GIGA
Working Papers

German ■ Institute of Global and Area Studies
Leibniz-Institut für Globale und Regionale Studien

The Geography of Intergenerational Mobility: Evidence of Educational Persistence and the "Great Gatsby Curve" in Brazil

Abstract

This paper explores the variation in intergenerational educational mobility across the Brazilian states based on Markov transition matrixes and univariate econometric techniques. The analysis of the national household survey (PNAD-2014) confirms a strong variation in mobility among the 27 federative units in Brazil and demonstrates a significant correlation between mobility and income inequality. In this sense, this work presents empirical evidence for the existence of the "Great Gatsby curve" within a single country: states with greater income disparities present higher levels of persistence in educational levels across generations. Finally, I investigate one specific mechanism behind this correlation – namely, whether higher income inequality might lead to a lower investment in human capital among children from socially vulnerable households. The paper delivers robust and compelling results showing that children born into families where the parents have not completed primary education have a statistically significant reduction in their chance of completing the educational system if they live in states with a higher level of income inequality.

Tharcisio Leone

is an economist, a research fellow at the GIGA Institute of Latin American Studies, and a member of the GIGA's Research Programme 3 "Growth and Development." He is currently working on his doctorate, which examines intergenerational social mobility, at the Free University of Berlin. His research focuses on educational economics, intergenerational mobility, equality of opportunity, and social inequality.

<tharcisio.leone@giga-hamburg.de>
<www.giga-hamburg.de/en/team/leone>
<sites.google.com/view/tharcisioleone/home>

The Geography of Intergenerational Mobility: Evidence of Educational Persistence and the "Great Gatsby Curve" in Brazil

Tharcisio Leone

Article Outline

1 Introduction

Empirical evidence from cross-country comparisons has revealed a negative correlation between intergenerational mobility and income inequality: countries with greater income disparity tend to have lower levels of economic mobility between generations (Björklund and Jäntti 2009; Blanden 2013; Corak 2006; Ermisch et al. 2012; Smeeding et al. 2011).[1] The so-called

---

"Great Gatsby curve" illustrates the transmission of income inequality across generations and underlines the fact that the higher the level of inequality in one generation, the more children's chances of economic success depend on whether they have poor or rich parents (Boudreaux 2014; Chetty et al. 2014b; Corak 2013a; Jerrim and Macmillan 2015; Mazumder et al. 2015).

The original Great Gatsby curve was based on research conducted at the international level, using cross-country comparisons. However, some authors have questioned the results, owing to the poor comparability of the data across countries (Andrews and Leigh 2009; Chetty et al. 2014a; Güell et al. 2018; Jantti and Jenkins 2013; Jerrim and Macmillan 2015). The demonstration of equivalence (lack of bias) is an important criteria for any cross-regional comparison in order to provide empirical findings free from differences in the data construction across countries. For this purpose, studies that address the lack of suitable data represent an important and beneficial contribution to international research (Andrews and Leigh 2009; Boudreaux 2014).

This paper is intended primarily to expand the available literature by providing a Great Gatsby curve free of comparability bias, in which the correlation between income inequality and intergenerational mobility is analysed across different regions within a single country, using observations recorded and consolidated in a single database.[2]

The investigation of intergenerational mobility within Brazilian states in this study is based on the educational attainment of children and their parents and applies the recently published Mobility Supplement from the nationally representative Brazilian household survey (PNAD-2014).[3] The case of Brazil, with its continental dimensions and widespread regional and social inequalities, is a very promising area for research. The country has one of the highest levels of income inequality in the world and at the same time a significant variation in inequality across the 27 states (see Figure 1).[4] The income inequality – as measured by the Gini coefficient – varied in the year 2014 from 0.416 in Santa Catarina to 0.577 in Distrito Federal.

I focus on state-level variation because in Brazil the responsibility for the provision of primary and secondary education lies with the states. According to the Law of Directives and Bases of National Education (LDB), the current legislation that regulates the education system in Brazil, the tasks of the federal government in relation to primary and secondary public education are restricted to providing technical and financial support to the states and municipalities, thereby guaranteeing the equalisation of opportunities and a minimum level of quality.[5]

---

2 This approach has already been adopted by Güell et al. (2018) for Italy and Bradbury and Triest (2016), Chetty et al. (2014a), and Kearney and Levine (2014) for the United States. They analysed single data sets and found that the correlation between intergenerational mobility and inequality also holds true across provinces in these countries.

3 The microdata from PNAD-2014 were only made freely available for research in November 2016.

4 To be more precise, Brazil comprises 26 states plus a federal district (Distrito Federal), where the federal capital is located. See the figures in the Appendix for a clear picture of the variation in income inequality across the Brazilian states.

5 The Appendix provides a more comprehensive and detailed overview of the Brazilian educational system.

Despite the increasing scientific interest in the Great Gatsby curve, far too little is known about the causal link between inequality and intergenerational mobility, because only limited research has been undertaken on the determinants of this correlation (Jerrim and Macmillan 2015). In the final part of this paper, I seek to fill this research gap by focusing on a possible mechanism through which inequality might affect intergenerational mobility – namely, curtailed investment in education. Kearney and Levine (2014) propose that a greater level of inequality could lead to an underestimation of the return on investment in human capital for children from socially vulnerable families, which would increase their school drop-out rates, thereby decreasing their chances of mobility.

The paper is structured as follows. Section 2 reviews the related literature and presents the econometric models used as the theoretical basis for the investigation. Section 3 presents the database. In Section 4, I then describe the three different empirical approaches applied in the paper. Section 5 deals with the empirical findings. I first estimate the level of intergenerational educational mobility in the 27 Brazilian states, then I correlate the results from mobility with income inequality. Finally, I apply an econometric model to investigate whether (socially) vulnerable children living in states with a higher gap between the middle and the bottom of the income distribution have a greater probability of leaving school without a certificate. Section 6 concludes with a summary of the key findings.[6]

2 Theoretical Background and Literature Review

The term "intergenerational mobility" describes the ability of children to move beyond their social origins and achieve a socio-economic status that is not dictated by that of their parents (Fox et al. 2016; Ribeiro 2007). In the mobility literature, the focus of the economic investigations is the measurement of the correlation between parents' and children's economic outcomes in terms of income, education, or occupation (Blanden and Macmillan 2014; Corak et al. 2014; Hills et al. 2015). The greater this association, the greater the economic advantages and disadvantages inherited from the family background (Schneebaum et al. 2016).

The scientific community has been working for a long time on a framework for understanding the transmission of economic outcomes from parents to their offspring (Blanden et al. 2014; Black and Devereux 2010). The studies of Solon (1992) and Zimmerman (1992) were the precursors to the modern empirical estimations of intergenerational correlation of outcomes (Björklund and Jäntti 2009; Blanden et al. 2014; Ichino et al. 2011). In the subsequent years, motivated mainly by Solon's (2004) theoretical contribution, several researchers around

---

6   This paper is supplemented by a comprehensive Appendix with relevant information concerning the educational system in Brazil, the data harmonisation, the codification process for the variables, additional figures, and the formal description of the underlying theoretical models.

the world have begun to investigate the persistence of income, wealth, consumption, and education between parents and their children (see e.g. Ayala and Sastre 2008; Blanden 2013; Bratsberg et al. 2007; Chen 2009; Corak et al. 2014; Dunn 2007; Roemer 2004; Ueda 2009). In a second stage of the literature, researchers have focused on the variation of intergenerational mobility over time (see e.g. Aaronson and Mazumder 2008; Björklund et al. 2009; Hertz et al. 2007; Hout and Guest 2013; Lee and Solon 2009; Mazumder et al. 2012) and across countries (see e.g. Aaberge et al. 2002; Ayala and Sastre 2008; Blanden 2013; Blanden et al. 2014; Corak 2006; Jäntti et al. 2006).[7]

An analysis of these works shows that two different research methodologies have primarily been used to measure intergenerational mobility in the economic literature: the first approach focuses on income and the second on educational attainment.[8] Given the limited availability of lifetime income data – especially in developing countries (Azam and Bhatt 2015; Ferreira and Veloso 2006) – an increasing number of authors have used the strong positive correlation between education and income to measure mobility across generations. This approach is justified by the solid set of studies and empirical evidence which indicate that educational inequality plays a determining role in the transmission of inequalities across generations, making it a robust indicator for future trends in income inequality (Blanden and Macmillan 2014).

Only more recently has the economic literature dealt with the mechanisms behind the intergenerational persistence in outcomes (Black and Devereux 2010; Rothwell and Massey 2015). Corak (2006) was the first to provide empirical evidence of a negative correlation between intergenerational mobility and income inequality (Kearney and Levine 2014). Based on cross-country comparisons and Solon's (2004) theoretical approach, the author showed that countries with greater income disparity tend to exhibit lower levels of economic mobility between generations.

It didn't take long for Corak's (2006) findings to enter the political debate. In his speech as chairman of President Barack Obama's Council of Economic Advisers, economics professor Alan Krueger (2012) introduced the "Great Gatsby curve," and within a short space of time this curve gained a prominent position in the international economic community (Jerrim and Macmillan 2015). It has been mentioned by Nobel Prize winners (see e.g. Heckman 2013) and has been extensively addressed by the mainstream press (see e.g. Economist 2013; The Guardian 2012) and high-ranking policymakers (see e.g. Obama 2013; White House 2013). Furthermore, the Great Gatsby curve has also been addressed in a long list of recent publications in peer-reviewed journals (see e.g. Boudreaux 2014; Brahim and McLeod 2016; Chetty et al.

---

7   Black and Devereux (2010), Fox et al. (2016), Jantti and Jenkins (2013), and Hills et al. (2015) offer a detailed discussion of recent developments in the literature on intergenerational mobility.

8   A third approach, found especially in sociological studies, measures the degree of intergenerational mobility using the professional occupations of parents and their children (see e.g. Pastore and do Valle Silva 2000; Reddy 2015; Xie and Killewald 2013).

2014a,b; Corak 2013a,b; Fan et al. 2015; Güell et al. 2018; Jerrim and Macmillan 2015; Lefgren et al. 2015; Mazumder et al. 2015).

The negative relationship between inequality and intergenerational mobility illustrated by the Great Gatsby curve is also supported by the economic theory. Becker and Tomes (1986), Breen and Jonsson (2005), Corak (2013a), Duncan and Murnane (2011), and Solon (2004) are just some examples of authors who have argued that the disparities in the investment in children's human capital across families increase with the growth of income inequality. Solon (2004), for example, adapted the classical model of Becker and Tomes (1979, 1986) in a detailed theoretical model presenting the intergenerational transmission of inequality and demonstrated on the basis of a mathematical approach that higher-income parents have a higher capacity to invest in the human capital of their children, and they are also more inclined to make this investment if the expected earnings return on human capital increases.[9]

However, Solon's (2004) model has been used in the economic literature only as a starting point for understanding the variation in the intergenerational persistence of outcomes across countries and over time. The Great Gatsby curve does not present a causality link between inequality and mobility, but rather a summary of all mechanisms reflecting the outcome of a host of ways that income inequality affects children's development (Corak 2013a; Kearney and Levine 2014).

Recent research has offered a vast amount of evidence that childhood development has direct effects on adult economic productivity (Cunha et al. 2006; Knudsen et al. 2006; Phillips et al. 2000). Socially vulnerable families lack the socio-economic resources to provide effective early development for their children. Therefore, these children are exposed from a very young age to adverse environments, leading to skill and ability deficits that result in low productivity in the future (Lawrence et al. 2005; Shonko and Meisels 2000). Also, during adult life, children continue to benefit from the resources of their family. Social connections, for example, play an important role in mobility chances. Children from wealthy families can use the extensive network of their parents to climb the economic ladder, which means they have an advantage relative to children from low-income households (Corak 2013a).

