

Estimating the returns to education using a parametric control function approach: Evidences for a developing country[★]

ANDRÉ PORTELA SOUZA^{*}
EDUARDO ZYLBERSTAJN^{**}

Abstract

This paper investigates the causal effect of education on earnings in Brazil by employing a method, proposed by Klein and Vella (2010), that obtains identification on the presence of conditional heteroskedasticity. In contrast to traditionally used IV methods, this approach yields unbiased estimates in the absence of instruments. Results indicate that the average return to education in Brazil was relatively stable at around 15% from 1995 to 2003, declined afterwards reaching 11.1% in 2014, but has bounced back to 12.1% after the economic crisis in 2015. The results suggest that the OLS estimations are downward biased and we interpret this as a sign of under-education premiums that are likely to occur in environments where the more talented drop out from school and move into the labor market earlier in life.

Keywords: Return to education, wage equation, control function approach, under-education premium

JEL Codes: C3, I21, J31

1 Introduction

Earnings inequality has declined during the past decade in most of Latin American countries. In Brazil, the Gini coefficient declined from 0.567 in 2001 to 0.497 in 2014¹. Still, Brazil remains a very unequal country, and recent

Submitted on 13 June 2019; Reviewed on 23 December 2019

[★] The authors are thankful to Bruno Oliva, Cristine Pinto, Guilherme Stein, Hélio Zylberstajn, Naércio Menezes-Filho, Priscilla Albuquerque Tavares, Reynaldo Fernandes and Vladimir Ponczek for helpful comments and suggestions. The authors are also greatly thankful to Leandro Alexandre da Silva for technical assistance in writing the optimization code.

^{*} São Paulo School of Economics (FGV/EESP) and Center for Applied Microeconomic Studies

^{**} Sao Paulo School of Economics (FGV/EESP) and FIPE

✉ andre.portela.souza@fgv.br ✉ ezylberstajn@fipe.org.br

¹ According to Instituto Brasileiro de Geografia e Estatística (IBGE).

developments suggest that this downward trend might be at risk: the Gini coefficient has been bouncing back in the last years.

The inequality level in a given country is the result of the distribution of the individual productivity characteristics and the returns of these attributes in the population. In particular, the returns to education play a key role to determine earnings (and thus its inequality) and are likely one of the most explored issues in the empirical economics literature. However, to these days several empirical challenges make it difficult to consistently estimate the returns to schooling. This is, in turn, a serious problem researchers and policy makers must address to better understand income inequality dynamics and to foster investments in human capital.

This paper employs a novel technique developed by Klein and Vella (2010), henceforth KV (2010), to estimate the causal effect of schooling on wages in Brazil for the period 1995 to 2015. The difficulty in identifying the causal effects of education on earnings arise from the endogeneity of educational choices to wages. Since Mincer (1974) established a methodology to estimate wage equations, several authors have documented and tried to deal with the OLS bias, usually by employing instrumental variables (IV) strategies—see, for instance, Griliches (1977), Angrist and Krueger (1991), Duflo (2001), and Heckman, Urzua, and Vytlacil (2006); Card (1999), in particular, presents a detailed survey on this subject.

Although large in quantity, the empirical literature lacks in robustness of results. Angrist and Krueger (1991) estimate returns to schooling in the US labor market between 6 and 10%, above the OLS estimates of 5–7%. Oreopoulos (2007) estimates that one extra year of education yields an average increase of 13% in wages, compared to a downward biased OLS estimate of 8%. Carneiro and Lee (2008) and Chen (2008) also make use of IV to estimate an average return to education between 13 and 15% for American men, but the validity of the instruments they use (proxies of the costs of school attendance) is questioned by authors such as Cameron and Taber (2004).

In Brazil, several studies have estimated the returns to education—see, for instance, Langoni (1973), Senna (1976), Tannen (1991), and Barros and Reis (1991)—but the difficulty in dealing with endogeneity has been a constant issue. More recently, Teixeira and Menezes-Filho (2010) use a national survey with data from 1997 to 2007 and employ an IV strategy to find that returns were between 5.5 and 9.4%, lower than their OLS estimate of 11.6%.

The KV (2010) strategy uses a control function approach. Identification

relies on heteroskedasticity of the error terms, which yields nonlinearity to the control term, allowing for non-biased estimates of the coefficient of interest. As [Saniter \(2012\)](#) points out, this approach makes use of second moment restrictions (variance) instead of first moment exclusion restrictions (inherent to any IV approach) and therefore has the key advantage of allowing inference to be made at any moment in time (as long as our identifying assumptions hold). This is relevant, since IV estimates are often bounded by the time of the variation induced by the instrument. An example of such case is found in [Ichino and Winter-Ebmer \(2004\)](#), who use own and father's World War II involvement as instruments for schooling; thus, their results are estimates for a particular cohort of Germans and cannot be replicated for previous or subsequent cohorts.

[Farré, Klein, and Vella \(2013\)](#), henceforth [FKV \(2013\)](#), implement a parametric estimator for the [KV \(2010\)](#) identification strategy. Using a sample drawn from the National Longitudinal Survey of Youth (NLSY 1979), they find that accounting for endogeneity increase the estimate returns to schooling from 6.8 to 11.2%. They interpret their findings as compatible with “the over education penalty” hypothesis: factors associated with some individuals becoming over educated (actual education level above the predicted education) are penalized in the labor market. This same approach is used in this paper.

Although we follow [FKV \(2013\)](#), it is worth mentioning that the authors are not alone in the pursuit of identification through heteroskedasticity. [Lewbel \(2012\)](#) proposes a somewhat similar approach, used in recent works such as [Fortin and Raguéd \(2017\)](#) and [Ben-Moshe, D'Haultfœuille, and Lewbel \(2017\)](#).

The contributions of this paper to the literature are at least twofold. First, we consistently estimate the returns to education in a developing country for a long period and thus document its trends and evolution over time. To provide a preview of the results found in this paper, the estimated average return to education in Brazil fell from 15.6% in 1995 to 12.1% in 2015. The OLS estimates are downward biased in the entire period by 3.0 percentage points on average. We contribute further by providing some evidence of an “under education premium” hypothesis. In this context, underlying factors and/or individual unobserved characteristics associated with under education are positively associated with higher wages in the labor market. Developing countries are characterized by factors associated with lower education attainment such as the presence of credit constraints, low school quality, high returns to labor market experience for young adults etc. [Horowitz and Wang \(2004\)](#), for instance, show that, under credit constraint, families specialize the time allocation of their children. It is possible

that children with higher ability are selected out of school and into the labor market earlier in life. We provide some indirect evidence of this hypothesis.

Second, our findings contribute to the debate of the causes of the inequality decrease in Latin America in general, and in Brazil in particular. It has been well documented that Latin American countries have experienced a decrease in inequalities in labor earnings recently and such trends are mainly explained by the decrease in the returns to education (e.g., [Azevedo, Inchaust, & Sanfelice, 2013](#); [López-Calva & Lustig, 2010](#); [Lustig, López-Calva, & Ortiz-Juarez, 2013](#); and [Manacorda, Sanchez-Paramo, & Schady, 2010](#)). In general, their estimates assume exogeneity of schooling. Our findings provide robust estimates of declining returns to education from 1995 to 2015 and thus corroborate their conclusions.

