DOI http://dx.doi.org/10.20947/S0102-3098a0304

ORIGINAL ARTICLE

Healthy habits and the labor market: a quantile analysis for Brazil*

Rafael Mesquita Pereira**
Lívia Madeira Triaca***
Cristiano Aguiar de Oliveira****
Guilherme da Fonseca Vilela*****

This study investigates the effects of healthy behaviors on labor income in Brazil. While existing specialized literature generally points to a positive association between such behaviors and wages, this relationship remains underexplored in developing countries. Using microdata from the 2019 National Health Survey (PNS), conducted by IBGE, we constructed indicator variables for health-related behaviors and estimated Mincerian wage equations. The analysis uses mean and quantile regressions, with correction for sample selection bias due to labor market participation. Contrary to conventional findings, results indicate a negative association between physical activity and labor income across the income distribution. This result is explained by the definition of physical activity adopted in the study, broadened to include not only leisure activities, but also work related physical activities, commuting, and household chores — more frequent among lower-income individuals. Furthermore, engagement in preventive health behaviors is negatively associated with earnings, particularly among women, suggesting complex interactions between health investment, time constraints, and labor market outcomes.

Keywords: Physical activities. Healthy habits. Labor market. Quantile regression. Sample selection correction.

* This article is based on the study of Pereira *et al*. (2024).

** Universidade Federal do Rio Grande (FURG), Rio Grande-RS, Brazil (rafaelmesquita@furg.br; https://orcid.org/0000-0002-6080-0208).

*** Universidade Federal do Rio Grande (FURG), Rio Grande-RS, Brazil (liviamtriaca@gmail.com; https://orcid.org/0000-0001-7192-6554).

**** Universidade Federal do Rio Grande (FURG), Rio Grande-RS, Brazil (cristiano.oliveira@furg.br; https://orcid.org/0000-0003-4933-6616).

***** Universidade Federal da Fronteira Sul (UFFS), Passo Fundo-RS, Brazil (ggfvilela@hotmail.com; https://orcid.org/0000-0002-4096-2402).

Introduction

The benefits of physical exercise and healthy habits are widely recognized in modern society. Academic literature consistently provides evidence of the various health benefits associated with physical activity, particularly for the cardiovascular/metabolic and immune systems (Blair *et al.*, 1992; Lin *et al.*, 2016; Krinski *et al.*, 2010). More recently, new evidence suggests that not only regular physical activity, but also the reduction of sedentary behavior – defined as the time spent sitting, lying, or reclining during the day, excluding sleeping hours – is linked to better health outcomes. (Tremblay *et al.*, 2017).

In addition to these studies in health sciences, healthy behaviors have long been analyzed within economics as well. Since Grossman (1972), health stock has been considered a form of capital that enhances workers' productivity and, consequently, increases their labor earnings. Furthermore, Grossman emphasizes that investing in health is positively associated with reduced production time lost to illness.

In general, existing studies on this topic indicate that physical exercise positively affects workers' productivity and thus their wages (Kosteas, 2012; Lechner, 2015; Lechner; Sari, 2015; An; Liu, 2012; Kari, 2018). Specifically for Brazil, Godoy and Triches (2017) found that workers who engaged in physical activities achieved better social and economic outcomes than sedentary individuals. The positive estimated impact of physical activity on wages ranged from 15% to 31%.

Beyond physical activity, the literature shows that other lifestyle behaviors also affect individuals' socioeconomic conditions and should be considered when analyzing labor market outcomes. Evidence highlights the harmful effects of smoking, alcohol consumption, and inadequate dietary habits on the prevalence of non-communicable diseases (NCDs), identifying them as major risk factors for premature mortality.

Every year, more than 7 million people die due to direct tobacco use, while approximately 1.3 million deaths result from non-smokers' exposure to secondhand smoke. Regarding alcohol consumption, around 3 million deaths annually are attributable to harmful use, representing about 5.3% of total deaths (IHME, 2020). According to the WHO (2024), a healthy diet is crucial for preventing NCDs such as diabetes, cardiovascular diseases, strokes, and cancer. These unhealthy behaviors generate significant economic costs, which extend beyond the treatment of diseases to include losses in human capital due to morbidity and mortality.

Given this context, the goal of this study is to analyze the effects of physical activity, healthy dietary habits, and abstention from alcohol and smoking on labor earnings in Brazil. Using data from the 2019 National Health Survey (Pesquisa Nacional de Saúde – PNS), conducted by the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística – IBGE, 2021), indicator variables for these healthy behaviors are constructed and included in the estimation of the labor earnings equation, following Grossman's (1972) concept of health capital. These estimations are conducted using mean and quantile regressions, accounting for potential sample selection bias via the Heckman (1979) and Arellano and Bonhomme (2017) models, respectively.

In this context, Lechner and Sari (2015) emphasize that although the relationship between physical activity and the labor market is a relevant topic, the literature remains limited. To our knowledge, only Godoy and Triches (2017) addressed this issue in the Brazilian context, using data from 2008, which does not reflect the country's current social and economic arrangements. Moreover, this research advances the literature by incorporating not only leisure activities, but also work-related and commuting activities into the physical activity variables, as well as domestic chores and leisure commuting, thus providing novel evidence to this body of literature.

Therefore, to address the proposed research question, the paper is structured as follows: in addition to this introduction, the second section presents a literature review on the topic. The subsequent section details the data and methodologies used for the econometric estimations. Section 4 presents the results, followed by concluding remarks.

Literature review

Grossman (1972), in his seminal contribution, extended Mincer's (1957) classical wage equation by incorporating a health capital variable, under the assumption that individuals allocate their time not only between labor and leisure, but also among various consumption activities, including physical exercise. In this context, as emphasized by Godoy and Triches (2017), Grossman's model posits that: (i) age is negatively associated with health capital but positively related to healthcare investment; (ii) wages are positively correlated with both health and healthcare demands; (iii) education is positively associated with health capital and negatively associated with healthcare investment; (iv) individuals with higher educational level tend to demand a higher optimal health stock; and (v) health is conceptualized as a good produced by individuals themselves.

According to this framework, individuals invest in their health through healthy behaviors, education, and medical care, with the goal of maximizing utility over the life cycle. A greater stock of health capital reduces absenteeism (lost workdays due to illness) and mitigates the effects of presenteeism (lower productivity when working while ill), thus directly contributing to increased labor earnings (Currie; Madrian, 1999; Bhattacharya; Hyde; Tu, 2013).

Moreover, healthy behaviors support the long-term preservation of work capacity, maintaining productivity levels at older ages (Case; Deaton, 2005). The sustained preservation of physical and cognitive functions enables higher returns on investments in human capital, as healthier individuals tend to remain active in the labor market for longer periods and with greater efficiency (Contoyannis; Rice, 2001). Health, therefore, operates as a complementary input to education and work experience, reinforcing human capital accumulation throughout the life course (Grossman, 2000).

According to Lechner (2015), some individuals are aware that current investments in health may yield future labor market returns by enhancing technical and behavioral skills. However, awareness of these labor market benefits may be less widespread than recognition of direct health-related advantages.

In this regard, evidence from Kosteas (2012), An and Liu (2012), and Lechner (2015) indicates that regular physical activity positively influences individual well-being by improving cardiovascular health, weight management, cognitive performance, psychological conditions, and energy levels – all factors that indirectly increase job satisfaction. These benefits may translate into higher wages by boosting productivity. Additionally, healthy habits such as regular exercise and balanced nutrition may signal desirable traits to employers, such as self-discipline, dedication, and commitment, while also enhancing social skills and social capital, both of which positively influence productivity (Cutler; Lleras-Muney, 2010; Lechner; Sari, 2015).

Furthermore, Lechner (2015) notes that rational individuals allocate time to sports and physical exercise when the expected and discounted returns, combined with the consumption value, exceed the opportunity costs, including monetary expenses, injury risk, and any disutility associated with physical effort. Given the finite nature of time, individuals face trade-offs among competing activities, such as work, leisure, and rest.

Nevertheless, as Lechner (2015) also points out, increased time devoted to physical activity necessarily displaces time spent on other productive or restorative activities (such as sleep), which may also yield positive labor market outcomes. Hence, engagement in physical activity is subject to inherent trade-offs among work, leisure, and rest, and may even present negative correlations with sedentary occupational routines.

From an empirical perspective, Lechner and Sari (2015) analyze the effects of participation in sports and exercise on labor market outcomes among adult workers in Canada. Their findings indicate a positive earnings effect exceeding 10% for individuals engaged in such activities. Similarly, Lechner (2015) observes that, in Germany, individuals who engage in sports earn at least 5% more than non-participants. In the United States, Kosteas (2012) identifies significant earnings premiums associated with regular physical activity, estimating increases of 10.5% for men and 12.9% for women.

In the Brazilian context, although empirical evidence is still relatively scarce, Godoy and Triches (2017), using data from the 2008 National Household Sample Survey (Pesquisa Nacional por Amostragem de Domicílios – PNAD), find that returns to physical activity increase steadily across the labor income distribution. Among men, estimated returns range from 15% in the first quartile (Q25) to 31.56% in the upper quantile (Q90). Among women, these returns vary from 15.16% (Q25) to 19.34% (Q90).

Moreover, the labor economics literature provides robust evidence regarding the relationship between health behaviors and occupational outcomes. Risky behaviors, such as smoking and excessive alcohol consumption, are consistently associated with lower labor force participation and reduced wages (French; Zarkin, 1995; Macdonald; Shields, 2001; Van Ours, 2004; Terza, 2002; Johansson *et al*., 2007; Böckerman *et al*., 2017; Cawley; Ruhm, 2011). Likewise, obesity – often linked to poor diet and physical inactivity – is correlated with lower productivity, reduced wages, and limited human capital accumulation, thus impairing employability (Morris, 2007; Rooth, 2009; Hammond; Levine, 2010; Cawley; Ruhm, 2011).

When evaluating the impact of healthy behaviors on productivity, it is crucial to acknowledge that such behaviors reflect consumption choices that directly influence occupational outcomes. In this regard, the observed negative associations are consistent with evidence that risky health behaviors impair physical and cognitive capacities, ultimately undermining individuals' ability to attain favorable positions in the labor market (Böckerman *et al.*, 2018).

Data and methodology

The present research uses data from the 2019 National Health Survey (PNS). According to the IBGE (2021), the main objective of this survey is to provide national data on the health status and lifestyle habits of the Brazilian population, as well as to analyze the use of and access to healthcare services and preventive actions within this context, thereby providing inputs for the implementation of public policies in this area.

To conduct the econometric analysis of the proposed research question, it is first necessary to define the variables that will represent physical activity (PA) and healthy behaviors. These characteristics constitute health capital, according to Grossman's (1972) framework.

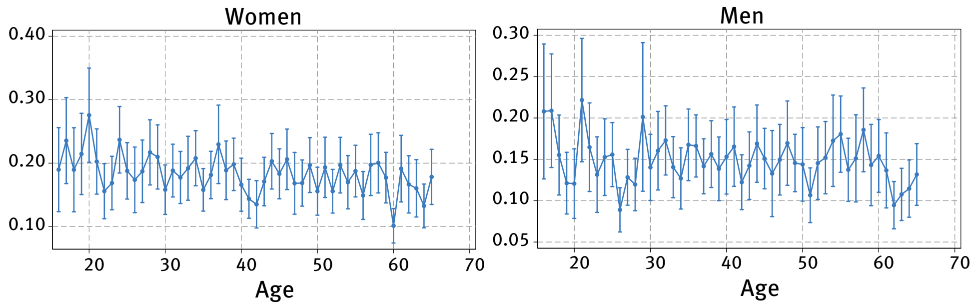
The variables concerning physical activity are constructed based on the frequency of days and time spent on PA in the following domains, as highlighted by Lechner (2015), Benedetti *et al.* (2021), and Cruz *et al.* (2022): leisure (exercise or sports type and intensity), work (lifting weights, doing strenuous tasks, or walking extensively), commuting (walking or cycling to work or for daily activities), and domestic tasks (cleaning, lifting weights, or doing strenuous chores). Additionally, the intensity of physical activity is differentiated: moderate physical activity is defined as that which requires more physical effort, causing the person to breathe faster than normal and moderately increase their heart rate. In contrast, vigorous physical activity requires substantial physical exertion, making the individual breathe much faster than normal and significantly increasing their heart rate. On a scale of 0 to 10, the perceived effort is 7 or 8, during which it is not possible to talk while moving. Therefore, in this study, we consider the following as vigorous physical activity: running/jogging, running on a treadmill, weight training, aerobic gymnastics/spinning/step/jump, water aerobics, swimming, martial arts and combat sports, cycling/exercise bike, soccer, basketball, volleyball, tennis, and dancing (when performed for the purpose of physical activity) (Benedetti *et al.*, 2021; Cruz *et al.*, 2022).