Despite this complexity, the variation in the intergenerational persistence of economic outcomes presented by the Great Gatsby curve calls for us to reflect on the reasons for the different levels of mobility, and how these underlying drivers can influence the ultimate outcomes. To address these questions, it is important to bear in mind the three fundamental institutions that play a strong role in children's chances of mobility: the family, the labour market, and the state (Corak 2013a).

---

9   The Appendix provides a formal description of Solon's (2004) model, which shows how income inequality can affect the chances of intergenerational mobility.

As described in Solon's model (2004), the income inequality resulting from the labour market impacts the financial capacity and the incentives for investment in the human capital of children across families. The individual capabilities of children are also strongly influenced by non-monetary resources, such as the behavioural patterns, motivations, and social connections which are transmitted in the family environment and play an important role in mobility potential. Finally, the importance of public policy for intergenerational mobility relates to all key aspects that affect the interaction between families and the labour market, such as taxation and regulatory structures (Björklund and Jäntti 2009; Corak 2013a).

Assuming that the investigation of the causal effects of mobility is viewed with increasing mistrust in the academic community – due to the methodological difficulties of measuring causation within the intergenerational persistence framework (Björklund and Jäntti 2009; Chetty et al. 2014a; Fessler and Schneebaum 2012) – this paper isn't, in principle, looking for causal relationships, but rather aims to identify stylised facts and trends, thereby improving our understanding of the mechanisms behind the correlation between income inequality and the persistence of economic outcomes across generations represented by the Great Gatsby curve. Consequently, the empirical approach applied in the second part of this paper resembles that of Kearney and Levine's (2014) renowned paper.

Kearney and Levine (2014) proposed curtailed investment in human capital as an important channel via which an increase in income inequality may adversely affect the mobility chances of the younger generations. According to the authors, an increase in the gap between the bottom and the top of the income distribution could change the expected return on human capital investment for children from socially disadvantaged families. In this case, children born into poverty generally do not believe that a school-leaving qualification will help them move up the economic ladder, which thus reinforces their economic marginalisation.[10] Based on a formal econometric model and five sources of individual-level data for the USA, the paper confirmed the hypothesis that low-income youths are more likely to drop out of school if they live in a place with greater income inequality.

3 Data

The data for this study stems from the Brazilian National Household Sample Survey (PNAD), which is a representative household survey conducted annually by the Brazilian Institute of Geography and Statistics (IBGE) to collect socio-economic and demographic information about the Brazilian population, including household composition, education, labour, income, migration, and fertility.[11]

---

10 See Appendix for a formal description of Kearney and Levine's (2014) model concerning the decision to drop out of school.

11 In 2014 the PNAD's sample consisted of 151,291 households with 362,627 individuals.

To investigate mobility, I use the data wave from PNAD's Socio-Occupational Mobility Survey. Every year the PNAD investigates an additional topic on the basis of the "Supplementary Survey," and in year 2014 its focus was socio-occupational mobility. For the survey, respondents 16 years and older were asked to provide information about their parents' professional occupation and level of education.[12]

The two main outcomes of interest in this paper are years of schooling and levels of education, for both children and parents.[13] The educational levels are classified into four identical categories: no school certificate and primary, secondary, and tertiary education, with primary education referring to the compulsory education.[14]

Given that the PNAD does not provide the number of years of schooling for the parents, I calculated this variable according to the information about the highest level of education attended.[15] In addition, information about gender, year of birth, location of residence (rural or urban areas), and whether the respondent grew up in a two-parent family are used as control variables. Finally, I use variables related to income to estimate the indicators of inequality.

I excluded individuals under 25 years old from the sample, given that approximately 42 per cent of them were still attending school, training, or university in 2014. Similarly, I excluded persons over 75 years of age due to the positive correlation between education and life expectancy. Consequently, this paper considers people born between 1940 and 1989 in the empirical analysis and works with a sample of 46,051 individuals.

I inserted a dummy variable for "economic marginalisation," which refers to children from parents with no school certificate, in the sample. Finally, the observations were categorised into 10-year birth cohorts (1940–1949, 1950–1959, 1960–1969, 1970–1979, and 1980–1989) in order to minimise the lifecycle bias resulting from the variation in average years of schooling and in education dispersion over time (see figures 2 and 3).[16]

For the indicator of income inequality, I created the continuous variable "75/10 ratio," which represents the relation between the income of the richest 25 per cent and the poorest 10 per cent of the income distribution. Given that the measures of inequality were determined

---

12 The information about the education and occupation of parents refers to the level when the respondents were 15 years old.

13 In those cases where the educational level of the father and mother is known, this paper will use the educational attainment of the most educated parent in the empirical estimations.

14 In practical terms, primary education in this paper means the minimum years of schooling required by law when the children and parents were of school age. Currently, the compulsory education in Brazil ends at the age of 17. See Appendix for overview of the changes in compulsory education over time.

15 Please see the Appendix for a detailed description of the codification process.

16 For the empirical estimations, this paper applies the weights presented in the sample, representing the inverse of the probability of an observation being selected into the sample.

retrospectively for the year in which the individuals (should) have completed compulsory education, I used the earlier PNAD sample surveys for the calculations.[17]

Table 1 reports the summary statistics on income distribution, educational attainment, average age, and share of rural population divided by the states and macro-regions of Brazil. Figures 1 and 2 visually display the development over time of average schooling and its standard deviation, respectively. Figures 3 and 4 focus on the sample as a whole, showing the variation in schooling and educational levels across the 27 Brazilian states.

Table 1 shows the average educational attainment of children and parents. Note that in all states the average schooling of children is higher than that of their parents, and that the mothers are almost always more educated than their spouses.[18] Columns (7) to (9) list the proportions of people who were enrolled in school in 2014. According to data from PNAD-2014, all states in Brazil are close to the objective of achieving universal education for children between 7 and 14 years old.[19] However, beyond the age of primary education, the deviation in the net enrolment ratio across states increases significantly. The proportion of children aged 15–17 enrolled in school is lowest in Roraima (0.758) and highest in the Distrito Federal (0.895) and in the south-eastern states, such as São Paulo (0.864), Minas Gerais (0.867), and Rio de Janeiro (0.874). Moreover, the variation in the share of adults between 18 and 24 years who are still attending school, training, or university is even greater. This ratio ranges from 0.263 in Pernambuco to 0.414 in Distrito Federal.

Figure 3 indicates significant differences in average educational attainment across Brazilian states. In southern states such as São Paulo, Rio de Janeiro, and Santa Catarina, children have higher average years of schooling relative to those in the states in the north-east. Finally, Figure 4 illustrates the main reason for these differences in average education: The share of individuals with no school-leaving certificate in the north-eastern states is greater than in the other macro-regions of Brazil.

4 Conceptual Framework

This paper employs a two-step empirical framework. I start by measuring intergenerational persistence in education, using linear regression models (Checchi et al. 2013) and transition matrices (Jäntti et al. 2006). I then apply an econometric method to investigate whether children from disadvantaged families have a lower chance of completing secondary education (Kearney and Levine 2014).

---

17 Please see Section 4.2 for the empirical background and the Appendix for a full description of the harmonisation process that needed to be undertaken in order to fit the data over time.

18 The exceptions are Amapá, Espírito Santo, Rio de Janeiro, São Paulo, Paraná, Santa Catarina, Rio Grande do Sul, and Mato Grosso do Sul, where the average education of fathers is higher than that of mothers.

19 The estimated values varied between 0.965 in Acre and 0.994 in São Paulo.

4.1 Intergenerational Educational Mobility

A. Mobility Matrices

Following Daouli et al. (2010), this section classifies the educational outcomes of children (generation t) and parents (generation t+1) into four categories: no school certificate and primary, secondary, and tertiary education. Thereafter, I estimate the intergenerational transition matrices $\mathbb{P}$ with the number of states S, such that:

$$p_{ij} = \mathbb{P}(X_{t+1} = j \mid X_t = i) \qquad \text{for} \qquad i,j \in S, \quad t = 0,1,2,\dots \qquad (1)$$

The estimated transition matrices present two important properties:

$$\forall \ i,j \ \varepsilon \ \mathbb{R}, \quad p(i,j) \geq 0, \qquad (2)$$

and

$$\sum_{j=1}^{N} p_{ij} = \sum_{j=1}^{N} \mathbb{P}(X_{t+1} = j \mid X_t = i) = \sum_{j=1}^{N} \mathbb{P}_{\{X_t = i\}}(X_{t+1} = j) = 1. \qquad (3)$$

In transition matrix $\mathbb{P}$, the value of $p_{i,j}$ denotes the proportion of children from parents with the educational attainment j who achieved the education level i. Given that the estimations are based on identical education levels for children and their parents, the diagonal cells from the square matrices $\mathbb{P}$ represent immobility or inheritance in the intergenerational transition from state j to state i (Altham and Ferrie 2007; Reddy 2015; Xie and Killewald 2013). Consequently, the "immobility ratio" (ImR) can be calculated as a percentage of the sum total of all entries on the main diagonal of the matrix $\mathbb{P}$ and its number of states S (Heineck and Riphahn 2007):

$$\text{ImR} = \frac{Tr(\mathbb{P})}{S} = \frac{\sum_{i=1}^{N} p_{ij}}{S} \qquad (4)$$

Following Corak et al. (2014), I describe upward and downward mobility – UpM and DoM, respectively – as the probability that the children's level of education exceeds or is less than the parents' educational level l.[20]

$$UpM = \Pr(X_t > l \mid X_{t+1} = l) \quad \text{and} \quad DoM = \Pr(X_t < l \mid X_{t+1} = l) \qquad (5)$$

In order to summarise the degree of mobility intrinsic in transition matrix $\mathbb{P}$, allowing for a ranking of the Brazilian states according to mobility levels, I follow Checchi et al. (1999) and Daouli et al. (2010) and calculate the Prais–Shorrocks indicator based on the trace ($Tr(\mathbb{P})$) and the number of states in the transition matrix.

---

20 From a graphical point of view, the downward (upward) mobility is derived from the values of the elements below (above) the main diagonal of the square matrix $\mathbb{P}$ (Heineck and Riphahn 2007).

$$M_{PS}(\mathbb{P}) = \frac{S - Tr(\mathbb{P})}{S - 1} \qquad \text{with} \quad M_{PS} \in [0,1] \qquad\qquad (6)$$

The M_PS ($\mathbb{P}$) provides a measure of the normalised distance between the identity matrix and the independent matrix. It ranges from 0 to 1, with values closer to one indicating a higher level of intergenerational educational mobility.[21]

B. Linear Regression Models

Following the standard empirical model presented in the economic literature on intergenerational mobility (e.g. Black and Devereux 2010; Blanden 2013; Hertz et al. 2007), this paper estimates the educational persistence between parents and children with the regression equation:

$$educ_{is}^{c} = \alpha + \beta\, educ_{is}^{p} + \varepsilon_i \qquad \text{for} \qquad i = 1,2,\dots,N \qquad\qquad (7)$$

where $educ_{is}^{c}$ is the years of schooling of a child i resident in the state s, and $educ_{is}^{p}$ denotes the same variable for his or her parents. The error term ε_i reflects the combined effects on a child's education of factors orthogonal to parental education, and the slope coefficient β is the parameter of interest, representing the elasticity of children's education with respect to their parents' education. The coefficient β is commonly known in the economic literature as the "regression coefficient" and gives the value of each 1 per cent difference in parental education across families that will be transmitted as an educational difference to their children (Blanden 2013).

