Revista Brasileira de Economia. 2023, V. 77: e082023

A bivariate approach to the Mincerian earnings equation

Danúbia R. Cunha
Universidade Católica de Brasília, Brasil and Universidade Federal de Goiás, Brasil

Helton Saulo
Universidade de Brasília, Brasil

Sandro E. Monsueto
Universidade Federal de Goiás, Brasil

José Angelo Divino
Universidade Católica de Brasília, Brasil

This paper estimates bivariate regressions for wages and hours worked as an alternative to the univariate Mincerian earnings equation. The bivariate vector of dependent variables included both common and specific covariates. Using individual level data from the Brazilian National Household Sample Survey (PNAD), the Student t distribution produced the best fit to the data according to information criteria and Mahalanobis distance. The bivariate estimation accounts for correlation between the dependent variables, identifies antagonistic effects from common covariates and allows assuming different bivariate distributions. Education, type of employment contract and geographical region affect wages and hours worked in opposite directions.

Keywords. Bivariate distributions, Mincerian equation, Wages, Hours worked, Bivariate regression.

JEL classification. J20, C50.

1. Introduction

The Mincerian earnings equation introduced by Mincer (1974) is the baseline for a broad empirical literature on labor economics, including contributions by Senna (1976), Card (1999), Resende and Wyllie (2006) and Aali-Bujari et al. (2019). These studies generally seek to estimate the returns to education and experience on the wage rate earned by

Danúbia R. Cunha : danubiarodriguess@hotmail.com
Helton Saulo : heltonsaulo@gmail.com
Sandro E. Monsueto : monsueto@ufg.br
José Angelo Divino : jangelo@p.ucb.br

The authors thank participants in the XLVII meeting of the Brazilian Economic Association (ANPEC) for valuable comments and suggestions. Danúbia Rodrigues thanks CAPES/PROEX for the doctorate scholarship. Helton Saulo acknowledges financial support from CNPq and FAP-DF. José A. Divino thanks CNPq for the financial support through grant number 302632-2019-0. The authors declare that they have no conflict of interest that could have appeared to influence the work reported in this paper. All remaining errors are the authors’ sole responsibility.

https://doi.org/10.5935/0034-7140.20230008

the worker.[1] Mincer proposed that the distribution of wages among different occupations is positively correlated to the amount of investment in human capital, which positively affects productivity and economic growth.[2]

The Mincerian earnings equation was originally represented by a linear regression in which the wage rate was explained by education and experience. Following this approach by Mincer (1974), other explanatory variables were included in the regression, such as individual characteristics of gender and race that are used to assess the presence of discrimination in the labor market.

When deciding to join the labor market, a worker chooses the quantity of hours that he will supply to the market. Sedlacek and Santos (1991) used data from the Brazilian National Household Sample Survey (PNAD) to analyze the relationship between the husband's income and the labor supply by the spouse. They found that the higher the husband's income, the higher the reservation wage and less likely the wife will work. Moreover, the younger and more children the family has, the less likely they are to join the labor market or, when they do so, they will supply fewer hours of work.

As far as estimation methods are concerned, since Mincer (1974), the literature has used the traditional ordinary least squares (OLS) method and its variants with instrumental variables, quantile regression, sample selection, and procedures based on maximum likelihood estimation [Chatterjee and Price (1991), Heckman (1979), Buchinsky (2001)]. In Brazil, the greater availability of microdata and the improvement of the computational capacity contributed to the expansion of the empirical evidence, as highlighted by Maciel et al. (2001), Giuberti and Menezes-Filho (2005) and Madalozzo (2010).

A common feature in the literature is the use of earnings per hour as the dependent variable in the Mincerian equation. This variable, in general, is obtained by simple division of wage earned by hours worked in the period. Such an approach, however, implies the agglutination, in a single variable, of two distinct components, represented by earnings and hours worked, which should be modeled separately. The determinants of earnings and hours worked are not necessarily the same and those that enter in both regressions might differ in either quantitative (magnitudes) or qualitative (signals) terms.

This feature is not captured by traditional estimates of the Mincerian equation that uses wage rate as the sole dependent variable. The stock of human capital, measured by formal education and experience, for instance, tends to increase the workers' remuneration, but it might also reduce the willingness to supply working hours in the labor market. Those who are more qualified might receive higher remuneration by working less hours than those who are less qualified. These antagonistic effects of education on wages and hours worked are not captured by the univariate estimates of the Mincerian earnings equation.

Therefore, there is a gap in the literature that this study seeks to fill. The common practice of using the earnings per hour dependent variable might hide effects of

[1] The wage rate is usually defined as the wage per hour earned by workers.

[2] Human capital is understood as the set of attributes acquired by a worker by means of education, skill, and experience that improve productivity. This term was introduced by Mincer (1958) and later explored by Becker (1993) and Heckman et al. (2000), among others.

covariates that would be distinct if separately assessed by regressions on wage and hours worked. In contrast to the classical approach, this paper aims to estimate a bivariate regression for the Mincerian equation considering earnings and hours worked as a bivariate vector of dependent variables. The regressions include both common and specific covariates for the bivariate vector of earnings and hours worked. The bivariate Normal, Student t, and Birnbaum-Saunders (BS)[3] distributions are used in the estimation. For the sake of comparison, the univariate Mincerian earnings equation will also be estimated, considering a single dependent variable represented by earnings per hour worked. Estimates will be made for the Brazilian economy using data extracted from the Brazilian National Household Sample Survey (PNAD) for the period from 2013 to 2015.

Advantages of the bivariate regression approach include the possibility of modeling a correlation structure among the dependent variables. If there is correlation, the estimation of univariate regressions separately for earnings and hours worked might provide biased results [Marchant et al. (2016)]. The bivariate framework allows to identify antagonistic effects of common covariates on the two different dependent variables. Finally, there is flexibility to assume different bivariate distributions for the earnings and hours-worked model. As in Heckman (1976), the parameters will be estimated by maximum likelihood, which is efficient according to Mittelhammer et al. (2000). Thus, the bivariate model emerges as an important alternative to the univariate equation that is traditionally estimated for the Mincerian earnings equation.

The results indicate that some common explanatory variables have different signals and magnitudes of the estimated coefficients in the bivariate regression of earnings and hours-worked. Specifically, the estimated coefficients for education, type of employment contract, and geographical region have distinct signals and different magnitudes for wage and hours worked regressions. Considering education, for instance, more years in school imply in higher average wage and lower supply of hours to the labor market. In the univariate regression, however, only the positive effect of an additional year of study on the wage rate is observed. Furthermore, the bivariate model captures the correlation between the two dependent variables, which increases robustness in relation to the estimation of separate univariate regressions. Thus, there are important advantages associated to the bivariate approach when compared to the univariate regression, suggesting that the former is more suitable for the estimation of the Mincerian earnings equation.

The paper is organized as follows. Section 2 describes the empirical model, presents the database, reports, and discusses the results. Finally, the third section is dedicated to the concluding remarks.

[3] See Johnson et al. (1995), Balakrishnan and Lai (2009), Santos-Neto et al. (2012) and Saulo et al. (2021; 2020)

2. Econometric approach

2.1 *Empirical model*

The Mincerian earnings equation is typically described by the following univariate regression:

$$\log(\boldsymbol{w}) = \boldsymbol{X}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{1}$$

where $\log(\boldsymbol{w})$ is a vector with the logarithm of the wage per hour (dependent variable), $\boldsymbol{\gamma}$ is a vector of coefficients, $\boldsymbol{X}$ is a matrix of explanatory variables, such as education, experience, race, gender and others, and $\boldsymbol{\varepsilon}$ is a random error vector, usually assumed to follow a normal distribution.[4]

The differential of the present paper is to model the earnings equation (1) as a bivariate regression of wages and hours worked separately in order to capture different effects of the common explanatory variables on wages and labor supply. Furthermore, as earnings and hours worked are correlated, the bivariate regression is more appropriate than the univariate estimation of separate regressions.