Since the 2019 PNS contains information on the time spent in minutes on these activities per day, and asks respondents whether the activities required intense physical effort, it is feasible to construct PA variables across all these domains. Thus, these controls in the four domains (leisure, work, domestic, and commuting) are defined by calculating the weekly frequency (days) multiplied by the average duration (minutes) of moderate and vigorous physical activities. Time spent on vigorous activities is doubled, and only activities performed for at least ten consecutive minutes per week are considered valid (WHO, 2010).

Therefore, individuals who engage in 150 or more minutes per week of moderate physical or vigorous activity (which counts double the number of minutes)[1] across the assessed domains are considered sufficiently active. Those who engage in between 10 and 149 minutes per week at the same intensity and in the same domains are considered insufficiently active, and individuals who do not achieve at least 10 minutes of physical activity per week under these conditions are classified as inactive (WHO, 2010; Hallal *et al.*, 2003; Lechner, 2015; An; Liu, 2012).

According to data from the 2019 PNS, it is observed that, on average, 17.10% of the Brazilian population engages in insufficient PA per week. Figure 1 shows the levels of insufficient physical activity by age, practiced by women and men in Brazil. In general, it is observed that, with increasing age, fewer women and men engage in this level of physical activity. Furthermore, it is important to highlight that women engage in moderate or vigorous physical activities within this classification in greater proportions (19.25% on average) than men (15.37% on average).

FIGURE 1
Estimated means of insuficient PA by age for women and men with 95% CI
Brazil – 2019



Source: Own elaboration based on data from the 2019 PNS (IBGE, 2021).
Note: The average estimates obtained are weighted by sampling weights. CI = Confidence Intervals.

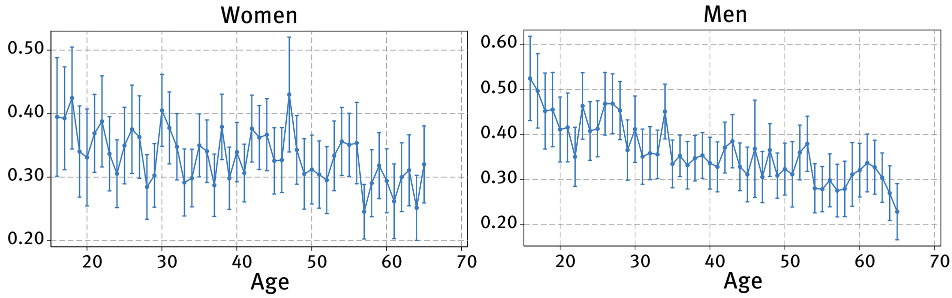
In terms of sufficient PA, statistics show that 38.76% of the Brazilian population practices more than 150 minutes per week of moderate or vigorous PA across all domains considered. Furthermore, Figure 2 presents the levels of sufficient physical activity by age for Brazilian men and women. The average practice of sufficient PA among women declines at lower rates than among men as age increases. Furthermore, 37.98% of men practice sufficient PA, while 39.72% of women meet this level of PA, following the same trend observed for insufficient PA.

Despite a slight downward trend in the practice of PA with increasing age, it is worth highlighting the significant proportion of people engaged in these activities, even at older

[1] Each minute of vigorous physical activity performed is worth two minutes, that is, if the individual ran for 10 minutes, this would be considered 20 minutes of vigorous physical activity. Therefore, if the amount of these weekly activities and the moderate activities adds up to more than 150 minutes, the individual is considered sufficiently active. If it adds up to between 10 and 149 minutes, the individual will be considered insufficiently active.

ages. In this regard, Cruz *et al.* (2022) highlight the relevant proportion of the Brazilian adult population that engages in PA at the workplace.

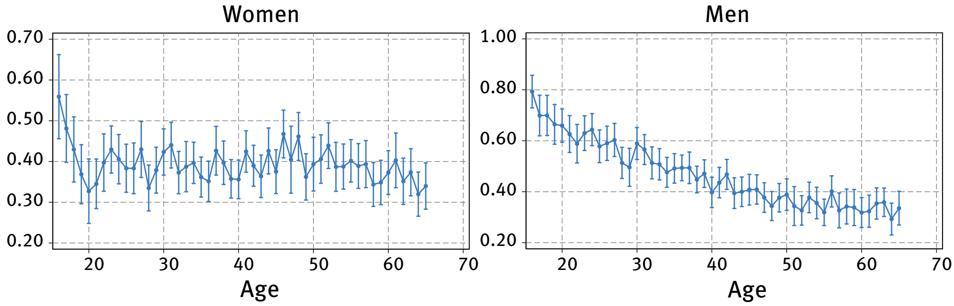
FIGURE 2
Estimated means of suficient PA by age for women and men with 95% CI
Brazil – 2019



Source: Own elaboration based on data from the 2019 PNS (IBGE, 2021).
Note: The average estimates obtained are weighted by sampling weights. CI = Confidence Intervals.

To separately analyze the effect of PA on leisure, we constructed a dummy variable that assumes the value 1 if the individual practices some activity[2] other than work, whether at an insufficient or sufficient level, and 0 otherwise. On average, 45% of Brazilians practice leisure-time PA. Among men, this proportion is 47.62% and, among women, 41.74%. It is worth noting that, unlike the pattern observed for insufficient and sufficient PA, a greater proportion of men practice PA during leisure time compared to women, although the rate of decline with age is steeper among men.

FIGURE 3
Estimated means of leisure PA by age for women and men with 95% CI
Brazil – 2019



Source: Own elaboration based on data from the 2019 PNS (IBGE, 2021).
Note: The average estimates obtained are weighted by sampling weights. CI = Confidence Intervals.

[2] These activities include walking (not valid for work), walking on a treadmill, running/jogging, running on a treadmill, weight training, aerobic gymnastics/spinning/step/jump, water aerobics, gymnastics/localized exercise/pilates/stretching/yoga, swimming, martial arts and fighting, cycling/exercise bike, soccer, basketball, volleyball, tennis, dancing (with the aim of practicing physical activity).

To control for other healthy behaviors that may affect labor market outcomes, two lifestyle indices are included in the analysis, considering other healthy behaviors that the literature highlights as important for individuals' social and economic outcomes: the healthy habits index and the healthy eating habits index. The main objective of this strategy is to isolate ("clean up") the effects that some behaviors have on others.

The healthy habits index includes tobacco and alcohol consumption. Smoking is identified as the leading cause of preventable death worldwide, and associated with more than 50 types of diseases (WHO, 2013). The 2019 PNS examined tobacco use (both smoked and non-smoked) among individuals aged 18 or older. In this dataset, a consumption variable classifies all individuals who consume tobacco daily or occasionally, whether smoked or not, as smokers. This aligns with the WHO's stance that all forms of tobacco use pose health risks, with no safe level of exposure.

Alcohol consumption is another behavior with direct effects on increasing morbidity and mortality rates. Over 200 diseases are attributed to alcohol consumption, in addition to their association with accidents and violence (WHO, 2018). The effects of alcohol consumption vary depending on consumption patterns and quantities; therefore, its impact is analyzed using the variable of alcohol abuse, defined as the intake of five or more doses among men, and four or more doses among women, on a single occasion in the last 30 days.

A diet high in fats and insufficient in fruits and vegetables represents an unhealthy eating pattern that contributes to the development of many non-communicable diseases (NCDs) (WHO, 2013). The 2019 PNS data allows the analysis of eating habits through a series of indicators on healthy and unhealthy consumption patterns. Thus, to calculate the healthy eating habits index, five dietary variables are included: fruits and vegetables, beans, soda, sweets, and salt.

Adequate intake of fruits and vegetables is assessed based on the weekly frequency of consuming vegetables, legumes, and fruits or fruit juices during meals. For the purpose of calculation, regular consumption is met when an individual reports consuming both types of foods at least five times a week. Regular consumption of beans is also considered a healthy eating pattern and is evaluated according to IBGE (2014) guidelines, which define regular consumption as five or more days per week. Additionally, unhealthy eating patterns such as the consumption of soda, salt, and sweets (cakes, pies, chocolates, candies, cookies, or sweet biscuits) are also considered. Individuals who consume soda or artificial juices at least five days a week are classified as regular consumers. The same standard applies to the regular consumption of sweets, for individuals who consume sweets five or more days a week.

Another unhealthy pattern analyzed is the high intake of salt, which is associated with an increased risk of NCDs such as hypertension and kidney diseases. To assess participants' perception of salt consumption, the 2019 PNS asked whether they considered their salt intake – both in home-cooked meals and processed foods, to be very high, high, adequate, low, or very low. High and very high consumption levels are classified as unhealthy (IBGE, 2014).

The two indices mentioned above were generated using the Principal Component Analysis (PCA) method. For the calculation of the indices, all lifestyle variables were dichotomized so that a unit represents healthy behavior (Table 1). After calculating the indices, continuous variables were obtained, with higher values indicating healthier lifestyles.

In this study, these indicators are divided into two categories: eating habits and non-consumption habits of tobacco and alcoholic beverages. Table 1 describes the variables that comprise each of these indices.

TABLE 1
Description of the variables used to compute the healthy eating habits and healthy habits indexes

Index	Variable	Description
Healthy Eating Habits Index	Regular consumption of fruits and vegetables	Binary variable that assumes the value 1 if the individual consumed both foods at least five times a week
	Regular consumption of beans	Binary variable that assumes the value 1 if the individual consumed beans on five or more days of the week.
	Adequate consumption of soft drinks	Binary variable that assumes the value 1 if the individual did not consume soft drinks or artificial juices on five or more days of the week
	Adequate consumption of sweets	Binary variable that assumes the value 1 if the individual did not consume sweets on five or more days of the week
	Adequate salt intake	Binary variable that assumes the value 1 if the individual reported that their salt consumption is adequate, low or very low
Healthy Habits Index	Non-smoker	Binary variable that assumes the value 1 when the individual does not consume tobacco
	Non-abusive alcohol consumption	Binary variable that assumes the value 1 when the individual has not consumed 5 or more doses of alcoholic beverage on a single occasion in the last 30 days if being a man, and 4 or more doses of alcoholic beverage on a single occasion in the last 30 days if being a woman

Source: Own elaboration based on data from the 2019 PNS (IBGE, 2021).

Bartlett's test for sphericity and the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy were conducted (Kaiser, 1958) to assess the validity and adequacy of these healthy habits indicators. Bartlett's test for sphericity calculates the determinant of the sum of squares and cross-products matrix (S) from which the correlation matrix is derived. The determinant of matrix S is converted into a chi-square statistic and tested for significance. In this test, the healthy eating habits index yielded a correlation matrix determinant value of 0.904, which was statistically significant at the 1% level. A similar result was obtained for the healthy habits indicator, which had a correlation matrix determinant value of 0.970, also statistically significant at the 1% level. These findings indicate that the variables comprising the indicators are not intercorrelated, thus validating their use in constructing the indexes.

The KMO test assesses whether the variables used in the composition of the indicators are sufficiently correlated, such that if the obtained value falls between 0.5 and 1, the indicator is deemed appropriate; otherwise (below 0.5), the PCA is not suitable (Soares; Dalboni; Teixeira, 2022). Accordingly, the KMO index for healthy eating habits was 0.5382, and 0.500 for healthy habits, indicating that both were considered adequate. Thus, both tests validate the construction of the two indicators presented in Table 1.

To account for the impact of chronic diseases on labor market participation and worker productivity, a binary indicator variable is included, taking the value 1 if the individual reports having a chronic condition or illness, and 0 otherwise. As noted by Polanco *et al.* (2024), individuals with chronic illnesses tend to make greater use of health services and are often subject to medical limitations that reduce their available working time or hinder their ability to interact in shared work environments. Consequently, these individuals are also more likely to face significant barriers to engaging in physical activity.