Given the variation in standard deviations across states and time in Brazil, as shown in Figure 2, I follow Azam (2016) and Checchi et al. (2013) and normalise the years of schooling in equation (7) by the corresponding standard deviation. The OLS estimate of β is given by:

$$\hat{\beta} = p_s^{cp} = \frac{\sigma_s^c}{\sigma_s^p}, \qquad \text{with} \qquad \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_1 - \mu)^2} \qquad\qquad (8)$$

where σ_s^c and σ_s^p correspond to the standard deviation in education for children and parents in state s, while the coefficient σ_s^{cp} captures the association between children's and parents' education, respectively. Based on equations (7) and (8), the resulting empirical model can be summarised as:

$$\frac{educ_{is}^{c}}{\sigma_s^c} = \delta + \rho\left(\frac{educ_{is}^{p}}{\sigma_s^p}\right) + \varepsilon_i \qquad \text{with} \qquad \rho \in [0,1] \qquad\qquad (9)$$

In this regard, the coefficient ϱ is defined in the economic literature as the "relative" measure of intergenerational mobility or the "correlation coefficient." The higher its value, the stronger the correlation between the educational attainment of children and parents.

---

21 A $M_{PS}(\mathbb{P}) = 1$ would mean that the probability that children will end up with education level i is independent of the parents' educational attainment j (full "equality of opportunity"). In contrast, a $M_{PS}(\mathbb{P}) = 0$ corresponds to an "identity matrix" in which all the main diagonal elements are one and all the remaining elements are zero, indicating a perfectly immobile society (Chevalier et al. 2003).

Given that the estimations are based on the pooled sample, equation (9) includes a vector of dummy variables UF with the state of residence i. Moreover, I use a vector X comprising controls for gender, race, and year of birth. Finally, some interaction terms between the variables are assumed. Thus, the resulting fully interacted model takes the following form:

$$\frac{educ_{is}^c}{\sigma_s^c} = \delta + \rho \left(\frac{educ_{is}^p}{\sigma_s^p} \right) + \eta \left(\frac{educ_{is}^p}{\sigma_s^p} \text{ x } UF_i \right) + \lambda UF_i + \gamma(X_i \text{ x } UF_i) + \varepsilon_i \tag{10}$$

4.2 Linking Inequality and School Dropouts

In this section, I follow Kearney and Levine (2014) and apply a probit model aimed at investigating whether children from marginalised socio-economic backgrounds living in states with greater income inequality levels have a lower chance of completing secondary education.

In this underlying latent model, the observed binary response $ComSec_{i,t}$ assumes the value 1 if the ith individual born in year t has completed secondary education and this is a function of socio-economic background, income inequality in the state of residence, and individual characteristics. Thus, the empirical probit model can be written as

$$\begin{aligned} ComSec_{i,t} = {} & \pi_0 + \pi_1(MSB_i \text{ x } Ineq_{s,t+14}) + \pi_2 \, MSB_i + \pi_3 \, Ineq_{s,t+14} + \gamma_1 male_i \\ & + \gamma_2 rural_i + \gamma_3 bothP_i + \gamma_4 race_i + \gamma_5 birth_i + \varepsilon_i \end{aligned} \tag{11}$$

The (marginalised) socio-economic background is summarised in the variable MSB_i, which represents individuals from (two) parents with no school certificate. The variable $Ineq$ refers to income inequality, measured by the 75/10 ratio, in the individual's state (s) of residence 14 years after their birth ($t + 14$).[22] The model also includes controls for gender ($male$), location of residence ($rural$), self-declared race/ethnicity ($race$), and birth year ($birth$), as well as a dummy indicating whether the children lived with both parents in the same household at age 15 ($bothP$).

The parameter π_1 estimated from the interaction term between the continuous variable $Ineq_{s,t+14}$ and the discrete (binary) variable MSB_i is the main coefficient of interest and indicates whether individuals with a lower family-education background living in states with high income inequality have a lower probability of completing secondary education. In order to present a more informative view of the expected changes in the educational outcome of children as a function of changes in the explanatory variables (economic background and income inequality), the marginal effects are estimated from equation (11) as:

---

22 Following Solon's (2004) theoretical model, what is particularly relevant for the accumulation of human capital is the level of income inequality when children have completed their compulsory education and are facing a decision about whether or not to pursue more years of schooling. Given that until the year 2009, education in Brazil was compulsory for children aged 7 to 14 years, the equation (11) uses the 75/10 ratio from the year in which the individual turned 14 as a measure of inequality.

$$\frac{\partial E(ComSec \mid x)}{\partial x}\Big| \; = \; \frac{\partial F(x\beta)}{\partial x}\Big| \; = \; f(\tilde{x}\beta)\beta \qquad (12)$$

For the categorical variables, the marginal effects indicate how $ComSec_{i,t}$ is predicted as MSB_i changes from 0 to 1, holding all the other covariates constant at their average values, while for the continuous variable $Ineq_{s,t+14}$, the results from the marginal effects indicate how much the increase in the inequality ratio will change children's probability of achieving a secondary education.

5 Empirical Results

This section presents the study's empirical findings. I start with the estimation of intergenerational educational mobility based on the transition matrix and the linear regression model. This is followed by the results on whether mobility at the state level is correlated with income inequality. Section 5.3 deals with one important mechanism behind the relationship between inequality and mobility illustrated by the Great Gatsby curve – namely, whether greater income inequality contributes to a higher school-dropout rate for economically marginalised children.

5.1 Intergenerational Educational Mobility

Mobility matrices and linear regression models have been widely used in the economic literature to measure the extent of intergenerational educational mobility. These two empirical approaches complement each other and together provide a more detailed picture of mobility. The regression model takes into account the variation in standard deviation across both generations and presents a degree of mobility free from bias that can be caused by an increase in average education over time. The transition matrix approach, in comparison, has the advantage of providing a more comprehensive overview of the direction of the mobility (Corak and Heisz 1999; Dearden et al. 1997; Fields 2002).

A. Mobility Matrices

Figure 6 measures children's probability of attaining a certain educational level as a function of parents' education. If we analyse the four charts together, we see only minimal changes over time in the intergenerational persistence of education in Brazil. Note that regardless of birth year, the chance of attaining higher education is strongly correlated with the parents' educational background. In summary, it is possible to state that the children of more highly educated parents tend to become more highly educated adults, while the children of less educated parents tend to become adults with less education.

However, the data clearly show that the probability of attaining the compulsory level of education has increased considerably over time. As can be seen in Figure 6, the proportion of people with no school certificate and only primary education is becoming increasingly smaller.

Following on this brief description of the development of mobility over time, I now turn to the variation in the intergenerational persistence of education across the Brazilian states. Figure 7 presents the direction of mobility, displaying the results of equations (4) and (5). Figure 8 places the states in increasing order, according to the degree of mobility estimated from equation (6).

Figure 7 illustrates the two different directions in mobility. Individuals who achieve a higher educational level than their parents move upward on the educational scale, while downward mobility refers to the cases where the children's level of schooling remains lower than that of their parents. In Brazil, 38.8 per cent of children have achieved a higher level of education than their parents, while only around 15 per cent have experienced downward mobility.

However, these values vary strongly across the states. Paraiba is the state in Brazil with the highest level of intergenerational immobility in education (49.1 per cent), approximately 12 percentage points more than the results obtained in Rio Grande do Norte, the state with the lowest level of persistence in education across generations (37.3 per cent). The levels of upward mobility exhibit even greater variation across the states, from 30.4 per cent in Pará to 52.1 per cent in Distrito Federal.

Figure 8 ranks the Brazilian states on the basis of the Prais-Shorrocks indicator from equation (6) and provides more detailed information on the movement of children within the education distribution. The red circles indicate the ratio of children from parents with no school-leaving certificate who have successfully completed tertiary education, representing the maximum possible degree of upward mobility. The indicator "top (bottom) persistence" displays the proportion of children from parents with tertiary education (no certificate) who have achieved the same educational level as their parents.

The bottom persistence shows the lack of mobility at the lowest extreme of the transition matrix. In Brazil, nearly half of children (49.2 per cent) from parents without a school certificate have not completed (primary) education, highlighting once again the strong intergenerational persistence in educational levels. This value also varies strongly across the states, from 32.2 per cent in Distrito Federal to 69.7 per cent in Piauí. Figure 8 indicates that the chances of ascending from the bottom of the education distribution are especially low for individuals living in the north-eastern states.[23]

With a Prais-Schorrocks index equal to 0.836, Rio Grande do Norte (RN) leads the Brazilian rankings for intergenerational mobility. The main reason for this is that RN exhibits very low persistence at the top of distribution. Only 15.1 per cent of children from parents with a tertiary education achieved a college degree. By way of comparison, this value is 92.4 per cent in Distrito Federal, 82.3 per cent in Roraima, and 78.5 per cent in São Paulo.

---

23 The seven states with the greatest educational persistence at the bottom of the distribution are all located in north-eastern Brazil: Rio Grande do Norte (57.8 per cent), Bahia (58.2 per cent), Paraíba (60.3 per cent), Maranhão (60.6 per cent), Alagoas (61.1 per cent), Sergipe (62.6 per cent), and Piauí (69.7 per cent).

Finally, Figure 8 illustrates how extremely difficult it is to climb the educational ladder in Brazil. In only four of the 27 states do the chances of moving from the bottom to the top of the educational distribution exceed 10 per cent: Mato Grosso 10.9 per cent, Amapá 11 per cent, Roraima 16.4 per cent, and Distrito Federal 16.6 per cent.

B. Linear Regression Model

In this section, I estimate the educational persistence between children and parents for each state based on equation (10). Figure 9 presents the results of this exercise based on a geographical breakdown. The lighter areas denote states with lower levels of educational persistence across generations (or higher mobility values).

For the sample as a whole, the correlation coefficient generated a value of 0.475, while the variation in intergenerational educational persistence across Brazilian states reached a maximum of 0.257, which represents the difference between Rio de Janeiro (0.510) and Roraima (0.253). Among the top five in educational mobility apart from Roraima, we find the states of Amapá (0.351), Goiás (0.356), Tocantins (0.370), and Maranhão (0.377).[24] Bahia (0.488), Distrito Federal (0.492), Alagoas (0.497), Acre (0.502), and Rio de Janeiro (0.510) located at the other end of the scale.

As already indicated in Figure 6, children's chances of attaining primary education have increased significantly over time in Brazil. Accordingly, figures 2 and 3 report a strong variation in average years of schooling and standard deviation across the birth cohorts. These findings are strong indications that the degree of intergenerational mobility may have changed in recent decades. In order to test this hypothesis, I divided the full sample into five birth cohorts, each of which covered 10 consecutive birth years, and subsequently estimated the predictive margins from equation (10) with a two-way interaction (education by birth cohort) to investigate how children's chances of mobility change according to their year of birth.

The results of this exercise are plotted in Figure 10 and confirm a decrease in the association between parents' and children's education over time. Note that for all birth cohorts, as parents' schooling increases, the linear prediction for children's education also increases. However, the increase (slope) is greater for children born between 1940 and 1949 than for the 1980–1989 cohort. At low levels of parental education, there is virtually no difference across birth cohorts (the children of parents with a low educational level don't achieve a high level of education no matter when they were born). As parents' educational level increases, the education gap between children becomes increasingly larger, because children born between 1940 and 1949 benefit more from the greater human capital of their parents than the younger generations. Given this variation of correlation coefficients over time, Table 2 displays the levels of mobility (separately) across birth cohorts.

---

24 As can be seen in Figure 8, educational mobility is higher in the northern region than in the other regions of the country. Among the seven states in this region, only Pará and Acre aren't among the top-10 states with the highest degree of mobility.