In the bivariate environment, the model can be estimated as a vector of dependent variables $\boldsymbol{Y}_i = (Y_{1i}, Y_{2i})^\top$, where Y_{1i} is the wage in the main job and Y_{2i} represents the hours dedicated to the main job by each individual i. This vector might be modeled by a set $\boldsymbol{x}$ of explanatory variables using one of the bivariate distributions described in the Appendix, such that:

i) Bivariate Normal distribution:

$$\begin{aligned} \mathrm{E}[\log(Y_{1i})\,|\,\boldsymbol{X}_i = \boldsymbol{x}_i] &= \mu_{1i} = \boldsymbol{x}_i^\top \boldsymbol{\beta}_1, \quad i = 1,\ldots,n, \\ \mathrm{E}[\log(Y_{2i})\,|\,\boldsymbol{X}_i = \boldsymbol{x}_i] &= \mu_{2i} = \boldsymbol{x}_i^\top \boldsymbol{\beta}_2, \quad i = 1,\ldots,n; \end{aligned} \tag{2}$$

ii) Bivariate t distribution:

$$\begin{aligned} \mathrm{E}[\log(Y_{1i})\,|\,\boldsymbol{X}_i = \boldsymbol{x}_i] &= \mu_{1i} = \boldsymbol{x}_i^\top \boldsymbol{\beta}_1, \quad i = 1,\ldots,n, \\ \mathrm{E}[\log(Y_{2i})\,|\,\boldsymbol{X}_i = \boldsymbol{x}_i] &= \mu_{2i} = \boldsymbol{x}_i^\top \boldsymbol{\beta}_2, \quad i = 1,\ldots,n; \end{aligned} \tag{3}$$

iii) Bivariate BS distribution:

$$\begin{aligned} \mathrm{E}[Y_{1i}\,|\,\boldsymbol{X}_i = \boldsymbol{x}_i] &= \mu_{1i} = \exp(\boldsymbol{x}_i^\top \boldsymbol{\beta}_1), i = 1,\ldots,n, \\ \mathrm{E}[Y_{1i}\,|\,\boldsymbol{X}_i = \boldsymbol{x}_i] &= \mu_{2i} = \exp(\boldsymbol{x}_i^\top \boldsymbol{\beta}_2), i = 1,\ldots,n. \end{aligned} \tag{4}$$

Notice that in the cases of the Normal and t, we assume that the dependent variables have bivariate log-normal and log-t distributions, which implies that the logarithm of the variables follow the Normal and t bivariate distributions, respectively [Vanegas

[4] The variable wage per hour is commonly calculated as

$$\text{wage-hour} = \text{monthly wage}/(\text{hours worked} \times 4.33).$$

and Paula (2016)]. For the bivariate BS distribution, it is not necessary to apply the logarithm due to the parameterization as a function of averages of this distribution [Saulo et al. (2021; 2020)]. Based on the literature, we defined the set of covariates used in the estimations and separated covariates that affect both earnings and hours worked simultaneously from those that affect only one of them separately.

The common covariates, which affect both earnings and hours worked, are:

- Gender: dummy variable that assumes value 1 for men and 0 for women;
- Race: dummy variable that assumes value 1 for Caucasians and 0 for non-Caucasians;
- Marital status: dummy variable that assumes value 1 for married and 0 for unmarried individuals;
- Age and age^2: age, measured in years, and age square are used as proxy for the labor market experience of the individual, following the literature;
- Years of schooling: is a proxy for education, ranging from 0 to 16 years of study in the sample;
- Category (high, high mean, mean, low mean, low): binary variables used to designate occupancy category, segmented according to socioeconomic criteria and having the low category as a reference [5];
- Type of employment contract (with employment record card, without employment record card, autonomous, civil servant): dummy variables that seek to capture the type of occupation of the individual in the labor market, having "with employment record card" as the base category;
- Metropolitan region (Belém-PA, Fortaleza-CE, Recife-PE, Salvador-BA, Belo Horizonte-MG, Rio de Janeiro-RJ, Curitiba-PR, Porto Alegre-RS, Brasília-DF and São Paulo-SP): dummy variables that designate the metropolitan regions of residence of the individuals, taking São Paulo as the reference category;
- Year (2013, 2014, and 2015): time dummies for the years of the sample, having 2013 as the reference year;
- Sector of activity (agriculture, industry, construction, commerce, food and others, education, health, and social services): dummy variables used to capture cluster effects by sector of activity of the individuals, having individuals working in the public sector as reference.

The covariates that affect only earnings are:

- Labor union: dummy variable that assumes value of 1 for individuals affiliated to any labor union and 0 for those who were not affiliated;

[5] The occupational classification is based on Jannuzzi (2001).

- Social Security: dummy variable that assumes value of 1 for individuals who were taxpayers for social security in the reference period and 0 for those who were not taxpayers;
- Time in job: number of years employed in the current main job, ranging from 0 to 56 years in the sample.

The covariates that affect only hours worked are:

- Head: dummy variable that assumes value 1 if the reference individual in the household is head of the family and 0 otherwise (non-head);
- Minor: dummy variable used to capture if there are children under 10 years old in the household;
- Inactivity: dummy variable that assumes value 1 if there are unemployed individuals in the household and 0 if there are no unemployed individuals in the household.

The database was collected from the Brazilian National Household Sample Survey (PNAD) in the period from 2013 to 2015. This survey is annual, produced and published by the Brazilian Institute of Geography and Statistics (IBGE). It provides a wide set of demographic and socioeconomic information about the Brazilian population at the individual and household levels. We considered a sample of individuals aged between 18 and 65 years with complete information on earnings and hours worked, totalizing 167,271 observations. The sample refers to the 10 major metropolitan regions of the country, namely Belém-PA, Fortaleza-CE, Recife-PE, Salvador- BA, Belo Horizonte-MG, Rio de Janeiro-RJ, Curitiba-PR, Porto Alegre-RS, Brasília-DF, and São Paulo-SP. The nominal values of earnings were deflated by the National Consumer Price Index (INPC). There is no control for groups of individuals in each year, characterizing the data set a pooled cross-section. All results were obtained in the R statistical software [https://www.r-project.org/].

Table 1 provides some descriptive statistics for earnings and hours worked at level and logarithmic scales, including sample size, average (avg), median, standard deviation (SD), coefficient of variation (CV), asymmetry (CA), and kurtosis (CK). These statistics indicate that earnings in level has a high asymmetry and a significant kurtosis, suggesting that an asymmetric distribution with heavy tails is better to fit the data. On the other hand, hours worked in level show low asymmetry and moderate kurtosis. The application of the logarithm tends to produce symmetry, especially in the case of earnings. Figure 1 shows histograms of earnings and hours worked at level and logarithmic scales.

2.2 *Investigation of the best fit*

Initially, we estimate the Normal, t, and BS univariate regressions for earnings and hours worked, as well as their bivariate counterparts, to investigate the best fit to the data in each case. Table 2 reports the values of the Akaike (AIC) and Bayesian (BIC) information

Table 1. Descriptive statistics for wage and hours worked in level and logarithmic scales

Data	Sample Size	Median	Average	SD	CV	CA	CK
Wages	167,271	1344.54	2278.51	3087.26	135.49%	7.06	116.51
Hours worked	167,271	40.00	39.91	11.55	28.95%	−0.55	3.44
log(Wages)	167,271	7.20	7.20	0.80	10.89%	0.41	2.01
log(Hours worked)	167,271	3.69	3.61	0.51	14.59%	−4.13	21.87

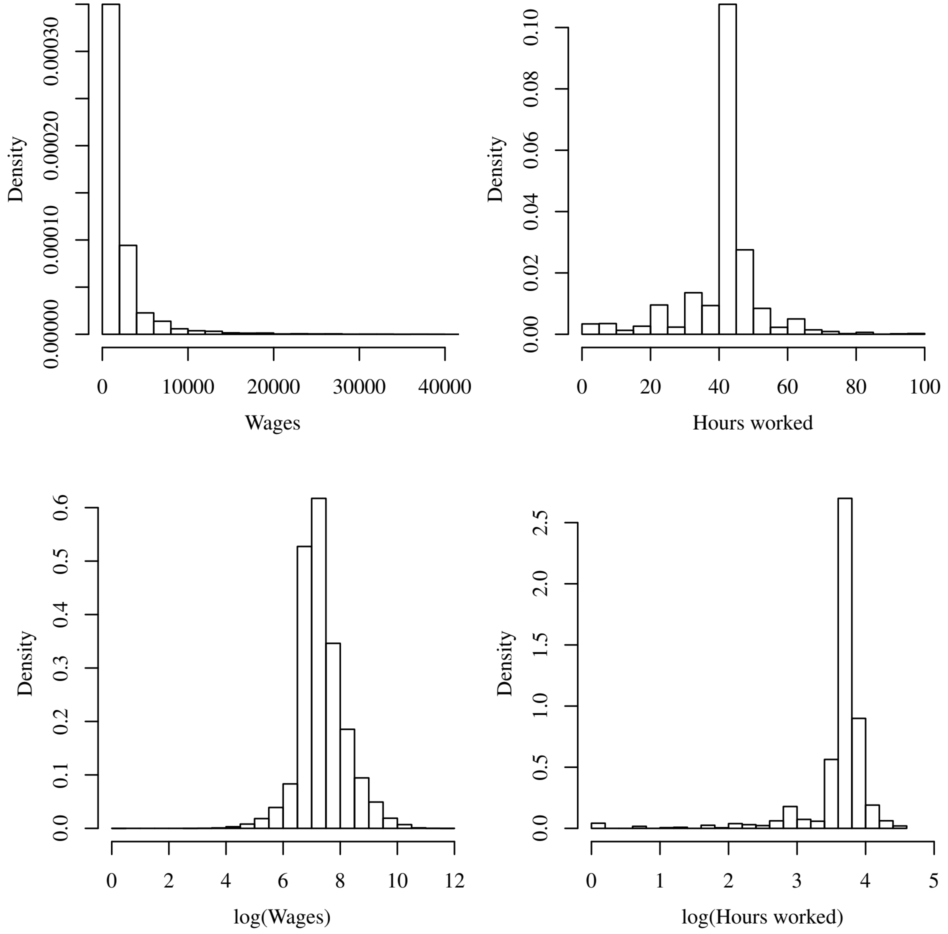

Figure 1. Histogram for earnings and hours worked (level and logarithmic scales).

criteria, calculated as:

$$\text{AIC} = -2\ell + 2k \quad \text{and} \quad \text{BIC} = -2\ell + k\log(n),$$

where ℓ is the value of the log-likelihood function, k denotes the number of parameters, and n indicates the number of observations. According to Table 2, the univariate and bivariate models based on the t distribution yielded the best adjustments, as they resulted the lowest values for both AIC and BIC. Thus, among the 3 distributions tested, the univariate and bivariate models of the t distribution shall be used according to the information criteria. Notice that the t distribution has heavier tails than the normal distribution, implying robustness against outlying observations [see Lucas (1997)].