The remainder of the paper is organized as followed. First, [section 2](#) presents the econometric model and the implementation procedure of the estimator. [Section 3](#) describes the data used in the estimation. [Section 4](#) presents the main results. [Section 5](#) discusses in greater depth the OLS-IV bias and [section 6](#) presents some evidence to support the robustness of our findings. [Section 7](#) concludes.

2 Empirical exercise

2.1 Econometric model and identification

In this section, we describe the KV ([2010](#)) approach. The ultimate goal of any linear model that estimates the causal effect of education on earnings is to obtain consistent estimates for the parameter δ in a wage equation of the following kind:

$$w_i = x_i\beta + \delta s_i + u_i, \quad (1)$$

where w_i represents hourly log wages, s_i represents years of schooling, x_i is a vector of exogenous regressors, and u_i the error term for individual i .

OLS estimation of (1) is inconsistent because of the endogeneity bias caused by omitted variables correlated to both wage and schooling (e.g., individual abilities). To better illustrate the strategy proposed by KV ([2010](#)), let us first write the education equation:

$$s_i = x_i\varphi + v_i, \quad (2)$$

where x_i may be (but not necessarily is) identical for both equations and v_i is the error term for individual i . From this point on, we refer to (1) as the wage

equation and to (2) as the education equation.

If there is endogeneity, the covariance between u_i and v_i is different from zero. Then, we can write $u_i = \lambda v_i + e_i$, where $\text{cov}(v_i, e_i) = 0$. Replacing this expression in equations (1) and (2) allows us to and then rewrite the model in a control function setting:

$$w_i = x_i\beta + \delta s_i + \lambda v_i + e_i, \tag{3}$$

where s_i is the control term and δ is a measure of the degree of endogeneity. Because s_i is a perfect combination of x_i and v_i , the regressors are collinear and OLS is thus infeasible.

However, we can also write

$$\lambda = \frac{\text{cov}(u_i, v_i)}{\text{var}(v_i)} = \frac{\text{cov}(u_i, v_i)}{\sigma_v \sigma_v} \frac{\sigma_u}{\sigma_u} = \rho \frac{\sigma_u}{\sigma_v}, \tag{4}$$

where σ_j ($j = u, v$) denotes the standard deviations of the error terms u and v and ρ is the correlation coefficient between them.

One additional assumption we must impose regards the structure of the error terms. Formally, the error terms u_i and v_i are assumed to have a multiplicative structure composed by an heteroskedastic part and an homoskedastic part. Let u_i^* and v_i^* be the respective homoskedastic error part. The error terms are defined as follows:

$$u_i = H_u(x_i^u)u_i^* \quad \text{and} \quad v_i = H_v(x_i^v)v_i^*. \tag{5}$$

If the assumption that the errors are heteroskedastic holds, then $\sigma_j = H_j(x_i^j)\sigma_j^*$, where H_j is the heteroskedasticity function ($j = u, v$) and $x_i^j \subseteq x_i$, meaning that the set of variables in the heteroskedasticity function is not necessarily equal to the set of variables in the wage equation—because heteroscedasticity might come from only a subset of such variables.

The correlation coefficient ρ can now be written as

$$\rho = \frac{\text{cov}(u_i, v_i)}{\sigma_u \sigma_v} = \frac{H_u(x_i^u)H_v(x_i^v)\text{cov}(u_i^*, v_i^*)}{H_u(x_i^u)\sigma_u^*H_v(x_i^v)\sigma_v^*} = \frac{\text{cov}(u_i^*, v_i^*)}{\sigma_u^* \sigma_v^*}. \tag{6}$$

In this setting, KV (2010) interpret δ and λ as measures of unobserved abilities.

Finally, combining the expressions above, equation (3) can now be rewritten:

$$w_i = x_i\beta + \delta s_i + \rho \frac{H_u(x_i^u)}{H_v(x_i^v)} v_i + e_i, \quad (7)$$

which is the final equation to be estimated. Notice that x_i^u and x_i^v may be different or identical in both equations.

There are two important identification assumptions to properly estimate the equation above. First, $H_u(x_u)/H_v(x_v)$ needs to vary across x , so that regressors are not collinear. KV (2010) has named this assumption as the *variable impact property* (VIP), which simply requires that the heteroskedasticity is present in either one of both equations in such a way that the quotient of the two functions is not constant across x . A good candidate (but not the only one) for providing this is age. What is needed is that the heteroskedasticity due to age in the education equation is different than the one in the wage equation, which is economically plausible: on the one hand, schooling in Brazil has been expanding throughout cohorts and being universalized; on the other hand, we expect heteroskedasticity in the wage errors to increase as age increases due to heterogeneous experience and human capital accumulation (other than education).

The second assumption is that the correlation of the homoscedastic part of the errors must be constant and independent of the regressors, i.e. $\rho = \text{corr}(u_i^*, v_i^* | x_i) = \text{corr}(u_i^*, v_i^*) = \text{const}$. KV (2010) refer to this as the *constant correlation condition* (CCC). Put differently, the CCC requires that, once the differences in observed socioeconomic characteristics are accounted for, the return to unobserved ability is constant. Thus, the heterogeneity of the returns to unobserved abilities comes entirely from the heterogeneity of the individuals' socioeconomic characteristics described in the x_i 's.

It is important to highlight that, as FLK (2013) explain, this setting does not imply that the returns to unobserved abilities are constant. The assumption is that once all the socioeconomic factors are accounted for, then the returns to unobserved ability are constant. This seems to be a plausible assumption, particularly in the Brazilian context, where inequality of opportunities is large (Bourguignon, Ferreira, & Menéndez, 2007).

2.2 Implementation

The implementation of the KV (2010) estimator can be done either parametrically and non-parametrically. Notably, a non-parametric approach has the advantage

of not imposing normality of errors for consistency nor a functional form for the heteroskedasticity functions. KV (2010) describe a semi-parametric method, for which they also provide proof of consistency; their method is implemented by Saniter (2012), Klein and Vella (2009), Schroeder (2010). Nonetheless, all these papers describe the computational burden of adopting such strategy. Saniter (2012), for instance, who estimated the returns to schooling in Germany, reports having used 500 computer cores at the same time.

Aside from the studies mentioned above, Schroeder (2010) estimates the impact of microcredit borrowing on per-capita household consumption in Bangladesh using conditional second moments. Wang (2010) adapts the KV (2010) estimator to a Chinese database and finds that the returns to education in urban China are in the range 2.3–4.6%, below the OLS estimates of 3.9–7.3%. Figueirêdo, Nogueira, and Santana (2014) use the KV (2010) approach to assess the influence of family background in student achievement in ENEM, a Brazilian national exam that is undertaken when students graduate from high school.

FKV (2013) follow a different path. They choose a functional form for the heteroskedasticity functions (that take an exponential form), thus largely simplifying the implementation of the estimator. They also simulate a Monte Carlo exercise in which the true heteroskedastic functions are exponential but the estimation is done parametrically assuming either heteroskedasticity in only one of the equations or a quadratic form for the heteroskedastic functions. In both cases, results show that the estimator performs quite well.²

The heavy computational demand in the nonparametric estimation arises for reasons that become clear below and we will highlight them when the method is described. In order to avoid the nonparametric burden, we adopt the same parametric approach proposed by FKV (2013), treating $H_u(x_i^u)$ and $H_v(x_i^v)$ as exponential functions of the individuals' observable characteristics. More specifically, we define the following heteroskedastic function:

$$H_j^2(x_j) = \exp(z\theta_j). \tag{8}$$

This approach makes the estimation feasible, with the risk of losing efficiency due to misspecification of the functional form. It should be noted that z may or may not be equal to x . In our benchmark specification, we use all individual characteristics in both equations.