5.2   The Great Gatsby Curve

Independent of the indicator used to measure intergenerational mobility, the findings presented in Section 5.1 allow us to establish this paper's first important result: The chances of attaining educational mobility vary significantly from one Brazilian state to another.

This section addresses the question of why intergenerational persistence in education varies so widely across Brazilian states, investigating the effect of income mobility on educational mobility. As discussed in Section 2, Solon (2004) has concluded that the current level of income inequality between families can affect the investment in their children's human capital and, consequently, these children's chances of intergenerational mobility. It can therefore be expected that the variation in mobility presented in Figure 9 can be explained by the significant variation in inequality across Brazilian states.

According to Solon's (2004) theoretical model, what is particularly relevant for the accumulation of human capital is the level of inequality when children have completed their compulsory education and face a decision about whether or not to pursue further schooling. Therefore, this paper has used – as a measure of inequality – the Gini coefficients for the years in which the individuals should have concluded their compulsory schooling.[25]

Given the variation over time in mobility shown in Table 2, I focused the investigation on one single birth cohort containing individuals born between 1970 and 1979 in order to minimise the lifecycle bias.[26] Consequently, the measures of inequality are based on the PNAD samples between 1984 and 1993, and in order to eliminate possible short-term fluctuations in inequality across these years, I average the Gini coefficients throughout the period under consideration.

Figure 10 plots the Great Gatsby curve for the Brazilian states. On the y-axis we find the level of intergenerational persistence in education estimated from equation (10),[27] while income inequality is plotted on the horizontal axis. The findings confirm the statistically significant relationship between the Gini coefficient and intergenerational mobility:[28] States with a higher level of income disparity, such as Paraíba (PB) and Ceará (CE), presented higher values of persistence in education (or low levels of mobility), while the correlation coefficients tended

---

25 A child born in 1970, for example, started school at age seven in 1977 and presumably concluded their compulsory (primary) education in 1984.

26 The youngest cohort (1980–1989) has not been chosen for the investigation because approximately 9.2 per cent of the individuals in this group were enrolled in the educational system in 2014. The oldest birth cohorts (1940–1949 and 1950–1959) needed to be excluded from the analysis because there are no data available for the Gini coefficient for the years before 1976.

27 As already observed by Chetty et al. (2014a), the cross-regional studies have exclusively used the measure of regression coefficients to investigate the variation in intergeneration mobility. In this section I therefore use the correlation coefficient as a measure of mobility in order to enable an accurate comparison of the conclusions with the results from the cross-country literature.

28 The Pearson correlation coefficient (r) achieves a value of 0.4245 and indicates a moderate positive linear relationship between persistence in education and income inequality.

to be lower in states with a more equal distribution of income, such as Santa Catarina (SC) and Amazonas (AM).

However, the negative correlation between income inequality and mobility does not hold true for all states. Amapá (AP) can be considered a "point outside the curve," because the state had the most equal distribution of income in the country but presented a relatively high persistence in education across generations. In addition, some states with similar Gini coefficients, such as Bahia (BA), Goiás (GO), and Rio Grande do Norte (RN), present very different indices of intergenerational mobility.

5.3 Linking Inequality and School Dropouts

In this section, I move away from the analysis of intergenerational persistence in education via the correlation hypothesis to an investigation of the determinants which could better explain the association between inequality and mobility illustrated by the Great Gatsby curve. At this point, it is important to introduce the concept of "economic marginalisation" presented by Kearney and Levine (2014), which can be described as the process of a person setting aside participation in the educational system given their very low expected-earnings premium. In this case, young individuals do not believe that an investment in human capital can increase their chances of mobility, which leads them to leave school early.[29]

According to the human capital approach developed by Kearney and Levine (2014), the marginality arises as a consequence of higher income inequality. An increase in the 75/10 ratio of income distribution might lead to direct social exclusion, particularly for children from socially vulnerable families that do not see the possibility of climbing up the social ladder via education. The marginalised population often lives in disadvantaged areas with negative neighbourhood behavioural patterns and notably restricted access to high-quality schools, thus reducing their belief in personal advancement through schooling and consequently making social mobility more difficult (Rothwell and Massey 2015).[30]

With this problem in mind, the empirical objective of this section is to investigate whether children from socially disadvantaged households living in states with greater income inequality have a lower chance of completing (secondary) education. Figures 12 and 13 provide descriptive statistics in order to illustrate the main variables used in the identification strategy.

Figure 12 illustrates the variation across Brazilian states in the ratio of the income of the upper-bound value of the third quartile (i.e. the 25 per cent of individuals with the highest income) to that of the first decile. For this exercise the 27 units of the federation have been classified into three inequality groups (low, middle, and high) according to the 75/10 ratio for the year 2014. The resulting visual presents an almost perfect geographic distribution of inequality and a significant variation across states: In Rio de Janeiro the income of the richest

---

29 The Appendix provides a formal description of Kearney and Levine's (2014) model.

30 See Rothwell and Massey (2015) for a large and rich literature overview concerning the channels through which neighborhoods can affect future earnings.

25 per cent corresponds to 2.76 times the income of the poorest 10 per cent, while this ratio is 8.32 in Piauí.

Figure 13 provides the first empirical evidence for the subsequently applied econometric model. It presents the proportion of the population with secondary-school education, divided by the inequality groups introduced in the previous figure and the educational achievement of parents, which is used as a proxy for "economic marginalisation." The findings highlight the effect of marginalisation on the decision to leave school early. Note that independent of the inequality level, less than 20 per cent of children from illiterate parents have completed secondary education. In contrast, more than 80 per cent of children of parents with a graduate degree have a secondary school-leaving qualification.

In addition, Figure 13 confirms that for vulnerable children, dropping out of school is associated with income inequality: The children of illiterate parents and parents with no (primary) education living in states with lower income inequality have a higher chance of completing secondary education than vulnerable children from high-inequality states.

5.3.1   Probit Latent Variable Model

In order to empirically test the assumption regarding economic marginalisation, I run equation (11) and present the results in Table 3.[31] The first column contains the results for the whole sample, and the subsequent columns contain the values for the five-year birth cohorts.[32]

Parental educational level, gender, location of residence, race, year of birth, and whether a child has been living with both parents have a statistically significant effect on the chance of completing secondary education. Being male, for example, decreases the probability of achieving a (secondary) school-leaving certificate by 20.3 percentage points. As expected, children of parents with no school certificate have a lower chance of completing secondary education (40.5 per cent) compared to children of parents with at least a primary education.

The interaction term between the categorical variable "socio-economic marginalization" and the continuous variable "income inequality" is the focus of this investigation and confirms the statistically significant effect of income disparity on educational attainment. The negative coefficient indicates that children of parents with no school certificate are more disadvantaged by an increase in income inequality. Specifically, each additional point in the 75/10 ratio decreases the likelihood of achieving secondary education by 5.4 per cent for children of parents without education.

For a better overview of the interaction between income inequality and economic marginalisation, I estimate the marginal effects from equation (11) and display the predicted probabilities

---

31 The parameters in the probit model were estimated using maximum likelihood methods.

32 Because there is no nationally representative database for the period prior to 1981 that could be harmonised in a reliable way with the most recent samples of PNAD, this section limited the estimates to individuals born from the year 1965 onwards, thereby using the income inequality after the year 1981. See the Online Appendix for a detailed description of the data harmonisation.

for all 10th values of the 75/10 ratio (from 3 to 12) in Figure 14.[33] Note that independent of the level of inequality, children of parents with no education have an even lower chance of completing secondary school. Moreover, both curves have different shapes and slopes: The slope of the no-education curve is higher, indicating that the effects of an increase in income inequality are disproportionately higher for children of parents with no education. As a consequence, at a low level of income inequality, there is a relatively small difference in the probability of achieving a secondary school certificate between children from educated and uneducated parents. However, as the 75/10 ratio increases, the gap between these two groups becomes increasingly bigger.[34]

5.3.2   Robustness Checks

As described in detail by Neumayer and Plümper (2017), econometric inferences become more credible and effective if they are sufficiently independent from the model specification. For that reason, this section tests the same economic marginalisation hypothesis using alternative model specifications and alternative econometric approaches in order to improve the validity of the empirical evidence presented in the previous section.

A. Alternative Econometric Approaches

As previously described, the objective of Section 5.3.1 was to identify whether children from (socio-economically) marginalised households living in states with greater income inequality are more disadvantaged in their school careers. As a proxy for socio-economic marginalisation, I used a dummy variable indicating children from parents with no primary education ($NoEduc_i$) in equation (11).

As usual in such circumstances, the empirical model assumed that the correlations between the residual and the predictors are zero. But now, based on Wooldridge's (2010) theoretical approach, I relax this assumption and consider the case where the probit model contains a binary explanatory variable that is endogenous. The "feeling of marginalisation" varies according to the parents' economic situation, and having both parents in the household can shift the family's budget constraints, providing higher socio-economic status for the family, similarly to a higher level of parental education. I therefore use for the variable responsible for the socio-economic marginalisation ($NoEducP_i$) the instrumental variable "both parents" ($bothP_i$), which is a binary variable equal to one if the individual lived with both parents in the household at the age of 15.

---

33 In effect, the adjusted predictions at representative values (APRs) are comparing two hypothetical populations – children of parents with and without a (primary) education – that have exactly the same values for all the other independent variables in the model, with the exception of the income inequality level in the state of residence (75/10 ratio). Since the only difference between these two populations is the inequality, inequality must be the cause of the differences in their likelihood of achieving a secondary education (Williams et al. 2012).

34 Figure 14 shows that both curves have non-overlapping confidence intervals, demonstrating a statistically significant difference between the estimations.

In this section, I continue to use equation (11) to study the effects of economic marginalisation on the chances of completing secondary education, but the empirical investigations have been conducted on the basis of three different empirical approaches: ordinary least squares (OLS) estimations of a linear probability model (LPM), two-stage least squares (2SLS) estimations of the LPM, and a bivariate probit that drops the variable "both parents" ($bothP_i$) from the probit for MSB_i.[35]

Table 4 provides the results of the robustness checks using the whole sample and confirms that the estimates from Section 5.3.1 are also robust to alternative econometric approaches. For brevity's sake, the table reports only the coefficients π_1 from the interaction term between income inequality ($Ineq_s$) and the proxy for socio-economic marginalisation (MSB_i). Next, I have used margins to obtain the predicted probabilities for this interaction and have also displayed the adjusted predictions of educational chances at representative values of income inequality (APRs) – that is, for every 10th value for the distribution of the 75/10 ratio.[36]

As in the main model specification, all three expanded models presented negative and statistically significant values for the interaction term, indicating that the higher the inequality level in the state, the lower the share of students with a secondary school-leaving qualification. The nonlinear models (columns 1 and 4) give larger estimated coefficients for this interaction than the linear model (columns 2 and 3): -0.0540 and -0.0487 versus -0.0179 and -0.0175, respectively, suggesting that the nonlinearity in the probit models plays a decisive role in determining the chances of formal educational achievement.

With the estimations of marginal effects for different inequality levels, it is possible to observe that the effects of economic marginalisation differ greatly according to the level of inequality. When (MSB_i) is assumed to be exogenous, the probit and LPM models provide very similar average partial effects by increasing income disparity. Children of parents with no formal education in the lowest inequality decile have, for example, a 22 per cent lower chance of achieving a secondary education certificate than pupils from parents with at least primary education. The same difference at the top decile is approximately 40 per cent. This empirical evidence remains practically unchanged when $bothP_i$ is used as IV in the LPM estimation.

