Concentration and Intensity of Knowledge in Sectors and their Effects on the Returns of Education in Brazil

Concentração e intensidade de conhecimento nos setores e seus efeitos nos retornos da educação no brasil

Michele Aparecida Nepomuceno Pinto[1]
Joilson Giorno[2]

Resumo: O objetivo deste trabalho é triplo. Primeiro, estimar os retornos para a educação a nível setorial a fim de verificar a hipótese desses retornos serem crescentes. Segundo, usar o estoque de conhecimento superior presente nos setores de serviços de alta tecnologia para testar a hipótese do efeito transbordamento. Terceiro, verificar se existe o efeito *threshold* educacional em nível setorial. A primeira hipótese está ligada aos trabalhos de Acemoglu (1996), Lucas (1988) e Romer (1990). A segunda é derivada do trabalho de Dias *et al.* (2013). A terceira é proposta por Dias e McDemortt (2012) e Dias (2017), que afirmam que os benefícios dos retornos crescentes ocorrem somente após um determinado nível educacional na América Latina. Os dados utilizados são da PNAD de 2006 e 2015, o que permite comparar os efeitos no tempo. As estimativas mostraram que os setores apresentam retornos crescentes e o efeito espraiamento está presente em quase todos eles. A hipótese de que existe um *threshold* educacional foi confirmada e o valor é em torno de 9,0 anos de escolaridade, em média, nos setores. Os benefícios de ganhos de produtividade crescente advindos da acumulação de educação nos setores ainda estão por ocorrer na maioria dos setores analisados.

Palavras-chave: Capital humano. Diferenciais de rendimentos. Retorno da escolaridade. Modelo não linear.

Abstract: The objective of this study is threefold. First, to estimate the returns to education at sectoral level in order to verify the hypothesis that these returns are increasing. Second, to use the stock of high knowledge present in the high-tech service sectors to test the spillover effect hypothesis. Third, if the threshold effect of education is verified at the sectoral level. The first hypothesis is linked to the works done by Acemoglu (1996), Lucas (1988) and Romer (1990). The second one is due to the paper by Dias *et al.* (2013). The third one was proposed by Dias and McDermott (2012) and Dias (2017) for whom the increasing returns to education must cross a threshold level in Latin America. The PNAD 2006 and

[1] Autor correspondente. Universidade Federal de Mato Grosso do Sul (UFMS), departamento de Administração, Campus de Chapadão do Sul, Chapadão do Sul, Mato Grosso do Sul, Brasil.

[2] Universidade Estadual de Maringá (UEM), Programa de Pós-Graduação em Economia, Maringá, Paraná, Brasil.

2015 data were used, which allows to compare the effects over time. The estimates showed that increasing returns to education are present in all sectors and there is a threshold effect in almost all of them. The hypothesis that there is an educational threshold has been confirmed and the value is around 9.0 years of schooling, on average, in the sectors. The benefits of increasing productivity gains from the accumulation of education across sectors are yet to occur in most sectors analyzed.

Keywords: Human capital. Income differentials. Return to schooling.
Nonlinear model.

JEL Classification: I28; O41; O47.

1 INTRODUCTION

The effect of education on economic growth has been the goal of several studies and dates back to the 1960s[3]. The importance of human capital for economic growth, however, reached a new stage with the works of Lucas (1988) and Romer (1990). The authors showed that the average schooling of the countries positively affects the productivity of workers and demonstrated the importance of human capital in generating ideas that lead to innovations that cause cost reduction and/or increased productivity in the long run.

Despite the importance of human capital to increase the productivity in a macroeconomic level, there was the need for a microeconomic foundation to explain increasing returns at this level, this microeconomic foundation only appeared with Acemoglu (1996). According to the author, the decision of individuals to invest in education occurs *a priori*, that is, they observe the labor market and make their decisions to accumulate human capital and at which educational level to stop. Companies, for their part, invest in physical capital, aiming capturing the best human capital. This non-Walsarian process, called matching, takes the investments to become attractive enough to capture the desired human capital, then the match occurs. This matching process leads to the occurrence of increasing returns in the human capital market, benefiting

[3] Schultz (1962) and Uzawa (1965) are examples.

the economy as a whole (DIAS *et al.,* 2013). This non-Walrasian mechanism was replaced by a Walrasian one in Dias and McDemortt (2012), where entrepreneurs generate the demand for human capital and establish incentives for the accumulation of human capital in the long run.

Empirical studies initiated by Mincer (1958; 1974), which proposed an empirical model for calculating rates of social and private returns from a wage equation, did not initially have microeconomic support, but the empirical results were robust. Therefore, the models were initially tested empirically and linearly. One of the first works to realize that the returns were not linear, but increasing for some countries was made for Brazil by Psacharopoulos (1987). The author found that the returns of education to Brazil did not obey international standards that were decreasing. In contrast, using data from the 1980 Census, the author showed that returns were increasing and the estimated rate was 15%. Lam and Schoeni (1993) also estimated the returns for Brazil, using the 1982 National Household Survey, and obtained that the returns to education were increasing.

More recently, studies have started to use the Heckman (1979) model to correct selection bias when investigating the hypothesis of increasing returns to Brazil. Among the works stand out Kassouf (1998), the first author to consider this method to estimate the Brazilian educational returns, which was followed by Soares and Gonzaga (1999), Sachsida, Loureiro and Mendonça (2004), among others. These studies, in the most part, have made analyzes for Brazil as a whole, controlling for some individual characteristics, without considering regional, sectoral and/or increasing returns specificities. The paper of Blom, Holm-Nielsen and Verner (2001), whose main objective was the problem of income inequality, in its estimates using National Household Survey data from 1982-1998, ended up finding convexity for higher education, that is, increasing returns.

Increase returns at the aggregate level of countries and states means that increases in schooling levels are associated with increasing income. In simple words, companies employ individuals with higher educational levels that translate into increased productivity, which in turn end up receiving compatible wage increases. Estimates of returns to education in wage functions are assessing whether this hypothesis is present in the labor market. This concern to measure the increasing returns in the literature has returned effectively in an international level with the paper of Trostel (2004). The author has confirmed the hypothesis of increasing returns to several European countries and to the United States.

Dias *et al.* (2013) estimated human capital functions for Brazil and its states, based on the microeconomic foundation proposed by Acemoglu (1996). Using data from the 2009 National Household Survey, the authors find that, at the Brazilian level, the rate of return of schooling grows after 4.7 years of schooling. These increasing returns become even greater from those who obtain higher degrees of education, especially the second and third grades. The results clearly indicate a sub-supply of human capital at higher levels. This effect of increasing returns also appears in the work of Russo and Dias (2016), where the effect of health quality on the empirical estimates of education returns was considered. Education returns have continued to increase in estimates, even considering this effect of health quality.

More recently, Dias (2017) reviewed a historical analysis of growth and development comparing Brazil and the USA. According to the author, the data show that the gains of scale in the USA occurred between 1820 and 1900 when the country grew above all other countries and crossed 4.5 years of schooling. According to the author, Brazil will benefit from increasing returns when it crosses 9.0 years of schooling on average, and currently this average is 8.0 years. One of the objectives of this paper is to test this hypothesis for the sectors of the Brazilian economy.

The literature on the evaluation of the individual returns to education at the sectoral level of the economy and the spreading effect is scarce. Heuermann (2009) is one of the authors who sought to analyze the returns of education by sector. Using data from Germany, the author found that returns to human capital accumulation are higher in industry than in services. The author estimated returns of 1.8% for each one percentage point increased in the proportion of highly skilled workers and 0.6% for an equivalent increase in the proportion of workers who are not highly skilled and who highly skilled workers tend to benefit of the intra-sectoral effects of human capital spillover, while non-highly skilled workers tend to benefit, above all, from the pecuniary externalities that arise between sectors.