Table 2. Information criteria for the univariate and bivariate models

		Univariate distributions		
		Normal	Student t	BS
AIC	Wages	272,072.00	255,565.00	2,753,495.00
	Hours worked	249,342.00	60,438.00	1,539,515.00
BIC	Wages	272,423.00	255,936.00	2,753,846.00
	Hours worked	249,693.00	60,809.00	1,539,866.00
		Bivariate distributions		
		Normal	Student t	BS
AIC	Wages and hours worked	515,257.00	451,026.00	4,281,716.00
BIC	Wages and hours worked	515,969.00	451,738.00	4,282,428.00

Once the best univariate and bivariate models were chosen, we applied the Mahalanobis distance to evaluate the quality of the fit to the data, as proposed by Marchant et al. (2016). In the case of the bivariate t distribution, this distance is given by:

$$D = \frac{1}{2}(\boldsymbol{U} - \boldsymbol{\mu})^{\top} \boldsymbol{\psi}^{-1}(\boldsymbol{U} - \boldsymbol{\mu}) \sim F_{2,\nu}, \qquad (5)$$

where $\boldsymbol{U} \sim t\mathrm{Biv}(\mu_1, \mu_2, \sigma_1, \sigma_2, \nu, \rho)$ according to equation (15) in the Appendix and $\boldsymbol{\psi}$ is the covariance matrix. According to (5), the Mahalanobis distance for the bivariate t distribution follows a $F_{2,\nu}$ distribution. That is, F distribution with 2 and ν degrees of freedom. In the univariate case, we have a $F_{1,\nu}$. In order to obtain the estimated values of the Mahalanobis distance, the parameters are replaced by their maximum likelihood estimates, which asymptotically results in the same distribution as (5) [Vilca et al. (2014)]. The Wilson-Hilferty approximation might then be applied to the Mahalanobis distance to obtain a standard Normal distribution approximation in (5). Thus, the quality of the fit of the univariate and bivariate t regression models might be evaluated by the transformed distances with the Wilson-Hilferty approximation [Ibacache-Pulgar et al. (2014)]. In this case, the distances in (5) are adapted to accommodate the regressive structure and the univariate or bivariate condition.

Figure 2 displays the probability-probability (PP) plots of the transformed Mahalanobis distance for the univariate t regressions of earnings and hours worked. The PP plot is commonly used to assess how close 2 sets of data are by plotting the 2 corresponding cumulative distribution functions. The closer the points are from the 45^o line in the (0.0) to (1.1) area, the best is the fit. Figure 2, shows the cumulative distribution function of the standard Normal versus the empirical cumulative distribution function of the transformed Mahalanobis distance. The results reveal an excellent fit of the univariate models.

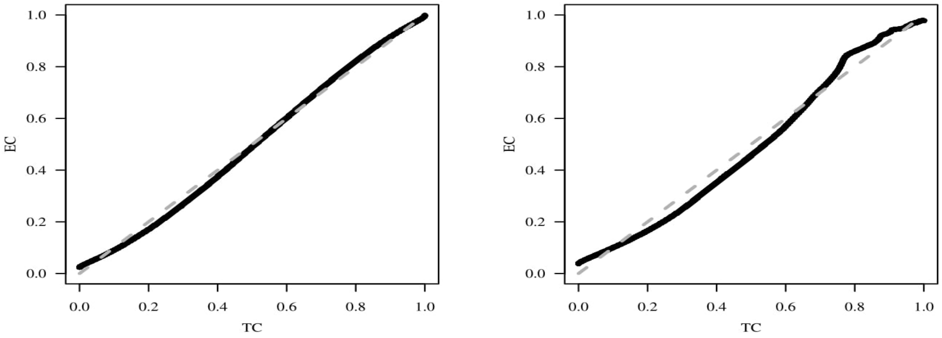

Figure 2. PP plots of the transformed distances for the univariate t regression models of earnings (left) and hours worked (right). Legend: EC = empirical probability, TC = theoretical probability.

Figure 3 shows the PP plot of the transformed Mahalanobis distance for the bivariate t regressions of earnings and hours worked. The results also suggest an excellent fit to the data for the bivariate case. Thus, for both univariate and bivariate cases, the t regression models provided excellent adjustments to the data and might therefore be used.

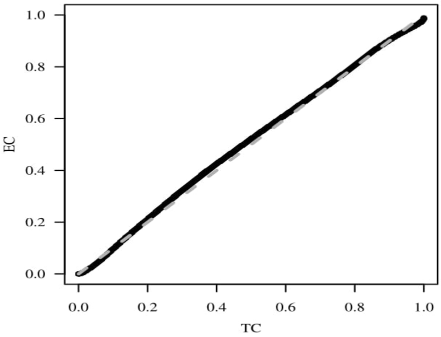

Figure 3. PP plots of the transformed distance for the bivariate t regression model of earnings and hours worked. Legend: EC = empirical probability, TC = theoretical probability.

2.3 *Estimations and analysis*

Table 3 reports the results of the maximum likelihood estimation for the bivariate t distribution regression model of earnings and hours worked, with the respective standard errors, Wald statistics, and p-values. The model based on the t distribution presented the best fit to the data according to the AIC and BIC information criteria and the PP plot of the Mahalanobis distance reported in the previous section. The Wald statistic is used to test the following hypotheses: $\mathrm{H}_0 : \theta = \theta_0$ versus $\mathrm{H}_1 : \theta \neq \theta_0$. The Wald statistic is defined by:

$$W = \frac{\widehat{\theta} - \theta_0}{\text{Standard Error}(\widehat{\theta})},$$

which is approximately distributed as a standard Normal under H_0, in which $\widehat{\theta}$ and θ_0 are the estimate and its proposed value under H_0, respectively. In this case, our interest lies in knowing if $\theta_0 = 0$ or $\mathrm{H}_0 : \theta = 0$ versus $\mathrm{H}_1 : \theta \neq 0$, at a significance level of $\alpha = 0.05\,$(or 5%).

Regarding to the interpretation of the estimated coefficients, the following cases deserves special attention:

- When the independent variable x is quantitative (for instance, number of years in school) and the value of the coefficient estimated is: (i) out of range: $-0.05 \leq \widehat{\beta} \leq 0.05$, there is an increase (or decrease if the estimate is negative) of $(\exp(\widehat{\beta}) - 1) \times 100\%$ in the expected value (mean) of the dependent variable due to an increase of 1 unit in x; (ii) within the range $-0.05 \leq \widehat{\beta} \leq 0.05$, there is an increase (or decrease if the estimate is negative) of $(\exp(\widehat{\beta}) - 1) \approx \widehat{\beta} \times 100\%$ in the expected value (mean) of the dependent variable when x increases by one unit.

- When the independent variable x is a dummy (for instance, gender) and the coefficient value is: (i) out of range $-0.05 \leq \widehat{\beta} \leq 0.05$, there is an increase (or decrease if the estimate is negative) of $(\exp(\widehat{\beta}) - 1) \times 100\%$ in the expected value (mean) of the dependent variable when x changes from 0 (women) to 1 (men); (ii) within the range $-0.05 \leq \widehat{\beta} \leq 0.05$, we have an increase (or decrease if the estimate is negative) of $(\exp(\widehat{\beta}) - 1) \approx \widehat{\beta} \times 100\%$ in the expected value (mean) of the dependent variable when x changes from 0 (women) to 1 (men).

Table 3 indicates that the estimated correlation of 0.1877 between earnings and hours worked is statistically significant at the 5% significance level. This means that the bivariate model is more appropriate than the univariate estimation of independent regressions, which might lead to biased results due to the untreated correlation between the two dependent variables.

Considering the estimated coefficients, the variable "Gender" indicates that men have an average income that is 34.58% higher than women and they supply 8.62% more hours worked, on average, than women. On the other hand, the variable "Race" reveals that the wages of Caucasian individuals are, on average, 10.31% higher than the wages of non-Caucasians. However, when it comes to hours worked, Caucasians only supply 0.02% more hours than non-Caucasians. This results confirm that there is discrimination in the Brazilian labor market. Cavalieri and Fernandes (1998), for instance, found wage discrimination using data from the PNAD of 1989. They also found higher wages for men than for women and for Caucasian individuals than for non-Caucasians, even after controlling for age, years in school, and geographical region of residence.