Lastly but by no means least importantly, the use of the bivariate probit, assuming that MSB_i and $bothP_i$ are correlated, presents substantially lower estimated APRs than the (normal) probit model.[37] However, the estimates continue to indicate the same direction and statistical significance.

---

35 To facilitate comparison, Table 4 also contains the estimation results from the probit model in Section 5.3.1, in which the variable $bothP_i$ was treated as exogenous.

36 It is important to note that the estimated marginal effects refer to the component terms of socio-economic marginalisation and income inequality, and not necessarily the marginal effect of the interaction term. As explained in greater detail by Williams et al. (2012), because the value of the interaction terms cannot change independently of the values of the component terms, it is not possible to estimate a separate effect for the interaction.

37 The adjusted predications range from 0.223 and 0.401 for the probit model, while in the bivariate probit the APRs vary between 0.175 and 0.323.

B. Alternative Model Specifications

In the following, I explore the dependence of parameter π_1, estimated from equation (11), on four specific changes in model specification: In column 5, the estimations were limited to individuals who have never lived in another Brazilian state or another country. Column 6 used the ratio 90/10 as an indicator of income inequality, instead of the 75/10 ratio. In column 7, I changed the variable responsible for socio-economic marginalisation, substituting parents with no (primary) education for illiterate parents. Finally, in column 8 the dummy variable representing children with illiterate parents has been added to the empirical model and estimated in combination with $(NoEduc_i)$.[38]

All four expanded models generated robust results, demonstrating the significantly negative impact of income inequality on educational attainment, as already indicated in Section 5.3.1. In this context, it is hardly surprising that the results for column 5, with only individuals who have never lived in another state, indicated a higher effect of inequality on educational outcomes than the other specifications. As already noted by Kearney and Levine (2014), boys and girls who have been born into a region with an extremely uneven distribution of wealth and have never seen another reality tend to underestimate the returns on schooling given their lower belief in social mobility through education.

Once again, the estimations of marginal effects for different inequality levels pointed to an increase in the gap in educational attainment as a result of income disparity. According to the model with only the local population, for example, the advantage of having parents with primary education is 21.0 per cent at the bottom of the distribution and 42.3 per cent at the other extreme of the inequality scale. These results are consistent with the findings presented in Figure 14 and show that – keeping all the other variables constant – the adverse effect of socio-economic marginalisation on the chance of completing secondary education tends to be stronger in states with greater income disparity.

6 Conclusion

The estimates presented in this paper are based on data from the mobility supplement from the PNAD-2014, which is a nationally representative survey from Brazil detailing the educational attainments for two generations within the same family. The empirical findings provided here have shown for the first time that intergenerational persistence in education varies substantially across Brazilian states. For example, the probability that a child born to parents without a school certificate will achieve a university degree is 3.2 per cent in Pará, but 16.6 per cent in Roraima. Together with findings from other countries (Azam and Bhatt 2015; Chetty et

---

38 For the specification in column 8, the empirical model assumes the following form: $ComSec_{i,t} = \pi_0 + \pi_1(NoEducP_i \ x \ Ineq_{s,t+14}) + \pi_2(IlliteP_i \ x \ Ineq_{s,t+14}) + \pi_3 \ NoEducP_i + \pi_4 \ IlliteP_i + \pi_5 \ Ineq_{s,t+14} + \gamma_1 male_i + \gamma_2 rural_i + \gamma_3 bothP_i + \gamma_4 race_i + \gamma_5 birth_i + \varepsilon_i.$

al. 2014a; Güell et al. 2018), this work strengthens the assumption that mobility levels can vary considerably within a single country.

In addition, this study contributes to the literature that is presenting new findings on the Great Gatsby curve. I have found compelling empirical evidence for a statistically significant association between intergenerational mobility and income inequality, thus confirming the existence of the Great Gatsby curve at the national level as well: persistence in educational levels across generations tends to be stronger in states with a more unequal distribution of income.

Finally, this work has aimed directly at illuminating the mechanisms underlying the link between inequality and mobility presented in the Great Gatsby curve – currently the biggest gap in this field of research. Thanks to the empirical approach proposed by Kearney and Levine (2014), it was possible to study the effects of an increase in income inequality on the chances of education for children from socially vulnerable families. I have found compelling evidence that children born into families with no education are more likely to leave school early if they live in states where the gap between the bottom and the middle of the income distribution is wider. These findings are particularly relevant for the literature because they are independent of the econometric model and remain robust to different model specifications and alternative econometric approaches.

Bibliography

Aaberge, R., Björklund, A., Jäntti, M., Palme, M., Pedersen, P. J., Smith, N., and Wennemo, T. (2002). Income inequality and income mobility in the Scandinavian countries compared to the United States. *Review of Income and Wealth*, 48(4):443-469.

Aaronson, D. and Mazumder, B. (2008). Intergenerational economic mobility in the United States, 1940 to 2000. *Journal of Human Resources*, 43(1):139-172.

Altham, P. M. and Ferrie, J. P. (2007). Comparing contingency tables tools for analyzing data from two groups cross-classied by two characteristics. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 40(1):3-16.

Andrews, D. and Leigh, A. (2009). More inequality, less social mobility. *Applied Economics Letters*, 16(15):1489-1492.

Ayala, L. and Sastre, M. (2008). The structure of income mobility: empirical evidence from UE-Countries. *Empirical Economics*, 35(3):451-473.

Azam, M. (2016). Intergenerational educational persistence among daughters: Evidence from India. *The BE Journal of Economic Analysis & Policy*, 16(4).

Azam, M. and Bhatt, V. (2015). Like father, like son? Intergenerational educational mobility in India. *Demography*, 52(6):1929-1959.

Becker, G. S. and Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of political Economy*, 87(6):1153-1189.

Becker, G. S. and Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of labor economics*, 4(3, Part 2):1-39.

Björklund, A. and Jäntti, M. (2009). *Intergenerational income mobility and the role of family background. Oxford Handbook of Economic Inequality*, Oxford University Press, Oxford, pages 491-521.

Björklund, A., Jäntti, M., and Lindquist, M. J. (2009). Family background and income during the rise of the welfare state: brother correlations in income for Swedish men born 1932-1968. *Journal of Public Economics*, 93(5):671-680.

Black, S. E. and Devereux, P. J. (2010). *Recent developments in intergenerational mobility. Technical report*, National Bureau of Economic Research.

Blanden, J. (2013). Cross-country rankings in intergenerational mobility: A comparison of approaches from economics and sociology. *Journal of Economic Surveys*, 27(1):38-73.

Blanden, J., Haveman, R., Smeeding, T., and Wilson, K. (2014). Intergenerational mobility in the United States and Great Britain: A comparative study of parent-child pathways. *Review of Income and Wealth*, 60(3):425-449.

Blanden, Jo and Macmillan, Lindsey (2014). *Education and intergenerational mobility: help or hindrance?* DoQSS Working Papers 14-01, Department of Quantitative Social Science - UCL Institute of Education, University College London.

Boudreaux, C. J. (2014). Jumping off of the Great Gatsby curve: How institutions facilitate entrepreneurship and intergenerational mobility. *Journal of Institutional Economics*, 10(2):231-255.

Bradbury, K. and Triest, R. K. (2016). Inequality of opportunity and aggregate economic performance. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 2.2 (2016): 178-201.

Brahim, S. A. and McLeod, D. (2016). Inequality and mobility: Gatsby in the Americas. *Modern Economy*, 7(05):643.

Bratsberg, B., Roed, K., Raaum, O., Naylor, R., Eriksson, T., et al. (2007). Nonlinearities in intergenerational earnings mobility: Consequences for cross-country comparisons. *The Economic Journal*, 117(519).

Breen, R. and Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annu. Rev. Sociol.*, 31:223-243.

Checchi, D., Fiorio, C. V., and Leonardi, M. (2013). Intergenerational persistence of educational attainment in Italy. *Economics Letters*, 118(1):229-232.

Checchi, D., Ichino, A., and Rustichini, A. (1999). More equal but less mobile? Education financing and intergenerational mobility in Italy and in the US. *Journal of public economics*, 74(3):351-393.

Chen, W.-H. (2009). Cross-national differences in income mobility: Evidence from Canada, the United States, Great Britain and Germany. *Review of Income and Wealth*, 55(1):75-100.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014a). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics*, 129(4):1553-1623.

Chetty, R., Hendren, N., Kline, P., Saez, E., and Turner, N. (2014b). Is the United States still a land of opportunity? Recent trends in intergenerational mobility. *The American Economic Review*, 104(5):141-147.

Chevalier, A., Denny, K., and McMahon, D. (2003). *A multi-country study of intergenerational educational mobility*. University College Dublin. Institute for the Study of Social Change (Geary Institute).

Corak, M. (2006). Do poor children become poor adults? Lessons from a cross-country comparison of generational earnings mobility. In: *Dynamics of inequality and poverty*, pages 143-188. Emerald Group Publishing Limited.

Corak, M. (2013a). Income inequality, equality of opportunity, and intergenerational mobility. *The Journal of Economic Perspectives*, 27(3):79-102.

Corak, M. (2013b). Inequality from generation to generation: The United States in comparison. In: Rycroft, R. S. (editor), *The Economics of Inequality, Poverty, and Discrimination in the 21st Century* [2 Volumes]. ABC-CLIO.

Corak, M. and Heisz, A. (1999). The intergenerational earnings and income mobility of Canadian men: Evidence from longitudinal income tax data. *Journal of Human Resources*, pages 504-533.

Corak, M., Lindquist, M. J., and Mazumder, B. (2014). A comparison of upward and downward intergenerational mobility in Canada, Sweden and the United States. *Labour Economics*, 30:185-200.

Cunha, F., Heckman, J. J., Lochner, L., and Masterov, D. V. (2006). *Interpreting the evidence on life cycle skill formation. Handbook of the Economics of Education*, 1:697-812.

Daouli, J., Demoussis, M., and Giannakopoulos, N. (2010). Mothers, fathers and daughters: Intergenerational transmission of education in Greece. *Economics of Education Review*, 29(1):83-93.

Dearden, L., Machin, S., and Reed, H. (1997). Intergenerational mobility in Britain. *The Economic Journal*, pages 47-66.

Duncan, G. J. and Murnane, R. J. (2011). *Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children*. Russell Sage Foundation.

Dunn, C. E. (2007). The intergenerational transmission of lifetime earnings: Evidence from Brazil. *The BE Journal of Economic Analysis & Policy*, 7(2).

Economist (2013). The Great Gatsby Curve: Don't worry, old sport. Online: <www.economist.com/blogs/democracyinamerica/2013/07/great-gatsby-curve> (26 June 2019).

Ermisch, J., Jantti, M., and Smeeding, T. M. (2012). *From parents to children: The intergenerational transmission of advantage*. Russell Sage Foundation.

Fan, Y., Yi, J., and Zhang, J. (2015). *The Great Gatsby Curve in China: Cross-sectional inequality and intergenerational mobility. Technical report*, Working Paper, CUHK, Hongkong.

Ferreira, S. G. and Veloso, F. A. (2006). Intergenerational mobility of wages in Brazil. *Brazilian Review of Econometrics*, 26(2):181-211.

Fessler, P. and Schneebaum, A. (2012). Gender and educational attainment across generations in Austria. *Feminist Economics*, 18(1):161-188.

Fields, G. S. (2002). *Distribution and development: A new look at the developing world*. MIT press.

Fox, L., Torche, F., and Waldfogel, J. (2016). Intergenerational mobility. In: *The Oxford Handbook of the Social Science of Poverty*. Oxford Handbooks Online.