² Also using simulation techniques, Klein and Vella (2010) show that even with little heteroskedasticity the OLS bias is eliminated. These results are encouraging to the use of the estimation technique described in this section.

The only subtle change that we make regards state dummies. In the wage equation we include indicator variables for the current state of residence. Local labor market conditions vary across states and this may affect not only the wage levels but the wage dispersion as well. In the schooling equation, in the other hand, we add indicator variables for the state of birth. They are proxies for different school attainment costs both direct (school accessibility) and indirect (opportunity costs to attend school such as child labor wages). Again, they may affect both education levels and dispersion. We show evidence in support of this in [section 3](#).

Following FKV (2013), estimation is done following a three-step procedure:

1. Estimate \hat{v} through OLS in the education equation (2): $\hat{v} = s - x\hat{\varphi}_{\text{OLS}}$.
2. Using a Poisson regression,³ estimate θ_v from equation (8) by regressing \hat{v}^2 on z . Then, compute the standard deviation of the reduced form error as $\hat{H}_v = \sqrt{\exp(z\hat{\theta}_v)}$.
3. Using \hat{v} and \hat{H}_v and the assumed form for $H_u = \exp(z\theta_u)$, solve the following nonlinear least squares problem:⁴

$$\min_{\beta, \rho, \delta, \theta_u} \sum_i^n \left(e_i = w_i - x_i\beta - \delta s_i - \rho \frac{\sqrt{\exp(z_i\theta_u)}}{\hat{H}_{vi}} \hat{v}_i \right)^2. \quad (9)$$

Solving equation (9) yields the coefficient estimates of interest, particularly $\hat{\delta}$.⁵ Estimation was performed for several years of data. Standard errors are defined as square roots of the diagonal elements of the covariance matrix, which in turn is obtained numerically (see [Nocedal, 1996](#)).

To summarize and facilitate the understanding of our empirical strategy, the final estimation models are outlined below, in the order in which they are

³ FKV (2013) estimate $\ln(\hat{v}^2)$ using OLS. In our case, the error is not normally distributed, making the traditional log-linear approach inappropriate to recover \hat{H}_v^2 . To overcome this issue, we use a Poisson regression that fits a model $H_j^2(x_j) = \exp(z\theta_j + \varepsilon)$, in line with our specification in equation (8).

⁴ Once again, we remind that in our benchmark specification, we have $z = x$ except for state of birth dummies.

⁵ The optimization used the L-BFGS-B algorithm (Zhu, Byrd, Lu, & Noceda, 1997) and the routine was programmed in Python. Notice that if the estimation was to be done non- (or semi-) parametrically, step 3 implies that instead of estimating θ_u , one needs to estimate the unknown functions H_j at each iteration, which is precisely the main computational burden in the non-parametric strategy. On the other hand, by imposing the functional form $H_{ij} = \sqrt{\exp(z_i\theta_j + \varepsilon_i)}$ the optimization of equation (9) is done in approximately seven minutes using a common server with standard CPU and RAM specifications.

estimated:

$$s_i = \varphi_0 + \varphi_1 Age_i + \varphi_2 AgeSq_i + \varphi_3 Female_i + \varphi_4 White_i + \sum_{m=1}^{26} \tau_m dsb_{im} + v_i \quad (10)$$

$$H_v^2(x_i) = \hat{v}_i^2 = \exp\left(\theta_{v0} + \theta_{v1} Age_i + \theta_{v2} AgeSq_i + \theta_{v3} Female_i + \theta_{v4} White_i + \sum_{m=1}^{26} \psi_m dsb_{im} + \varepsilon_{vi}\right) \quad (11)$$

$$LnWage_i = \beta_0 + \delta s_i + \beta_1 Age_i + \beta_2 AgeSq_i + \beta_3 Female_i + \beta_4 White_i + \sum_{m=1}^{26} \gamma_m dsr_{im} + \rho \frac{\sqrt{\exp(z_i \theta_u)}}{\hat{H}_{vi}} \hat{v}_i + e_i, \quad (12)$$

where dsb_{im} and dsr_{im} are state of birth and state of residence dummies respectively and $z_i \theta_u = \theta_{u0} + \theta_{u1} Age_i + \theta_{u2} AgeSq_i + \theta_{u3} Female_i + \theta_{u4} White_i + \sum_{m=1}^{26} \varsigma_m dsb_{im}$. Note that state of birth dummies were used in the education equation and in the $H_j^2(x_i)$ functions (we consider them exogenous), while state of residence were included as controls for local labor markets in the wage equation. equation (10) was estimated using OLS, equation (11) was estimated by fitting a Poisson regression (see footnote 3 above) and the equation (12) was estimated by the NLP described above.

3 Data

The analysis presented in the next sections are based on the *Pesquisa Nacional por Amostra de Domicílios* (National Household Sample Survey, henceforth PNAD) for the years of 1995 to 2015. PNAD is an annual household survey conducted by the *Instituto Brasileiro de Geografia e Estatística* (IBGE) and it is a nationally representative sample of the population. In the most recent wave (2015), 356,904 individuals were surveyed.

As explained before, we adopt a parsimonious specification with strictly exogenous covariates as our benchmark,⁶ so we use few variables and compati-

⁶ We include some control variables as a robustness check in section 6.

lization of different waves with different questionnaires is straightforward.⁷ Our concern is that by including more control variables, we would also introduce additional sources of endogeneity.

Because our interest lies on returns to schooling, we restrained our sample to individuals between 25 and 55 years of age—to minimize censored schooling and selection into the labor market biases. Also, we have excluded individuals that did not report any labor income in the week of reference. Individuals with missing values in any of the variables used in the estimation were also dropped. We considered the log hourly wage from the main job as the labor income variable and dropped the top and bottom 1% observations to exclude outliers. Table 1 provides an overview of all the variables used in this study and Table 2 displays some summary statistics for the 2015 dataset.

Table 1. Variables description.

Variable	Description
LnWage	Log of hourly wage
YrsEduc	Years of education
Age	Years of age at interview
Female	Dummy indicator with value 1 if female and 0 otherwise
White	Dummy indicator for white individuals

Source: PNAD waves from 1995 to 2015 (IBGE).

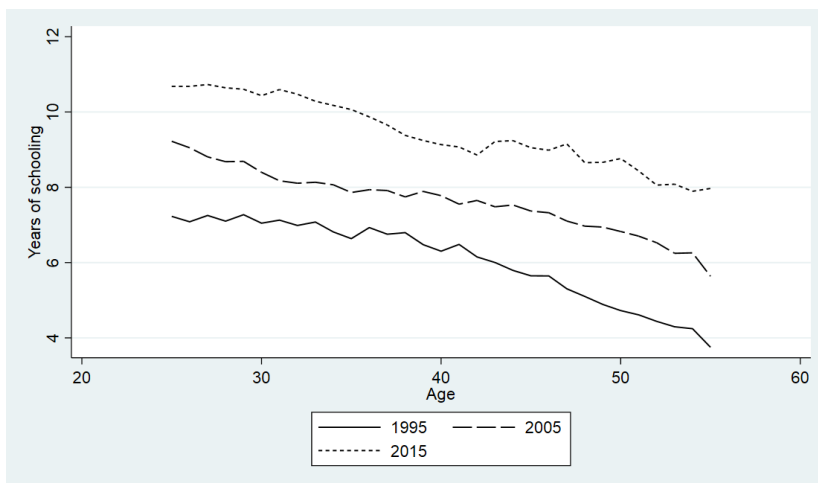
Table 2. Descriptive statistics.