The "Age" variable indicates that one additional year of experience in the labor market increases wage by 5.27%, while hours worked raise only by 1.28%. Considering the variable "Education", an increase of one year of study raises in 4.68% the average wage. However, this same increase in schooling leads to an average decrease of 0.09% in hours worked. Thus, the higher the individual’s schooling, the higher his average wage and the lower his supply of hours in the labor market. This finding illustrates a fundamental advantage of the bivariate regression, since the effects of "Education" go in opposite directions in the bivariate regressions and this cannot be captured by the traditional univariate estimation that considers wage per hour as the unique dependent variable. Lau et al. (1993) also found a positive effect of "Education" on earnings (*per capita*) due to the higher level of schooling. Gonzaga et al. (2002) argued that, in Brazil, level of schooling is inversely related to hours worked.

Taking into account the metropolitan regions, Brasília-DF presents an average wage 8.95% higher and workers supply 2.70% less hours of work than São Paulo-SP. Again, it is also possible to identify distinct effects of an explanatory variable in the bivariate regression that cannot be captured by the traditional univariate model. In order to explain this finding, the unobservable characteristics of the workers, such as skill and motivation, as well as specific differences among the sectors of activity and the geographical regions of the country should be taken into account. In the specific case of Brasília, the differential is due to the location of the federal public administration in Brasília, which pays higher average wages than the private sector.

Regarding the types of labor contracts, the estimates point out that those with "no employment record card" have an average wage that is 17.73% lower than the wages of individuals "with employment record card". In addition, they offer about 15.08% less hours worked than their peers "with employment record card". The "civil servant" category incorporates, on average, an increase of 26.90% in wage while supplying an average of 4.10% less hours worked in relation to the workers "with employment record card". Meanwhile, those who are "autonomous" have an average wage 12.82% lower and supply 16.04% less hours to the labor market than the workers "with employment record card". It is worth mentioning that the the "civil servant" category also presents antagonistic effects on wages and hours worked that might be captured only by bivariate estimation.

For variables that affect only wages or hours worked separately, Table 3 illustrates that individuals who contribute to social security have an average wage 42.70% higher than those who do not contribute. The "head" variable, which affects only hours worked, confirms that the head of the household supplies 1.80% more hours to the labor market on average than those who are not in this condition.[6]

For comparison purposes, Table 4 presents the results of the traditional univariate regression in which the dependent variable is the wage rate (or wage per hour). In principle, some results show similarity in terms of magnitude with the estimates of the bivariate regression model. However, the univariate model cannot disentangle the effects

[6] As a sensitivity analysis, we also estimated the bivariate regression including only common explanatory variables. The results, reported in Appendix, indicate that there is no significant change in the previous findings.

of a given explanatory variable on wages and hours worked, as were the cases of "Education", "Brasília-DF", and "Civil Servant" discussed above. These variables displayed different signals in the estimated coefficients for the wage and hours worked regressions. For "Education", for instance, the higher the individual's level of schooling, the higher is his average wage and the lower is his supply of hours of work. However, in the univariate regression reported in Table 4, only the positive effect of an additional year of study on the average wage rate can be estimated. In addition, the bivariate model captured a positive and statically significant correlation between wage and hours worked, allowing for a more robust estimation than the simple adjustment of two independent regressions. Therefore, there are important advantages coming from the bivariate model, including the evidence that the determinants of wages and hours worked might not be the same in both quantitative and qualitative terms. In this environment, the bivariate estimation emerges as an important alternative for the estimation of the Mincerian earnings equation.

3. Conclusion

This paper proposed an alternative approach to estimate the Mincerian earnings equation based on bivariate regression modeling. The combination of wages and hours worked in a single dependent variable, as traditionally is done in the empirical literature, prevents capturing distinct effects of common covariates on those dependent variables separately. On the other hand, the univariate estimation of independent regressions for earnings and hours worked is not indicated due to the correlation between these variables, which might bias the estimates. We proposed the estimation of a regression for wages and hours worked as a bivariate vector of dependent variables, including common and specific covariates among the explanatory variables and using the Normal, Student t and BS bivariate distributions. The estimates used data at the individual level extracted from the Brazilian National Household Sample Survey (PNAD) for the years from 2013 to 2015.

In the bivariate case, the Normal, t and BS distributions were used to jointly model wages and hours worked. The AIC and BIC information criteria and the Mahalanobis distance indicated that the Student t distribution yielded the best fit to the data. In addition, a positive and statistically significant correlation between wages and hours worked justified the use of the bivariate regression in detriment of two separate regressions for those variables, which would yield in biased estimates.

The bivariate estimation indicated that a given common covariate might have distinct effects on wages and hours worked. The results for "Education", for instance, indicated that an additional year of study leads to an average increase of 4.68% in wages and an average decrease of 0.9% in hours worked. This suggests that individuals with more years of schooling, on average, have higher wages and work less hours than those with less years of schooling. Other covariates common to the bivariate vector, such as type of employment contract and geographic region of residence, also had antagonistic effects on earnings and hours worked. This evidence illustrates a fundamental advantage of the bivariate regression, which allows to disentangle the distinct effects of a given common covariate on wages and hours worked. This cannot be done by the traditional estimation of the univariate regression that considers the wage per hour as the dependent variable.

Thus, the bivariate regression might be considered as an alternative approach for the estimation of the Mincerian earnings equation. As further work, one might implement the Heckman two-step correction for selection bias (Heckman, 1979), since the PNAD survey refers to individuals who were actually working in the sample period. However, the individual's earnings are associated with the decision to supply work, which ultimately depends on their opportunity cost. It is advantageous to work if the wage (or potential earnings) is greater than the opportunity cost (reservation wage). In addition, other bivariate probability distributions might be adjusted to model wages and hours worked, such as Pareto and its extensions, which are commonly used in income modeling. Finally, a bivariate logistic regression model might be used to estimate the influence of individual characteristics on the probability of a given worker to belong to a particular income group and type of work. Some of these extensions are object of our ongoing research.

It is also worth mentioning that, in further research, the study might benefit by moving towards a structural approach, with a careful modeling strategy of the labor market and the resulting wage equation. Here, our focus was just on the application of an alternative bivariate approach to estimate the traditional Mincerian wage equation by using Brazilian micro data. In addition, due to well-known distortions in the Brazilian labor market, further extensions should consider an empirical analysis by sector of activity and type of occupation. We leave these issues for future research.

Table 3. Bivariate t distribution regression models for wages and hours worked ($\nu = 4$)