Güell, M., Pellizzari, M., Pica, G., and Rodríguez Mora, J. V. (2018). Correlating social mobility and economic outcomes. *The Economic Journal*, 128(612):F353-F403.

Heckman, J. J. (2013). *The economics and the econometrics of human development*. Presentation given at the Eighth Vienna Seminar of Nobel Laureates. Online: <www.wu.ac.at/start/james heckman.pdf (26 June 2019).

Heineck, G., and Riphahn, R. T. (2007). Intergenerational Transmission of Educational Attainment in Germany - the Last Five Decades. *Journal of Economics and Statistics* (Jahrbücher für Nationalökonomie und Statistik), Discussion(October):36-60.

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., and Verashchagina, A. (2007). The inheritance of educational inequality: International comparisons and fifty year trends. *The BE Journal of Economic Analysis & Policy*, 7(2).

Hills, J. et al. (2015). *New research evidence on social mobility and educational attainment. Technical report*, Centre for Analysis of Social Exclusion, LSE.

Hout, M., and Guest, A. M. (2013). Intergenerational occupational mobility in Great Britain and the United States since 1850: Comment. *The American Economic Review*, 103(5):2021.

Ichino, A., Karabarbounis, L., and Moretti, E. (2011). The political economy of intergenerational income mobility. *Economic Inquiry*, 49(1):47-69.

Jäntti, M., Bratsberg, B., Roed, K., Raaum, O., Österbacka, E., Björklund, A., and Eriksson, T. (2006). *American exceptionalism in a new light: a comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States*. IZA Discussion Paper, (1938):1-40.

Jäntti, M., and Jenkins, S. P. (2013). *Income mobility*. IZA Discussion Paper Series No. 7730.

Jerrim, J., and Macmillan, L. (2015). Income inequality, intergenerational mobility, and the Great Gatsby Curve: is education the key? *Social Forces*, 94(2):505-533.

Kearney, M. S., and Levine, P. B. (2014). *Income inequality, social mobility, and the decision to drop out of high school. Technical report*, National Bureau of Economic Research.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., and Shonko, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building america's future workforce. *Proceedings of the National Academy of Sciences*, 103(27):10155-10162.

Krueger, A. (2012). *The rise and consequences of inequality*. Presentation made to the Center for American Progress, 12 January. Online: <www.americanprogress.org/events/2012/01/12/17181/the-rise-and-consequences-of-inequality/> (26 June 2019).

Lawrence, J., Schweinhart, J. M., Zongping, X., William, S., Clive, R., and Milagros, N. (2005). Lifetime effects: The high/scope perry preschool study through age 40. *The Academy of Experimental Criminology*, 4:2-3.

Lee, C.-I., and Solon, G. (2009). Trends in intergenerational income mobility. *The Review of Economics and Statistics*, 91(4):766-772.

Lefgren, L., McIntyre, F., and Sims, D. P. (2015). Beyond education and fairness: A labor market taxation model for the Great Gatsby Curve. *Economic Inquiry*, 53(2):962-978.

Mazumder, B. et al. (2012). *Is intergenerational economic mobility lower now than in the past?* Chicago Fed Letter, (Apr).

Mazumder, B. et al. (2015). *Inequality in skills and the Great Gatsby Curve*. Chicago Fed Letter, (Jan).

Neumayer, E., and Plümper, T. (2017). *Robustness tests for quantitative research*. Cambridge, University Press.

Obama, B. (2013). *Full transcript: President Obama's December 4 remarks on the economy*. Speech on economic mobility and the Affordable Care Act, reported by the Washington Post. Online: <www.washingtonpost.com/politics/running-transcript-president-obamas-december-4-remarks-on-theeconomy/2013/12/04/7cec31ba-5cff-11e3-be07-006c776266ed_story.html> (26 June 2019).

Pastore, J., and do Valle Silva, N. (2000). *Mobilidade social no Brasil*. Makron Books.

Phillips, D. A., Shonko, J. P., et al. (2000). *From neurons to neighborhoods: The science of early childhood development*. National Academies Press.

Reddy, A. B. (2015). Changes in intergenerational occupational mobility in India: Evidence from national sample surveys, 1983-2012, *World Development*, 76:329-343.

Ribeiro, C. A. C. (2007). *Estrutura de classe e mobilidade social no Brasil*. Edusc Sao Paulo.

Roemer, J. E. (2004). Equal opportunity and intergenerational mobility: going beyond intergenerational income transition matrices. *Generational income mobility in North America and Europe*, pages 48-57.

Rothwell, J. and Massey, D. (2015). Geographic effects on intergenerational income mobility. *Economic Geography*, 91(1):83-106.

Schneebaum, A., Rumplmaier, B., and Altzinger, W. (2016). Gender and migration background in intergenerational educational mobility. *Education Economics*, 24(3):239-260.

Shonko, J. P. and Meisels, S. J. (2000). *Handbook of early childhood intervention*. Cambridge University Press.

Smeeding, T., Erikson, R., and Jäntti, M. (2011). *Persistence, privilege, and parenting: The comparative study of intergenerational mobility*. Russell Sage Foundation.

Solon, G. (1992). Intergenerational income mobility in the United States. *The American Economic Review*, pages 393-408.

Solon, G. (2004). A model of intergenerational mobility variation over time and place. *Generational income mobility in North America and Europe*, pages 38-47.

The Guardian (2012). The Great Gatsby and the American Dream. *The Guardian*.

Ueda, A. (2009). Intergenerational mobility of earnings and income in Japan. *The BE Journal of Economic Analysis & Policy*, 9(1).

White House (2013). *What is the Great Gatsby Curve?* Online: <https://obamawhitehouse.archives.gov/blog/2013/06/11/what-great-gatsby-curve> (26 June 2019).

Williams, R. et al. (2012). Using the margins command to estimate and interpret adjusted predictions and marginal effects. *Stata Journal*, 12(2):308.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Xie, Y. and Killewald, A. (2013). Intergenerational occupational mobility in Great Britain and the United States since 1850: Comment. *The American Economic Review*, 103(5):2003-2020.

Zimmerman, D. J. (1992). Regression toward mediocrity in economic stature. *The American Economic Review*, pages 409-429.

Figures

Figure 1. Income Inequality across States
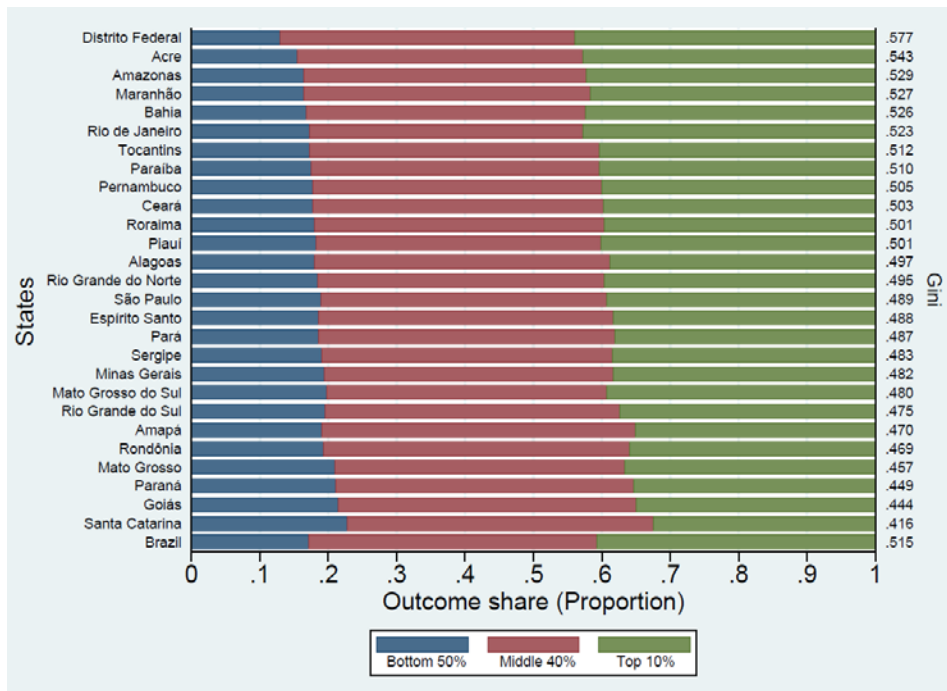
Note: Estimations based on self-declared per capita domiciliary income.
Source: PNAD-2014, author's own estimates.

Figure 2. Development of Average Schooling, per State
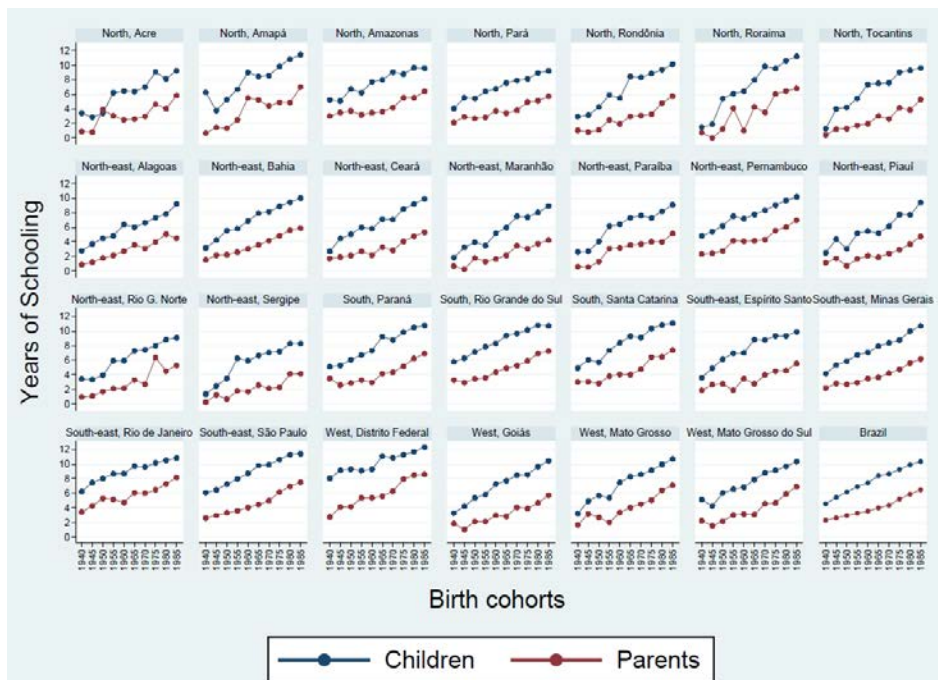
Notes: Children's education for boys and girls. Estimations of parent's education based on educational attainment
of the most educated parent.
Source: PNAD-2014, author's own estimates.