Variable	Mean	S.D.	Min	Max
LnWage	3.54	0.77	1.44	6.32
YrsEduc	9.54	4.31	0	17
Age	39.04	8.55	25	55
Female	0.4309	–	0	1
White	0.4271	–	0	1
N		109,433		

Source: PNAD 2015 (IBGE).

Overall, during the period of analysis (1995 to 2015), Brazil witnessed important demographic changes. Not only the labor force got older, but there was a significant expansion in access to education. Figure 1 illustrates this

⁷ We used the DataZoom Stata package from PUC-Rio in order to construct each year's dataset. Available at <http://www.econ.puc-rio.br/datazoom>



Source: PNAD 2015 (IBGE).

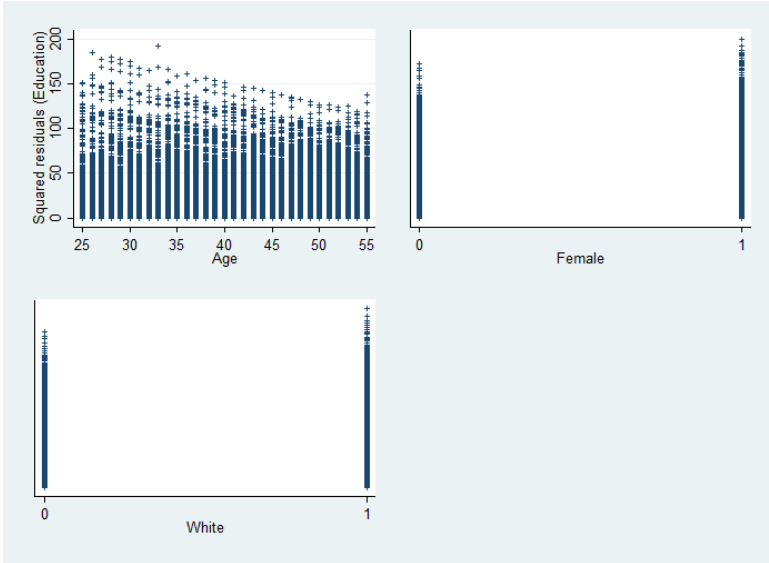
Figure 1. Education equation: graphical analysis of heteroskedasticity.

phenomenon, by graphing average years of schooling by age for different years from our sample (1995, 2005 and 2015).

Figure 2 and Figure 3 provide visual evidence for the presence of heteroskedasticity and, more importantly, also suggest that heteroskedasticity is not constant across x . See, for instance, that the slope of the relationship between age and the squared residuals \hat{v}^2 (Figure 2) and \hat{u}^2 (Figure 3) are different, as is the case with the female or white dummies. This is crucial because, as highlighted earlier, identification relies on these differences.

4 Results

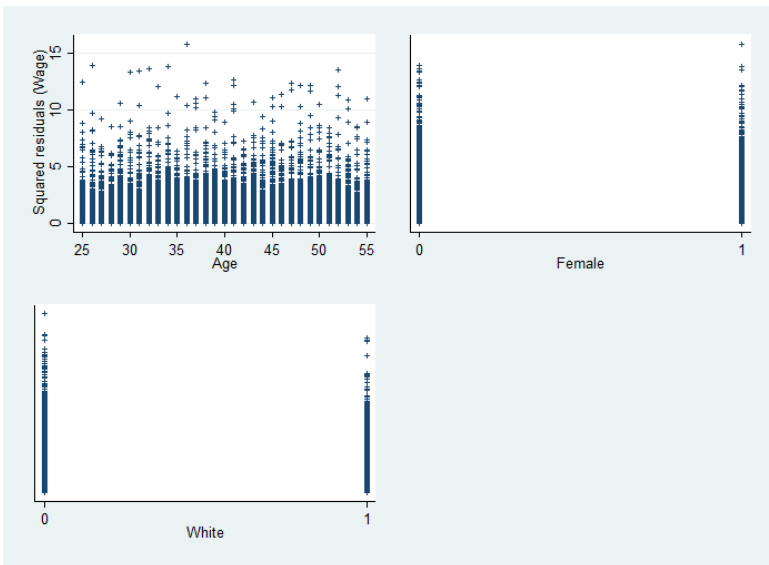
The three-step procedure outlined in section 2 yields the results displayed in Table 3 and Table 4. We begin displaying in Table 3 the results for the estimation of the education equation, which yielded estimates for \hat{v} . The results show that older people are less educated than younger people in 2015. This is a consistent finding for all the years and greatly reflects the fact that younger cohorts were more exposed to the expansion of the public-school system in the end of the second half of the 20th century. Moreover, older and non-white individuals are less educated. Also of interest is the formal test for the presence of heteroskedastic errors in the education equation, displayed in the Table A1 in the Appendix. The test strongly rejects the hypothesis of homoscedastic errors, which is in favor of the identification strategy adopted in this paper. The



Source: Author's calculations based on PNAD 2015 (IBGE).

Note: Residuals are obtained after the OLS estimation of equation (2).

Figure 2. Education equation: graphical analysis of heteroskedasticity.



Source: Author's calculations based on PNAD 2015 (IBGE).

Note: Residuals are obtained after the optimization of equation (9).

Figure 3. Wage equation: graphical analysis of heteroskedasticity.

Table 3. OLS Estimates: Education equation (2015).

Variable	β_{OLS}
Age/100	-7.8534*** (1.4100)
(Age/100) ²	-2.1196 (1.7690)
Female	1.4382*** (0.0246)
White	1.4635*** (0.0271)
Constant	13.8015*** (0.2857)
N	109,433
Adjusted R^2	0.1315

Source: Authors' calculations based on PNAD (IBGE).

Notes: OLS estimates for equation (10). State of birth dummies are also included in the regression. Standard errors in parentheses.

*significant at 10%; **significant at 5%; ***significant at 1%.

Table 4. OLS and Control function estimates: Wage equation (2015).

Variable	β_{OLS}	β_{CF}
YrsEduc	0.0886*** (0.0005)	0.1211*** (0.0024)
Age/100	4.2036*** (0.2195)	4.1059*** (0.0511)
(Age/100) ²	-3.5601*** (0.2754)	-3.0398*** (0.0568)
Female	-0.2361*** (0.0039)	-0.2807*** (0.0052)
White	0.1271*** (0.0043)	0.0777*** (0.0060)
Constant	2.0457*** (0.0442)	1.6954*** (0.0366)
ρ	-	-0.0591*** (0.0044)
N	109,433	109,433
Adjusted R^2	0.3344	0.3400

Source: Authors' calculations based on PNAD (IBGE).

Notes: OLS estimates for equation (1) and CF estimates for equation (12). State of birth dummies are also included in the regression. Standard errors in parentheses. *significant at 10%; **significant at 5%; ***significant at 1%.

estimates for $\hat{\theta}_v$ are also displayed in the Appendix, in the [Table A2](#).