Initially, to estimate the wage equation at the mean level, based on the framework proposed by Grossman (1972), the Heckman (1979) two-step model will be employed. This strategy is adopted because the 2019 PNS includes individuals both inside and outside the labor market, suggesting the possibility of non-random self-selection into employment, which introduces sample selection bias. In general, according to Heckman (1979), this bias may arise from two sources: self-selection among individuals or data units under investigation; or sample selection decisions made by analysts or data processors, which operate similarly to self-selection mechanisms.

Coelho, Veszteg, and Soares (2010) emphasize that individuals participate in the labor market whenever the offered wage exceeds their reservation wage. Consequently, the "zeros" in the outcome variable generate a sample selection problem, which can be addressed using Heckman's (1979) methodology.

Econometrically, the classic earnings equation serves as the starting point for the Heckman (1979) model:

$$W_i = x'_i\beta + \varepsilon_i \tag{1}$$

Where W_i represents labor earnings (such as wage or income), x'_i denotes observed variables related to the productivity of individual i-th, and ε_i is the error term. Importantly, W is observed only for those who participate in the labor force. According to Cameron and Trivedi (2005), sample participation is modeled as:

$$W_{1i} = \begin{cases} 1 \; if \;\; W^*_{1i} > 0 \\ 0 \; if \;\; W^*_{1i} \leq 0 \end{cases} \tag{2}$$

leading to the following truncated outcome equation:

$$W_{2i} = \begin{cases} W^*_{2i} \;\; if \;\; W^*_{1i} > 0 \\ \;\; . \; if \;\; W^*_{1i} \leq 0 \end{cases} \tag{3}$$

Cameron and Trivedi (2005) argue that W_2 is observed only when $W^*_1 > 0$, implying that W_2 has no meaningful value when $W^*_1 \leq 0$. So, the observed density is then $f^*\,(W^*_2 \mid W^*_1 > 0) \times \Pr[W^*_1 > 0]$ when $W_2 > 0$. Thus, the sample selection model leads to the following likelihood function:

$$L = \Pi^{n}_{i=1}\,\{\Pr\,[W^*_{1i} \leq 0]\}^{1-W_{1i}}\{\mathrm{f}\,(W^*_2\,|W^*_{1i} > 0)\; x\; \Pr[W^*_{1i} > 0]\}^{W_{1i}} \tag{4}$$

where the first term represents the discrete contribution when $W^*_{1i} \leq 0$, and the second term captures the continuous contribution when $W^*_{1i} > 0$.

In practice, exclusion variables are applied in the first stage; these variables influence labor supply decisions but do not directly affect the outcome variable. Consequently, these controls are not included in the second-stage estimation. In this study, receiving retirement income, pension benefits, and other sources of income are used as factors negatively influencing labor supply decisions, while being head of household is considered a positive factor (Vieira; Pereira, 2024).

Once the first stage (selection equation) is completed, the second stage is estimated via Ordinary Least Squares (OLS). According to Cameron and Trivedi (2005), Heckman's two-step model augments the OLS regression by including the estimated omitted regressor, the inverse Mills ratio $\lambda_i(Z_i' y)$, using only the positive observations of W_2:

$$W_{2i} = x'_{2i}\beta_2 + \lambda(x'_{1i}\hat{\beta}_1)\sigma_{12} + \varepsilon_i \qquad (5)$$

Where ε is the error term, $\hat{\beta}_1$ represents the coefficients from the first-stage Probit model of W_1 on x'_1, given that $\Pr[W^*_{1i} > 0] = \varphi(x'_{1i}\beta_1)$ e $\lambda(x'_{1i}\hat{\beta}_1) = \varphi(x'_{1i}\hat{\beta}_1)/\Phi(x'_1\hat{\beta}_1)$ is the estimated inverse Mills ratio, commonly referred to as the "Heckman Lambda" (Tauchmann, 2010; Vermeulen, 2001).

An alternative is to estimate the system by full maximum likelihood, taking into account the nonlinearity of the selection equation. This approach yields asymptotically efficient and consistent estimators, whereas the two-step procedure, while simpler, provides consistent estimators only (Lee, 1979; Maddala, 1983).

Practically, the effects of Insufficient (IPA) and Sufficient Physical Activity (SPA), the Healthy Eating Habits Index (HEHI), and the Healthy Habits Index (HHI) on Brazilian workers' earnings will be estimated through the following function:

$$\ln w_i = \alpha_i + \beta_1 X_i + \beta_2 IPA_i + \beta_3 SPA_i + \beta_4 HEHI_i + \beta_5 HHI_i + u_i \qquad (6)$$

Where w is the log of hourly earnings, X is a vector of controls related to individuals' personal productivity and employment characteristics (education level, gender, race, age, chronic disease, formal employment, public sector employment, urban/rural residence, blue-collar occupations, federative units (states), and CNAE industry sectors). IPA and SPA are dummy variables indicating insufficient and sufficient levels of physical activity, respectively. $HEHI$ and HHI are indicators for healthy eating habits and overall healthy habits. u is the model's error term. The subscript i denotes individual-level estimation.

To estimate the earnings function across the distribution quantiles, it is necessary to recognize that sample selection bias behaves heterogeneously throughout the distribution, as higher wages may require distinct credentials to participate in corresponding labor markets. To address this issue, Arellano and Bonhomme (2017) propose a quantile regression model with sample selection correction, where percentile levels are adequately shifted as a function of selection intensity.

In practice, this adjustment involves rotating the "check" function, which is optimized in standard quantile regression. Consequently, the objective function becomes "discordantly tilted", with percentile perturbations observation-specific and dependent on selection strength. This rotation preserves the linear programming structure and computational simplicity of quantile regression methods.

In this quantile model, sample selection is modeled via the bivariate cumulative distribution function, or copula, of the outcome and selection errors. As in the Heckman model, exclusion variables (e.g., employment determinants unrelated to wages) are essential for reliable identification (Arellano; Bonhomme, 2017). Thus, the same exclusion variables used in the Heckman model are employed here.

According to Biewen and Erhardt (2021), the Arellano and Bonhomme (2017) quantile selection model can be represented as follows:

$$Y^* = \boldsymbol{X}' \beta (U) \quad (7)$$

$$D = 1\{V \leq p(\boldsymbol{Z})\} \quad (8)$$

$$Y = Y^* \text{ se } D = 1 \quad (9)$$

Where Y^* is the potential outcome variable, D is the selection indicator, and Y is the observed outcome, available only if D = 1. The vectors $\boldsymbol{X}$ and $\boldsymbol{Z}$ contain covariates, with $\boldsymbol{X}$ assumed to be a subset of $\boldsymbol{Z}$, which contain the exclusion variables.

The estimation method proceeds as follows: in the first stage, a Probit model estimates the propensity scores representing the selection equation. The second stage uses the generalized method of moments (GMM) to estimate the copula parameter ρ, which operates similarly to the Heckman Lambda. The third stage estimates the "rotated" quantile regressions with selection correction based on the estimated $\hat{\rho}$, for any desired quantile τ (Biewen; Erhardt, 2021).

Thus, Equation (4) is rearranged to the following quantile specification:

$$\ln w_i = \alpha_i(\tau) + \beta_1(\tau)\, X_i + \beta_2(\tau)\, IPA_i + \beta_3(\tau)\, SPA_i + \beta_4(\tau)\, HEHI_i + \beta_5(\tau)\, HHI_i + u_{i\tau} \quad (10)$$

Where τ represents the quantiles analyzed, such as 0.10, 0.25, 0.50 (median), 0.75, or 0.90.

Table 2 presents the definitions of the variables used in the estimations of the econometric models, along with their descriptive statistics. The variable hourly earnings is obtained by dividing the income reported by the individual from their primary job by the hours worked per week, multiplied by 4.2 (based on a 30-day month divided by seven days a week) (Pereira; Oliveira, 2016, 2017). On average, hourly earnings in Brazil amount to R$14.69 for the full sample, a value slightly higher among men, reaching around R$15.69. Conversely, women in the country earn, on average, 16.74% less than men.

TABLE 2
Descriptive statistics of the variables used in the econometric models
Brazil – 2019

Variables	Definition	Full sample	Women	Men
Hourly income	Earnings per hour worked (income from work)	14.6871	13.4425	15.6938
		(0.2962)	(0.2679)	(0.4505)
Educational level	Highest educational level achieved	4.4859	4.7799	4.2481
		(0.0192)	(0.0251)	(0.0243)
Age	Age in years	39.3554	39.4059	39.3145
		(0.1138)	(0.1696)	(0.1514)
Sex	1 = Male; 0 = Female	0.5528	-	-
		(0.0044)		
Race (ethnicity)	1 = White; 0 = Other	0.4447	0.4634	0.4297
		(0.0049)	(0.0069)	(0.0060)
Urban area	Residence area: 1 = Urban; 0 = Rural	0.8825	0.9200	0.8522
		(0.0024)	(0.0027)	(0.0036)
Statutory	Municipal, state or federal public servant	0.0839	0.1161	0.0578
		(0.0023)	(0.0040)	(0.0025)
Formal employment	Employment with work regime governed by the CLT (1)	0.3842	0.3655	0.3993
		(0.0043)	(0.0062)	(0.0060)
Blue-collars	Blue-collar occupation: 1 = yes; 0 = no	0.8702	0.8391	0.8953
		(0.0033)	(0.0048)	(0.0039)
Chronic diseases	Chronic diseases: 1 = yes; 0 = no	0.4362	0.5024	0.3828
		(0.0045)	(0.0066)	(0.0057)
Exclusion variables (First stage of sample selectivity correction)				
Works	1 = Yes; 0 = No	0.6521	0.5540	0.7598
		(0.4763)	(0.4971)	(0.4272)
Collects retirement	1 = Yes; 0 = No	0.0513	0.0641	0.0410
		(0.0018)	(0.0032)	(0.0019)
Collects pension benefits	1 = Yes; 0 = No	0.0212	0.0416	0.0047
		(0.0011)	(0.0022)	(0.0006)
Other sources of income	1 = Yes; 0 = No	0.0859	0.1311	0.0494
		(0.0023)	(0.0042)	(0.0022)
Responsible for the household	1 = Yes; 0 = No	0.6011	0.5657	0.6298
		(0.0049)	(0.0067)	(0.0066)
Variables related to healthy habits				
Insufficient physical activity	1 = insufficient level of physical activity 0 = inactive	0.1710	0.1925	0.1537
		(0.0034)	(0.0048)	(0.0044)
Sufficient physical activity	1 = sufficient level of physical activity 0 = inactive	0.3876	0.3972	0.3798
		(0.0041)	(0.0060)	(0.0054)
Leisure physical activity	1 = leisure level of physical activity 0 = inactive	0.4499	0.4174	0.4762
		(0.0047)	(0.0064)	(0.0060)
Healthy eating habits index	Indicator: higher levels indicate a greater proportion of good habits, vice versa	-0.1817	-0.1016	-0.2465
		(0.0115)	(0.0166)	(0.0160)
Healthy habits index	Indicator: higher levels indicate a greater proportion of good habits, vice versa	-0.1137	-0.0684	-0.1503
		(0.0091)	(0.0126)	(0.0123)

Source: Prepared by the authors based on data from the 2019 PNS (IBGE, 2021).
(1) Consolidation of Labor Laws (Consolidação das Leis do Trabalho – CLT). This Law can be seen at http://www.planalto.gov.br/ccivil_03/decreto-lei/del5452.htm
Note: Standard deviations are in parentheses. The average estimates obtained are weighted by sampling weights.

Approximately 55% of the individuals in the sample are male, and 44.5% identify as white. Regarding the labor market, on average, 65.21% of participants have a primary job income greater than zero, meaning they are employed.

Controls related to different employment contract schemes (statutory and formal employment) are included in the models because, as Almeida and Júnior (2017) point out, the productivity of public sector workers does not necessarily affect their wages, primarily due to the job security they enjoy. Therefore, it is necessary to distinguish them from private sector workers in the regressions, as the latter are subject to the laws of labor supply and demand. In general, the data show that around 38.42% work under the CLT regime, and 8.39% are public servants at the federal, state, or municipal level.

With regard to blue-collar[3] occupations, 87.02% of Brazilian workers, both male and female, are employed in jobs with these characteristics. It is important to note that blue-collar jobs involve manual labor which, in general, presents a greater risk of occupational accidents, as opposed to white-collar jobs, represented by administrative or managerial positions (Pereira; Almeida; Oliveira, 2020; Pereira; Oliveira; Machado, 2023). On average, 43.62% of Brazilians have chronic diseases, with this occurrence being higher among women (50.24%) compared to men (38.28%).