Figure 3. Development of Inequality in Schooling, per State
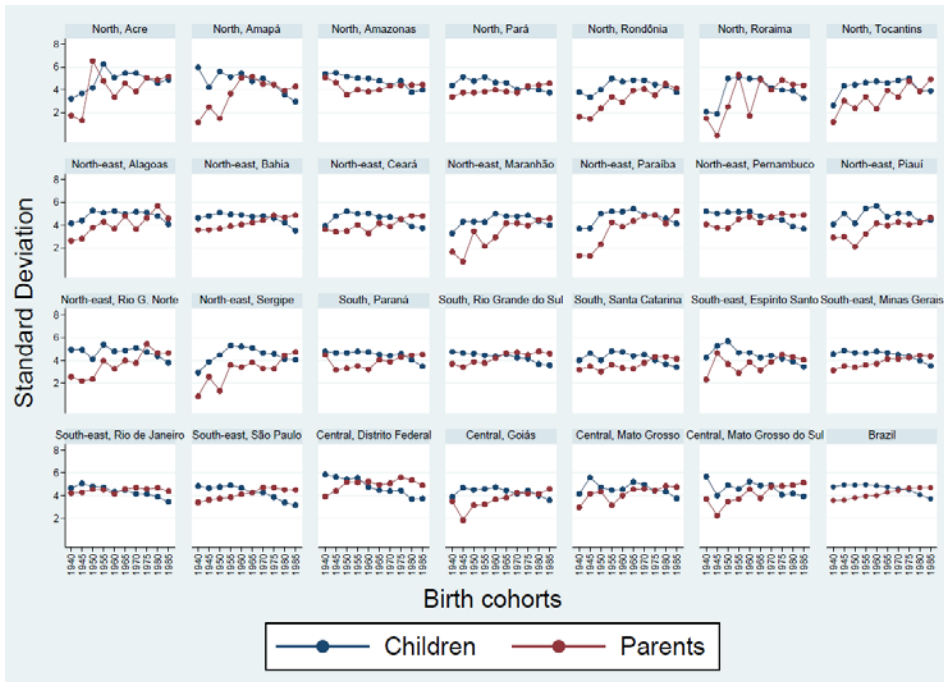
Notes: Children's education for boys and girls. Estimations of parent's education based on educational attainment
     of the most educated parent.
Source: PNAD-2014, author's own estimates.

Figure 4. Average Years of Schooling
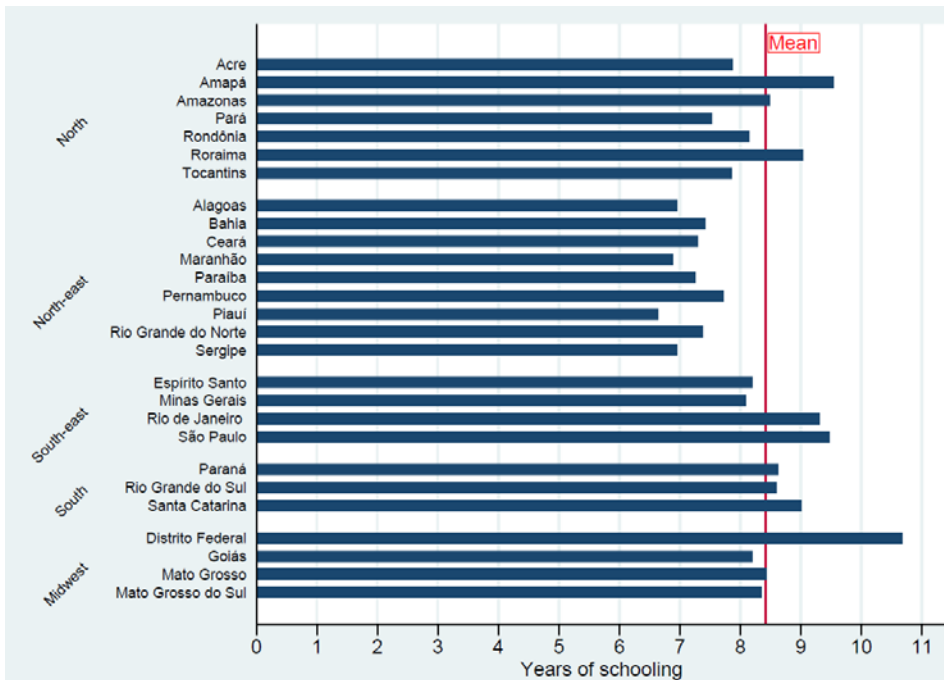
Note: Estimations for boys and girls.
Source: PNAD-2014, author's own estimates.

Figure 5. Levels of Education by Region and State
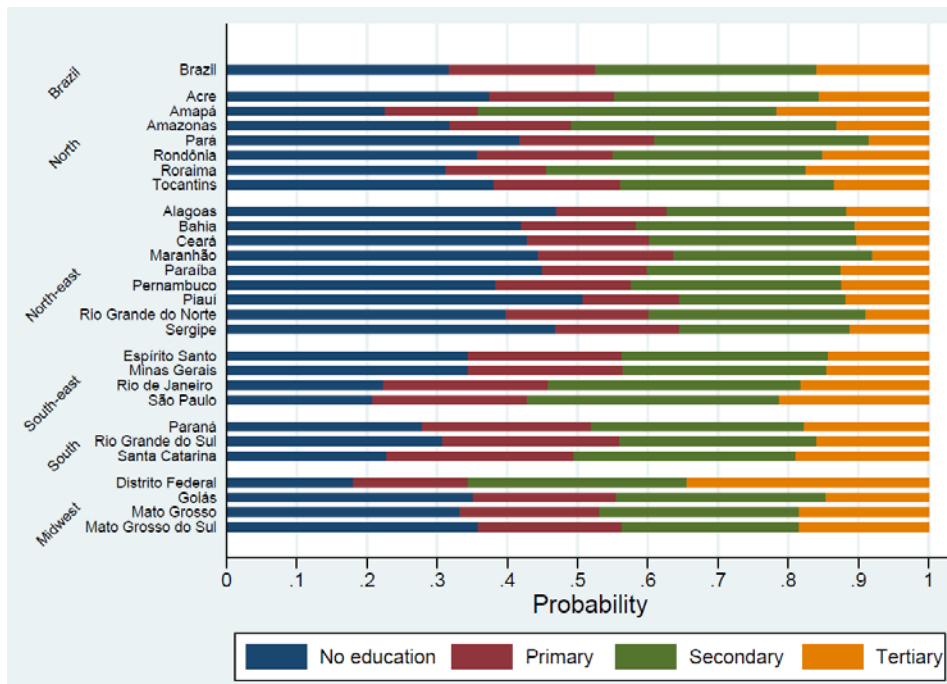
Note: Estimations for boys and girls.
Source: PNAD-2014, author's own estimates.

Figure 6. Immobility Ratio and Upward–Downward Mobility
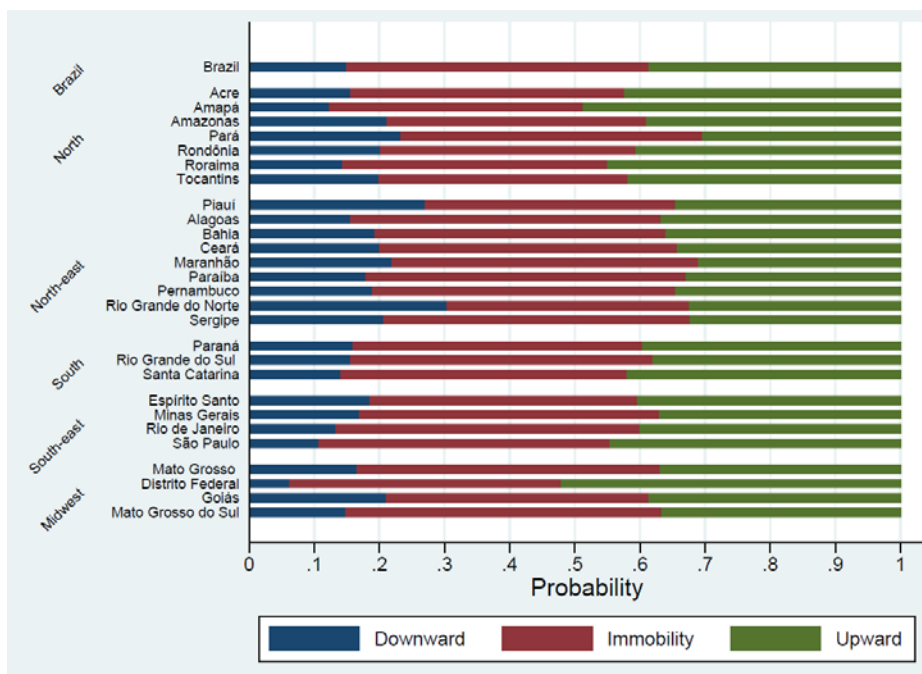
Note: Downward (upward) mobility represents the share of children who have achieved a lower (higher) level of
      education than their most educated parent.
Source: PNAD-2014, author's own estimates.

Figure 7. Children's Predicted Probabilities of Educational Attainment
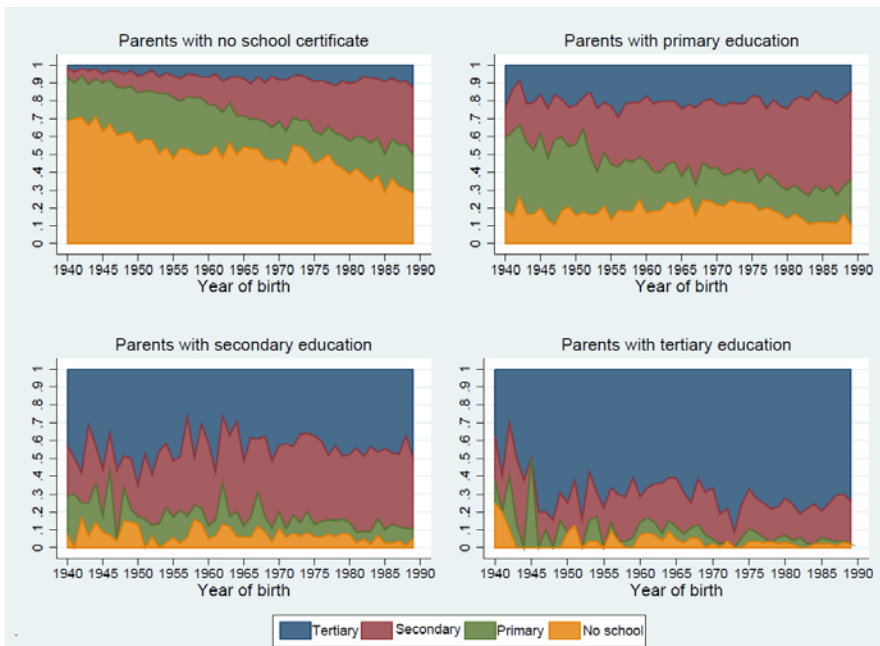
Notes: Children's education for boys and girls. Parents' schooling refers to the educational attainment of the better-
    educated parent.
Source: PNAD-2014, author's own estimates.

Figure 8. Intergenerational Mobility Indexes
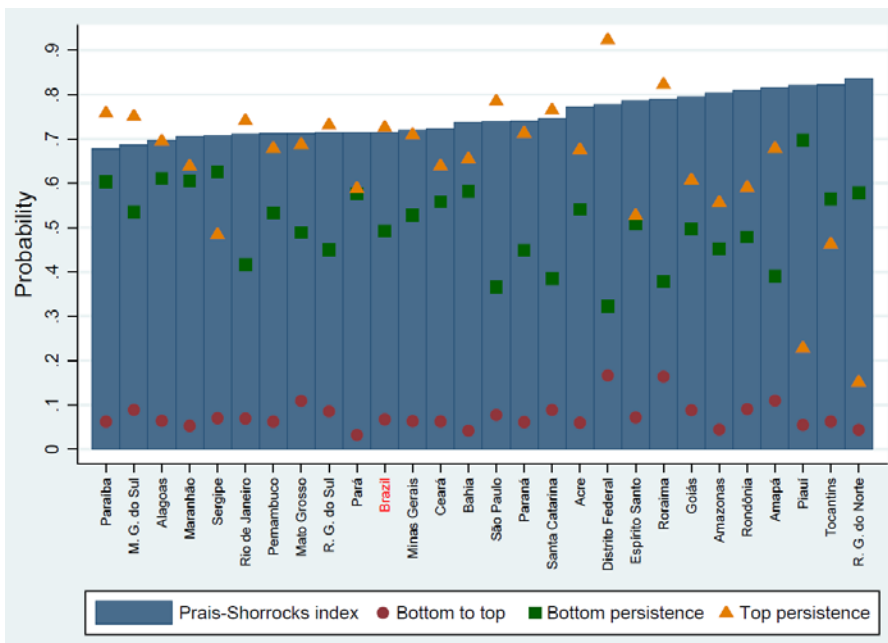
Notes: The Prais-Shorrocks index provides a measure of the normalised distance between the identity matrix and
    the independent matrix. It takes a value of zero (one) when no (all) children move away from the educational
    level of their parents. The bottom-to-top index reports the proportion of individuals born into families with
    no education that have achieved a university degree. The top (bottom) persistence shows the share of chil-
    dren born to parents with tertiary (no) education who have attained the same educational level as their
    parents.
Source: PNAD-2014, own estimates.