Next, we turn to the wage equation. [Table 4](#) shows the differences between the OLS and the control function estimates for β , with standard errors displayed in parenthesis. As for the coefficients, we see that age affects wages in a linear fashion in the CF setting and that men and white workers are expected to earn more. The coefficients are, in most cases, smaller in the CF setting.

The main result of this paper is displayed in the first row, where we find an average return of 9.2% for each extra year of education in the OLS case and a 11.4% in the controlled setting, meaning that the OLS model is thus biased downward by 2.2 percentage points. The estimates for are displayed in the appendix in the [Table 3](#). The other coefficients have none or small changes in the controlled function setting, except for the age coefficient, which is higher in the OLS estimate.

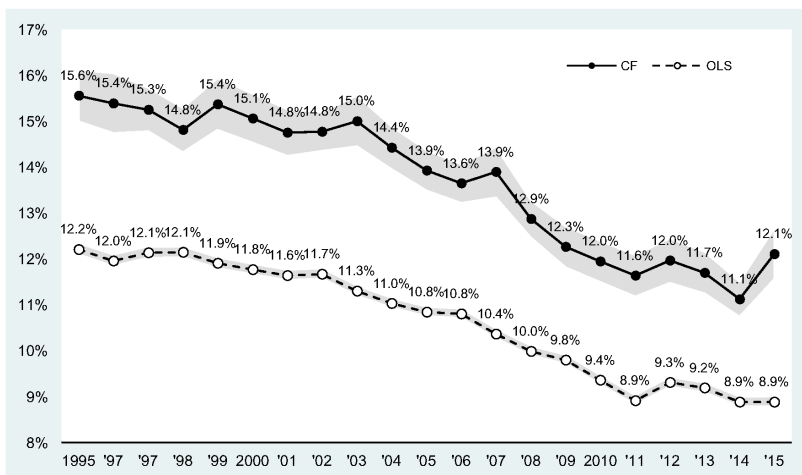
The same exercise is replicated for the years of 1995 to 2014. Remember that the KV (2010) approach allows to calculate the returns to schooling at any period (independent of any instrument) and for all individuals, not only compliers.

For simplicity, results are displayed graphically in [Figure 4](#).⁸ Each year represents the equivalent to the first row of [Table 4](#), which depicts the estimates for 2013. A clear pattern of declining returns to education emerges, either in the OLS or the CF setting. Moreover, when comparing 1995 to 2015 one finds similar declines in return rates: 3.0 and 2.8 respectively.

[Menezes-Filho, Fernandes, and Picchetti \(2006\)](#) and [Tavares and Menezes-Filho \(2011\)](#) have already documented the decline in returns to schooling in the past decade and its importance to the decline in earnings inequality, even though they are unable to properly account for endogeneity. [Fernandes and Menezes Filho \(2012\)](#) present evidence that this phenomenon might be related to the sharp increase in the relative supply of medium and high-skilled workers and [Manacorda et al. \(2010\)](#) find similar results. It seems a reasonable explanation, but arguably further research is needed, especially to identify the reasons for the recent increase in the returns.

The results presented thus far are not directly comparable to other studies that attempt to measure the causal impact of education on wages in Brazil. [Menezes-Filho et al. \(2006\)](#) report returns of 14% in 1997 and [Tavares and Menezes-Filho \(2011\)](#) report declining returns between 1995 and 2009, when they reached slightly less than 12%. Both studies rely on repeated cross-sections

⁸ All reported coefficients are significant at the 1% level.



Source: Authors' calculations based on PNAD (IBGE).

Notes: Each line depicts the evolution of δ in equations (1) for OLS and (12) for the control function estimates. Estimations are done separately for each year. All coefficients are significant at the 1% level in all years and shadowed areas indicate the 95% confidence interval. In the last years of each decade, IBGE does not carry PNAD: values for 2000 and 2010 are simple averages between immediately previous and following years.

Figure 4. Returns to education in Brazil: 1995–2015.

to estimate average returns to schooling. [Teixeira and Menezes-Filho \(2010\)](#), on the other hand, use an IV approach to estimate a much lower return of 5.5% per year of education, leading to a conclusion of an upward bias in OLS estimates. The instruments they use are the number of schools in the state and year when individuals were born and an educational law passed in 1971, which unified primary education in Brazil. While their results seem to contradict the ones presented in this paper, they can be reconciled considering their LATE interpretation. As argued by [Imbens and Angrist \(1994\)](#), the coefficients estimated by means of IV represent the causal effect only for the subsample of compliers (the individuals who were *actually* affected by variations in the instrument). It seems sound to assume that for [Teixeira and Menezes-Filho \(2010\)](#) the compliers are individuals with lower levels of education, so their estimates are not directly comparable to the ones presented in this section. To make results comparable, we would need to be able to identify the same compliers in our sample and then estimate the returns to education only for that subsample.

These results also seem reasonable when compared internationally. [Psacharopoulos and Patrinos \(2004\)](#) argue that, overall, the international average of the Mincerian return to schooling is 10% and that it is higher in middle- and

low-income countries due to diminishing marginal returns. Interestingly, three recent studies apply similar approaches to ours. Klein and Vella (2009) find a return of 10% for Australian workers and Saniter (2012) estimated an average return of 8.5% in Germany (both studies estimate H_j semiparametrically). Farré et al. (2013) employ the same parametric methodology we used in this paper and find an average return of 11.2% for the US using the NLSY79 database. All three studies find that the OLS estimates are biased downwards, as our results also suggest. As extensively documented in the literature, it seems reasonable that in Brazil, a middle-income country, returns to schooling are higher than those in Australia, Germany and the US. Less obvious, but also of interest, is the fact that the OLS estimates are biased downwards in these studies and in our results as well.

One last remark in this section is in order. Our estimates suggest an upward bounce in the declining average return to schooling in the years where economic downturns were more pronounced. This was the case in 2003 (when former president Lula da Silva took office), 2007 (just prior to the financial crisis), 2012 and 2015. One of the advantages of the control function approach, as argued before, is the possibility of measuring the coefficient of interest in several sequential years and observing patterns as this one. One possible explanation is that during downturns, (relative) demand for workers with higher productivity increases—but further research on the reasons for these spikes are needed.

5 Discussion: the OLS-IV gap

In this section, we shed some light on the direction of the OLS bias that was found and reported earlier. The negative estimate for ρ implies that the OLS estimate will be smaller than the one obtained in the controlled function setting.

The OLS-IV gap has been extensively discussed in the literature. Some early papers suggest that the omitted variable bias arises because of unaccounted ability, which in turn would produce an upward bias in OLS estimates. Other authors have claimed that measurement errors in the education variable would produce downward biased OLS estimates (see Angrist & Krueger, 1991; Card, 1995, 1999; Cameron & Taber, 2004).