Regarding the healthy eating habits index, on average, both women and men tend to have poorer eating habits, based on the elements considered in this indicator. The same occurs when analyzing healthy habits, with women and men, on average, more likely to smoke and consume alcohol excessively. However, it is worth noting that in both cases these magnitudes are lower among women.

The prevalence (mean) of healthy habits variables by gender and across quantiles is presented in Table 3. For both men and women, the highest prevalence of insufficient and sufficient physical activity occurs in the lower quantiles, particularly in the 10th percentile, which has the highest proportion of workers compared to the other quantiles analyzed for both genders. In other words, these statistics indicate the highest density of these practices in Brazil is concentrated in the lower income levels of the labor income distribution.

In contrast, for leisure-time PA, the proportion of practitioners increases with income level. This behavior is observed for both women and men, indicating a likely income effect related to the practice of leisure-time PA.

The healthy eating habits index reveals a similar pattern for women and men, with unhealthy consumption predominating, particularly in the intermediate quantiles, though to different extents. The same occurs when analyzing the healthy habits index, which includes alcohol and tobacco consumption, as these prevail for both sexes across nearly all quantiles of the distribution.

[3] A working-class person who performs non-agricultural manual labor, which may involve manufacturing, mining, sanitation, custodial work, oil field work, construction, mechanical, maintenance, warehousing, firefighting, technical installation, and many other types of physical labor. This nomenclature is used in English-speaking countries.

TABLE 3
Prevalence (mean) of insufficient, sufficien, and leisure physical activity, healthy eating habits index, healthy habits index, hourly wage, and blue-collar workers by income quantiles for women and men
Brazil – 2019

Variables	Women					Men				
	Quantiles					Quantiles				
	10	25	50	75	90	10	25	50	75	90
Insufficient physical activity	0.2196	0.2395	0.1978	0.1886	0.1450	0.1722	0.1681	0.1684	0.1589	0.1291
	(0.0137)	(0.0131)	(0.0095)	(0.0097)	(0.0110)	(0.0127)	(0.0106)	(0.0087)	(0.0088)	(0.0080)
Sufficient physical activity	0.4350	0.4223	0.4394	0.3718	0.3347	0.4806	0.4161	0.3984	0.3794	0.3257
	(0.0203)	(0.0155)	(0.0111)	(0.0134)	(0.0151)	(0.0217)	(0.0149)	(0.0121)	(0.0109)	(0.0125)
Leisure physical activity	0.3057	0.3737	0.3516	0.4230	0.5430	0.3959	0.4184	0.4177	0.4488	0.5274
	(0.0175)	(0.0168)	(0.0114)	(0.0130)	(0.0166)	(0.0231)	(0.0154)	(0.0115)	(0.0105)	(0.0124)
Healthy eating habits index	0.0174	-0.1267	-0.1549	-0.1819	0.0119	-0.0607	-0.2037	-0.2673	-0.3085	-0.2292
	(0.0437)	(0.0398)	(0.0333)	(0.0337)	(0.0420)	(0.0455)	(0.0407)	(0.0342)	(0.0310)	(0.0333)
Healthy habits index	-0.0082	-0.0164	0.0052	-0.0651	-0.2141	-0.1558	-0.1519	-0.1428	-0.1598	-0.1393
	(0.0402)	(0.0304)	(0.0264)	(0.0257)	(0.0323)	(0.0452)	(0.0295)	(0.0257)	(0.0262)	(0.0247)
Hourly income	4.4898	7.5454	7.4039	10.9091	20.6419	2.7427	6.1057	6.9645	10.0663	17.4329
	(0.1760)	(0.2558)	(0.1266)	(0.1620)	(0.3640)	(0.1191)	(0.4857)	(0.1249)	(0.1387)	(0.2149)
Blue-collars	0.9866	0.9566	0.8657	0.7750	0.7407	0.9603	0.9542	0.9240	0.9108	0.8727
	(0.0051)	(0.0056)	(0.0084)	(0.0111)	(0.0140)	(0.0215)	(0.0082)	(0.0066)	(0.0066)	(0.0080)

Source: Prepared by the authors based on data from the 2019 PNS (IBGE, 2021).
Note: Standard deviations are in parentheses. The average estimates obtained are weighted by sampling weights.

Finally, regarding to blue-collar occupations, the largest proportion of workers in these types of jobs is found in the lower quantiles of the distribution, a trend consistent in both female and male labor markets. Among women, for example, there is an approximate 25 percentage point difference in the density of the lowest quantile compared to the highest quantile. This suggests that, as income increases, women migrate more than men to jobs with lower exposure to occupational accident risks.

It is important to highlight that some sample adjustments were made to mitigate potential biases in the estimations of the econometric models. First, only individuals aged between 16 and 65 years old were kept, as well as workers who work at least 30 hours per week in their main job (full-time work). Additionally, workers in the top income bracket (corresponding to the richest 1% in the country) were excluded from the sample to prevent overestimation of coefficients in the labor income functions estimations.

Results

The results of the econometric model estimations for the mean (Heckman model) and for the quantiles (Arellano and Bonhomme model) are presented in Table 4. First, it is important to emphasize that, given the complex sampling design of the 2019 PNS, sample weights were incorporated into all estimations. Additionally, the results of the sample selection corrections from both the Heckman model and the Arellano and Bonhomme method (first stage) are reported in Table 1 of the Appendix. The exclusion restrictions associated with collecting pension or retirement benefits, or other non-labor income sources exhibited a negative sign, indicating that such income reduces the likelihood of labor market participation. Conversely, the variable representing family responsibility showed a positive coefficient, suggesting that this condition increases the probability of entering the labor force.

These findings are consistent with those reported by Pereira and Oliveira (2016) and Vieira and Pereira (2024) regarding the expected signs of the exclusion variables in both the Heckman and Arellano and Bonhomme models. The validity of the sample selection correction is supported by the statistical significance of the selection parameters, namely the lambda (Heckman) and rho (copula parameter), confirming the importance of correcting for selection bias in both the mean and quantile regressions as shown in Table 4.

Regarding the results of the earnings equation estimations for both the mean and the quantiles, a positive relationship is identified between educational attainment and labor earnings, indicating that higher levels of education are associated with greater income returns (Vieira; Pereira, 2024). The age and age squared variables, which serve as proxies for labor market experience, display the expected signs – positive for age and negative for age squared – consistent with the literature, and reflect increasing productivity at diminishing rates as individuals age (Pereira; Oliveira, 2016, 2017; Vieira; Pereira, 2024).

The gender variable exhibits a positive sign across the mean and all quantiles, as does the race variable, suggesting an ethnic and gender premium in the Brazilian labor market favoring white and male workers. These results align with those found by Silveira and Leão (2021). Additionally, residing in urban areas and holding a public sector job under a statutory regime are both positively correlated with labor income, as expected.

However, the behavior of the variable associated with formal employment – i.e., jobs governed by the CLT – warrants special attention. While it exhibits positive returns up to the median, this relationship reverses at the upper quantile (Q90), indicating a negative association between formal employment and the highest labor income levels.

The control variable for blue-collar occupations, as anticipated, displays a negative coefficient for the mean and across all quantiles. Notably, this negative effect intensifies at the top of the income distribution (Q90), reinforcing the inverse relationship between blue-collar employment and high earnings. In these upper-income brackets, the share of blue-collar workers is significantly lower, as shown in Table 3.

TABLE 4
Results of the estimation of the labor income equation at the mean and quantiles for the full sample
Brazil – 2019

Variables	Mean (Heckman)	Quantile (Arellano and Bonhomme)				
		Q10	Q25	Median	Q75	Q90
Insufficient PA	-0.0712***	-0.0419**	-0.663***	-0.0716***	-0.0850***	-0.0997***
	(0.0147)	(0.0182)	(0.0166)	(0.0147)	(0.0152)	(0.0206)
Sufficient PA	-0.0728***	-0.0320*	-0.0599***	-0.0753***	-0.0741***	-0.0860***
	(0.0114)	(0.0172)	(0.0125)	(0.0109)	(0.0130)	(0.0177)
Healthy eating habits index	-0.00517	-0.00442	-0.00354	-0.00155	-0.000627	-0.00477
	(0.00405)	(0.00558)	(0.00485)	(0.00408)	(0.00471)	(0.00590)
Healthy habits index	-0.00215	0.00621	0.00438	-0.00167	0.00330	-0.00967
	(0.00496)	(0.00683)	(0.00584)	(0.00505)	(0.00589)	(0.00768)
Educational level 2	0.0915***	0.203***	0.0964**	0.0594*	0.109***	0.0675
	(0.0316)	(0.0604)	(0.0474)	(0.0340)	(0.0388)	(0.0634)
Educational level 3	0.226***	0.281***	0.183***	0.174***	0.247***	0.181***
	(0.0342)	(0.0649)	(0.0500)	(0.0385)	(0.0436)	(0.0669)
Educational level 4	0.257***	0.341***	0.226***	0.185***	0.260***	0.217***
	(0.0346)	(0.0648)	(0.0491)	(0.0375)	(0.0429)	(0.0680)
Educational level 5	0.356***	0.391***	0.287***	0.286***	0.356***	0.342***
	(0.0328)	(0.0602)	(0.0482)	(0.0360)	(0.0404)	(0.0649)
Educational level 6	0.554***	0.478***	0.438***	0.459***	0.582***	0.593***
	(0.0390)	(0.0650)	(0.0533)	(0.0415)	(0.0465)	(0.0804)
Educational level 7	0.865***	0.624***	0.625***	0.742***	0.883***	1.002***
	(0.0383)	(0.0641)	(0.0535)	(0.0384)	(0.0452)	(0.101)
Age	0.0311***	0.0288***	0.0244***	0.0253***	0.0326***	0.0372***
	(0.00394)	(0.00485)	(0.00458)	(0.00415)	(0.00399)	(0.00501)
Age squared	-0.000278***	-0.000299***	-0.000233***	-0.000219***	-0.000281***	-0.000307***
	(4,85e-05)	(6,18e-05)	(5,74e-05)	(5,21e-05)	(5,08e-05)	(6,37e-05)
Sex	0.228***	0.142***	0.152***	0.175***	0.218***	0.260***
	(0.0157)	(0.0191)	(0.0168)	(0.0157)	(0.0154)	(0.0199)
Race (ethnicity)	0.0781***	0.0444***	0.0525***	0.0715***	0.0714***	0.0926***
	(0.0115)	(0.0155)	(0.0125)	(0.0119)	(0.0128)	(0.0167)
Urban area	0.166***	0.353***	0.242***	0.141***	0.126***	0.122***
	(0.0157)	(0.0351)	(0.0227)	(0.0176)	(0.0146)	(0.0179)
Statutory	0.363***	0.668***	0.413***	0.266***	0.220***	0.175***
	(0.0218)	(0.0312)	(0.0254)	(0.0235)	(0.0260)	(0.0519)
Formal employment	0.169***	0.486***	0.220***	0.0889***	0.00397	-0.0791***
	(0.0114)	(0.0250)	(0.0165)	(0.0136)	(0.0148)	(0.0179)
Blue-collar	-0.0364**	-0.0569***	-0.0440**	-0.0458***	-0.0295	-0.0639**
	(0.0169)	(0.0213)	(0.0196)	(0.0175)	(0,0184)	(0.0284)
Chronic diseases	-0.0156	-0.0237	-0.0260**	-0.0250**	0.0119	0.00787
	(0.0114)	(0.0153)	(0.0113)	(0.0105)	(0.0124)	(0.0167)
Lambda	-0.07847***					
	(0.02372)					
Rho						1.376***
						(0.363)

(continue)

(continued)

Variables	Mean (Heckman)	Quantile (Arellano and Bonhomme)				
		Q10	Q25	Median	Q75	Q90
Constant	0.500***	-0.224	0.441***	0.784***	0.848***	1.038***
	(0.102)	(0.147)	(0.144)	(0.122)	(0.113)	(0.142)
Control for Federative Units	Yes	Yes	Yes	Yes	Yes	Yes
Control for CNAE Sectors	Yes	No	No	No	No	No
Observations	60,041	59,576	59,576	59,576	59,576	59,576
Wald Chi² (49)	5315.08					
Prob > Chi²	0.0000					
Log pseudo maximum likelihood (coefficient)	-1.17e+08					

Source: Prepared by the authors based on data from the 2019 PNS (IBGE, 2021).
Note: The Arellano and Bonhomme (2017) model was estimated with correction for heteroscedasticity using the bootstrap method, with 500 repetitions. Standard errors are in parentheses. The statistical significance of the estimates is defined by: *** p<0.01, ** p<0.05, * p<0.1.