Response	Explanatory var.	Coefficient	Standard error	Wald stat.	p-value
	σ_1	0.55	0.0019	294.05	<0.0001
	σ_2	0.51	0.0018	292.05	<0.0001
	ρ	0.19	0.0131	14.53	<0.0001
Wage	(Intercept)	5.63	0.0194	290.41	<0.0001
	Gender	0.30	0.0031	96.25	<0.0001
	Age	0.05	8.0000×10^{-4}	66.78	<0.0001
	Age^2	-6.00×10^{-4}	8.0000×10^{-6}	−57.56	<0.0001
	Race	0.10	0.0031	31.85	<0.0001
	Marital status	0.01	0.0069	0.94	0.3498
	Education	0.05	5.0000×10^{-4}	97.66	<0.0001
	Belém-PA	−0.32	0.0063	−50.81	<0.0001
	Fortaleza-CE	−0.35	0.0059	−59.20	<0.0001
	Recife-PE	−0.35	0.0057	−61.28	<0.0001
	Salvador-BA	−0.31	0.0058	−53.88	<0.0001
	Belo Horizonte-MG	−0.09	0.0055	−17.04	<0.0001
	Rio de Janeiro-RJ	−0.08	0.0052	−15.34	<0.0001
	Curitiba-PR	0.01	0.0065	1.85	0.0643
	Porto Alegre-RS	−0.09	0.0051	−17.23	<0.0001
	Brasília-DF	0.09	0.0063	13.57	<0.0001
	2014	0.00	0.0033	0.72	0.4690
	2015	−0.06	0.0034	−17.51	<0.0001
	High	0.89	0.0084	105.53	<0.0001
	High mean	0.37	0.0073	50.53	<0.0001
	Mean	0.19	0.0069	26.85	<0.0001
	Low mean	0.05	0.0068	7.49	<0.0001
	Agricultural	−0.43	0.0180	−23.97	<0.0001
	Industry	−0.20	0.0087	−23.38	<0.0001
	Construction	−0.08	0.0093	−8.38	<0.0001
	Commerce, food and others	−0.21	0.0083	−25.46	<0.0001
	Education, health and soc. serv.	−0.25	0.0079	−32.17	<0.0001
	Other services	−0.21	0.0083	−25.86	<0.0001
	No employment record card	−0.20	0.0041	−47.21	<0.0001
	Autonomous	−0.14	0.0042	−32.68	<0.0001
	Civil servant	0.24	0.0072	33.29	<0.0001
	Labor Union	0.12	0.0038	32.46	<0.0001
	Social Security	0.36	0.0072	49.34	<0.0001
	Time in job	0.01	2.0000×10^{-4}	59.78	<0.0001
Hours worked	(Intercept)	3.28	0.0166	196.99	<0.0001
	Gender	0.08	0.0027	30.55	<0.0001
	Age	0.01	7.0000×10^{-4}	18.74	<0.0001
	Age^2	-1.00×10^{-4}	<0.0001	−15.15	<0.0001
	Race	2.00×10^{-4}	0.0027	0.09	0.9318
	Marital status	−0.01	0.0061	−1.51	0.1312
	Education	-9.00×10^{-4}	4.0000×10^{-4}	−2.23	0.0255
	Belém-PA	0.00	0.0055	0.66	0.5106
	Fortaleza-CE	0.02	0.0052	3.33	9.0000×10^{-4}
	Recife-PE	−0.01	0.0050	−2.99	0.0028
	Salvador-BA	−0.04	0.0051	−8.68	<0.0001
	Belo Horizonte-MG	0.01	0.0047	2.55	0.0109
	Rio de Janeiro-RJ	−0.09	0.0045	−19.59	<0.0001
	Curitiba-PA	−0.01	0.0057	−2.29	0.0222
	Porto Alegre-RS	0.02	0.0044	3.85	1.0000×10^{-4}
	Brasília-DF	−0.03	0.0054	−5.01	<0.0001
	2014	0.01	0.0029	4.14	<0.0001
	2015	−0.04	0.0029	−13.45	<0.0001
	High	0.12	0.0073	16.89	<0.0001
	High mean	0.10	0.0066	15.07	<0.0001
	Mean	0.15	0.0063	23.30	<0.0001
	Low mean	0.16	0.0061	25.27	<0.0001
	Agriculture	0.16	0.0152	10.32	<0.0001
	Industry	0.02	0.0072	2.82	0.0048
	Construction	0.04	0.0078	5.47	<0.0001
	Commerce, food and others	0.07	0.0069	9.49	<0.0001
	Education, health and soc. serv.	−0.07	0.0065	−10.07	<0.0001
	Other services	−0.01	0.0069	−2.14	0.0323
	No employment record card	−0.16	0.0036	−45.28	<0.0001
	Autonomous	−0.17	0.0034	−50.85	<0.0001
	Civil servant	−0.04	0.0060	−6.86	<0.0001
	Head	0.02	0.0027	6.72	<0.0001
	Minor	0.00	0.0026	−1.10	0.2709
	Inactivity	−0.01	0.0037	−1.50	0.1348

Table 4. Univariate regression for wage per hour

Response	Explanatory var.	Coefficient	Standard error	Wald stat.	p-value
Wage per hour	(Intercept)	0.943	0.0229	41.23	<0.0001
	Gender	0.203	0.0037	54.42	<0.0001
	Age	0.035	0.0009	36.99	<0.0001
	Age^2	0.000	0.0001	−30.31	<0.0001
	Race	0.101	0.0037	26.91	<0.0001
	Marital status	0.010	0.0084	1.16	0.2477
	Education	0.048	0.0006	85.13	<0.0001
	Belém-PA	−0.316	0.0076	−41.37	<0.0001
	Fortaleza-CE	−0.367	0.0072	−51.15	<0.0001
	Recife-PE	−0.332	0.0069	−47.83	<0.0001
	Salvador-BA	−0.267	0.0070	−38.33	<0.0001
	Belo Horizonte-MG	−0.105	0.0067	−15.73	<0.0001
	Rio de Janeiro-RJ	0.006	0.0062	0.97	0.3332
	Curitiba-PR	0.021	0.0080	2.57	0.0101
	Porto Alegre-RS	−0.110	0.0062	−17.58	<0.0001
	Brasília-DF	0.114	0.0075	15.23	<0.0001
	2014	−0.009	0.0040	−2.15	0.0312
	2015	−0.016	0.0041	−3.90	0.0001
	High	0.764	0.0097	78.59	<0.0001
	High mean	0.275	0.0088	31.24	<0.0001
	Mean	0.040	0.0083	4.79	<0.0001
	Low mean	−0.106	0.0081	−13.04	<0.0001
	Agriculture	−0.583	0.0200	−29.11	<0.0001
	Industry	−0.224	0.0101	−22.24	<0.0001
	Construction	−0.119	0.0109	−11.00	<0.0001
	Commerce, food and others	−0.278	0.0095	−29.25	<0.0001
	Education, health and soc. serv.	−0.188	0.0090	−20.87	<0.0001
	Other services	−0.200	0.0094	−21.24	<0.0001
	No employment record card	−0.034	0.0049	−7.00	<0.0001
	Autonomous	0.040	0.0047	8.51	<0.0001
	Civil servant	0.288	0.0083	34.59	<0.0001
	Labor union	0.120	0.0047	25.71	<0.0001
	Social security	0.332	0.0081	40.86	<0.0001
	Time in job	0.010	0.0002	42.06	<0.0001
	Head	0.064	0.0036	17.92	<0.0001
	Minor	0.012	0.0036	3.35	0.0008
	Inactivity	−0.062	0.0052	−11.93	<0.0001

Appendix A: Distributions and bivariate regression models

In the symmetric context, the bivariate Normal distribution has been intensely used in the literature [Johnson et al. (1995)]. However, an alternative symmetric to the bivariate Normal distribution is the Student t model, as in Johnson et al. (1995) and Balakrishnan and Lai (2009), which has heavier tails than the Normal bivariate distribution. This flexibility is important to accommodate observations with more outliers, which makes the t an alternative of interest. On the other hand, in the univariate asymmetric context, a distribution that has received considerable attention is the BS model, which was introduced by Birnbaum and Saunders (1969) whereby its genesis is motivated by material fatigue problems [Johnson et al. (1995)]. Recently, Saulo et al. (2020; 2021) proposed a bivariate BS distribution and its corresponding regression model, based on the univariate BS distribution reparameterized by the mean proposed by Santos-Neto et al. (2012). In this reparameterization, there is no need to transform the dependent variable to a logarithmic scale, which is an advantage since it can lead to difficulties in interpretation. In general terms, Normal, Student t, and BS bivariate distributions can be considered as good candidates in the context of modeling earnings and hours worked, since in the the Normal and t cases the logarithm of the data is used, i.e. the log-normal and log-t are considered for the level variables [Vanegas and Paula (2016)], and in the BS case, the data (asymmetric on the right) are used in the original scale. The Normal, Student t, and BS bivariate distributions and their respective regression models are presented in sequence.

Bivariate Normal distribution

Let $\boldsymbol{Y} = (Y_1,Y_2)^\top$ be a bivariate random vector following a bivariate normal distribution with means μ_1 e μ_2, and standard deviations σ_1 e σ_2. In addition to these 4 parameters there is a correlation coefficient ρ between Y_1 and Y_2 defined by $-1 < \rho < 1$. Therefore, denoting $\boldsymbol{Y} \sim \mathrm{NBiv}(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$. The joint probability density function (PDF)of Y_1 and Y_2 can be written as:

$$
\begin{aligned}
&f(y_1,y_2;\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)\\
&\quad = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(y_1-\mu_1)^2}{\sigma_1^2}+\frac{(y_2-\mu_2)^2}{\sigma_2^2}-\frac{2\rho(y_1-\mu_1)(y_2-\mu_2)}{\sigma_1\sigma_2}\right]\right\}.
\end{aligned} \tag{6}
$$

The joint PDF of the random vector $\boldsymbol{Z} = (Z_1,Z_2)^\top$ following a bivariate standard Normal distribution (means zero and variances one) is given by:

$$
\phi_2(z_1,z_2;\rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(y_1^2+y_2^2-2\rho y_1 y_2)\right\}. \tag{7}
$$

The corresponding joint cumulative distribution function (CDF) associated with (6) is given by:

$$
\Phi_2(z_1,z_2;\rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{z_1}\int_{-\infty}^{z_2} \exp\left\{-\frac{1}{2(1-\rho^2)}(u^2+v^2-2\rho uv)\right\}. \tag{8}
$$

When $\rho = 0$, i.e., when the Normal variables are uncorrelated, (6) can be expressed as the product of 2 Normal CDFs.