This intra-sectoral and spreading effect at the international level was also investigated by Schumacher, Dias and Tebaldi (2014). The authors innovated when, comparing the returns of education in Brazil and the United States, they constructed two variables: SKC - Sectoral Knowledge Concentration and KSI - Knowledge Spread Intensity. The objective was to evaluate the returns to education and the effects of these two variables in the spreading of knowledge in the economy in line with the hypothesis proposed by Lucas (1988) and Romer (1990). The results showed that the spreading effects of SKC and KSI variables occur in both Brazil and the US. The SKC variable when tested for interaction with educational level is been shown to be significant and it positively affects the return of education in both countries. The KSI variable was significant, but with the opposite effect for both countries. For Brazil, this variable has a negative effect while in the US the effect is positive. Therefore, they concluded that the intensity of the spreading of sectoral knowledge in the US positively influences the earnings of other workers in the economy. In Brazil there is an inverse effect. The reason is probably because the intensity of knowledge still does not improve the productivity gains of other workers. This is another hypothesis that this work will investigate for the sectors of the Brazilian economy.

Thus, the present study seeks to complement the existing literature estimating human capital functions by sector using the methodologies present in Dias *et al.* (2013) and Schumacher, Dias and Tebaldi (2014) for two periods using National Household Survey data from 2006 and 2015. The main innovation is to consider the spillover effect as coming from the following high-tech subsectors: telecommunications, advisory services and research. The hypothesis is to verify if these sectors, when employing people with education at master's level and higher, cause an increase of productivity in the other sectors of the economy with which they maintain a relationship.

In addition to this introduction, the paper presents four more sections. Section 2 presents the methodology used in this work, as well as the database. Section 3 presents the descriptive statistics of the variables and discusses the profile of the large sectors of the economy according to some individual characteristics of each worker. Section 4 presents the empirical estimates of the work. Finally, in the last section, the final considerations are made.

2 METHODOLOGY

2.1 Database

The database used in this study is the National Household Sample Survey (PNAD), collected by IBGE (Brazilian Institute of Geography and Statistics), in the years 2006 and 2015. The methodological procedure adopted by the IBGE implies that each person in the sample represents a certain number of people in the population. Thus, the population estimated in this work is obtained with the weight or expansion factor of each individual available in the research, which allows the data to be elaborated by weighing each observation by the respective weight.

The year 2006 contains a total of 410,241 observations, representing a population of 187 million people and 2015 contains a total of 356,904 observations, representing a population of 204 million. Of this total, it ended up using 200,970 observations in 2006 and 191,103 observations in 2015 that referred to individuals aged between 18 and 64 years and with hourly wages below R$ 600.00[4]. Individuals who declared themselves as employers, as well as public servants, were withdrawn from both samples. The 2006 income was

[4] Due to the research being the Brazilian case, the Brazilian real was used as the unit of measurement.

deflated using the PNAD income deflator available on the IPEADATA website, which allows comparisons to be made between the periods.

The variables used in the research were: logarithm of the monthly income of the individual's main labor per hour as a dependent variable (*Lnhourwage*), and education, race, location (urban or rural), gender, formal work, region and the 10 sectors of the economy as explanatory variables. Experience and squared experience variables were also created, which capture the depreciation of human capital and SKC and KSI, which estimate the sectoral concentration of knowledge and the intensity of diffusion of knowledge, respectively.

2.2 A theoretical model for the return of education

The econometric model used in this work is based on the idea of the general wage equation characterized by the non-Walsarian equilibrium of Acemoglu (1996), added by variables that capture the effects of spillover of knowledge proposed by Schumacher, Dias and Tebaldi (2014), which were inspired by the work of Romer (1990) on sectorial externalities and Lucas (1988) on externalities of human capital.

Thus, following the methodology proposed by Schumacher, Dias and Tebaldi (2014) is presented a generic wage equation characterized by the non-Walsarian equilibrium in which the wage is a proportion of the product,

$$w_{ij} = \rho f\left[H_i(S,X), A_j\left(R_j\right), Z\right] \tag{1}$$

which describes the real salary of the individual *i*, employed in the sector *j*, as a function of his human capital, *Hi*, the level of knowledge in the sector where he is employed, *Aj*, and a vector of variables describing the labor market, *Z*. The level of human capital of each individual is a function of the individual's schooling, *S*, and his experience in the labor market, *X*. The level of knowledge in the sector j is a function of the proportion of people involved in research activities in this sector, *Rj*. The first-order partial derivatives $\partial f/\partial H_i$ and $\partial f/\partial A_j$ are both positives.

The growth of sectorial knowledge occurs as proposed by Romer (1990): $g_A = \delta R_j$. The level of knowledge in any sector at time t is given by the exponential equation $A_j = A_0 e^{g_A t}$, where A_0 is the level of initial knowledge. The individual human capital is $H_i = \gamma e^{\Omega(S)+\Phi(X)}$.

Assuming the functional form

$$[H_i(S,X), A_j(R_j), Z] = A_j{}^{\alpha} H_i{}^{\beta} e^{xZ} \tag{2}$$

for the production function, the salary of the individual *i*, employed in the sector *j*, is

$$w_{ij} = \rho A_j{}^{\alpha} H_i{}^{\beta} e^{xZ} \tag{3}$$

From this wage equation it is possible to arrive at an estimable income equation of the form

$$ln\, w_{ij} = \gamma_0 + \gamma_1 S + \gamma_2 X + \gamma_3 Z + \gamma_4 R_j \tag{4}$$

For the hypotheses assumed to be valid, some propositions must be assumed:

$$w_{ij} = f[H_i(S,X), A_j(R_j), Z]\ com\ \frac{\partial f}{H_i} > 0\ e\ \frac{\partial f}{\partial A_j} > 0 = \rho A_j{}^{\alpha} H_i{}^{\beta} e^{xZ} \tag{5}$$

$$A_j = A_0 e^{g_A t} \tag{6}$$

$$g_A = \delta R_j \tag{7}$$

$$H_i = \gamma e^{\Omega(S)+\Phi(X)} \tag{8}$$

Replacing (7) into (6), applying the logarithm and deleting the time indicator:

$$ln\, A_j = ln\, A_0 + \delta R_j \tag{9}$$

Applying logarithm in (8) and (5):

$$ln\, H_i = ln\, \gamma + \Omega(S) + \Phi(X) \tag{10}$$

$$ln\, w_{ij} = ln\, \rho + \alpha\, ln\, A_j + \beta\, ln\, H_i + \chi Z \tag{11}$$

Replacing (9) and (10) into (11):

$$
\begin{aligned}
ln\, w_{ij} &= ln\, \rho + \alpha \left(ln\, A_0 + \delta R_j\right) + \beta\, [ln\, \gamma + \Omega(S) + \Phi(X)] + \chi Z \\
&= (ln\, \rho + \alpha\, ln\, A_0 + \beta\, ln\, \gamma) + \alpha\delta R_j + \beta\Omega(S) + \beta\Phi(X) + \chi Z \\
&= a + bS + cX + dZ \\
&+ eR_j \qquad (12)
\end{aligned}
$$

Where

$$
\begin{gathered}
a \equiv ln\, \rho + \alpha\, ln\, A_0 + \beta\, ln\, \gamma \\
b \equiv \beta\Omega \\
c \equiv \beta\Phi \\
d \equiv \chi \\
e \equiv \alpha\delta
\end{gathered}
$$

To estimate equation (4) was defined *S* as a vector of years of schooling [5] and its quadratic and cubic variation, *X* as a vector of labor market experience and its quadratic variation [6] and *Z* as a vector of the variables indicative of the type of work, if formal or not (*formal*), race, if the individual declared himself as white or not (*race*), urban area, if the individual lives in a urban area or not (*urban*), gender, if the individual declares himself as a man or a woman (*sex*), the region of the country where the individual lives (*North, Northeast, South, Southeast and Midwest*), woman with a child under 14 years (*womanchild*14) and discrete variables: number of family members (*nmembers*) and family income *per capita* (*incomepc*). To get closer to *Rj* it was considered the fraction (in log) of people in each sector that has at least a master's degree (from now on, "sectoral concentration of knowledge", represented by *SKC)*, and the fraction (in log) of people in each state employed in the business services sector (henceforth

[5] The PNAD survey does not allow to discriminate between different levels of graduate studies. Question 4803 of the survey assumes values in the range 1-17, where 1 means less than one year of schooling, 2 means one year of schooling and so on up to 16, which means 15 years or older, and 17, which means undetermined. The education variable was therefore combined with question 6003, which is about the current level of education of the individual, and 6007, which is about the highest level of education that the individual has attained.