With respect to the variables associated with physical activity, a negative relationship is observed between insufficient physical activity and labor income. Specifically, workers who engage in insufficient physical activity earn, on average, 7.12% less in the Brazilian labor market. This adverse effect persists across the entire earnings distribution, ranging from 4.19% in the lowest decile (Q10) to 9.97% in the highest decile (Q90). It is noteworthy that the magnitude of this effect increases progressively along the earnings distribution, being more pronounced at the upper quantiles.

A similar pattern is found in the estimates for sufficient physical activity. Workers classified as engaging in sufficient levels of physical activity also experience lower earnings, with an average negative effect of 7.28%. Across the quantiles, this impact ranges from 3.2% at the lower end of the distribution (Q10) to 8.6% at the upper end (Q90). As with the estimates for insufficient activity, the negative association becomes more substantial as one moves toward higher income brackets.

This pattern is likely driven by the occupational composition of income levels in Brazil. Higher-income jobs are typically associated with administrative and managerial (white-collar) positions, which tend to involve lower levels of physical exertion (Pereira; Almeida; Oliveira, 2020). In contrast, blue-collar occupations – characterized by manual labor – tend to require greater physical effort. As noted by Dang, Maitra, and Menon (2019), energy expenditure is generally higher in blue-collar jobs than in white-collar positions for both men and women. Given that physical activity metrics capture energy expenditure from labor-related tasks, blue-collar jobs contribute more significantly to physical activity levels. Since these occupations are generally associated with lower earnings – as shown in Table 3 – this helps explain the negative relationship between physical activity and labor income.

Moreover, components such as commuting, domestic chores, and leisure-related travel are included in measures of moderate to vigorous physical activity. These activities are more common among lower-income individuals, which further contributes to the negative association between both insufficient and sufficient physical activity and earnings in the Brazilian labor market.

Regarding the estimates related to the healthy eating habits index, no consistent or statistically significant patterns were identified in either the mean or quantile regressions. A similar result is observed for the composite healthy habits index, which also does not display significant associations with labor income when considering the full sample.

To explore potential heterogeneities in the relationship between healthy behaviors and labor income, Table 5 presents the results of the estimates for the healthy behavior variables in regressions on the mean and quantiles, focusing on the subsamples of women and men. First, in the sample restricted to women, both insufficient and sufficient physical activity continue to show negative associations with labor income, consistent with the results observed in the full sample. Again, the magnitude of these effects progressively intensifies across the income distribution.

Regarding insufficient physical activity, female workers who engage in such levels of activity earn, on average, 8.69% less in the labor market. As in the full sample, the most pronounced negative effect is observed in the upper end of the income distribution, with the highest estimated penalty reaching 15.3% in the top quantile (Q90).

Similarly, women who engage in sufficient physical activity also experience lower labor earnings, with an average negative return of 6.60%. In the quantile regression, the most substantial negative effect is found in the penultimate quantile (Q75), with an estimate of 10.9%. This suggests that the negative impact of sufficient physical activity on female labor income becomes more pronounced as earnings increase.

The indicator for healthy eating habits among women shows statistically significant effects only in the third quartile (Q75), indicating a positive association between income and these habits. However, it is important to emphasize that this result does not allow for robust conclusions regarding the relationship between healthy eating habits and earnings across the full female income distribution.

Similarly, the healthy habits index – which captures the absence of smoking and alcohol consumption – shows a negative relationship with women's labor income at the mean and upper quantiles (Q75 and Q90). On average, women adhering to these habits earn 1.53% less. In the remaining quantiles, the estimated coefficients are not statistically significant, suggesting that the relationship between these health behaviors and labor income is indeterminate in the lower strata of the female labor market in Brazil.

TABLE 5
Results of the estimation of the labor income equation at the mean and quantiles for women and men
Brazil – 2019

Variables	Mean (Heckman)	Quantile (Arellano and Bonhomme)				
		Q10	Q25	Median	Q75	Q90
Women						
Insufficient PA	-0.0869***	-0.0569**	-0.0612***	-0.0727***	-0.124***	-0.153***
	(0.0190)	(0.0258)	(0.0234)	(0.0195)	(0.0243)	(0.0343)
Sufficient PA	-0.0660***	-0.0299	-0.0473**	-0.0558***	-0.109***	-0.0987***
	(0.0161)	(0.0227)	(0.0188)	(0.0162)	(0.0210)	(0.0284)
Healthy eating habits index	0.00535	0.000872	0.00133	0.00729	0.0129*	0.00637
	(0.00639)	(0.00847)	(0.00790)	(0.00625)	(0.00732)	(0.00911)
Healthy habits index	-0.0153**	-0.0139	-0.00512	-0.00766	-0.0302***	-0.0292**
	(0.00779)	(0.00978)	(0.00817)	(0.00767)	(0.00997)	(0.0127)
Lambda	-0.11418					
	(0.02623)					
Rho						3.078*** (0.818)
Constant	0.742***	0.460*	1.167***	1.387***	1.542***	1.759***
	(0.146)	(0.260)	(0.279)	(0.221)	(0.217)	(0.268)
Observations	31,309	31,072	31,072	31,072	31,072	31,072
Men						
Insufficient PA	-0.0589***	-0.0247	-0.0642***	-0.0764***	-0.0704***	-0.0862***
	(0.0212)	(0.0257)	(0.0209)	(0.0202)	(0.0234)	(0.0291)
Sufficient PA	-0.0716***	-0.0485**	-0.0676***	-0.0821***	-0.0570***	-0.0637***
	(0.0155)	(0.0227)	(0.0173)	(0.0165)	(0.0188)	(0.0208)
Healthy eating habits index	-0.0105**	-0.00869	-0.00802	-0.00498	-0.00546	-0.0134*
	(0.00519)	(0.00717)	(0.00602)	(0.00560)	(0.00632)	(0.00695)
Healthy habits index	0.00751	0.0190**	0.0116	0.00524	0.0153**	0.00728
	(0.00642)	(0.00897)	(0.00772)	(0.00682)	(0.00769)	(0.00864)
Lambda	-0.02714					
	(0.03568)					
Rho						0.727 (0.515)
Constant	0.513***	-0.458***	0.144	0.589***	0.848***	0.957***
	(0.125)	(0.177)	(0.157)	(0.140)	(0.141)	(0.173)
Observations	28,732	28,504	28,504	28,504	28,504	28,504
Control for Federative Units	Yes	Yes	Yes	Yes	Yes	Yes
Control for CNAE Sectors	Yes	No	No	No	No	No

Source: Prepared by the authors based on data from the 2019 PNS (IBGE, 2021).
Note: The Arellano and Bonhomme (2017) model was estimated with correction for heteroscedasticity using the bootstrap method, with 500 repetitions. Controls related to individuals' personal, productivity and employment characteristics (education level, race, age, formal employment, public sector employment, urban/rural residence, chronic diseases, blue-collar occupations, federative units (states) and CNAE industrial sectors (only in regression to the mean)) were included in the estimations but suppressed from the table. Standard errors are in parentheses. The statistical significance of the estimates is defined by: *** p<0.01, ** p<0.05, * p<0.1.

Regarding the estimates restricted to the sample of men, it is important to highlight that the estimated values for lambda and rho are not statistically significant, indicating either the absence of sample selection bias among men or that the exclusion restrictions used in the first stage failed to capture the non-random labor supply decision. However, as emphasized by Vermeulen (2001), the use of a single valid exclusion restriction is sufficient to identify selection models. Given that five exclusion variables were employed, the hypothesis of no selection bias can be reasonably ruled out. This, in turn, supports the validity of the results of the second stage.

As in the results for the full sample and for women, the findings for men confirm a negative association between both insufficient and sufficient levels of moderate-to-vigorous physical activity and labor income. Lechner and Sari (2015) also report similar patterns across genders, although their study finds a positive relationship between physical activity and wages.

For men, the average effect of insufficient physical activity is a 5.89% reduction in labor earnings. In the quantile analysis, no statistically significant relationship is found at the lower quantile (Q10). However, the magnitude of the negative impact increases with income, peaking at 8.62% in the upper quantile (Q90).

Similarly, male workers who engage in sufficient physical activity earn, on average, 7.16% less in the Brazilian labor market. In the quantile regressions, the most pronounced negative effect appears at the median (Q50), where male workers practicing sufficient physical activity earn 8.21% less than their counterparts who do not engage in this level of vigorous activity.

With regard to the healthy eating habits indicator, individuals who predominantly consume the recommended foods earn, on average, 1.05% less. However, this relationship is statistically significant only in the highest quantile (Q90), where a negative return of 1.34% is observed in the earnings of workers who adhere to healthy eating patterns.

In contrast, the indicator reflecting healthy habits – defined as abstaining from smoking and consuming alcohol – shows a positive association with men's income only at the Q10 and Q75 quantiles. At these points in the distribution, modest returns of 1.9% and 1.53% are observed, respectively. However, at the mean and in the remaining quantiles, the coefficients are not statistically significant, suggesting an indeterminate relationship between this set of behaviors and male labor earnings.

Overall, the results of this study contrast with existing literature linking physical activity to labor market outcomes, which typically identifies a positive return to time spent in physical activity (Lechner, 2015; Lechner; Sari, 2015; Kosteas, 2012; Godoy; Triches, 2017). Nevertheless, Lechner (2015) acknowledges that the relative importance of the various mechanisms through which sports and exercise impact labor market outcomes remains poorly understood, a factor that has historically cast doubt on the robustness of empirical findings regarding the "sports premium." Along similar lines, Rooth (2011) notes that it is empirically very difficult to isolate these channels, and that individuals who participate in sports may represent a selective group in various unobservable dimensions, complicating efforts to identify the causal effects of physical activity.

This discrepancy may arise because the aforementioned studies consider only leisure-time physical activity, disregarding physical exertion performed in other domains. An and Liu (2012) highlight that measurement errors are common in outcome variables capturing monthly hours of physical activity, especially in dichotomous variables restricted to leisure-time activity, which fail to account for other forms of physical activity such as occupational and domestic tasks.

To demonstrate that controlling solely for leisure-time physical activity obscures the true relationship between income and physical activity, we re-estimated the same equations used in the previous analyses, this time including only the leisure-time physical activity control, rather than the controls for insufficient and sufficient activity. As presented in Table 6, in all estimations – whether for the full sample, women, or men – leisure-time physical activity is consistently and positively associated with earnings. Moreover, the magnitude of this positive association increases across the income distribution, suggesting that higher-income individuals benefit more, in relative terms, from engaging in physical activity during leisure time.