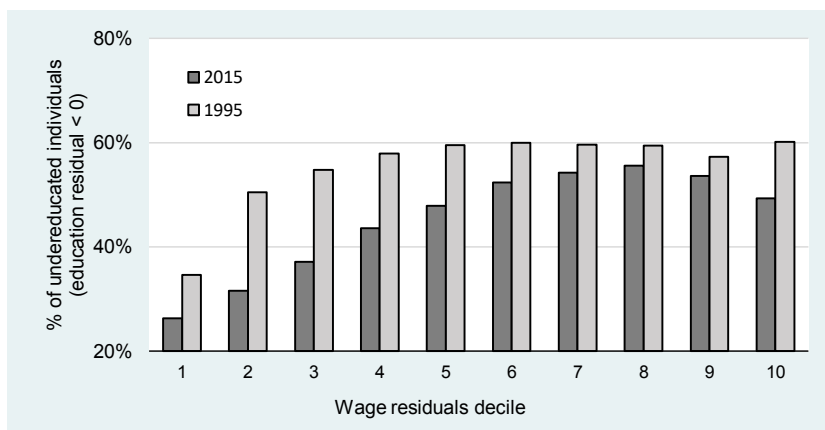
A third trend in the literature, however, claims that aside from ability, the error component of the education equation captures other factors, such as motivation, that would lead individuals to obtain what Vella and Gregory (1996) call “over-education”. This over achievement would, in turn, yield lower returns to schooling because returns to over education are lower than the average returns

to education (Dolton & Vignoles, 2000; Groot & van den Brink, 2000; Rubb, 2002; Farré et al., 2013), yielding what has been called the "over education penalty". This penalty is an interpretation for the negative value of ρ . If the correlation between u and v is negative, then if one has a higher than expected education (as predicted by the education model), she will likely have a lower than expected wage. In a developing country context, however, it might be useful to look at the other half of the coin.

If the claim that there is an over-education penalty is valid, then for the same reasons one might expect an "under-education premium". This means that if the education of one individual is lower than expected, she probably earns more than her expected wage.

In fact, we do observe in our data that individuals with education levels above the predicted levels (the undereducated) are the ones with greater underpredicted wages. A negative (positive) residual means that the expected education or wage is higher (lower) than what is observed. Over- (under-) education means that $v > 0$ ($v < 0$), and a wage premium (penalty) means that $u > 0$ ($u < 0$). Figure 5 presents the share of undereducated individuals by wage residual deciles for 1995 and 2015. The lower deciles have a disproportionately greater share of overeducated individuals, whereas the higher deciles have a disproportionately share of undereducated individuals.

As FKV (2013), we interpret u_i^* and v_i^* as measures of unobserved abilities.



Source: Authors' calculations based on PNAD (IBGE).

Notes: Education residuals are the \hat{v} 's; the wage residuals are given by \hat{u} , which is calculated as the difference between observed and predicted log wages when estimating δ .

Figure 5. Share of undereducated individuals by wage residual deciles (1995 and 2015).

The contribution of the unobserved abilities to wages and schooling depends on the individual's socioeconomic characteristics. However, after conditioning on the socioeconomic factors, the returns to unobserved ability is constant. These assumptions are plausible in developing country contexts. There is a large body of empirical literature for developing countries showing that families specialize the time allocation of their children across activities such as working and schooling (e.g., Emerson & Souza, 2007, 2008, Horowitz & Souza, 2010, among others). These findings are consistent with models of intrahousehold decisions among poor families. Dahan and Gaviria (2003) construct a model that shows that families may treat their children unequally even when they are identical if returns to human capital increase with the level of human capital, and parent's decisions are based on efficiency consideration. Horowitz and Wang (2004) develop a model of heterogeneous children and show that when the ability differences across children are great, families may reverse specialize such that the more talented child is allocated in the labor market earlier and the less talented one goes to school. In contexts such as this, one may observe "over education penalties" and "under education premiums" as presented in Figure 5.

Interestingly, one possible explanation to the change in the shares of under-educated by wage residuals deciles from 1995 to 2015 might be the schooling attainment expansion observed across cohorts as shown in Figure 1 from section 3.

6 Robustness check

In this section, we replicate the estimation of the returns to schooling using different approaches to examine the robustness of the findings presented in previous sections.

First, we perform two different tests, with results displayed in Table 5. The estimates presented in the second columns use a less rigid form for H_v and H_u , replacing the age and age squared variables by 5-year interval dummies. In the third column, we relax the multiplicative structure imposed by equation (6) by estimating a different version of equation (8), $H_j^2(x_j) = \exp(z\theta_j) + \varphi$, which can be estimated by OLS instead of Poisson regression. In this alternative setting, however, we lose the previous interpretation of ρ .

These tests are relevant because the identification strategy employed earlier in the estimation requires that the variables used in each specification of the equations generate enough and consistent heteroskedasticity. Thus, changing the functional form and the set of variables used in each heteroskedasticity function are key tests for the robustness of our findings.

Table 5. Control function estimates with alternative specifications: Wage equation (2015).

Variable	β_{CF} (benchmark specification)	β_{CF} (dummies specification)	β_{CF} (exponential heteroskedasticity)
YrsEduc	0.1211*** (0.0024)	0.1268*** (0.0031)	0.1221*** (0.0022)
ρ	-0.0591*** (0.0044)	-0.0474*** (0.0055)	-0.0599* (0.0374)
N	109,433	109,433	109,433
Adjusted R^2	0.3400	0.3398	0.3395

Source: Authors' calculations based on PNAD (IBGE).

Notes: Column 1 replicates results from Table 4, reporting estimates for equation (12). Results reported in column 2 uses a more flexible specification for H_j with only dummy variables, while in column 3 H_j is modeled as an exponential function of age, age squared, gender and race. Standard errors in parentheses. *significant at 10%; **significant at 5%; ***significant at 1%.

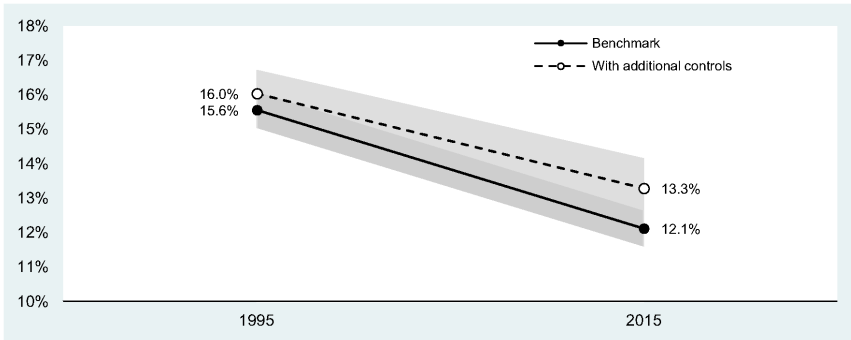
The estimates for these alternate specifications are 12.6% and 12.2%, respectively above and below the 12.1% estimate obtained under the benchmark specification. Note, however, that the difference between all these estimates are within the 95% confidence interval of each other. Thus, we can state that the results presented earlier are unaffected when using different functional forms for heteroskedasticity.

Second, we relax our “strictly exogenous covariates” premise. To be sure, we chose our parsimonious specification as a benchmark because adding controls that are (potentially) endogenous would weaken our assumptions. Variables such as labor market participation, married status, number of children etc. are all correlated with income, as well established in the literature.

On the other hand, however, PNAD has a large set of variables that are relevant controls in traditional set ups to wage and education equations. Therefore, we estimated our model for the years 1995 and 2015, with results displayed in Figure 6. The additional control variables were four dummies: one controlling for household status (chief of household equaled one), a second dummy controlling for migration, a third controlling for rural areas and a fourth as an additional race control (Asian).

The point estimates for the coefficient of interest in both years were higher than what was obtained in the benchmark specification. The declining slope, on the other hand, was lower, i.e. returns to schooling declined by a lesser amount when adding controls (-2.8pp versus -3.4 in the benchmark specification).