Normal bivariate regression model

Consider $\boldsymbol{Y}_1,\ldots,\boldsymbol{Y}_n$ such that $\boldsymbol{Y}_i = (Y_{1i},Y_{2i})^\top$ follows a Normal bivariate model, i.e., $\boldsymbol{Y}_i \sim \mathrm{NBiv}(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$. Consider that there are r and s covariates, let´s say $\boldsymbol{x}_i^{(1)} = (x_{i1}^{(1)},x_{i2}^{(1)},\ldots,x_{ir}^{(1)})^\top$ and $\boldsymbol{x}_i^{(2)} = (x_{i1}^{(2)},x_{i2}^{(2)},\ldots,x_{is}^{(2)})^\top$, associated with Y_{1i} and Y_{2i}, respectively, such that

$$\mathrm{E}[Y_{1i}\,|\,\boldsymbol{X}_i^{(1)} = \boldsymbol{x}_i^{(1)}] = \mu_{1i} = \boldsymbol{x}_i^{(1)\top}\boldsymbol{\beta}_1, \quad i = 1,\ldots,n, \tag{9}$$

$$\mathrm{E}[Y_{2i}\,|\,\boldsymbol{X}_i^{(2)} = \boldsymbol{x}_i^{(2)}] = \mu_{2i} = \boldsymbol{x}_i^{(2)\top}\boldsymbol{\beta}_2, \quad i = 1,\ldots,n, \tag{10}$$

with

$$\boldsymbol{x}_i^{(j)\top}\boldsymbol{\beta}_j = \beta_{j1}x_{i1}^{(j)} + \beta_{j2}x_{i2}^{(j)} + \cdots + \beta_{jl}x_{il}^{(j)}, \quad j = 1,2, \quad i = 1,\ldots,n, \tag{11}$$

where $\boldsymbol{\beta}_k = (\beta_{k1},\beta_{k2},\ldots,\beta_{kl})$ is a vector of l unknown parameters, and $\boldsymbol{x}_i^{(k)}$ is the i-th line of matrix $\boldsymbol{X}^{(k)}$, whose dimension is $n \times l$, for $k = 1,2$ and $l = r,s$. Thus, we have the following Normal bivariate model:

$$\begin{pmatrix} Y_{1i} \\ Y_{2i} \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_i^{(1)\top}\boldsymbol{\beta}_1 \\ \boldsymbol{x}_i^{(2)\top}\boldsymbol{\beta}_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \quad i = 1,\ldots,n, \tag{12}$$

where $(\varepsilon_{1i},\varepsilon_{2i}) \sim \mathrm{NBiv}(0,0,\sigma_1,\sigma_2,\rho)$, and they are independently distributed.

To estimate the unknown parameters σ_1, σ_2, $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and ρ based on a random sample of size n, i.e., $\{(y_{1i},y_{2i},\boldsymbol{x}_i^{(1)},\boldsymbol{x}_i^{(2)}); i = 1,\ldots,n\}$, the maximum likelihood method is used. The likelihood and log-likelihood functions of the observed sample can be written respectively as

$$L = \prod_{i=1}^{n} f(y_{1i},y_{2i};\mu_{1i},\mu_{2i},\sigma_1,\sigma_2,\rho), \tag{13}$$

$$\ell = \sum_{i=1}^{n} \log(f(y_{1i},y_{2i};\mu_{1i},\mu_{2i},\sigma_1,\sigma_2,\rho)), \tag{14}$$

where f is the joint PDF of the bivariate normal distribution. The model parameter estimates must be obtained by maximizing the log-likelihood function (14). This is done by solving a nonlinear iterative optimization process, particularly the quasi-Newton Broyden-Fletcher-Goldfarb-Shanno (BFGS) method can be used. The BFGS method is implemented in R software (`http://cran.r-project.org`), using the `optim` and `optimx` functions.

Bivariate t distribution

Let $\boldsymbol{U} = (U_1,U_2)^\top$ be a bivariate random vector following a bivariate t distribution with location parameters μ_1 and μ_2, scale parameters σ_1 e σ_2, degrees of freedom ν, and correlation coefficient $-1 < \rho < 1$ between U_1 and U_2, denoted by $\boldsymbol{U} \sim t\mathrm{Biv}(\mu_1,\mu_2,\sigma_1,\sigma_2,\nu,\rho)$. Therefore, the joint PDF of U_1 and U_2 is given by:

$$\begin{aligned} &f(u_1,u_2;\mu_1,\mu_2,\sigma_1,\sigma_2,\nu,\rho) \\ &= \tfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \left[1+\tfrac{1}{\nu(1-\rho^2)}\left(\tfrac{(u_1-\mu_1)^2}{\sigma_1^2}+\tfrac{(u_2-\mu_2)^2}{\sigma_2^2}-\tfrac{2\rho(u_1-\mu_1)(u_2-\mu_2)}{\sigma_1\sigma_2}\right)\right]^{-(\nu+2)/2}. \end{aligned} \tag{15}$$

The joint PDF of the random vector $\boldsymbol{U} = (U_1,U_2)^\top$ following a standard bivariate t distribution is given by:

$$f(u_1,u_2,\nu,\rho) = \frac{1}{2\pi\sqrt{1-\rho^2}}\left[1+\frac{1}{\nu(1-\rho^2)}\left(u_1^2+u_2^2-2\rho u_1u_2\right)\right]^{-(\nu+2)/2}.$$

The corresponding joint CDF associated to (16) is:

$$F(u_1,u_2;\nu,\rho) = \frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{u_1}\int_{-\infty}^{u_2}\left[1+\frac{1}{\nu(1-\rho^2)}\left(u^2+v^2-2\rho uv\right)\right]^{-(\nu+2)/2}. \tag{16}$$

Bivariate t regression model

Consider $\boldsymbol{U}_1,\ldots,\boldsymbol{U}_n$ such that $\boldsymbol{U}_i = (U_{1i},U_{2i})^\top$ follows a bivariate t distribution, i.e., $\boldsymbol{U}_i \sim t\mathrm{Biv}(\mu_1,\mu_2,\sigma_1,\sigma_2,\nu,\rho)$ with PDF (15). Assume that there are r and s covariates, $\boldsymbol{x}_i^{(1)} = (x_{i1}^{(1)},x_{i2}^{(1)},\ldots,x_{ir}^{(1)})^\top$ and $\boldsymbol{x}_i^{(2)} = (x_{i1}^{(2)},x_{i2}^{(2)},\ldots,x_{is}^{(2)})^\top$ say, associated with U_{1i} and U_{2i}, respectively, such that

$$\mathrm{E}[U_{1i}\,|\,\boldsymbol{X}_i^{(1)} = \boldsymbol{x}_i^{(1)}] = \mu_{1i} = \boldsymbol{x}_i^{(1)\top}\boldsymbol{\beta}_1, i=1,\ldots,n, \tag{17}$$

$$\mathrm{E}[U_{2i}\,|\,\boldsymbol{X}_i^{(2)} = \boldsymbol{x}_i^{(2)}] = \mu_{2i} = \boldsymbol{x}_i^{(2)\top}\boldsymbol{\beta}_2, i=1,\ldots,n, \tag{18}$$

with $\boldsymbol{x}_i^{(j)\top}\boldsymbol{\beta}_j$ as in (11), where $\boldsymbol{\beta}_k = (\beta_{k1},\beta_{k2},\ldots,\beta_{kl})$ is a vector of l unknown parameters, and $\boldsymbol{x}_i^{(k)}$ is the i-th line of matrix $\boldsymbol{X}^{(k)}$, whose dimension is $n\times l$, for $k=1,2$ and $l=r,s$. Therefore, we have the following bivariate t regression model

$$\begin{pmatrix} U_{1i} \\ U_{2i} \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_i^{(1)\top}\boldsymbol{\beta}_1 \\ \boldsymbol{x}_i^{(2)\top}\boldsymbol{\beta}_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \quad i=1,\ldots,n, \tag{19}$$

where $(\varepsilon_{1i},\varepsilon_{2i}) \sim t\mathrm{Biv}(0,0,\sigma_1,\sigma_2,\nu,\rho)$ are independently distributed.

The parameters of the model are estimated as the bivariate normal, that is, given as likelihood and log-likelihood function,

$$L = \prod_{i=1}^{n} f(u_{1i},u_{2i};\mu_{1i},\mu_{2i},\sigma_1,\sigma_2,\nu,\rho), \tag{20}$$

$$\ell = \sum_{i=1}^{n}\log(f(u_{1i},u_{2i};\mu_{1i},\mu_{2i},\sigma_1,\sigma_2,\nu,\rho)), \tag{21}$$

respectively, where f is the joint PDF of the bivariate t distribution, the model parameter estimates, σ_1, σ_2, $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ e ρ, are obtained by maximizing the log-likelihood function by solving an iterative nonlinear optimization process, particularly the quasi-Newton BFGS method. The parameter ν is estimated according to Barros et al. (2008). The profiled log-likelihood and the following steps are used:

1) Let $\nu_k = k$ be for each k, $k = 1,\ldots,20$, compute the parameter estimates of $(\boldsymbol{\beta}_1^\top,\boldsymbol{\beta}_2^\top,\sigma_1,\sigma_2,\nu,\rho)^\top$ using the maximum likelihood method. Moreover, compute the log-likelihood function;

2) The final estimate of ν is the one which maximizes the log-likelihood function and the associated estimates of $(\boldsymbol{\beta}_1^\top,\boldsymbol{\beta}_2^\top,\sigma_1,\sigma_2,\nu,\rho)^\top$ are the final ones.