[6] Following the literature, the variable was calculated as follows: exp = (age) - (schooling) - (6). In which number 6 refers to the average age at which the individual starts the studies and the variable exp^2 is the experience squared, which seeks to capture the depreciation of human capital.

"knowledge diffusion intensity", represented by *KSI*[7]), which followed the construction proposed by Schumacher, Dias and Tebaldi (2014).

Thus, the equation to be estimated is

$$ln\, w_{ij} = \gamma_0 + \Gamma_1 S + \Gamma_2 \mathrm{X} + \gamma_3 SKC + \gamma_4 KSI + \Gamma_5 \mathrm{Z} + \epsilon, \quad (13)$$

where the last term is the error term. The parameters γ_0, γ_3, γ_4 and the coefficient vectors Γ_1, Γ_2 and Γ_5.

However, a sample selection bias can occur if only the salaries of the working individuals are observed, in order to correct this bias, the Heckman procedure (1979) is used, which estimates the regression with all individuals in the sample and not only those who are working and considers the following aspects of the participants or not in the job market in the reference week (*yi*): schooling (*school*); number of family members (*nmembers*); family income *per capita* (*incomepc*). The others are dummy variables: whether the individual is of white race (*race*); if woman with a child under 14 years (*womanchild*14) and whether the individual lives in an urban area (*urban*).

The estimates considered the weights and stratification of the sample units. The results, at the Brazilian level, will be presented with the statistical test rho, which verifies the existence of a serial correlation between the wage equation and the selection equation. If this statistic is significant, it indicates the existence of sample selection bias in all specifications, which justifies the use of the method proposed by Heckman (1974; 1979).

To calculate the estimates of the rate of return, ρ, it was used the following model proposed by Trostel (2004): vector Γ_1 the wage equation returns the coefficients for education, education squared and education cubed variables. Thus, the vector can be established as:

$$\Gamma_1 = \theta_1 \mathrm{S} + \theta_2 \mathrm{S}^2 + \theta_3 \mathrm{S}^3 \quad (14)$$

[7] *SKC* is the abbreviation for sectorial concentration of knowledge, the logarithm of the fraction of people in each sector who hold at least one master's degree. *KSI* is the abbreviation for intensity of diffusion of knowledge, the log of the fraction of people in each state who are employed in the business services sector.

The estimated rate of return, ρ, is:

$$\hat{\rho}(S) = \hat{\theta}_1 + 2\hat{\theta}_2 S + 3\hat{\theta}_3 S^2 \quad (15)$$

The marginal rate being as follows:

$$\frac{\partial^2 \ln (w)}{\partial S^2} = \frac{\partial \hat{\rho}}{\partial S} = 2\hat{\theta}_2 + 6\hat{\theta}_3 S \quad (16)$$

In this case, if $\partial \hat{\rho}/\partial S > 0$ (or <0), the result indicates increasing returns (or decreases) in the return rates of schooling. The estimated coefficients on education are polynomials, such as the implicit marginal rate of return and its change to several years of schooling. Estimates of these coefficients will be given according to the Heckman model (1974; 1979). Thus, this study uses this theoretical model, which presents the appropriate way of measuring the gains coming from each year of additional schooling (DIAS *et al.*, 2013). The estimation of this model allows to test the characteristics of returns in the sectors and, in aggregate, for Brazil.

3 RESULTS

3.1 Descriptive analysis of data

Table 1 shows the descriptive statistics of the sample for the years 2006 and 2015. The data show that the statistics, in general, improved in the period. The logarithm of the hourly wage increased by 20%, on average. The average schooling increased by 15% in the decade analyzed. The results indicate that the individuals in the sample had, on average, 22.8 years of experience in the labor market in 2006 and that this increased to 23.8 years in 2015, which shows that the most experienced individuals are remaining in the labor market, thus making the average increase.

In 2015, the fraction of people in each sector with education above graduation (SKC) ranges from 0.3% in the agricultural sector to 3.5% in the high-technology services sector, which involves postal, telecommunications, financial intermediation, computer science, research and development, education, business services and associative activities. The fraction of people in each state employed in the business services sector (KSI) varies from 0.89% in the state of Maranhão to 6.01% in the Federal District.

Regarding the dummies of category, which assume the value 1 for when the individual has that characteristic and 0 for when he does not have it, there was a reduction of the individuals who declared themselves white in the period (47% of the population in 2006 against 41% in 2015) and the proportion of men and women (51% of the population is composed of women) remained unchanged, those who declared that they lived in urban households had a reduction of 1 percentage point: from 89% of the population in 2006 to 88% by 2015.

The percentage of the population that had formal work increased from 46% in 2006 to 52% of the sample in 2015. The distribution of the sample by region shows that there were no large migrations in the period since only the North presented a certain growth in the decade, most population is found in the Southeast and Northeast regions (concentrating approximately 60% of the total population in 2015) and the lowest population is in the Midwest.

Table 1 – Descriptive statistics of variables – 2006 and 2015

Variable	2006			2015		
	N	Mean	Standard deviation	N	Mean	Standard deviation
Lnhourwage	126.593	1,61	0,76	118.482	1,94	0,76
School	200.970	7,43	4,25	191.103	8,56	4,22
School²	200.970	73,31	63,46	191.103	91,00	67,47
School³	200.970	802,78	914,28	191.103	1051,01	1027,93
Exp	200.970	22,83	14,57	191.103	23,82	14,89
Exp²	200.970	733,73	781,22	191.103	789,30	805,46
SKC	125.475	-5,24	0,47	116.435	-4,96	0,83
KSI	200.970	-3,75	0,47	191.103	-3,62	0,44
Race	200.970	0,47	0,50	191.103	0,41	0,49
Gender	200.970	0,48	0,50	191.103	0,48	0,50
Formal	126.590	0,46	0,50	118.482	0,52	0,50
Urban	200.970	0,89	0,31	191.103	0,88	0,32
North	200.970	0,12	0,46	191.103	0,15	0,36
Northeast	200.970	0,29	0,36	191.103	0,28	0,45
South	200.970	0,16	0,46	191.103	0,16	0,37

Southeast	200.970	0,31	0,32	191.103	0,31	0,46
Midwest	200.970	0,11	0,46	191.103	0,11	0,31
Agriculture	126.593	0,11	0,30	118.482	0,09	0,29
Energy and Mining	126.593	0,03	0,18	118.482	0,03	0,16
Industry	126.593	0,10	0,29	118.482	0,07	0,26
High-tech industry	126.593	0,05	0,21	118.482	0,04	0,19
Utilities	126.593	0,04	0,19	118.482	0,04	0,20
Construction	126.593	0,09	0,29	118.482	0,11	0,32
Trade	126.593	0,26	0,18	118.482	0,29	0,20
Transport	126.593	0,05	0,23	118.482	0,06	0,25
High-tech services	126.593	0,14	0,36	118.482	0,15	0,35
Other services	126.593	0,13	0,33	118.482	0,11	0,31

Source: Own elaboration based on the PNADs of 2006 and 2015.