TABLE 6
Results of the estimation of the labor income equation at the mean and quantiles for men considering only the control of physical activities for leisure (1)
Brazil – 2019

Variables	Mean (Heckman)	Quantile (Arellano and Bonhomme)				
		Q10	Q25	Median	Q75	Q90
Full sample						
Leisure PA	0.0785***	0.0371**	0.0479***	0.0703***	0.0692***	0.0863***
	(0.0108)	(0.0146)	(0.0119)	(0.0111)	(0.0130)	(0.0167)
Healthy eating habits index	-0.00845**	-0.00534	-0.00742*	-0.00403	-0.00442	-0.00771
	(0.00405)	(0.00544)	(0.00448)	(0.00430)	(0.00487)	(0.00600)
Healthy habits index	-0.00338	0.00623	0.00288	0.000219	-4,03e-05	-0.00791
	(0.00496)	(0.00687)	(0.00543)	(0.00500)	(0.00594)	(0.00754)
Lambda	-0.08083***					
	(0.02354)					
Rho						1***
						(0.374)
Constant	0.429***	-0.341**	0.338**	0.606***	0.705***	0.851***
	(0.101)	(0.147)	(0.142)	(0.109)	(0.107)	(0.136)
Women						
Leisure PA	0.1000***	0.0582***	0.0557***	0.0775***	0.114***	0.118***
	(0.0154)	(0.0204)	(0.0187)	(0.0157)	(0.0199)	(0.0294)
Healthy eating habits index	0.000441	-0.00442	-0.000516	-0.000303	0.00424	0.00715
	(0.00637)	(0.00844)	(0.00758)	(0.00640)	(0.00726)	(0.00921)
Healthy habits index	-0.0157**	-0.0148	-0.00630	-0.00838	-0.0183**	-0.0211*
	(0.00774)	(0.00984)	(0.00833)	(0.00723)	(0.00893)	(0.0116)

(continue)

(continued)

Variables	Mean (Heckman)	Quantile (Arellano and Bonhomme)				
		Q10	Q25	Median	Q75	Q90
Lambda	-0.11241***					
	(0.02585)					
Rho						1.963***
						(0.673)
Constant	0.666***	0.229	0.875***	1.101***	1.298***	1.550***
	(0.147)	(0.245)	(0.261)	(0.195)	(0.196)	(0.237)
Men						
Leisure PA	0.0716***	0.0574**	0.0449***	0.0727***	0.0598***	0.0751***
	(0.0151)	(0.0227)	(0.0163)	(0.0155)	(0.0162)	(0.0193)
Healthy eating habits index	-0.0130**	-0.00758	-0.00829	-0.00971*	-0.00844	-0.0142**
	(0.00518)	(0.00709)	(0.00599)	(0.00545)	(0.00621)	(0.00717)
Healthy habits index	0.00608	0.0169*	0.0122*	0.00609	0.0112	0.000217
	(0.00645)	(0.00932)	(0.00736)	(0.00668)	(0.00759)	(0.00918)
Lambda	-0.02878					
	0.03514					
Rho						0.510
						(0.532)
Constant	0.424***	-0.534***	0.175	0.486***	0.697***	0.756***
	(0.124)	(0.170)	(0.164)	(0.138)	(0.138)	(0.183)
Control for Federative Units	Yes	Yes	Yes	Yes	Yes	Yes
Control for CNAE Sectors	Yes	No	No	No	No	No
Observations	60,041	59,576	59,576	59,576	59,576	59,576

Source: Prepared by the authors based on data from the 2019 PNS (IBGE, 2021).
(1) The results of the estimation of the selection equations for these regressions are available in Table 2 of the Appendix.
Note: The Arellano and Bonhomme (2017) model was estimated with correction for heteroscedasticity using the bootstrap method, with 500 repetitions. Controls related to individuals' personal, productivity and employment characteristics (education level, gender (only in full sample regression), race, age, formal employment, public sector employment, urban/rural residence, chronic diseases, blue-collar occupations, federative units (states) and CNAE industrial sectors (only in regression to the mean)) were included in the estimations but suppressed from the table. Standard errors are in parentheses. The statistical significance of the estimates is defined by: *** p<0.01, ** p<0.05, * p<0.1.

In other words, the presence of an income effect is clearly evident in this case, which fails to capture the true relationship between what is conceptually defined as physical activity across all domains – as outlined by Benedetti *et al.* (2021) – and the earnings of Brazilian workers. In this context, Cruz *et al.* (2022) emphasize that longer working hours, Black race/skin color, female gender, lower *per capita* income, and lower educational attainment are all negatively associated with the likelihood of engaging in leisure-time physical activity. These findings suggest that participation in physical activity in Brazil is shaped by socioeconomic inequities. Specific population groups are more likely to engage in physical activity through work-related, commuting, or domestic tasks rather than through leisure, reflecting a socially reproduced pattern of inequality. This contrasts with the recommendations of the Global Action Plan on Physical Activity, which advocates for equitable access to opportunities for leisure-time physical activity. Therefore, this research contributes to the literature by incorporating relevant dimensions that likely play a critical role in this relationship and by offering a more realistic depiction of the link between physical activity and labor earnings.

Regarding healthy habits, the existing literature provides a coherent framework for interpreting these results. The evidence is unambiguous in the case of tobacco use – smokers consistently earn less across a wide range of contexts – whereas the effect of alcohol consumption tends to follow an inverted-U pattern: moderate consumption is often rewarded, as it facilitates the development of social capital during business meals and informal gatherings, while both abstinence and excessive drinking are associated with wage penalties (Bai; Grignon, 2024; Cawley; Ruhm, 2011; French; Zarkin, 1995; Macdonald; Shields, 2001). Marketing practices further reinforce this ambiguity by portraying alcoholic beverages as symbols of status and conviviality; thus, participating in or abstaining from these rituals may either open or limit professional opportunities.

Our estimates align with this interpretive framework and demonstrate that the healthy habits index behaves differently across the earnings distribution and between genders. Among women in the upper quantiles, the index is associated with lower earnings, consistent with the hypothesis that abstaining from social drinking may reduce access to influential networks. In the lower half of the female earnings distribution (Q10 to Q50), this mechanism appears to be largely absent, and the coefficients are statistically indistinguishable from zero. Among men, the index exhibits a positive association with income at Q10 and Q75, where the benefits of moderate alcohol consumption likely dominate and the networking advantage documented in the literature materializes. At the top of the male earnings distribution, however, the benefits of moderation and the costs of social exclusion may counterbalance each other, resulting in no net effect.

Most of the coefficients estimated for the healthy eating habits index are statistically insignificant, indicating that, on average, adherence to a higher-quality diet does not translate into wage gains for the majority of workers. Two points in the distribution, however, deviate from this general pattern: the index is positively associated with income for women at the 75th percentile and negatively associated for men at the 90th percentile.

The literature offers plausible mechanisms to explain these outliers. Diets rich in fruits, vegetables, and whole grains have been shown to improve cognitive function and physical performance – traits that are rewarded in the labor market (Muth; Park, 2021; Slavin; Lloyd, 2012). Among upper-middle-income women, these benefits may combine with reduced penalties related to weight stigma, thereby producing the observed wage premium. In contrast, among high-earning men, strict adherence to healthy eating norms may conflict with socially constructed masculine consumption patterns (Mahalik; Burns; Syzdek, 2007; Wardle *et al.*, 2004), potentially contributing to the negative coefficient observed at the 90th percentile.

The descriptive findings underscore substantial heterogeneity across the income distribution. The healthy eating habits index records its lowest mean values in the middle quantiles (Q50 and Q75) for both men and women. Conversely, the healthy habits index reveals a distinct gender pattern: women display the healthiest behaviors in the lower quantiles, but this gradient reverses from Q75 onward, with less healthy behaviors concentrated among the highest-earning women. Among men, the index remains uniformly low across the distribution, with the steepest decline observed in the lowest quantiles. Taken together, these results suggest that, although higher income affords better access to health information and the means to adopt healthier routines, workplace social norms can offset these advantages, leading to heterogeneous health profiles even among high-income individuals.

Concluding remarks

This study aimed to examine the effects of physical activity, healthy eating habits, and overall health-related behaviors on the labor earnings of Brazilian workers. Drawing on data from the 2019 National Health Survey (PNS), indicator variables were constructed to represent these health behaviors and incorporated into the estimation of the earnings equation, serving as proxies for health capital. The analysis employed both mean and quantile regressions across the income distribution, incorporating corrections for potential sample selection bias as proposed by Heckman (1979) and Arellano and Bonhomme (2017). Overall, the estimates were consistent and robust, underscoring the reliability and significance of the analytical approach.

The primary findings reveal a negative relationship between indicators of healthy behaviors and labor income. Specifically, with regard to physical activity, the results show that both insufficient and sufficient levels of activity are associated with lower earnings. This pattern persisted across income quantiles and tended to be more pronounced at higher income levels. This suggests that occupations involving greater physical effort – such as manual labor – tend to be associated with lower wages, while administrative or managerial positions, typically offering higher pay, involve less moderate or vigorous physical activity.

For the indices capturing healthy eating habits and overall healthy behaviors, the results were more ambiguous. A consistent negative effect was observed only among women, for whom abstaining from smoking and alcohol consumption was associated with lower income, particularly in the upper quantiles of the earnings distribution.

In conclusion, the results highlight the complexity of the relationship between health and labor market outcomes and suggest that promoting healthy habits should take into account the diverse socioeconomic realities faced by Brazilian workers. Although approximately 56% of Brazilians engage in some form of physical activity – even at insufficient levels – a considerable portion of those who do so belong to the lowest income strata. Through this channel, they may achieve better health outcomes than some higher-income workers. While this specific analysis is beyond the scope of the present study, it opens avenues for future research on the topic.

Importantly, the central objective of this study is to demonstrate that restricting the definition of physical activity solely to leisure-time pursuits leads to a distorted understanding of its relationship with labor income. Both the WHO (2010) and Benedetti *et al.* (2021) stress that physical activity includes not only leisure but also work-related, commuting, and domestic tasks. In this context, the study seeks to highlight not only individuals' occupational roles but also, more fundamentally, their lifestyles – comprising a range of daily physical activities, including those related to work, transportation, household responsibilities, and, where feasible, leisure.

The key issue is that when only leisure-time physical activity is considered, a clear income effect emerges. Leisure-based activity is often inaccessible to lower-income individuals due to financial constraints and time scarcity. However, this limitation does not necessarily mean these individuals are sedentary or inactive. By including other domains of physical activity in the analysis, we show that many individuals remain physically active – not out of choice, but necessity. Nevertheless, their daily routines often meet the thresholds for moderate or even vigorous activity, as extensively discussed by Benedetti *et al.* (2021). In this regard, the present study advances the debate by expanding its scope and contributing new empirical evidence to inform public policies aimed at reducing sedentary behavior in Brazil.

Finally, it is important to acknowledge certain limitations of the study. The primary limitation concerns the self-reported nature of physical activity measures by domain, which is inherent in the 2019 PNS design. As noted by An and Liu (2012), many studies indicate that self-reported physical activity is prone to measurement error, and the direction of bias – whether individuals overreport or underreport their activity levels – remains uncertain.

Moreover, healthy behaviors such as physical activity typically generate long-term effects that may not be fully captured in cross-sectional analyses. While results align with the structure of the available data and no longitudinal follow-up is currently feasible, this remains an important limitation. Consequently, future research that addresses these challenges could yield even stronger evidence on the relationship between health behaviors and labor market outcomes.

References

ALMEIDA, A. T. C. de; ARAÚJO Jr., I. T. de. Smoking and wage penalty in the Brazilian labor market. **Economia Aplicada**, v. 21, n. 2, p. 249-276, 11 abr. 2017.

AN, R.; LIU, J. Local labor market fluctuations and physical activity among adults in the United States, 1990-2009. **International Scholarly Research Notices**, v. 2012, Article ID 318610, 2012.

ARELLANO, M.; BONHOMME, S. Quantile selection models with an application to understanding changes in wage inequality. **Econometrica**, v. 85, n. 1, p. 1-28, 2017.

BAI, Y.; GRIGNON, M. Why do drinkers earn more? Job characteristics as a possible link. **Health Economics**, v. 33, n. 6, p. 1133-1152, 2024.

BENEDETTI, T. R. B. *et al*. Validade e clareza dos conceitos e terminologias do Guia de Atividade Física para a População Brasileira. **Revista Brasileira de Atividade Física & Saúde**, v. 26, p. 1-11, 2021.

BHATTACHARYA, J.; HYDE, T.; TU, P. **Health economics**. [S.l.]: Bloomsbury Publishing, 2014.

BIEWEN, M.; ERHARDT, P. arhomme: an implementation of the Arellano and Bonhomme (2017) estimator for quantile regression with selection correction. **The Stata Journal: Promoting Communications on Statistics and Stata**, v. 21, n. 3, p. 602-625, set. 2021.

BLAIR, S. N. *et al*. How much physical activity is good for health? **Annual Review of Public Health**, v. 13, p. 99-126, 1992.

BÖCKERMAN, P. *et al*. If you drink, don't smoke: joint associations between risky health behaviors and labor market outcomes. **Social Science & Medicine**, v. 207, p. 55-63, 2018.

BÖCKERMAN, P.; HYYTINEN, A.; MACZULSKIJ, T. Alcohol consumption and long-term labor market outcomes. **Health Economics**, v. 26, n. 3, p. 275-291, 2017.

CAMERON, A. C.; TRIVEDI, P. K. **Microeconometrics**: methods and applications. [S.l.]: Cambridge University Press, 2005.