Interpretation of these results are not clear, but it is important to highlight the facts that (i) the overall pattern of declining returns persisted and (ii) confidence



Source: Authors' calculations based on PNAD (IBGE).

Notes: Each line depicts δ from equation (12) for the control function estimates, but the upper line equation included additional controls. Estimations are done separately for each year. All coefficients are significant at the 1% level in all years and shadowed areas indicate the 95% confidence interval.

Figure 6. Returns to education in Brazil: 1995 and 2015 in different specifications.

intervals for each year's estimates, in both specifications, overlapped. Our preferred approach, as highlighted before, is using strictly exogenous variables—but we recognize the importance of reporting these report these alternative estimates.

7 Conclusion

This paper estimates the causal returns to education for the Brazilian population during the period 1995–2015. The naïve OLS regression of earnings on years of schooling yields estimates with the well-known endogeneity bias. Klein and Vella (2010) developed a control function setting in which heteroskedasticity provides identification without the need for exclusion restrictions. The key advantage of this approach is the fact that estimation can be done for the entire population at any point in time, allowing for a more general application than the IV's LATE.

On the other hand, some additional assumptions need to be made. First, that heteroskedasticity is present in either the wage equation and/or in the schooling equation in different ways, such that the quotient of the two heteroscedastic functions is not constant. Second, the return on unobserved ability is constant after the differences in socioeconomic characteristics is accounted for. This simply means that the difference in returns to education arises from the inequality of the socioeconomic background. We argue that these two assumptions seem plausible, particularly in a developing country context, where the inequality of opportunities is latent.

One possible drawback of the KV (2010) method is the computational demands that arise due to their semiparametric estimators. Farré et al. (2013) propose a fully parametric approach that allows for the implementation of the KV (2010) estimator in practice. We apply this parametric approach and find that the average return to education have declined in Brazil from 15.6% in 1995 to 11.1% in 2014, and then bounced back to 12.1% in 2015, when the latest economic crisis began.

These estimates are higher than the OLS estimated coefficients suggest, pointing to a downward bias in the OLS estimation. We interpret this bias as a sign of under-education premiums that are likely to occur in environments where the more talented children are dropped from school and moved into the labor market earlier in life.

Finally, we also find a decline in the returns to schooling during the period, which seems to be associated with the well-documented increase in the supply of more educated workers observed in the past two decades in Brazil. This decline could be related to educational policies implemented by the Federal government during the period, such as PROUNI, FIES and other affirmative policies, which makes formal education a weaker signal of a worker's productivity.

References

- Angrist, J., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings. *Quarterly Journal of Economics*, 106(4), 979–1014. <http://dx.doi.org/10.2307/2937954>
- Azevedo, J. P., Inchaust, G., & Sanfelice, V. (2013). *Decomposing the recent inequality decline in Latin America* (Policy Research Working Paper No. WPS6715). The World Bank. <http://documents.worldbank.org/curated/en/597661468054543060/Decomposing-the-recent-inequality-decline-in-Latin-America>
- Barros, R., & Reis, J. G. A. (1991). Wage inequality and the distribution of education: A study of the evolution of regional differences in inequality in metropolitan Brazil. *Journal of Development Economics*, 36(1), 117–143.
- Ben-Moshe, D., D'Haultfœuille, X., & Lewbel, A. (2017). Identification of additive and polynomial models of mismeasured regressors without instruments. *Journal of Econometrics*, 200(2), 207–222. <http://dx.doi.org/10.1016/j.jeconom.2017.06.006>
- Bourguignon, F., Ferreira, F. H. G., & Menéndez, M. (2007). Inequality of opportunity in Brazil. *Review of Income and Wealth*, 53(4), 585–618. <http://dx.doi.org/10.1111/j.1475-4991.2007.00247.x>

- Cameron, S., & Taber, C. (2004). Estimation of educational borrowing constraints using returns to schooling. *Journal of Political Economy*, 112(1), 132–182. <http://dx.doi.org/10.1086/379937>
- Card, D. (1995). Earnings, schooling and ability revisited. In S. W. Polachek (Ed.), *Research in labor economics* (pp. 23–48). Greenwich, Connecticut: JAI Press.
- Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter & C. D. (Eds.), *Handbook of labor economics* (Vol. 3A, chap. 30). Amsterdam: Elsevier Science/North-Holland.
- Carneiro, P., & Lee, S. (2008). *Changes in college enrolment and wage inequality: Distinguishing price and composition effects*.
- Chen, S. H. (2008). Estimating the variance of wages in the presence of selection and unobserved heterogeneity. *The Review of Economics and Statistics*, 90(2), 275–289. <http://dx.doi.org/10.1162/rest.90.2.275>
- Daha, M., & Gaviria, A. (2003). Parental actions and siblings' inequality. *Journal of Development Economics*, 72(1), 281–297. [http://dx.doi.org/10.1016/S0304-3878\(03\)00077-4](http://dx.doi.org/10.1016/S0304-3878(03)00077-4)
- Dolton, P., & Vignoles, A. (2000). The incidence and effects of overeducation in the U.K. graduate labour market. *Economics of Education Review*, 19(2), 179–198. [http://dx.doi.org/10.1016/S0272-7757\(97\)00036-8](http://dx.doi.org/10.1016/S0272-7757(97)00036-8)
- Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4), 795–813.
- Emerson, P., & Souza, A. P. (2007). Child labor, school attendance and intra-household gender bias in Brazil. *The World Bank Economic Review*, 21(2), 301–316. <http://dx.doi.org/10.1093/wber/lhm001>
- Emerson, P., & Souza, A. P. (2008). Birth order, child labor, and school attendance in Brazil. *World Development*, 36(9), 1647–1644. <http://dx.doi.org/10.1016/j.worlddev.2007.09.004>
- Farré, L., Klein, R., & Vella, F. (2013). A parametric control function approach to estimating the returns to schooling in the absence of exclusion restrictions: An application to the NLSY. *Empirical Economics*, 44(1), 111–113. <http://dx.doi.org/10.1007/s00181-010-0376-5>
- Fernandes, R., & Menezes Filho, N. A. (2012). Educação, salários e a alocação de trabalhadores entre tarefas: Teoria e evidências para o Brasil. *Pesquisa e Planejamento Econômico*, 42(3), 311–328. <http://ppe.ipea.gov.br/index.php/ppe/article/view/1218>
- Figueirêdo, E., Nogueira, L., & Santana, F. L. (2014). Igualdade de oportunidades: Analisando o papel das circunstâncias no desempenho do ENEM. *Revista Brasileira de Economia*, 68(3), 373–392. <http://dx.doi.org/10.1590/S0034-71402014000300005>