The sector variables were constructed from the National Classification of Economic Activities (CNAE / IBGE), and their definitions are presented in Table 6 of the Appendix. Descriptive statistics for the sectors show that the highest percentage of the population is allocated to the commerce and services sectors (high technology services and other services, which include domestic services and urban cleaning, for example). The sectors of industry and other services were the ones that lost most workers, compared to the others, while the Construction sector was the one that increased the most population (growth of 28% in the period).

3.2 The composition of employment in the Brazilian sectors

The data available in the PNAD about Brazilian workers allows the researcher a detailed analysis of the individuals' characteristics of those who are in the labor market. Since the current study analyzes the characteristics of the returns of education in the Brazilian sectors and the income differentials of these sectors, it is important to analyze how the sample is composed according to the individual characteristics of each worker by sector. Table 2 presents the sectorial composition of workers in the ten Brazilian sectors that are being analyzed in this study for the years 2006 and 2015.

Table 2 – Composition of employment in the Brazilian sectors – 2006 and 2015 (%)

Sector	*Sample*		*Lnhourwage*		*School*		*SKC*		*Exp*	
	2006	2015	2006	2015	2006	2015	2006	2015	2006	2015
Agriculture	14.444	11.146	1,16	1,49	3,61	5,02	-6,13	-5,76	29,05	30,15
Energy and Mining	3.914	3.037	1,84	2,17	7,96	9,11	-4,86	-5,13	20,89	21,69
Industry	12.087	8.727	1,50	1,81	7,50	8,69	-5,16	-5,05	22,13	23,54
High-tech industry	5.788	4.656	2,01	2,27	9,74	10,52	-5,13	-4,30	17,55	19,60
Utilities	4.615	4.737	2,06	2,30	10,98	11,85	-4,95	-3,47	18,83	19,31
Construction	11.162	13.435	1,51	1,93	5,80	7,09	-5,08	-5,37	25,55	25,57
Trade	32.766	34.687	1,58	1,86	8,50	9,38	-5,12	-5,32	19,68	20,69
Transport	6.819	7.621	1,82	2,09	7,70	8,80	-5,40	-5,67	23,84	25,08
High-tech services	18.110	17.492	2,07	2,34	10,86	11,91	-5,47	-3,58	17,57	18,37
Other services	16.888	12.944	1,31	1,70	6,03	7,04	-4,79	-5,13	25,55	28,63

Source: Own elaboration based on the PNADs of 2006 and 2015.

The columns of the table called Sample show the sectorial division of the sample according to the number of workers employed in each sector. In 2006, more than 25.88% of the sample was allocated to the commercial sector, followed by high technology services (14.31% of the sample), other services (13.34%) and Agriculture (11.41%). In 2015, the sectors of Commerce and high technology services increased the percentage of workers allocated in these sectors, to 29.28% and 14.76%, respectively; while the other services and agriculture sectors saw their percentages decrease (10.92% and 9.40%, respectively). Analyzing the variation in the decade, it can be seen that the sector that lost most workers in the period was the industrial sector (22% reduction), while the one that grew the most was Construction (22% increase in the period).

As for the variable *Lnhourwage*, which presents the logarithm of the hourly wage per sector, it is possible to verify a great disparity of income between the sectors. In 2006, the highest average salary was in the high-tech services sector - which involved sectors such as mail, telecommunications, financial intermediation, IT, research and development - and the lowest in the agricultural sector, with the difference between the two sectors reaching 44%. In 2015, there was a growth in the variable in all sectors, with the largest increases in Agriculture and Other Services sectors (approximately 29%). The high-tech services sector, although not showing great growth in the period, continued to have the highest average wage among all sectors, however, the difference between this sector and the sector with the lowest income (Agriculture) decreased in 2015 to 36%.

The variables *School* and *Exp*, which represent the average years of schooling and experience, respectively, presented growth in the period. Average schooling grew by 13% in the period, with the sectors with the highest levels of average schooling being those of public utility and high technology services (more than 11 years of study, on average in 2015), while sector with the lowest level of schooling is Agriculture, which despite having increased its average by 38% in the period, in 2015 presented an average of 5 years of schooling. On the other hand, experience has also grown, on average, in all sectors, with the highest experience rates being in the Agriculture and Other Services sectors, showing that individuals belonging to these sectors start early in the activity and stay in the same for very long time, cbining this factor with low schooling and low wages, it is possible to verify that there is much to be improved in these sectors.

Finally, the SKC variable, which estimates the sectoral concentration of knowledge through the logarithm of the fraction of people in each sector with at least one master's degree, varies in 2015 from -5.76 (or 0.3%), in the agricultural sector to -3.47 (or 3.12%) in the sector of public utility services.

4 EMPIRICAL ESTIMATES

Based on the analysis of the sectorial composition of the workers in ten Brazilian sectors, this subsection seeks to meet the second objective of this article, which is to find the returns of education according to the different Brazilian sectors in the period of 2006 and 2015. For this, it was used the Ordinary Least Squares Method and Maximum Likelihood using the procedure developed by Heckman (1979) to run the regressions. It was initially verified if there was a correlation between the explanatory variables of the model and some specifications of the model were tested using the Hausman test (1978) to verify which would be the most correct and that best explained the model.

Table 3 presents the estimated coefficients of log wage regressions for Brazil in the years 2006 and 2015, with the standard errors presented in parentheses. The OLS columns refer to the coefficients estimated by ordinary least squares. The Heckman columns present the coefficients estimated by maximum likelihood using the procedure developed by Heckman (1979). The use of two estimation techniques brings robustness to the results.

The first analysis that must be done concerns the rho statistical test, which verifies the existence of a serial correlation between the wage equation and the selection equation, indicating or not the existence of sample selection bias in all specifications. As this statistic was significant in both years, the use of the method proposed by Heckman (1974; 1979) is pertinent, so all analyzes will follow considering this model.

The variables *School*, $School^2$ and $School^3$ presented significant coefficients in both periods, which is important, since it implies the existence of returns due to the accumulation of individual human capital. The estimates of the *Exp* and Exp^2 variables also showed significance, however, while the variable *School* showed an increase in its explanatory degree on the log of salaries in the period, the variable *Exp* presented a reduction. The variable Exp^2 presented a negative signal as expected, since over time, remuneration tends not to rise so significantly.

The *SKC* and *KSI* variables were defined by logarithm, so their coefficients show their wage elasticity. The results show that both variables are statistically significant and have importance in the explanation of wages. The *SKC* variable had positive coefficients in 2006 and 2015. In 2006, one percent increase in the proportion of highly educated people in a particular sector was associated with a 6.26% increase in salaries in that sector. By 2015, the effect is smaller: a one percent increase in the proportion of highly educated people was associated with a 2.53% increase in the sector's salary, which is largely explained by the increase in the average schooling of the sector as a whole, which tends to offset less individual additional years of study.

The effect of the *KSI* variable is higher and positive in both periods. A one percent increase in the state fraction of people involved in the business services sector is associated with an increase in the salary log of approximately 7.22% in 2006 and 11.20% in 2015. According to Schumacher, Dias and Tebaldi (2014), these effects are, respectively, evidence for sectorial externalities of human capital and for the role of knowledge diffusion performed by the services sector to companies.

Regarding the qualitative variables, estimates of the coefficients associated with the *Gender* variable show that, in 2006, a man tended to earn incomes 23.98% higher than a woman, a difference that showed a small increase in 2015, increasing to 24.85%[8]. Individuals who declared themselves to be white received 12.41% more than those who declared themselves to be non-whites in 2006 and 9.83% more in 2015, that is, the wage difference in this category decreased in the period.