Bivariate Birnbaum-Saunders (BS) distribution

Let $\boldsymbol{T} = (T_1,T_2)^\top$ be a bivariate random vector following a bivariate BS distribution parameterized by means with parameters μ_1, μ_2, δ_1, δ_2 e ρ. Therefore, the joint CDF of T_1 and T_2 can be written, for $t_1,t_2 > 0$, as (Saulo et al., 2020)

$$\begin{aligned} &F(t_1,t_2;\mu_1,\mu_2,\delta_1,\delta_2,\rho)\\ &\quad= \Phi_2\left(\sqrt{\frac{\delta_1}{2}}\left[\sqrt{\frac{(\delta_1+1)t_1}{\delta_1\mu_1}} - \sqrt{\frac{\delta_1\mu_1}{(\delta_1+1)t_1}}\right],\ \sqrt{\frac{\delta_2}{2}}\left[\sqrt{\frac{(\delta_2+1)t_2}{\delta_2\mu_2}} - \sqrt{\frac{\delta_2\mu_2}{(\delta_2+1)t_2}}\right];\rho\right), \end{aligned} \tag{22}$$

where $\mu_1 > 0$, $\delta_1 > 0$, $\mu_2 > 0$, $\delta_2 > 0$, $-1 < \rho < 1$, Φ_2 is the CDF of the standard bivariate distribution given in (8). Therefore, the joint PDF associated with (22) is given by

$$\begin{aligned} &f(t_1,t_2;\mu_1,\mu_2,\delta_1,\delta_2,\rho)\\ &\quad= \phi_2\left(\sqrt{\frac{\delta_1}{2}}\left[\sqrt{\frac{(\delta_1+1)t_1}{\delta_1\mu_1}} - \sqrt{\frac{\delta_1\mu_1}{(\delta_1+1)t_1}}\right],\ \sqrt{\frac{\delta_2}{2}}\left[\sqrt{\frac{(\delta_2+1)t_2}{\delta_2\mu_2}} - \sqrt{\frac{\delta_2\mu_2}{(\delta_2+1)t_2}}\right];\rho\right)\\ &\qquad\times \frac{\sqrt{\delta_1}}{2\sqrt{2}t_1}\left[\sqrt{\frac{(\delta_1+1)t_1}{\delta_1\mu_1}} + \sqrt{\frac{\delta_1\mu_1}{(\delta_1+1)t_1}}\right]\frac{\sqrt{\delta_2}}{2\sqrt{2}t_2}\left[\sqrt{\frac{(\delta_2+1)t_2}{\delta_2\mu_2}} + \sqrt{\frac{\delta_2\mu_2}{(\delta_2+1)t_2}}\right], \end{aligned} \tag{23}$$

where ϕ_2 is the PDF of a normal bivariate distribution given by (7). The bivariate BS distribution with PDF (23) is denoted by $\mathrm{BSBiv}(\mu_1,\mu_2,\delta_1,\delta_2,\rho)$.

Bivariate Birnbaum-Saunders (BS) regression model

Consider $\boldsymbol{T}_1,\ldots,\boldsymbol{T}_n$ such that $\boldsymbol{T}_i = (T_{1i},T_{2i})^\top$ follows a bivariate BS distribution, that is, $\boldsymbol{T}_i \sim \mathrm{BSBiv}(\mu_1,\mu_2,\delta_1,\delta_2,\rho)$. Assume that there is r and s covariates, $\boldsymbol{x}_i^{(1)} = (x_{i1}^{(1)},x_{i2}^{(1)},\ldots,x_{ir}^{(1)})^\top$ and $\boldsymbol{x}_i^{(2)} = (x_{i1}^{(2)},x_{i2}^{(2)},\ldots,x_{is}^{(2)})^\top$, associated with T_{1i} and T_{2i}, respectively. Therefore, from (22), we have, for $t_1,t_2 > 0$, and $i = 1,\ldots,n$, the joint CDF (Saulo et al., 2020)

$$\begin{aligned} &F(t_{1i},t_{2i};\mu_{1i},\mu_{2i},\delta_1,\delta_2,\rho)\\ &\quad= \Phi_2\left(\sqrt{\frac{\delta_1}{2}}\left[\sqrt{\frac{(\delta_1+1)t_{1i}}{\delta_1\mu_1^i}} - \sqrt{\frac{\delta_1\mu_1^i}{(\delta_1+1)t_{1i}}}\right],\sqrt{\frac{\delta_2}{2}}\left[\sqrt{\frac{(\delta_2+1)t_{2i}}{\delta_2\mu_2^i}} - \sqrt{\frac{\delta_2\mu_2^i}{(\delta_2+1)t_{2i}}}\right];\rho\right), \end{aligned} \tag{24}$$

where

$$\mathrm{E}[T_{1i}\,|\,\boldsymbol{X}_i^{(1)} = \boldsymbol{x}_i^{(1)}] = \mu_{1i} = \exp(\boldsymbol{x}_i^{(1)\top}\boldsymbol{\beta}_1), i = 1,\ldots,n, \tag{25}$$

$$\mathrm{E}[T_{2i}\,|\,\boldsymbol{X}_i^{(2)} = \boldsymbol{x}_i^{(2)}] = \mu_{2i} = \exp(\boldsymbol{x}_i^{(2)\top}\boldsymbol{\beta}_2), i = 1,\ldots,n, \tag{26}$$

with $\exp(\boldsymbol{x}_i^{(j)\top}\boldsymbol{\beta}_j)$ as in (11), where $\boldsymbol{\beta}_k = (\beta_{k1},\beta_{k2},\ldots,\beta_{kl})$ is a vector of unknown l parameters, and $\boldsymbol{x}_i^{(k)}$ is the i-th line of matrix $\boldsymbol{X}^{(k)}$, whose dimension is $n\times l$, for $k = 1,2$ and $l = r,s$. Here, differently from the BS regression model based on the classical parameterization Rieck and Nedelman (1991), there is no need for logarithmic transformation, that is, the data for the response are worked on in their original scale.

In order to estimate the parameters, as in the normal bivariate case, the maximum likelihood method is used. Consider a random sample of size n, $\{(t_{1i},t_{2i},\boldsymbol{x}_i^{(1)},\boldsymbol{x}_i^{(2)}); i = 1,\ldots,n\}$ say, therefore the likelihood and log-likelihood functions of the observed sample are given respectively by

$$L = \prod_{i=1}^{n} f(t_{1i},t_{2i};\mu_{1i},\mu_{2i},\delta_1,\delta_2,\rho), \tag{27}$$

$$\ell = \sum_{i=1}^{n} \log(f(t_{1i},t_{2i};\mu_{1i},\mu_{2i},\delta_1,\delta_2,\rho)), \tag{28}$$

where f is a joint PDF of the bivariate BS distribution. The parameter estimates $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, δ_1, δ_2 and ρ are obtained by maximizing the log-likelihood function (28) using an iterative non-linear optimization process, in this case, the BFGS quasi-Newton method.

Bivariate regression of wages and hours worked considering only common explanatory variables

Table A.1. Bivariate t distribution regression models for earnings and hours worked: including only common covariates ($\nu = 4$).

