	0.216***	0.107, 0.325	<0.001	0.204***	0.099, 0.309	<0.001
Adjusted ISEI	-0.050	-0.118, 0.018	0.150	0.090**	0.027, 0.153	0.005	-0.133***	-0.173, -0.093	<0.001	0.045*	0.007, 0.084	0.022
Observations	364			364			364			364		
Adjusted R ²	0.201			0.356			0.463			0.258		

1

2

Table 2 – Number of workers in the RAIS dataset, by social group and year.

Number of formal workers (millions)				
Year	Black women	white women	Black men	white men
2003	2.63	6.84	4.81	10.10
2004	2.84	7.30	5.29	10.63
2005	3.23	7.87	5.93	11.22
2006	2.67	6.58	5.81	10.68
2007	2.94	7.00	6.35	11.20
2008	3.23	7.40	6.81	11.54
2009	3.52	7.71	7.22	11.73
2010	3.96	8.18	7.91	12.28
2011	4.40	8.55	8.43	12.54
2012	4.77	8.76	8.76	12.54
2013	5.11	8.88	9.10	12.48
2014	5.47	8.87	9.34	12.17
2015	5.44	8.49	9.02	11.46
2016	5.29	8.09	8.61	10.81
2017	5.31	7.91	8.68	10.53
2018	5.51	7.81	9.00	10.28
2019	5.55	7.63	9.06	9.93

3

FIGURE LEGENDS/CAPTIONS

Figure 1 – Example of vertical and horizontal segregation in Brazil in terms of race/skin color and gender. The Y-Axis represents the occupation’s International Socioeconomic Index (see Methods). At the same time, the X-Axis illustrates the skills necessary to perform each occupation typically associated with female and male social groups. We present two occupations, “early childhood educators” and “domestic housekeepers”, both requiring “communication, collaboration, and creativity” and “assisting and caring” skills. As detailed in the Results section, predominantly women work in these occupations in Brazil. However, despite similar skill requirements, black women work more frequently in occupations with a lower socioeconomic status. Moreover, “mechanical engineering technicians” and “metal processing plant operators” are characterized by skills such as “constructing” and “working with machinery and specialized equipment”. As outlined in the Results section, predominantly men work in these occupations, but black men are more frequently in the occupations with the lower socioeconomic status.

Figure 2– The occupational space of Brazil. Panel A shows the occupational space with each node corresponding to an occupation ISCO-08 at four digits. Colors indicate ISCO 08 Major Group. **Panel B, C, D, and E** show which occupations are dominated by workers of different race/skin color (Black or White) and genders (women or men). Nodes are colored according to the socioeconomic status of the occupation: yellow represents high and dark-gray low values; white nodes indicate occupations in which that social group is underrepresented in terms of their relative specialization. See Supplementary Information section 9 for more details.

Figure 3 – The scatterplot shows how the occupations compare regarding their ISEI score (Y-Axis) and the one-dimensional Socio-Cognitive Skills projection (X-Axis), see Methods for more details. Each light-colored dot/cross is an occupation. Dots/Crosses correspond to occupations where men/women have the highest Relative Specialization and color, indicating the race/skin color (Black or White). The highlighted symbols indicate the centroids of each social group. The Socio-Cognitive Skills correspond to a one-dimensional projection obtained from applying the PCA (*Principal Component Analysis*) as a dimension reduction method. As shown, it essentially lays occupation along

a dimension that separates men's (negative values) and women's (positive values) relevant skills.

Figure 4 – Temporal evolution of the coefficient of skills associated positively with the female group (A and B) and male group (C and D). The colored surfaces correspond to a 95% Confidence Interval.

Figure 5 – Temporal evolution of the adjusted ISEI (beta) coefficient for the different social groups. The colored surfaces correspond to 95% Confidence Interval.