Table 3 – Regressions of the wage logarithm - 2006 and 2015

Variable	2006		2015	
	OLS	Heckman	OLS	Heckman
School	0,041***	0,0399***	0,0415***	0,0408***
	(0,00310)	(0,00311)	(0,00373)	(0,00374)
School ^{2}	-0,00249***	-0,00250***	-0,00603***	-0,00602***
	(0,00046)	(0,00046)	(0,00052)	(0,00052)
School ^{3}	0,00041***	0,00041***	0,00052***	0,00052***
	(0,00002)	(0,00002)	(0,000021)	(0,00002)
Exp	0,038***	0,038***	0,02590***	0,02590***
	(0,00048)	(0,00048)	(0,00052)	(0,00052)
Exp^2	-0,00047***	-0,00047***	-0,00033***	-0,00033***
	(0,00001)	(0,0000)	(0,00001)	(0,00001)
SKC	0,0627***	0,0626***	0,02530***	0,02530***
	(0,00571)	(0,00571)	(0,00502)	(0,00502)
KSI	0,0722***	0,0722***	0,113***	0,112***
	(0,00466)	(0,00466)	(0,00666)	(0,00666)
Gender	0,215***	0,215***	0,223***	0,222***

[8] Percentage values are calculated by following the equation $[(e^{\beta_i}-1) \times 100]$.

	(0,00395)	(0,00395)	(0,00435)	(0,00435)
Race	0,117***	0,117***	0,09390***	0,09380***
	(0,00369)	(0,00370)	(0,00415)	(0,00415)
Urban	0,0547***	0,0631***	0,118***	0,118***
	(0,00626)	(0,00658)	(0,00717)	(0,00718)
Formal	0,0768***	0,0767***	0,0355***	0,0355***
	(0,00369)	(0,00369)	(0,00409)	(0,00409)
North	0,217***	0,217***	0,192***	0,193***
	(0,00636)	(0,00636)	(0,00708)	(0,00708)
South	0,252***	0,252***	0,305***	0,305***
	(0,00569)	(0,00569)	(0,00666)	(0,00666)
Southeast	0,235***	0,235***	0,276***	0,276***
	(0,00477)	(0,00477)	(0,00582)	(0,00582)
Midwest	0,274***	0,274***	0,332***	0,332***
	(0,00610)	(0,00610)	(0,00685)	(0,00685)
Agriculture	-0,228***	-0,228***	-0,16***	-0,16***
	(0,00959)	(0,00959)	(0,00908)	(0,00908)
Energy and Mining	0,172***	0,172***	0,177***	0,177***
	(0,0103)	(0,0103)	(0,0121)	(0,0121)
Industry	-0,0450***	-0,0405***	-0,054***	-0,054***
	(0,00635)	(0,00635)	(0,00805)	(0,00805)
High-tech industry	0,208***	0,208***	0,158***	0,158***
	(0,0086)	(0,0086)	(0,0112)	(0,0112)
Utilities	0,203***	0,203***	0,154***	0,154***
	(0,00958)	(0,00958)	(0,0134)	(0,0134)
Construction	-0,0121**	-0,012**	0,0799***	0,0799***
	(0,00684)	(0,00683)	(0,00688)	(0,00688)
Transport	0,134***	0,134***	0,133***	0,133***
	(0,00817)	(0,00817)	(0,00839)	(0,00838)
High-tech services	0,178***	0,178***	0,15***	0,15***
	(0,006)	(0,006)	(0,0104)	(0,0104)
Other services	-0,041***	-0,041***	-0,00836	-0,00833
	(0,00629)	(0,00629)	(0,00696)	(0,00696)
Constant	0,0855**	0,125***	1,2***	1,22***
	(0,0362)	(0,0375)	(0,0396)	(0,041)
Selection Equation				
School		0,0304***		0,0315***
		(0,0007)		(0,00073)
Nmembers		-0,0271***		-0,0434***
		(0,00198)		(0,00213)
Womanchild14		-0,178***		-0,157***
		(0,0147)		(0,0179)
Race		0,0157***		0,00011
		(0,0059)		(0,00610)

Urban		-0,254***		-0,0361***
		(0,00935)		(0,00938)
Incomepc		-4,63E-12***		-4,59E-13***
		(2,54E-14)		(2,65E-14)
Constant		0,432***		0,215***
		(0,0124)		(0,0125)
Athrho		-0,104***		-0,0527**
		(0,0248)		(0,0223)
Lnsigma		-0,522***		-0,458***
		(0,00246)		(0,00218)
N	125.472	199.849	116.435	189.056

Source: Own elaboration based on the PNADs of 2006 and 2015.

Note: *** significant at 1%, ** significant at 5%. N refers to the number of observations.

As for the regions, the region considered as the base was the *Northeast*. Thus, as all coefficients of the regions were positive, this means that all had greater effects on the salary than the *Northeast*, with the highlights being the *MidWest*, *Southeast* and *South*, which, in 2015, were approximately 32% higher in the returns compared to the base region.

Regarding the sectoral coefficients, the base sector was the trade sector. Notably only three sectors present negative coefficients in 2015: agriculture, industry and other services, which means that these sectors have lower impacts on wages than the base sector. The high-tech industry sector stands out for having the highest percentage change, that is, in 2015, an individual in this sector tended to receive an average of 17% more than those in the base sector, while an individual to the agricultural sector tended to receive 14.78% less. It should be noted that the dummy of the Other services sector was not significant in 2015 in both specifications.

The results of the selection equation are presented in the second part of Table 3. With the exception of the *race* dummy, all other coefficients of the variables of the selection equation were significant in both periods. Signs of the *nmembers* and *womanchild14* variables show that these variables have a negative effect on the reserve wage, that is, they negatively affect the individual decision to migrate to paid labor.

After the estimation, some tests were performed. The test for heteroskedasticity proposed by Breusch and Pagan (1979) and Cook and Weisberg (1983) did not indicate the existence of this problem in the data, which was corroborated by the test of White (1980). The multicollinearity test, disregarding quadratic variables, indicated absence of this problem. The model specification was also tested, which proved to be correct. The test for omitted variables of Ramsey (1969) was not rejected. The tests of normality of residues proposed by Shapiro and Wilk (1965) and D'Agostino, Belanger and D'Agostino Junior (1990), showed that it is not possible to conclude by the non-normality of the residues.

Table 4 was elaborated based on Table 3, which allows analyzing the rate of return in Brazil for years of schooling, following the proposed methodology, as well as the years of schooling that show the beginning of the increasing returns to schooling (@). The columns $\partial\hat{\rho}(S)$ show that there are increasing returns to education in Brazil, since this occurred from 2.03 years of schooling in 2006 and 3.84 years of schooling in 2015. This would be the threshold effect of the level of education proposed by Dias and McDermott (2012), which consists, starting from a certain average level of education of the population, to begin observing the presence of increasing returns at the individual level.

Table 4 – Return of schooling in Brazil - 2006 and 2015

2006			2015		
Years of schooling	$\partial\hat{\rho}(S)$	$\partial\hat{\rho}/\partial S$	Years of schooling	$\partial\hat{\rho}(S)$	$\partial\hat{\rho}/\partial S$
4	0,0396	0,0048	4	0,0176	0,0004
8	0,0786	0,0014	8	0,0447	0,0130
11	0,1337	0,0220	11	0,0978	0,0147
15	0,2416	0,031	15	0,2125	0,0224
17	0,3104	0,0368	17	0,2886	0,0349
7,34	0,0707	0,0132	8,55	0,0524	0,0412
@	2,03	0,00	@	3,84	0,00

Source: Own elaboration based on the PNADs of 2006 and 2015.

The columns $\partial\hat{\rho}/\partial S$ represent the second-order derivatives of the log of wages in relation to years of schooling and are zero at the point where returns become increasing. Considering 2006, it is possible to see that returns are 7.86%

for those with primary education, 13.37% for those who have completed high school, 24.16% for those with higher education and 31.04% for those who have at least a master's degree.

Analyzing 2015, it can be seen that the returns are 4.47% for those with primary education, 9.78% for those who have completed high school, 21.25% for those with higher education and 28.86% for those who have at least a master's degree.