Dependent variable	Explanatory var.	Coefficient	Standard error	Wald stat.	p-value
	σ_1	0.560	0.0018	305.35	<0.0001
	σ_2	0.510	0.0018	305.35	<0.0001
	ρ	0.192	0.0118	16.22	<0.0001
Wage	(Intercept)	5.541	0.0198	280.01	<0.0001
	Gender	0.319	0.0031	101.35	<0.0001
	Age	0.055	8.0000×10^{-4}	69.95	<0.0001
	Age^2	-5.000×10^{-4}	0.0000	−54.89	<0.0001
	Race	0.103	0.0032	32.60	<0.0001
	Marital status	0.003	0.0071	0.42	0.6770
	Education	0.050	5.0000×10^{-4}	100.83	<0.0001
	Belém-PA	−0.319	0.0064	−49.56	<0.0001
	Fortaleza-CE	−0.347	0.0061	−57.23	<0.0001
	Recife-PE	−0.354	0.0059	−60.45	<0.0001
	Salvador-BA	−0.308	0.0059	−51.84	<0.0001
	Belo Horizonte-MG	−0.093	0.0056	−16.57	<0.0001
	Rio de Janeiro-RJ	−0.086	0.0053	−16.29	<0.0001
	Curitiba-PR	0.013	0.0067	1.87	0.0619
	Porto Alegre-RS	−0.086	0.0053	−16.22	<0.0001
	Brasília-DF	0.096	0.0065	14.80	<0.0001
	2014	0.004	0.0034	1.16	0.2475
	2015	−0.060	0.0034	−17.31	<0.0001
	High	0.945	0.0086	110.33	<0.0001
	High men	0.395	0.0075	52.70	<0.0001
	Mean	0.186	0.0071	26.22	<0.0001
	Low mean	0.035	0.0069	5.03	<0.0001
	Agriculture	−0.366	0.0181	−20.21	<0.0001
	Industry	−0.200	0.0090	−22.27	<0.0001
	Construction	−0.085	0.0096	−8.90	<0.0001
	Commerce, food and others	−0.229	0.0085	−26.86	<0.0001
	Education, health and soc. serv.	−0.265	0.0081	−32.53	<0.0001
	Other services	−0.225	0.0085	−26.31	<0.0001
	No employment record card	−0.227	0.0042	−53.91	<0.0001
	Autonomous	−0.122	0.0042	−29.19	<0.0001
	Civil servant	0.313	0.0073	42.91	<0.0001
Hours worked	(Intercept)	3.264	0.0164	198.60	<0.0001
	Gender	0.086	0.0026	32.60	<0.0001
	Age	0.014	7.0000×10^{-4}	20.35	<0.0001
	Age^2	-1.000×10^{-4}	0.0000	−16.19	<0.0001
	Race	-2.000×10^{-4}	0.0027	−0.06	0.9543
	Marital status	−0.008	0.0061	−1.34	0.1809
	Education	-9.000×10^{-4}	4.0000×10^{-4}	−2.22	0.0261
	Belém-PA	0.003	0.0055	0.46	0.6471
	Fortaleza-CE	0.017	0.0052	3.31	0.0009
	Recife-PE	−0.015	0.0050	−3.07	0.0021
	Salvador-BA	−0.044	0.0051	−8.62	<0.0001
	Belo Horizonte-MG	0.012	0.0047	2.54	0.0110
	Rio de Janeiro-RJ	−0.087	0.0045	−19.52	<0.0001
	Curitiba-PR	−0.012	0.0057	−2.18	0.0295
	Porto Alegre-RS	0.018	0.0045	3.99	0.0001
	Brasília-DF	−0.027	0.0054	−4.97	<0.0001
	2014	0.012	0.0029	4.10	<0.0001
	2015	−0.040	0.0029	−13.62	<0.0001
	High	0.123	0.0073	16.93	<0.0001
	High mean	0.099	0.0066	15.00	<0.0001
	Mean	0.146	0.0063	23.25	<0.0001
	Low mean	0.155	0.0061	25.22	<0.0001
	Agriculture	0.157	0.0151	10.37	<0.0001
	Industry	0.020	0.0073	2.75	0.0060
	Construction	0.042	0.0078	5.35	<0.0001
	Commerce, food and others	0.065	0.0069	9.38	<0.0001
	Education, health and soc. serv.	−0.067	0.0065	−10.19	<0.0001
	Other services	−0.015	0.0069	−2.21	0.0273
	No employment record card	−0.164	0.0036	−45.30	<0.0001
	Autonomous	−0.174	0.0034	−50.64	<0.0001
	Civil servant	−0.041	0.0060	−6.77	<0.0001

Bibliography

Aali-Bujari, A., F. Venegas-Martínez, and A. García-Santillán (2019): "Schooling levels and wage gains in Mexico," *Economics and Sociology*, 12, 74–83. [1]

Balakrishnan, N. and C-D. Lai (2009): *Continuous Bivariate Distributions*, New York: Springer. [3, 16]

Barros, M., G.A. Paula, and V. Leiva (2008): "A new class of survival regression models with heavy-tailed errors: robustness and diagnostics," *Lifetime Data Analysis*, 14, 316–332. [19]

Becker, G.S. (1993): *Human capital a theoretical and empirical analysis, with special reference to education*, New York: NBER. [2]

Birnbaum, Z.W. and S.C. Saunders (1969): "A new family of life distributions," *Journal of Applied Probability*, 6, 319–327. [16]

Buchinsky, M. (2001): "Quantile regression with sample selection: estimating women's return to education in the U.S," *Empirical Economics*, 26, 87–113. [2]

Card, David (1999): "The causal effect of education on earnings," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, Elsevier, vol. 3, Part A, chap. 30, 1801–1863, 1 ed. [1]

Cavalieri, C. H. and R. Fernandes (1998): "Diferenciais por gênero e cor: uma comparação entre as regiões metropolitanas brasileiras," *Revista de Economia Política*, 18. [10]

Chatterjee, S. and B. Price (1991): *Regression Analysis by Example*, New York: John Wiley. [2]

Giuberti, A. C. and N. Menezes-Filho (2005): "Discriminação de rendimentos por gênero: uma comparação entre o Brasil e os Estados Unidos," *Economia Aplicada*, 9, 369–384. [2]

Gonzaga, G., P.G.G.P. Leite, and D.C. Machado (2002): "Quem trabalha muito e quem trabalha pouco no Brasil?" in *Anais do XIII Encontro Nacional De Estudos Populacionais*, ABEP. [11]

Heckman, J. (1976): "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models," *Annals of Economic and Social Measurement*, 5, 475–492. [3]

——— (1979): "Sample selection bias as a specification error," *Econometrica*, 47, 153–161. [2, 13]

Heckman, J., J.L. Tobias, and E. Vytlacil (2000): "Simple estimators for treatment parameters in a latent variable framework with an application to estimating the returns to shooling," *NBER Working Paper*, 7950. [2]

Ibacache-Pulgar, G., G. Paula, and M. Galea (2014): "On influence diagnostics in elliptical multivariate regression models with equicorrelated random errors," *Statistical Modelling*, 16, 14–21. [8]

Jannuzzi, P. de M. (2001): "Status socioeconômico das ocupações brasileiras: medidas aproximativas para 1980, 1991 e anos 90," *Revista Brasiliera de Estatística*, 2, 47–74. [5]

Johnson, N.L., S. Kotz, and N. Balakrishnan (1995): *Continuous Univariate Distributions*, vol. 2, New York, US: Wiley. [3, 16]

Lau, L.J., D.T. Jamison, S. Liu, and S. Riukin (1993): "Education and economic growth: Some cross-sectional evidence from Brazil," *Journal of Development Economics*, 41, 45–70. [11]

Lucas, A. (1997): "Robustness of the student t based M-estimator," *Communications in Statistics: Theory and Methods*, 41, 1165–1182. [8]

Maciel, M.C., A.C. Campêlo, and M.C.F. Raposo (2001): "A dinâmica das mudanças na distribuição salarial e no retorno em educação para mulheres: Uma aplicação de regressão quantílica," in *Anais do XXIX Encontro Nacional de Economia*, Salvador: ANPEC. [2]

Madalozzo, R. (2010): "Occupational segregation and the gender wage gap in Brazil: an empirical analysis," *Economia Aplicada*, 14, 147–168. [2]

Marchant, C., V. Leiva, and F. J. A. Cysneiros (2016): "A multivariate log-linear model for Birnbaum-Saunders distributions," *IEEE Transactions on Reliability*, 65, 816–827. [3, 8]

Mincer, J. (1958): "Investment in human capital and personal income distribution," *Journal of Political Economy*, 66, 281–302. [2]

——— (1974): *Schooling, Experience, and Earnings*, National Bureau of Economic Research, Inc. [1, 2]

Mittelhammer, R.C., G.G. Judge, and D.J. Miller (2000): *Econometric Foundations*, New York, US: Cambridge University Press. [3]

Resende, Marcelo and Ricardo Wyllie (2006): "Retornos para educação no Brasil: evidências empíricas adicionais ," *Economia Aplicada*, 10, 349–365. [1]

Rieck, J.R. and J.R. Nedelman (1991): "A log-linear model for the Birnbaum-Saunders distribution," *Technometrics*, 3, 51–60. [20]

Santos-Neto, M., F.J.A. Cysneiros, V. Leiva, and S.E. Ahmed (2012): "On new parameterizations of the Birnbaum-Saunders distribution," *Pakistan Journal of Statistics*, 28, 1–26. [3, 16]

Saulo, H., J. Leão, R. Vila, V. Leiva, and V. Tomazella (2020): "On mean-based bivariate Birnbaum-Saunders distributions: Properties, inference and application," *Communications in Statistics – Theory and Methods*, 49, 6032–6056. [3, 5, 16, 19]

——— (2021): "A bivariate fatigue-life regression model and its application to fracture of metallic tools," *Brazilian Journal of Probability and Statistics*, 35, 119–137. [3, 5, 16]

Sedlacek, G. and E. Santos (1991): “A mulher cônjuge no mercado de trabalho como estratégia de geração da renda familiar,” *Pesquisa e Planejamento Econômico*, 21, 449–470. [2]

Senna, J.C. (1976): “Escolaridade, experiência no trabalho e salários no Brasil,” *Revista Brasileira de Economia*, 30, 163–193. [1]

Vanegas, L. H. and G. A. Paula (2016): “Log-symmetric distributions: statistical properties and parameter estimation,” *Brazilian Journal of Probability and Statistics*, 30, 196–220. [4, 16]

Vilca, F., N. Balakrishnan, and C.B. Zeller (2014): “The bivariate sinh-elliptical distribution with applications to Birnbaum-Saunders distribution and associated regression and measurement error models,” *Computational Statistics and Data Analysis*, 80, 1–16. [8]