REFERENCES

1. Neffke, F. M. H. The value of complementary co-workers. *Sci Adv* **5**, (2019).
2. Neffke, F. & Henning, M. Skill relatedness and firm diversification. *Strategic Management Journal* **34**, 297–316 (2012).
3. Alabdulkareem, A. *et al.* Unpacking the polarization of workplace skills. *Sci Adv* **4**, (2018).
4. Moro, E. *et al.* Universal resilience patterns in labor markets. *Nat Commun* **12**, (2021).
5. Frank, M. R. *et al.* Network constraints on worker mobility. *Nature Cities* **1**, 94–104 (2024).
6. Frank, M. R. *et al.* Toward understanding the impact of artificial intelligence on labor. *Proceedings of the National Academy of Sciences* **116**, 6531–6539 (2019).
7. Boschma, R. Relatedness as driver of regional diversification: a research agenda. *Reg Stud* **51**, 351–364 (2016).
8. Neffke, F., Henning, M. & Boschma, R. How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions. *Econ Geogr* **87**, 237–265 (2011).
9. Smith, A. An Inquiry into the Nature and Causes of the Wealth of Nations. in *The Glasgow Edition of the Works and Correspondence of Adam Smith, Vol. 2: An Inquiry into the Nature and Causes of the Wealth of Nations, Vol. 1* (ed. Todd, W. B.) (Oxford University Press, 1776). doi:10.1093/oseo/instance.00043218.
10. Reich, M., Gordon, D. M. & Edwards, R. C. A theory of labor market segmentation. *Am Econ Rev* **63**, 359–365 (1973).

- 69 11. Grusky, D. *Social Stratification, Class, Race, and Gender in Sociological*
70 *Perspective, Second Edition*. (Routledge, 2019). doi:10.4324/9780429306419.
- 71 12. Freeman, L. C. Segregation in Social Networks. *Sociological Methods &*
72 *Research* **6**, 411–429 (1978).
- 73 13. Blackburn, R. M., Brooks, B. & Jarman, J. The Vertical Dimension of
74 Occupational Segregation. *Work, Employment & Society* **15**, 511–538 (2001).
- 75 14. Dong, X. *et al.* Segregated interactions in urban and online space. *EPJ Data Sci* **9**,
76 (2020).
- 77 15. Boterman, W. R. & Musterd, S. Cocooning urban life: Exposure to diversity in
78 neighbourhoods, workplaces and transport. *Cities* **59**, 139–147 (2016).
- 79 16. Wang, Q., Phillips, N. E., Small, M. L. & Sampson, R. J. Urban mobility and
80 neighborhood isolation in America’s 50 largest cities. *Proceedings of the National*
81 *Academy of Sciences* **115**, 7735–7740 (2018).
- 82 17. Moro, E., Calacci, D., Dong, X. & Pentland, A. Mobility patterns are associated
83 with experienced income segregation in large US cities. *Nat Commun* **12**, 4633
84 (2021).
- 85 18. Gonzalez, L. *Por Um Feminismo Afro-Latino-Americano*. (Editora Schwarcz-
86 Companhia das Letras, 2020).
- 87 19. Collins, P. H. *Intersectionality as Critical Social Theory*. (Duke University Press,
88 2019). doi:10.1215/9781478007098.
- 89 20. Mandel, H. & Semyonov, M. Going Back in Time? Gender Differences in Trends
90 and Sources of the Racial Pay Gap, 1970 to 2010. *Am Sociol Rev* **81**, 1039–1068
91 (2016).
- 92 21. Wingfield, A. H. Racializing the Glass Escalator: Reconsidering Men’s
93 Experiences with Women’s Work. *Gender & Society* **23**, 5–26 (2009).
- 94 22. Chetty, R., Hendren, N., Jones, M. R. & Porter, S. R. Race and Economic
95 Opportunity in the United States: an Intergenerational Perspective*. *Q J Econ* **135**,
96 711–783 (2019).
- 97 23. Hartmann, D., Jara Figueroa, C., Kaltenberg, M. & Gala, P. Mapping
98 Stratification: The Industry-Occupation Space Reveals the Network Structure of
99 Inequality. *SSRN Electronic Journal* (2019) doi:10.2139/ssrn.3399239.
- 100 24. Ribeiro, R. & Araújo, G. S. Segregação ocupacional no mercado de trabalho
101 segundo cor e nível de escolaridade no Brasil contemporâneo. *Nova Economia* **26**,
102 147–177 (2016).