From these results, it is possible to see clearly that the growth of the rate of return is exponential as the years of study increase, leading to an important observation: that the relationship between years of schooling and individual income is not linear. Comparing the two years of analysis, it appears that there is a trend of diminishing returns, which may be due to the increase in average schooling of Brazilians because, over the years, by increasing the average education of the population as a whole, the premium for a one-year increase in schooling becomes lower. Dias *et al.* (2013), when comparing Brazil's returns to education with the United States, found that, in fact, the returns in the United States are lower because, according to them, this country seems to be at a stage of development where higher education has a limited impact on productivity and wages, while in Brazil there are still gains from increasing average schooling in general. Therefore, the results found in this study seem to point to the direction of general increases in the level of schooling in the country, generating lower returns to higher levels of education.

As the main objective of this study was to analyze the sectoral returns to education, a new specification of the regression of the year of 2015 was introduced where binary sector variables were interacted with the variables of schooling. From the coefficients found in the regression, the sectoral education returns were estimated, which are presented in Table 5.

Table 5 – Sector returns to schooling in 2015 (%)

Sector	N	4	8	11	15	17
Agriculture	11.146	6,29	8,75	12,28	19,25	23,70
Energy and Mining	3.037	0,37	-0,41	6,18	24,55	37,84
Industry	8.727	-0,70	-1,46	1,61	10,59	17,17
High-tech industry	4.656	0,17	-1,73	4,34	22,43	35,75
Utilities	4.737	-1,38	-0,38	6,85	25,15	38,00
Construction	13.435	0,02	-2,58	2,26	17,80	29,46
Trade	34.687	-0,50	-1,60	1,92	12,41	20,13
Transport	7.621	1,09	-3,45	0,01	13,78	24,58
High-tech services	17.492	-0,61	-0,46	5,61	21,62	33,02
Other services	12.944	0,05	-4,32	-3,81	1,94	6,97

Source: Own elaboration based on the PNAD of 2015.
Note: N refers to the number of observations.

The results show that some sectors have decreasing returns for the first years of schooling, that is, there is no interest by these sectors in hiring labor with little schooling. For 4 years of studies, the greatest return is in the agricultural sector, since returns in this sector are positive with 2.18 years of education. Much of this is due to the low average schooling of the sector, which is only 5 years of study, as verified in Table 1.

The highest returns are found in the upper or higher education brackets, highlighting the energy and mining sectors, high technology industry, utilities and high technology services. It should be noted that returns are much lower than the average for other services sector, due in large part to the type of work in this sector: household services, recycling, street cleaning, ie, manual and mechanical work, where having a higher-level degree, for example, does not bring a great return to the individual in order to improve his work productivity, for example.

The results found for the Brazilian sectors prove that the theory proposed by Dias (2017) is correct: in fact, positive and exponential returns occur when the sector passes the barrier of nine years of schooling on average, a fact verified in the energy and mining, high-tech industry, high-tech services, trade and utilities sectors. The other sectors need to be encouraged to increase the level of schooling of their workers in order to match them to the sectors already advanced in terms of returns to education.

The increasing returns of Agriculture start from 3.0 years in average education, while in the Energy and Mining Sector only from the 8.0 years. Figures 1.1 and 1.2 show that average schooling, represented by the vertical line, is above these values, but average education rates are lower than 10%, more specifically 6% and 4%, respectively. That is, the mechanism of increasing returns is present in the sectors, but the stimulus to the accumulation of knowledge for employment in these sectors is not great.

Figure 1.1 Agriculture

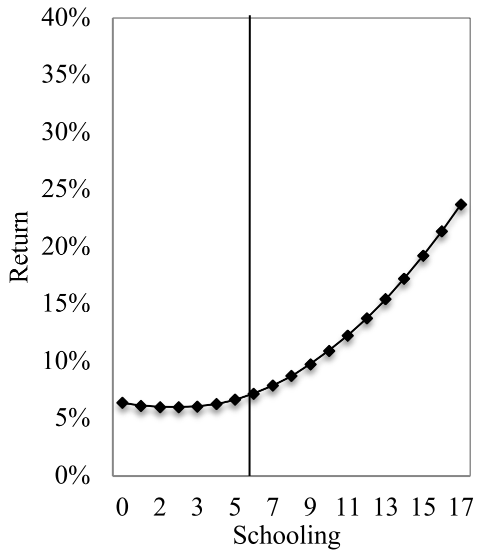
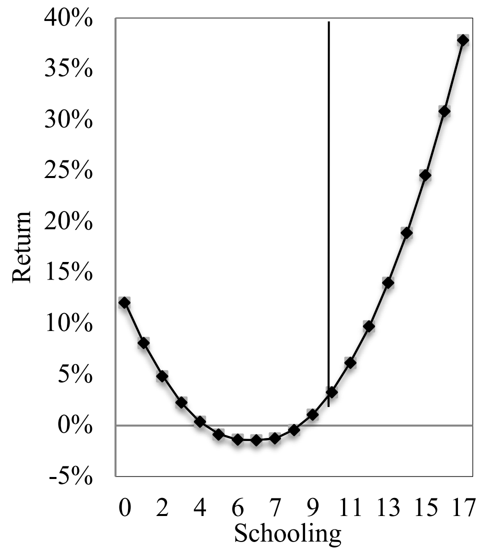


Figure 1.2 Energy and Mining



Source: Own elaboration based on the PNAD of 2015.

Figures 1.3 and 1.4 below show that industry sectors have also increasing returns associated with higher levels of education. However, industry in general, Figure 1.3, with average schooling of 8.0 years has the negative rate of return until 10 years. In the high-tech industry, this fact was exceeded with average schooling 11.2 years, but the rate of return per year of schooling is close to 3%. People employed in these sectors would only benefit from increasing wage gains and above 10% if their educational levels are more than 13 years.

Figure 1.3 Industry

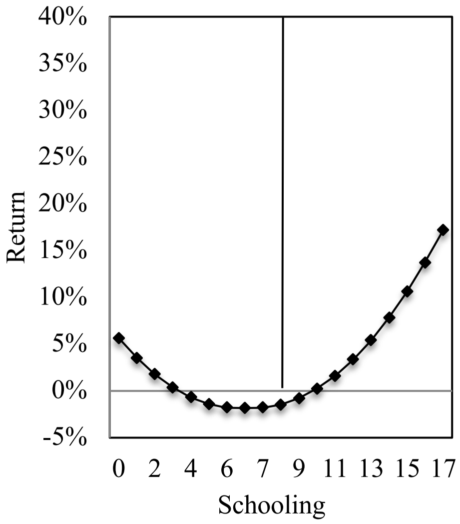
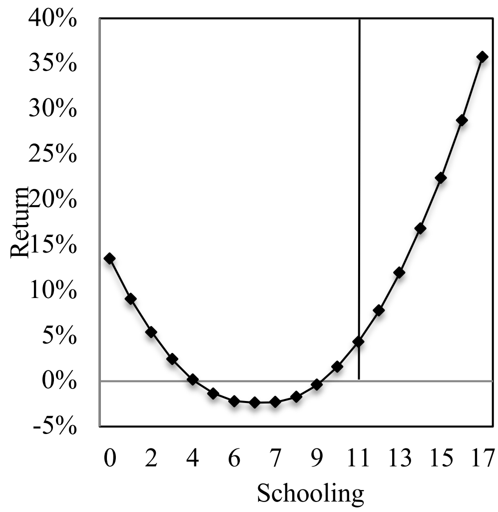


Figure 1.4 High-tech Industry



Source: Own elaboration based on the PNAD of 2015.

Figures 1.5 and 1.6 show that in the sectors of utilities and high-tech services the gains are increasing for the average of the workers. In these sectors, average schooling of 13 years generates gains above 10% and increasing to a higher educational level. These sectors demonstrate the potential gain to be measured for the others with the improvement of the average schooling of the workers.