- Fortin, B., & Raguéd, S. (2017). Does temporary interruption in postsecondary education induce a wage penalty? Evidence from Canada. *Economics of Education Review*, 58, 108–122. <http://dx.doi.org/10.1016/j.econedurev.2017.04.001>
- Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica*, 45(1), 1–22.
- Groot, W., & van den Brink, H. M. (2000). Overeducation in the labor market: A meta-analysis. *Economics of Education Review*, 19(2), 149–158. [http://dx.doi.org/10.1016/S0272-7757\(99\)00057-6](http://dx.doi.org/10.1016/S0272-7757(99)00057-6)
- Heckman, J. J., Urzua, S., & Vytlacil, E. (2006). Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics*, 88(3), 389–432. <http://dx.doi.org/10.1162/rest.88.3.389>
- Horowitz, A. W., & Souza, A. P. (2010). The impact of parental income on the intra-household distribution of school attainment? A measurement strategy and evidence. *The Quarterly Review of Economics and Finance*, 51(1).
- Horowitz, A. W., & Wang, J. (2004). Favorite son? Specialized child laborers and students in poor LDC households. *Journal of Development Economics*, 73(2), 631–642. <http://dx.doi.org/10.1016/j.jdeveco.2003.04.005>
- Ichino, A., & Winter-Ebmer, R. (2004). The long-run educational cost of World War II. *Journal of Labor Economics*, 22(1).
- Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 467–475.
- Klein, R., & Vella, F. (2009). A semiparametric model for binary response and continuous outcomes under index heteroscedasticity. *Journal of Applied Econometrics*, 24(5), 735–762. <http://dx.doi.org/10.1002/jae.1064>
- Klein, R., & Vella, F. (2010). Estimating a class of triangular simultaneous equations models without exclusion restrictions. *Journal of Econometrics*, 154(2), 154–164.
- Langoni, C. G. (1973). *Distribuição da renda e desenvolvimento econômico do Brasil: uma reafirmação* (Ensaio Econômico da EPGE No. 8). Rio de Janeiro: Instituto Brasileiro de Economia da Fundação Getúlio Vargas. <http://hdl.handle.net/10438/631>
- Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30(1), 67–80.
- López-Calva, L. F., & Lustig, N. (Eds.). (2010). *Declining inequality in Latin America: A decade of progress?* Washington, DC: Brookings Institution and UNDP.
- Lustig, N., López-Calva, L., & Ortiz-Juarez, E. (2013). Declining inequality in Latin America in the 2000s: The cases of Argentina, Brazil, and

- Mexico. *World Development*, 44, 129–141. <http://dx.doi.org/10.1016/j.worlddev.2012.09.013>
- Manacorda, M., Sanchez-Paramo, C., & Schady, N. (2010). Changes in returns to education in Latin America: The role of demand and supply of skills. *Industrial and Labor Relations Review*, 63(2), 307–326.
- Menezes-Filho, N. A., Fernandes, R., & Picchetti, P. (2006). Rising human capital but constant inequality: The education composition effect in Brazil. *Revista Brasileira de Economia*, 60(4), 407–424. <http://dx.doi.org/10.1590/S0034-71402006000400005>
- Mincer, J. (1974). *Schooling, experience, and earnings*. New York: Columbia University Press.
- Nocedal, J. (1996). Conjugate gradient methods and nonlinear optimization [Proceedings of the AMS-IMS-SIAM Summer Research Conference held at the University of Washington, July 1995]. In L. M. Adams & J. L. Nazareth (Eds.), *Linear and nonlinear conjugate gradient-related methods* (pp. 9–23). SIAM.
- Oreopoulos, P. (2007). Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics*, 91(11–12), 2213–2229. <http://dx.doi.org/10.1016/j.jpubeco.2007.02.002>
- Psacharopoulos, G., & Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education economics*, 12(2), 111–134.
- Rubb, S. (2002). Overeducation in the labor market: A comment and re-analysis of a meta-analysis. *Economics of Education Review*, 22(6), 621–629.
- Saniter, N. (2012). *Estimating heterogeneous returns to education in Germany via conditional heteroskedasticity* (IZA Discussion Paper No. 6813). Institute for the Study of Labor (IZA). <https://www.iza.org/de/publications/dp/6813/estimating-heterogeneous-returns-to-education-in-germany-via-conditional-heteroskedasticity>
- Schroeder, E. (2010). *The impact of microcredit borrowing on household consumption in Bangladesh*.
- Senna, J. J. (1976). Escolaridade, experiência no trabalho e salários no Brasil. *Revista Brasileira de Economia*, 30(2), 163–194. <http://bibliotecadigital.fgv.br/ojs/index.php/rbe/article/view/162/6590>
- Tannen, M. B. (1991). New estimates of the returns to schooling in Brazil. *Economics of Education Review*, 10(2), 123–135. [http://dx.doi.org/10.1016/0272-7757\(91\)90004-9](http://dx.doi.org/10.1016/0272-7757(91)90004-9)
- Tavares, P. A., & Menezes-Filho, N. A. (2011). Human capital and the recent decline of earnings inequality in Brazil. *Brazilian Review of Econometrics*, 31(2), 231–257. <http://dx.doi.org/10.12660/bre.v31n22011.7236>
- Teixeira, W. M., & Menezes-Filho, N. A. (2010). Estimando o retorno à educação do Brasil considerando a legislação educacional brasileira como

- um instrumento. *Revista de Economia Política*, 32(3), 479–496. <http://dx.doi.org/10.1590/S0101-31572012000300008>
- Vella, F., & Gregory, R. G. (1996). Selection bias and human capital investment: Estimating the rates of return to education for young males. *Labour Economics*, 3(2), 197–219. [http://dx.doi.org/10.1016/0927-5371\(96\)00006-1](http://dx.doi.org/10.1016/0927-5371(96)00006-1)
- Wang, Y. (2010). *A control function approach to estimating the returns to schooling in urban China*.
- Zhu, C., Byrd, R. H., Lu, P., & Noceda, J. (1997). Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software (TOMS)*, 23(4), 550–560. <http://dx.doi.org/10.1145/279232.279236>

Appendix

Table A1. Heteroskedasticity test: Education equation (2015).

Breusch-Pagan/Cook-Weisberg test for heteroskedasticity		
H_0 : Constant variance		
Variable	Chi ²	p
Age/100	2,647.31	0.0000
(Age/100) ²	2,556.01	0.0000
Female	125.37	0.0000
White	24.47	0.0000
Simultaneous	3,607.15	0.0000

Source: Authors' calculations based on PNAD 2013 (IBGE).

Note: State of birth dummies were also included in the regression.

Table A2. Heteroskedasticity test: Wage equation (2015).

Breusch-Pagan/Cook-Weisberg test for heteroskedasticity		
H_0 : Constant variance		
Variable	Chi ²	p
YrsEduc	698.05	0.0000
Age/100	816.62	0.0000
(Age/100) ²	789.74	0.0000
Female	0.30	0.5869
White	136.02	0.0000
Simultaneous	2,470.35	0.0000

Source: Authors' calculations based on PNAD 2013 (IBGE).

Note: State dummies were also included in the regression.

Table A3. Heteroskedasticity functions: education and wage equations (2015).

Variable	θ_v	θ_u
Age/100	8.4807*** (0.4729)	0.5529 (3.3782)
(Age/100) ²	-7.3240*** (0.5762)	-0.0426 (4.1958)
Female	-0.0954*** (0.0079)	-0.4248*** (0.0662)
White	0.01228 (0.0009)	-1.7035*** (0.1956)
Constant	0.5523*** (0.1030)	1.7985*** (0.0737)
N	109,433	109,433
Pseudo R^2	0.0423	0.3400

Source: Authors' calculations based on PNAD 2013 (IBGE).

Notes: Poisson regression estimates for θ_v (equation (11)) are depicted in column (I) and Nonlinear Least Squares for θ_u (equation (12)) are depicted in column (II). State of birth dummies were included in the regression. Standard errors in parentheses. *significant at 10%; **significant at 5%; ***significant at 1%.