CASE, A.; DEATON, A. **Broken down by work and sex**: how our health declines. [S.l.]: National Bureau of Economic Research Cambridge, MA, 2003.

CAWLEY, J.; RUHM, C. J. The economics of risky health behaviors. *In*: PAULY, M. V.; MCGUIRE, T. G.; BARROS, P. P. (Ed.). **Handbook of health economics**. [S.l.]: Elsevier, 2011. v. 2, p. 95-199.

COELHO, D.; VESZTEG, R.; SOARES, F. V. **Regressão quantílica com correção para a seletividade amostral**: estimativa dos retornos educacionais e diferenciais raciais na distribuição de salários das mulheres no Brasil. Brasília: Ipea, 2010. (Texto para Discussão, 1483). Available in: https://www.econstor.eu/handle/10419/91273. Access in: 21 jul. 2024.

CONTOYANNIS, P.; RICE, N. The impact of health on wages: evidence from the British Household Panel Survey. **Empirical Economics**, v. 26, n. 4, p. 599-622, 2001.

CRUZ, D. K. A. *et al*. Iniquidades socioeconômicas associadas aos diferentes domínios da atividade física: resultados da Pesquisa Nacional de Saúde 2019. **Epidemiologia e Serviços de Saúde**, v. 31, n. especial, e2021398, 2022.

CURRIE, J.; MADRIAN, B. C. Health, health insurance and the labor market. *In*: ASHENFELTER, O.; CARD, D. (Ed.). **Handbook of labor economics**. [S.l.]: Elsevier, 1999. v. 3, p. 3309-3416.

CUTLER, D. M.; LLERAS-MUNEY, A. Understanding differences in health behaviors by education. **Journal of Health Economics**, v. 29, n. 1, p. 1-28, 2010.

DANG, A.; MAITRA, P.; MENON, N. Labor market engagement and the body mass index of working adults: evidence from India. **Economics & Human Biology**, v. 33, p. 58-77, 2019.

FRENCH, M. T.; ZARKIN, G. A. Is moderate alcohol use related to wages? Evidence from four worksites. **Journal of Health Economics**, v. 14, n. 3, p. 319-344, 1995.

GODOY, M. R.; TRICHES, D. Effects of physical activity on earnings in the Brazilian labor market. **EconomiA**, v. 18, n. 2, p. 180-191, 2017.

GROSSMAN, M. Health capital and the demand for health. **Journal of Political Economy**, v. 80, n. 2, p. 223-255, 1972.

GROSSMAN, M. The human capital model. *In*: CULYER, A.; NEWHOUSE, J. (Ed.). **Handbook of health economics**. [S.l.]: Elsevier, 2000. v. 1, p. 347-408.

HALLAL, P. C. *et al*. Physical inactivity: prevalence and associated variables in Brazilian adults. **Medicine and Science in Sports and Exercise**, v. 35, n. 11, p. 1894-1900, 2003.

HAMMOND, R. A.; LEVINE, R. The economic impact of obesity in the United States. **Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy**, v. 3, p. 285-295, Aug. 2010.

HECKMAN, J. J. Sample selection bias as a specification error. **Econometrica: Journal of the Econometric Society**, v. 47, n. 1. p. 153-161, 1979.

IBGE – Instituto Brasileiro de Geografia e Estatística. **Pesquisa Nacional de Saúde 2013**: percepção do estado de saúde, estilos de vida e doenças crônicas. Brasil, grandes regiões e unidades da federação. Rio de Janeiro: IBGE, 2014.

IBGE – Instituto Brasileiro de Geografia e Estatística. **PNS – Pesquisa Nacional de Saúde 2019**. Rio de Janeiro: IBGE, 2021. Available in: https://www.ibge.gov.br/estatisticas/sociais/trabalho/9160-pesquisa-nacional-de-saude.html?edicao=30563. Access in: 2 May 2025.

IHME – Institute for Health Metrics and Evaluation. **Global Burden of Disease Study 2019 (GBD 2019) Data Resources | GHDx**. Available in: https://ghdx.healthdata.org/gbd-2019. Access in: 2 May 2025.

JOHANSSON, E. *et al*. The association of alcohol dependency with employment probability: evidence from the population survey ‘Health 2000 in Finland’. **Health Economics**, v. 16, n. 7, p. 739-754, 2007.

KAISER, H. F. The varimax criterion for analytic rotation in factor analysis. **Psychometrika**, v. 23, n. 3, p. 187-200, 1958.

KARI, J. **Lifelong physical activity and long-term labor market outcomes**. Jyväskylä, Finland: University of Jyväskylä, 2018. (Jyväskylä Studies in Business and Economics, n. 184).

KOSTEAS, V. D. The effect of exercise on earnings: evidence from the NLSY. **Journal of Labor Research**, v. 33, n. 2, p. 225-250, 2012.

KRINSKI, K. *et al*. Efeitos do exercício físico no sistema imunológico. **Revista Brasileira de Medicina**, v. 67, n. 7, 2010.

LECHNER, M. Sports, exercise, and labor market outcomes. **IZA World of Labor**, 2015.

LECHNER, M.; SARI, N. Labor market effects of sports and exercise: evidence from Canadian panel data. **Labour Economics**, v. 35, p. 1-15, 2015.

LEE, L.-F. Identification and estimation in binary choice models with limited (censored) dependent variables. **Econometrica: Journal of the Econometric Society**, v. 47, n. 4, p. 977-996, 1979.

LIN, X. *et al*. Leisure time physical activity and cardio-metabolic health: results from the Brazilian Longitudinal study of adult health (ELSA-Brasil). **Journal of the American Heart Association**, v. 5, n. 6, e003337, 2016.

MACDONALD, Z.; SHIELDS, M. A. The impact of alcohol consumption on occupational attainment in England. **Economica**, v. 68, n. 271, p. 427-453, 2001.

MADDALA, G. S. **Limited-dependent and qualitative variables in econometrics**. [S.l.]: Cambridge University Press, 1983.

MAHALIK, J. R.; BURNS, S. M.; SYZDEK, M. Masculinity and perceived normative health behaviors as predictors of men’s health behaviors. **Social Science & Medicine**, v. 64, n. 11, p. 2201-2209, 2007.

MINCER, J. **A study of personal income distribution**. [S.l.]: Columbia University, 1957.

MORRIS, S. The impact of obesity on employment. **Labour Economics**, v. 14, n. 3, p. 413-433, 2007.

MUTH, A.-K.; PARK, S. Q. The impact of dietary macronutrient intake on cognitive function and the brain. **Clinical Nutrition**, v. 40, n. 6, p. 3999-4010, 2021.

PEREIRA, R. M.; ALMEIDA, A. N. de; OLIVEIRA, C. A. de. O valor estatístico de uma vida: estimativas para o Brasil. **Estudos Econômicos**, v. 50, p. 227-259, 2020.

PEREIRA, R. M.; OLIVEIRA, C. A. de. Os diferenciais de salário por gênero no Rio Grande do Sul: uma aplicação do modelo de Heckman e da decomposição de Oaxaca-Blinder. **Redes**, v. 21, n. 1, p. 148-173, 2016.

PEREIRA, R. M.; OLIVEIRA, C. A. de. Discriminação por gênero no mercado de trabalho local: um estudo para os municípios do Rio Grande do Sul. **Análise Econômica**, ano. 35, n. 68, p. 87-116, 2017.

PEREIRA, R. M.; OLIVEIRA, C. A. de; MACHADO, G. C. Uma nota sobre o valor estatístico de uma vida por grupos etários e por idade: evidências a partir da teoria do ciclo de vida para o Brasil. **Revista Brasileira de Economia**, v. 77, e192023, 2023.

PEREIRA, R. M.; TRIACA, L. M.; OLIVEIRA, C. A.; VILELA, G. F. Hábitos saudáveis e mercado de trabalho: uma análise quantílica para o Brasil. *In*: 52º ENCONTRO NACIONAL DE ECONOMIA. **Anais [...]**. Natal-RN: Anpec, 2024.

ROOTH, D.-O. Obesity, attractiveness, and differential treatment in hiring: a field experiment. **Journal of Human Resources**, v. 44, n. 3, p. 710-735, 2009.

ROOTH, D.-O. Work out or out of work – the labor market return to physical fitness and leisure sports activities. **Labour Economics**, v. 18, n. 3, p. 399-409, 2011.

SILVEIRA, L. S.; LEÃO, N. S. Segregação ocupacional e diferenciais de renda por gênero e raça no Brasil: uma análise de grupos etários. **Revista Brasileira de Estudos de População**, v. 38, e0151, 2021.

SLAVIN, J. L.; LLOYD, B. Health benefits of fruits and vegetables. **Advances in Nutrition**, v. 3, n. 4, p. 506-516, 2012.

SOARES, L. de S. A.; DALBONI, F. M.; TEIXEIRA, E. C. Capital social e felicidade individual no Brasil. **Economia Aplicada**, v. 26, n. 3, p. 247-274, 2022.

TAUCHMANN, H. Consistency of Heckman-type two-step estimators for the multivariate sample-selection model. **Applied Economics**, v. 42, n. 30, p. 3895-3902, 2010.

TERZA, J. V. Alcohol abuse and employment: a second look. **Journal of Applied Econometrics**, v. 17, n. 4, p. 393-404, 2002.

TREMBLAY, M. S. *et al.* Sedentary behavior research network (SBRN) – terminology consensus project process and outcome. **International Journal of Behavioral Nutrition and Physical Activity**, v. 14, Article 75, 2017.

VAN OURS, J. C. A pint a day raises a man's pay; but smoking blows that gain away. **Journal of Health Economics**, v. 23, n. 5, p. 863-886, 2004.

VIEIRA, L. P.; PEREIRA, R. M. Smoking effects on labor income: new evidence for Brazil. **EconomiA**, v. 25, n. 1, p. 53-73, 2024.

WARDLE, J. *et al.* Gender differences in food choice: the contribution of health beliefs and dieting. **Annals of Behavioral Medicine**, v. 27, n. 2, p. 107-116, 2004.

WHO – World Health Organization. **Global recommendations on physical activity for health**. Geneva: World Health Organization, 2010.

WHO – World Health Organization. **Global action plan for the prevention and control of noncommunicable diseases 2013-2020**. Geneva: World Health Organization, 2013.

WHO – World Health Organization. **Global status report on alcohol and health 2018**. Geneva: World Health Organization, 2018.

WHO – World Health Organization. **Key facts**: healthy diet. Available in: https://www.who.int/news-room/fact-sheets/detail/healthy-diet. Access in: 2 May 2025.

About the authors

Rafael Mesquita Pereira is PhD in Applied Economics from the "Luiz de Queiroz" College of Agriculture at the University of São Paulo (ESALQ/USP). Adjunct Professor at the Institute of Economic, Administrative and Accounting Sciences (ICEAC) and faculty member of the Postgraduate Program in Applied Economics (PPGE) at the Federal University of Rio Grande (FURG).

Lívia Madeira Triaca is PhD in Development Economics from the Pontifical Catholic University of Rio Grande do Sul (PUC – RS). Adjunct Professor at the Institute of Economic, Administrative and Accounting Sciences (ICEAC) at the Federal University of Rio Grande (FURG) and faculty member of the Postgraduate Program in Organizations and Markets (PPGOM) at the Federal University of Pelotas (UFPEL).

Cristiano Aguiar de Oliveira is PhD in Applied Economics from Federal University of Rio Grande do Sul (UFRGS). Associate Professor at the Institute of Economic, Administrative and Accounting Sciences (ICEAC) and faculty member of the Postgraduate Program in Applied Economics (PPGE) at the Federal University of Rio Grande (FURG).

Guilherme da Fonseca Vilela is Bachelor of Medicine from the Federal University of Fronteira Sul (UFFS) and Master of Physical Education from the Federal University of Pelotas.