Figure 1.5 Utilities

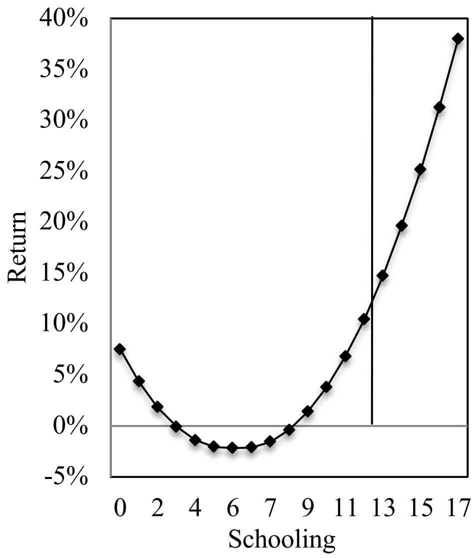
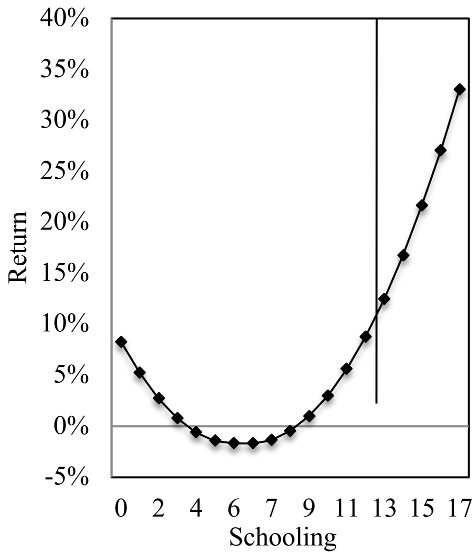


Figure 1.6 High-tech Services



Source: Own elaboration based on the PNAD of 2015.

Two important sectors of the economy in the absorption of labor with educational levels below the second degree are those of transport and construction. Figures 1.7 and 1.8 show that the mechanism of increasing returns is present, but the means of education of the two sectors are in the area where the highest immediate educational level, up to 11 years, does not increase wage gains.

Figure 1.7 Transport

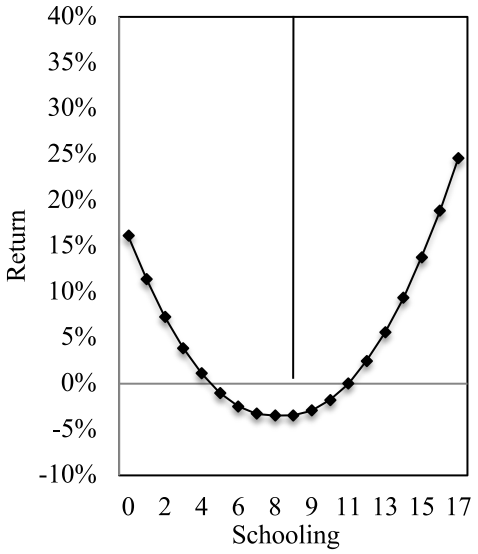
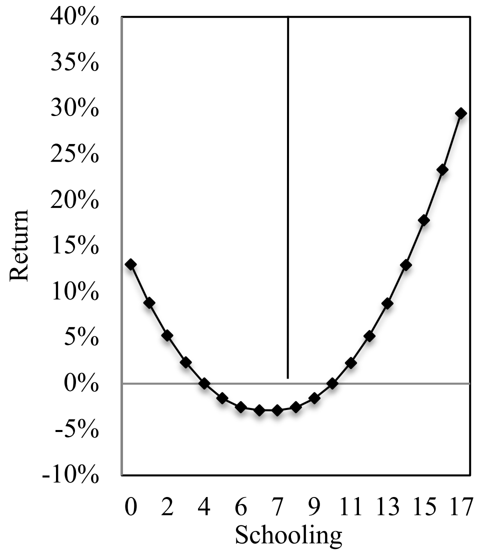


Figure 1.8 Construction



Source: Own elaboration based on the PNAD of 2015.

Finally, figures 1.9 and 1.10 show that the average schooling of the trade sector is crossing the line where more education means increases in the positive gains in productivity and wages, while the other services sector is still far from this fact.

Figure 1.9 Trade

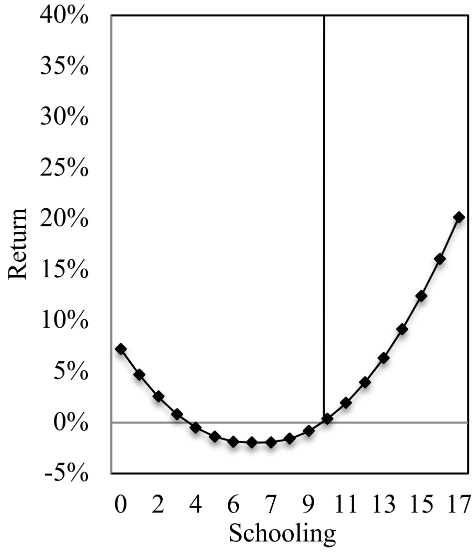
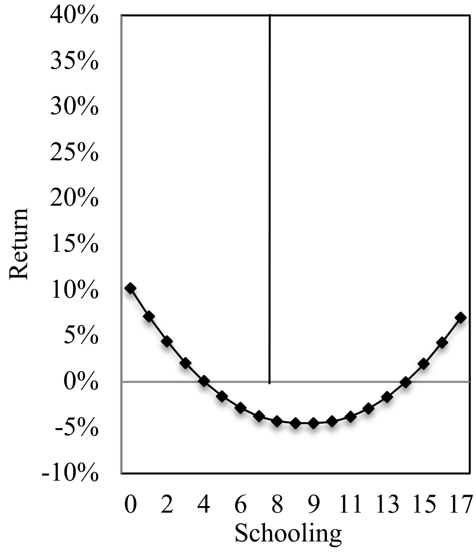


Figure1.10 Other services



Source: Own elaboration based on the PNAD of 2015.

5 CONCLUSIONS

The spreading effect is verified in the business support services sector. The hypothesis of this work was to verify if these sectors, when employing people with schooling at master's level and higher produce as a result an increase in productivity in the other sectors. The results showed that this variable (*SKC*) was statistically significant and important in explaining wages. However, the effect of the *KSI* variable (knowledge diffusion intensity) is greater and positive in both periods, that is, the intensity of knowledge diffusion has a positive and significant effect on the salary of the individuals of a given state. One percent increase in the state fraction of people engaged in the business services sector was associated with a wage increase of approximately 11.20% in 2015, showing that these effects are, respectively, evidence for sectoral human capital and the role of knowledge dissemination by the business services sector.

The variables *SKC* and *KSI* that have been added to the regression show that, in fact, there are sectoral externalities of human capital, that is, a greater fraction of people educated at master's and doctoral level in the sectors tend to increase the average salary of these sectors. One hypothesis to be tested in the future is the interaction of these variables with the average educational level of the sectors.

The results showed that there are increasing returns to education in Brazil, which occurred from 2.03 years of schooling in 2006 and 3.84 years of schooling in 2015. This would be the threshold effect of the level of education proposed by Dias and McDermott (2012). However, when estimated at the sectoral level, in general, schooling, in order to benefit increasing wage gains, should be above 9.0 years, as well observed by Dias (2017).

Regarding the returns to education by sectors, the results showed that these are, in general, increasing. The highest returns are found in the upper or higher education brackets, highlighting the energy and mining sectors, high technology industry, utilities and high technology services. These results, in general, also corroborate with the theory proposed by Dias (2017).

The results show that the returns of education occur in different ways in the various sectors of the economy, showing the importance of analyzing them separately, since the educational investment needs of each sector are visible, so that public policies can aim at sectors that most need to raise average levels of education. As the increasing gain generally occur from 9.0 years of average schooling in the sectors, educational policies should be focused on this level of schooling to be achieved.