25. Souza Silveira, L. & Siqueira Leão, N. Segregação ocupacional e diferenciais de renda por gênero e raça no Brasil: uma análise de grupos etários. *Rev Bras Estud Popul* **38**, 1–22 (2021).
26. Paul, M., Zaw, K. & Darity, W. Returns in the Labor Market: A Nuanced View of Penalties at the Intersection of Race and Gender in the US. *Fem Econ* **28**, 1–31 (2022).
27. de Almeida, T. M. C. Interseccionalidade de violências, discriminações e resistências: as trabalhadoras terceirizadas da limpeza e conservação. *Revista do CEAM* **4**, 73–95 (2018).
28. Cortes, P. & Pan, J. Occupation and Gender. *The Oxford Handbook of Women and the Economy* 424–452 Preprint at <https://doi.org/10.1093/oxfordhb/9780190628963.013.12> (2017).
29. Bettio, F., Verashchagina, A. & Camilleri-Cassar, F. Gender segregation in the labour market: Root causes, implications and policy responses in the EU. (2009).
30. Hegewisch, A. & Hartmann, H. Occupational Segregation and the Gender Wage Gap: A Job Half Done: (529142014-001). *PsycEXTRA Dataset* Preprint at <https://doi.org/10.1037/e529142014-001> (2014).
31. Ortiz, S. Y. & Roscigno, V. J. Discrimination, Women, and Work: Processes and Variations by Race and Class. *Sociol Q* **50**, 336–359 (2009).
32. Browne, I. & Misra, J. The Intersection of Gender and Race in the Labor Market. *Annu Rev Sociol* **29**, 487–513 (2003).
33. Ridgeway, C. L. & Correll, S. J. Unpacking the Gender System: A Theoretical Perspective on Gender Beliefs and Social Relations. *Gender & Society* **18**, 510–531 (2004).
34. Cardoso, B.-H., Kaltenberg, M. & Hartmann, D. How skills-related local labor markets shape occupational upward mobility prospects. *SSRN Electronic Journal* (2023) doi:10.2139/ssrn.4631303.
35. Ganzeboom, H. B. G. A new International Socio-Economic Index (ISEI) of occupational status for the International Standard Classification of Occupation 2008 (ISCO-08) constructed with data from the ISSP 2002–2007. in *annual conference of international social survey programme, Lisbon* vol. 1 (2010).
36. Ganzeboom, H. B. G., De Graaf, P. M. & Treiman, D. J. A standard international socio-economic index of occupational status. *Soc Sci Res* **21**, 1–56 (1992).

37. Reskin, B. Sex Segregation in the Workplace. *Annu Rev Sociol* **19**, 241–270 (1993).
38. Elliott, J. R. & Smith, R. A. Race, Gender, and Workplace Power. *Am Sociol Rev* **69**, 365–386 (2004).
39. Wilson, G., Sakura-Lemessy, I. A. N. & West, J. P. Reaching the Top: Racial Differences in Mobility Paths to Upper-Tier Occupations. *Work Occup* **26**, 165–186 (1999).
40. Piasna, A. & Drahokoupil, J. Gender inequalities in the new world of work. *Transfer: European Review of Labour and Research* **23**, 313–332 (2017).
41. Rendall, M. Structural Change in Developing Countries: Has it Decreased Gender Inequality? *World Dev* **45**, 1–16 (2013).
42. Kabeer, N. Gender equality, inclusive growth, and labour markets. in *Women's Economic Empowerment* 13–48 (Routledge, 2021).
43. Bertocchi, G. & Dimico, A. Slavery, education, and inequality. *Eur Econ Rev* **70**, 197–209 (2014).
44. Kelly, M. D. A. Racial inequality in the Anglophone Caribbean: comparing the cases of Jamaica and Trinidad and Tobago. *J Ethn Migr Stud* **49**, 1125–1153 (2022).
45. Walz, T. & Cuno, K. M. *Race and Slavery in the Middle East: Histories of Trans-Saharan Africans in Nineteenth-Century Egypt, Sudan, and the Ottoman Mediterranean*. (Oxford University Press, 2010).
46. Weisshaar, K. & Cabello-Hutt, T. Labor Force Participation Over the Life Course: The Long-Term Effects of Employment Trajectories on Wages and the Gendered Payoff to Employment. *Demography* **57**, 33–60 (2020).
47. Kaufman, R. L. Assessing Alternative Perspectives on Race and Sex Employment Segregation. *Am Sociol Rev* **67**, 547–572 (2002).
48. Souza Silveira, L. & Siqueira Leão, N. O impacto da segregação ocupacional por gênero e raça na desigualdade de renda no Brasil em três décadas (1986-2015). *Revista Latinoamericana de Población* **14**, 41–76 (2020).
49. Tomaskovic-Devey, D. et al. Documenting Desegregation: Segregation in American Workplaces by Race, Ethnicity, and Sex, 1966–2003. *Am Sociol Rev* **71**, 565–588 (2006).
50. Weeden, K. A., Newhart, M. & Gelbgiser, D. Occupational Segregation. *Pathways Magazine* 30–33 (2019).

- 170 51. Guimarães, A. S. A. The Brazilian system of racial classification. *Ethn Racial Stud*
171 **35**, 1157–1162 (2012).
172 52. Francis-Tan, A. & Tannuri-Pianto, M. Inside the black box: affirmative action and
173 the social construction of race in Brazil. *Ethn Racial Stud* **38**, 2771–2790 (2015).
174