Contact address

Rafael Mesquita Pereira
Universidade Federal do Rio Grande
Av. Itália, km 8, Carreiros
96203-900 – Rio Grande-RS, Brazil

Lívia Madeira Triaca
Universidade Federal do Rio Grande
Av. Itália, km 8, Carreiros
96203-900 – Rio Grande-RS, Brazil

Cristiano Aguiar de Oliveira
Universidade Federal do Rio Grande
Av. Itália, km 8, Carreiros
96203-900 – Rio Grande-RS, Brazil

Guilherme da Fonseca Vilela
Universidade Federal da Fronteira Sul – *campus* Passo Fundo
Rua Capitão Araújo, 20, Centro
99010-200 – Passo Fundo, RS - Brasil

CRediT

Funding: This work was carried out with funding from the National Council for Scientific and Technological Development (CNPq) – Brazil, under Call CNPq/MCTI No. 10/2023 – Track A/Emerging Groups – process n. 407980/2023-7.
Conflicts of interest: The authors certify that they have no personal, commercial, academic, political or financial interest that represents a conflict of interest in relation to the manuscript.
Ethical Approval: The authors declare that the study did not include human beings or animals.
Availability of data and material: The content is already available.
Authors' contributions:
Rafael Mesquita Pereira: conceptualization; data curation; project management; acquisition of funding; investigation; methodology; software; supervision; formal analysis; writing – original draft; writing – review and editing.
Lívia Madeira Triaca: conceptualization; data curation; investigation; methodology; software; formal analysis; writing – review and editing.
Cristiano Aguiar de Oliveira: writing – review and editing.
Guilherme da Fonseca Vilela: conceptualization; data curation; writing – review and editing.
Editors: Bernardo Lanza Queiroz, Júlia Almeida Calazans and Maria Carolina Tomas

Resumo

Hábitos saudáveis e mercado de trabalho: uma análise quantílica para o Brasil

Este estudo investiga os efeitos de comportamentos saudáveis sobre os rendimentos do trabalho no Brasil. Embora a literatura especializada geralmente aponte uma associação positiva entre tais comportamentos e os salários, essa relação ainda é pouco explorada em países em desenvolvimento. Utilizando microdados da Pesquisa Nacional de Saúde (PNS) de 2019, realizada pelo IBGE, construímos variáveis indicadoras para comportamentos relacionados à saúde e estimamos equações salariais do tipo mincerianas. A análise emprega regressões médias e quantílicas, com correção para viés de seleção amostral decorrente da participação no mercado de trabalho. Contrariando achados convencionais, os resultados indicam uma associação negativa entre atividade física e rendimentos do trabalho ao longo da distribuição de renda. Tal resultado é explicado pela definição mais ampla de atividade física adotada no estudo, que inclui não apenas práticas de lazer, mas também atividades físicas relacionadas ao trabalho, deslocamento e afazeres domésticos — mais frequentes entre indivíduos de menor renda. Além disso, o engajamento em comportamentos preventivos de saúde está negativamente associado aos rendimentos, sobretudo entre as mulheres, sugerindo interações complexas entre investimento em saúde, restrições de tempo e resultados no mercado de trabalho.

Palavras-chave: Atividades físicas. Hábitos saudáveis. Mercado de trabalho. Regressão quantílica. Correção de seleção de amostra.

Resumen

Hábitos saludables y mercado laboral: un análisis cuantil para Brasil

Este estudio investiga los efectos de los comportamientos saludables sobre los ingresos laborales en Brasil. Aunque la literatura especializada generalmente señala una asociación positiva entre dichos comportamientos y los salarios, esta relación aún está poco explorada en los países en desarrollo. Utilizando microdatos de la Encuesta Nacional de Salud (PNS) de 2019, realizada por el IBGE, construimos variables indicadoras de comportamientos relacionados con la salud y estimamos ecuaciones salariales de tipo minceriano. El análisis utiliza regresiones de medias y cuantiles, con corrección del sesgo de selección muestral resultante de la participación en el mercado laboral. Contrariamente a los hallazgos convencionales, los resultados indican una asociación negativa entre la actividad física y los ingresos laborales en toda la distribución del ingreso. Este resultado se explica por la definición más amplia de actividad física adoptada en el estudio, que incluye no solo las actividades de ocio, sino también las actividades físicas relacionadas con el trabajo, los desplazamientos y las tareas domésticas, más frecuentes entre los individuos con menores ingresos. Además, la participación en conductas preventivas de salud está asociada negativamente con los ingresos, particularmente entre las mujeres, lo que sugiere interacciones complejas entre la inversión en salud, las limitaciones de tiempo y los resultados en el mercado laboral.

Palabras clave: Actividades físicas. Hábitos saludables. Mercado de trabajo. Regresión cuantil. Corrección de selección de muestra.

Received for publication in 21/02/2025
Approved for publication in 20/06/2025

APPENDIX

TABLE 1
Estimation of the selection equation (correction of sample selection bias)
Brazil – 2019

Variables	Probit Decision to offer work [1 = job offer; 0 = no offer]					
	Full sample		Women		Men	
	Heckman	Arellano and Bonhomme	Heckman	Arellano and Bonhomme	Heckman	Arellano and Bonhomme
Educational level 2	0.262***	0.269***	0.0602	0.0626	0.448***	0.457***
	(0.0521)	(0.0525)	(0.0745)	(0.0747)	(0.0698)	(0.0703)
Educational level 3	0.387***	0.396***	0.217***	0.222***	0.587***	0.598***
	(0.0604)	(0.0608)	(0.0829)	(0.0832)	(0.0858)	(0.0863)
Educational level 4	0.291***	0.297***	0.193**	0.196**	0.391***	0.397***
	(0.0636)	(0.0640)	(0.0888)	(0.0891)	(0.0851)	(0.0856)
Educational level 5	0.702***	0.717***	0.587***	0.597***	0.823***	0.842***
	(0.0542)	(0.0545)	(0.0744)	(0.0746)	(0.0749)	(0.0750)
Educational level 6	0.597***	0.633***	0.612***	0.616***	0.549***	0.614***
	(0.0717)	(0.0707)	(0.0890)	(0.0894)	(0.106)	(0.105)
Educational level 7	0.918***	0.925***	0.856***	0.857***	0.893***	0.909***
	(0.0597)	(0.0598)	(0.0794)	(0.0797)	(0.0851)	(0.0854)
Age	0.158***	0.158***	0.136***	0.136***	0.180***	0.180***
	(0.00538)	(0.00543)	(0.00734)	(0.00737)	(0.00753)	(0.00762)
Age squared	-0.00188***	-0.00187***	-0.00167***	-0.00166***	-0.00208***	-0.00208***
	(6,55e-05)	(6,60e-05)	(9,06e-05)	(9,10e-05)	(9,14e-05)	(9,22e-05)
Sex	0.737***	0.751***	-	-	-	-
	(0.0222)	(0.0223)	-	-	-	-
Race (ethnicity)	-0.00722	-0.0136	-0.0191	-0.0158	0.00706	-0.0102
	(0.0244)	(0.0244)	(0.0321)	(0.0323)	(0.0363)	(0.0363)
Urban area	0.101***	0.102***	0.313***	0.315***	-0.0430	-0.0446
	(0.0231)	(0.0233)	(0.0332)	(0.0333)	(0.0357)	(0.0360)
Collects retirement	-1.128***	-1.139***	-0.786***	-0.804***	-1.564***	-1.575***
	(0.0434)	(0.0435)	(0.0579)	(0.0579)	(0.0554)	(0.0552)
Collects benefits	-0.424***	-0.421***	-0.195***	-0.191***	-1.399***	-1.411***
	(0.0621)	(0.0624)	(0.0633)	(0.0637)	(0.138)	(0.139)
Other sources of income	-0.765***	-0.758***	-0.635***	-0.619***	-0.947***	-0.946***
	(0.0305)	(0.0305)	(0.0354)	(0.0356)	(0.616)	(0.0601)
Responsible for household	0.308***	0.300***	0.132***	0.134***	0.502***	0.490***
	(0.0221)	(0.0222)	(0.0293)	(0.0298)	(0.0339)	(0.0329)
athrho	-0.140***		-0.216***		-0.0470	
	(0.0425)		(0.0504)		(0.0617)	
lnsigma	-0.574***		-0.624***		-0.547***	
	(0.0129)		(0.0173)		(0.0172)	
Constant	-3.556***	-3.531***	-3.192***	-3.181***	-3.225***	-3.161***
	(0.123)	(0.129)	(0.166)	(0.174)	(0.171)	(0.182)
Control for Federative Units	Yes	Yes	Yes	Yes	Yes	Yes
Observations	60,041	59,576	31,309	31,072	28,732	28,504

Source: Prepared by the authors based on data from the 2019 National Health Survey (IBGE, 2021).
Note: Standard errors in parentheses. Statistical significance of estimates defined by: *** p<0,01, ** p<0,05, * p<0,1.

TABLE 2
Estimation of the selection equation (correction of sample selection bias) considering only leisure-time physical activities
Brazil – 2019

Variables	Probit Decision to offer work [1 = job offer; 0 = no offer]					
	Full sample		Women		Men	
	Heckman	Arellano and Bonhomme	Heckman	Arellano and Bonhomme	Heckman	Arellano and Bonhomme
Educational level 2	0.262***	0.269***	0.0602	0.0626	0.448***	0.457***
	(0.0521)	(0.0525)	(0.0745)	(0.0747)	(0.0698)	(0.0703)
Educational level 3	0.386***	0.396***	0.217***	0.222***	0.587***	0.598***
	(0,0604)	(0,0608)	(0,0829)	(0.0832)	(0.0858)	(0.0863)
Educational level 4	0.291***	0.297***	0.193**	0.196**	0.391***	0.397***
	(0.0636)	(0.0640)	(0.0888)	(0.0891)	(0.0851)	(0.0856)
Educational level 5	0.702***	0.717***	0.587***	0.597***	0.823***	0.842***
	(0.0542)	(0.0545)	(0.0744)	(0.0746)	(0.0749)	(0.0750)
Educational level 6	0.597***	0.633***	0.612***	0.616***	0.549***	0.614***
	(0.0716)	(0.0707)	(0.0890)	(0.0894)	(0.106)	(0.105)
Educational level 7	0.919***	0.925***	0.856***	0.857***	0.893***	0.909***
	(0.0597)	(0.0598)	(0.0795)	(0.0797)	(0.0851)	(0.0854)
Age	0.158***	0.158***	0.136***	0.136***	0.180***	0.180***
	(0.00538)	(0.00543)	(0.00734)	(0.00737)	(0.00753)	0.00762)
Age squared	-0.00188***	-0.00187***	-0.00167***	-0.00166***	-0.00208***	-0.00208***
	(6,55e-05)	(6,60e-05)	(9,06e-05)	(9,10e-05)	(9,14e-05)	(9,22e-05)
Sex	0.737***	0.751***				
	(0.0222)	(0.0223)				
Race (ethnicity)	-0.00726	-0.0136	-0.0192	-0.0158	0.00704	-0.0102
	(0.0244)	(0.0244)	(0.0321)	(0.0323)	(0.0363)	(0.0363)
Urban area	0.100***	0.102***	0.313***	0.315***	-0.0429	-0.0446
	(0.0231)	(0.0233)	(0.0332)	(0.0333)	(0.0356)	(0.0360)
Collects retirement	-1.128***	-1.139***	-0.786***	-0.804***	-1.564***	-1.575***
	(0.0434)	(0.0435)	(0.0579)	(0.0579)	(0.0554)	(0.0552)
Collects benefits	-0.425***	-0.421***	-0.195***	-0.191***	-1.399***	-1.411***
	(0.0621)	(0.0624)	(0.0633)	(0.0637)	(0.138)	(0.139)
Other sources of income	-0.765***	-0.758***	-0.635***	-0.619***	-0.947***	-0.946***
	(0.0305)	0.0305)	(0.0355)	(0.0356)	(0.0614)	0.0601)
Responsible for household	0.308***	0.300***	0.132***	0.134***	0.502***	0.490***
	(0.0221)	(0.0222)	(0.0294)	(0.0298)	(0.0338)	(0.0329)
athrho	-0.145***		-0.213***		-0.0498	
	(0.0422)		(0.0497)		(0.0608)	
lnsigma	-0.574***		-0.626***		-0.547***	
	(0.0129)		(0.0173)		(0.0171)	
Constant	-3.556***	-3.531***	0.666***	-3.192***	-3.225***	-3.161***
	(0.123)	(0.129)	(0.147)	(0.166)	(0.171)	(0.182)
Control for Federative Units	Yes	Yes	Yes	Yes	Yes	Yes
Observations	60,041	59,576	31,309	31,072	28,732	28,504

Source: Prepared by the authors based on data from the 2019 National Health Survey (IBGE, 2021).
Note: Standard errors in parentheses. Statistical significance of estimates defined by: *** p‹0,01, ** p‹0,05, * p‹0,1.