REFERÊNCIAS

ACEMOGLU, D. A microfoundation for a social increasing returns in human capital accumulation. *The Quarterly Journal of Economics*, v. 111, n. 3, 1996.

BLOM, A.; HOLM-NIELSEN, L.; VERNER, D. Education, earnings, and inequality in Brazil, 1982-1998: implications for education policy. *Peabody Journal of Education*, v. 76, n. 3, p. 180-221, 2001.

BRAZILIAN INSTITUTE OF GEOGRAPHY AND STATISTICS – IBGE. *National sample and household surveys*. Brazil: PNAD, 2006; 2015. (Methodological notes). Available at https://www.ibge.gov.br/estatisticas/sociais/trabalho/9127-pesquisa-nacional-por-amostra-de-domicilios.html?=&t=downloads. Accessed at May 15, 2018.

BRAZILIAN INSTITUTE OF GEOGRAPHY AND STATISTICS – IBGE. *National classification of economic activities.* Brazil: CNAE/IBGE. Available at http://www.ibge.gov.br/concla/. Accessed at May 11, 2018.

BREUSCH, T. S.; PAGAN. A. R. A simple test forheteroscedasticity and random coefficient variation. *Econometrica*, v. 47, p. 1287-1294, 1979.

COOK, R. D.; WEISBERG, S. Diagnostics for heteroscedasticity in regression. *Biometrika*, v. 70, p. 1-10. 1983.

D'AGOSTINO, R. B.; BELANGER A. J.; D'AGOSTINO JUNIOR, R. B. A suggestion for using powerful and informative tests of normality. *American Statistician*, v. 44, p. 316-321, 1990.

DIAS, J. Can human capital alone close the Brazil-US income gap? Expected contributions of human capital to economic development in Brazil and the US. *In:* LAWRENCE, M.; MURRAY, F. (Ed.). *Human capital*: perspectives, challenges and future directions. 1. ed. New York: Nova Science Publisher, v. 1, p. 70-90, 2017.

DIAS, J.; MCDERMOTT, J. Aggregate threshold effects in the generation of human capital: the cases of Latin America and Europe. *In:* PUNZO, L. F.; FEIJO, C. A.; ANYUL, M. P. (Ed.). *Beyond the global crisis*: structural adjustments and regional integration in Europe and Latin America. Londres: Routledge, 2012.

DIAS, J.; MONTEIRO, W. F.; DIAS, M. H. A.; RUSSO, L. X. Função de capital humano dos estados brasileiros: retornos crescentes ou decrescentes da educação? *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 43, p. 333-380, 2013.

DIAS, J.; SCHUMACHER, F. I.; TEBALDI, E. Geographic and sector externalities from highly qualified human capital: the importance of the business service sector. *Applied Economics Letters (Print)*, v. 21, p. 329-334, 2014.

HAUSMAN, J. A. Specification tests in econometrics. *Econometrica*, v. 46, p. 1251-1272, 1978.

HECKMAN, J. Sample selection bias as a specification error. *Econometrica*, v. 47, n. 1, p. 153-161, 1979.

HECKMAN, J. Shadow prices, market wages, and labor supply. *Econometrica*: Journal of the Econometric Society, p. 679-694, 1974.

HEUERMANN, D. F. *Human capital externalities in Western Germany.* Trier: Institute of Labour Law and Industrial Relations in the European Community (IAAEG), 2009. (Discussion papers, n. 03/2009).

INSTITUTE OF APPLIED ECONOMIC RESEARCH – IPEADATA. Brazil: PNAD income deflator: INPC. Available at: http://ipeadata.gov.br. Accessed at May 17, 2018.

KASSOUF, A. Wage gender discrimination and segmentation in the Brazilian labor market. *Economia Aplicada*, v. 2, n. 2, p. 243-269, 1998.

LAM, D.; ROBERT, F. S. Effects of family background on earnings and returns to schooling: evidence from Brazil. *Journal of Political Economy*, v. 101, n. 4, p. 710-740, 1993.

LUCAS, R. E. On the mechanics of economic development. *Journal of Monetary Economics*, v. 22, n. 1, p. 3-42, jul. 1988.

MINCER, J. Investment in human capital and personal income distribution. *The Journal of Political Economy*, p. 281-302, 1958.

MINCER, J. Schooling, experience, and earnings. *National Bureau of Economic Research*. New York: Columbia University Press, 1974.

PSACHAROPOULOS, G. *Earnings and education in Brazil:* evidence from the 1980 census. Washington: The World Bank, 1987. (Discussion paper series, v. 90).

RAMSEY, J. B. Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society*, p. 350-371, 1969. (Series B 31).

ROMER, P. M. Endogenous technological change. *The Journal of Political Economy*, v. 98, n. 5, p. S71-S102, out. 1990.

RUSSO, L. X.; DIAS, J. The health influence on returns to education in Brazil: a nonlinear approach. *Economia,* Brasília, v. 17, p. 210-220, 2016.

SACHSIDA, A.; LOUREIRO, P.; MENDONÇA, M. Um estudo sobre retorno em escolaridade no Brasil. *Revista Brasileira de Economia*, v. 58, p. 249-265, 2004.

SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality. *Biometrika*, v. 52, p. 591-611, 1965. (Complete samples).

SCHULTZ, T. W. Reflections on investment in man. *Journal of Political Economy*, v. 70, n. 5, p. 1-8, 1962.

SCHUMACHER, F. I.; DIAS, J.; TEBALDI, E. Two tales on human capital and knowledge spillovers: the case of the US and Brazil. *Applied Economics,* v. 46, p. 1-11, 2014.

SOARES, R.; GONZAGA, G. Determinação de salários no Brasil: dualidade ou não linearidade no retorno à educação. *Revista de Econometria*, v. 19, n. 2, p. 367-404, 1999.

TROSTEL, P. A. Returns to scale in producing human capital from schooling. *Oxford Economic Papers*, v. 56, n. 3, p. 461-484, jul. 2004.

UZAWA, H. Optimum technical change in an aggregative model of economic growth. *American Economic Review,* v. 6, n. 1, p. 18-31, 1965.

WHITE, H. A. Heteroskedasticity: consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica,* v. 48, p. 817-838, 1980.

APPENDIX

Table 6 – Definition of sector variables

Variable	CNAE Code	Description
Agriculture	11, 12, 13, 14, 20, 50	Agriculture and related services, forestry and logging, fishing and aquaculture.
Energy and Mining	100, 110, 230, 234, 130, 140, 260, 270, 280	Extraction and processing of energy and minerals (excluding machinery and equipment), and related services.
Industry	150, 160, 170, 180, 190, 200, 360	Food, textiles, confection, leather and footwear, wood and furniture etc.
High-tech industry	210, 220, 240, 250, 290, 300, 310, 320, 330, 340, 350	Pulp, paper and publishing, chemicals, plastic, machinery and equipment, vehicles etc.
Utilities	400, 410, 850	Electricity, gas, water, health, social services etc.
Construction	459	Building.
Trade	500, 530, 550, 700, 450, 710, 930	Trade, repair, housing, real estate, rental of furniture, personal services.
Transport	600, 610, 620, 630	Transportation and related activities.

High-tech services	640, 650, 660, 670, 720, 730, 740, 800, 910, 920	Mail, telecommunications, financial intermediation, IT, research and development, education, business services, associative and recreational activities.
Other services	370, 531, 900, 950, 990, 998	Domestic services, recycling, street vending, urban cleaning and other activities.

Source: Own elaboration based on the PNADs of 2006 and 2015.
Note: The codes refer to the first three digits of the CNAE Domiciliar/IBGE classification. The codification of the CNAE can be obtained at http://www.ibge.gov.br/concla/.

Autor correspondente:
Michele Aparecida Nepomuceno Pinto
E-mail: mi_nepomuceno@hotmail.com

Recebido em: 30/09/2019.
Aceito em: 31/05/2021.