Figure 9. Intergenerational Persistence in Education
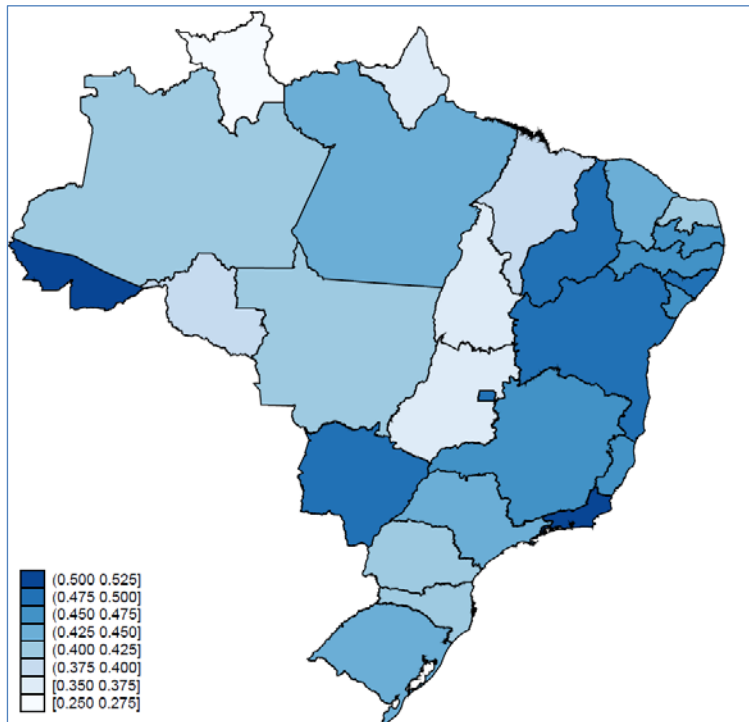
Note: The closer the estimated value is to one, the stronger the association between parents' and children's educational attainment and, consequently, the lower the intergenerational mobility.
Source: PNAD-2014, author's own estimates.

Figure 10. Adjusted Predictions of Birth Cohorts
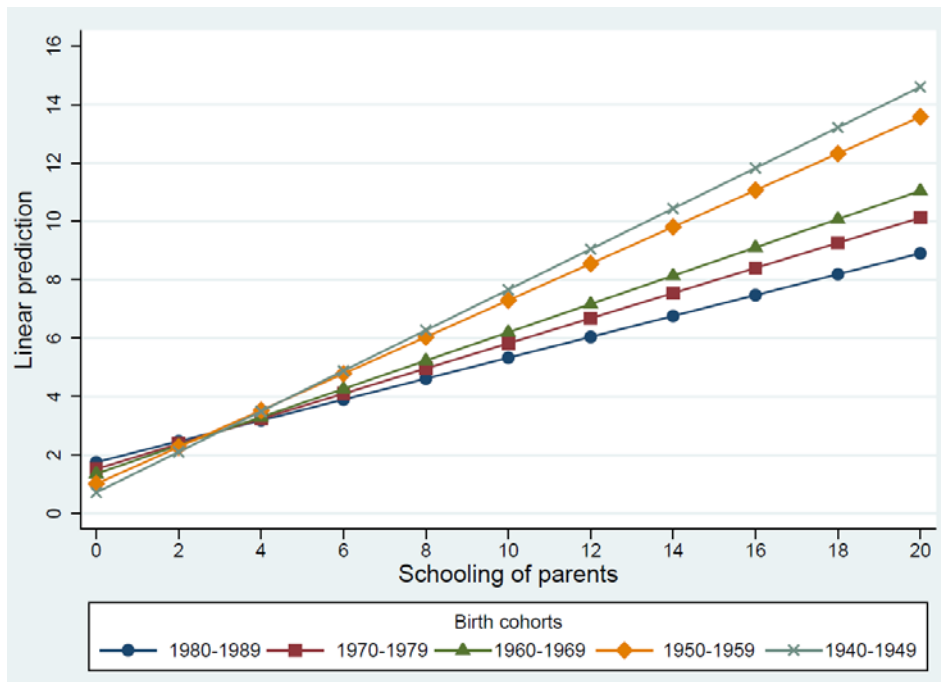
Source: PNAD-2014, author's own estimates.

Figure 11. The Great Gatsby Curve
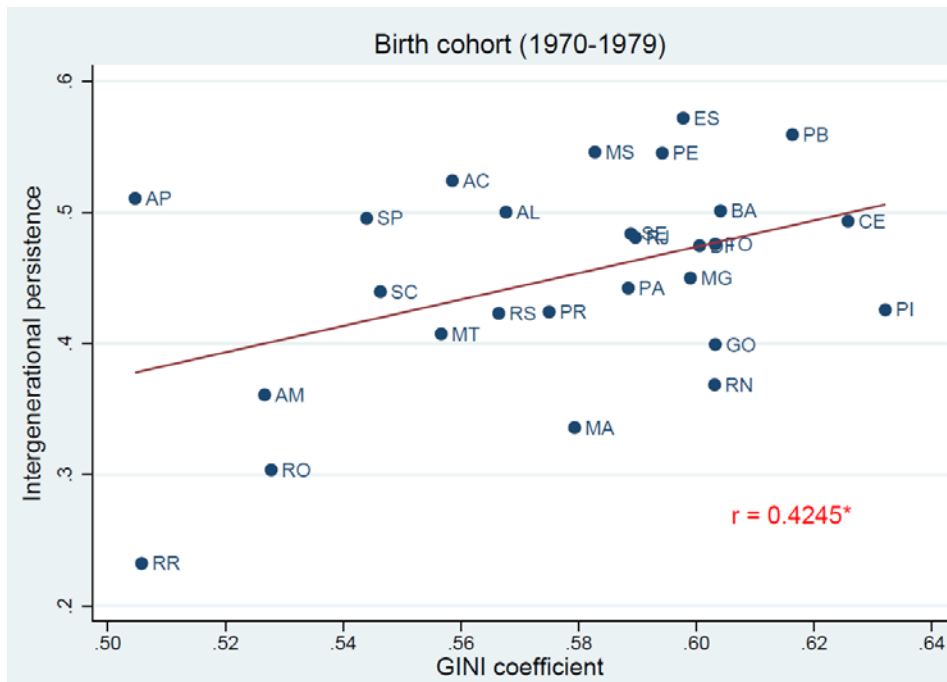
Notes: r = Pearson's correlation. Asterisk indicates correlation coefficients with p-values of 0.1 or lower. Gini coefficients refer to the average values between 1984 and 1993.
Source: PNADs, author's own estimates.

Figure 12. 75/10 Ratio of Income Distribution
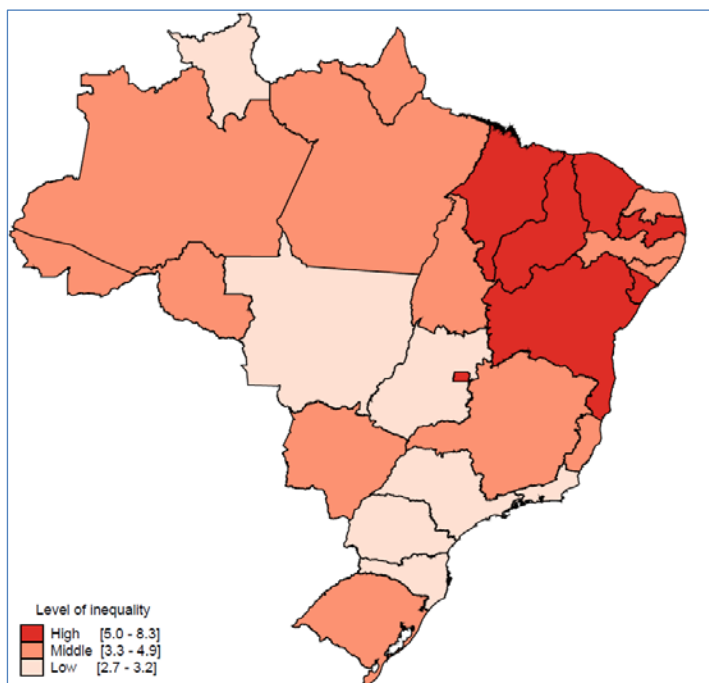
Notes: The 75/10 ratio represents the relation between the income earned by individuals in the 75th percentile and the earnings of individuals in the 10th percentile. Estimations based on total income of the economically active population aged 15 and over with earnings greater than zero.
Source: PNAD-2014, author's own estimates.

Figure 13. Educational Attainment and Inequality
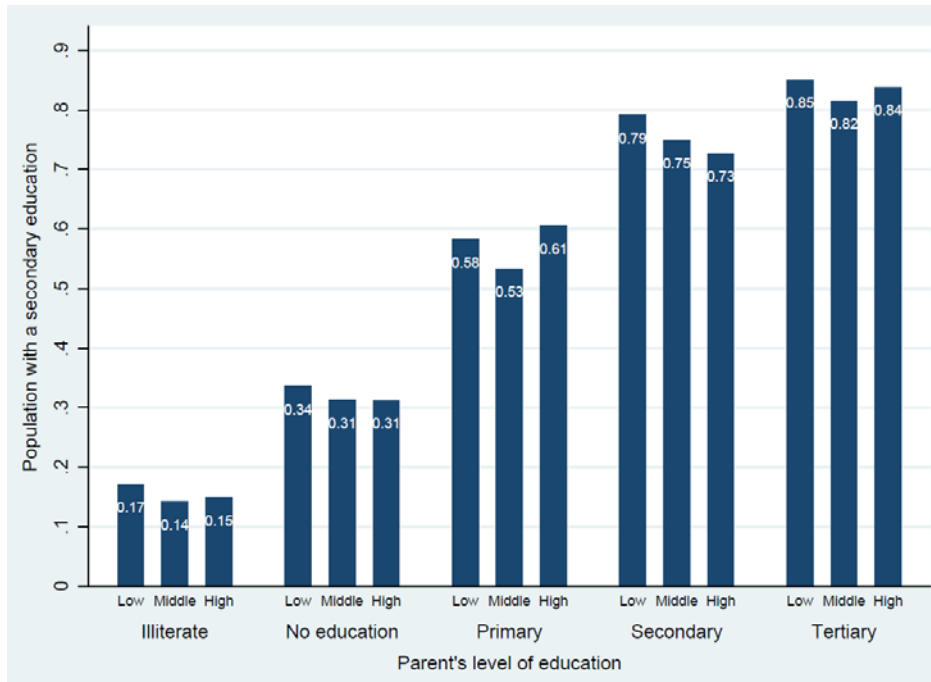
Note: Estimations of income inequality based on 75/10 ratio of total income of the economically active population
aged 15 and over and with earnings greater than zero.
Source: PNAD-2014, author's own estimates.

Figure 14. Adjusted Predictions for Secondary Education
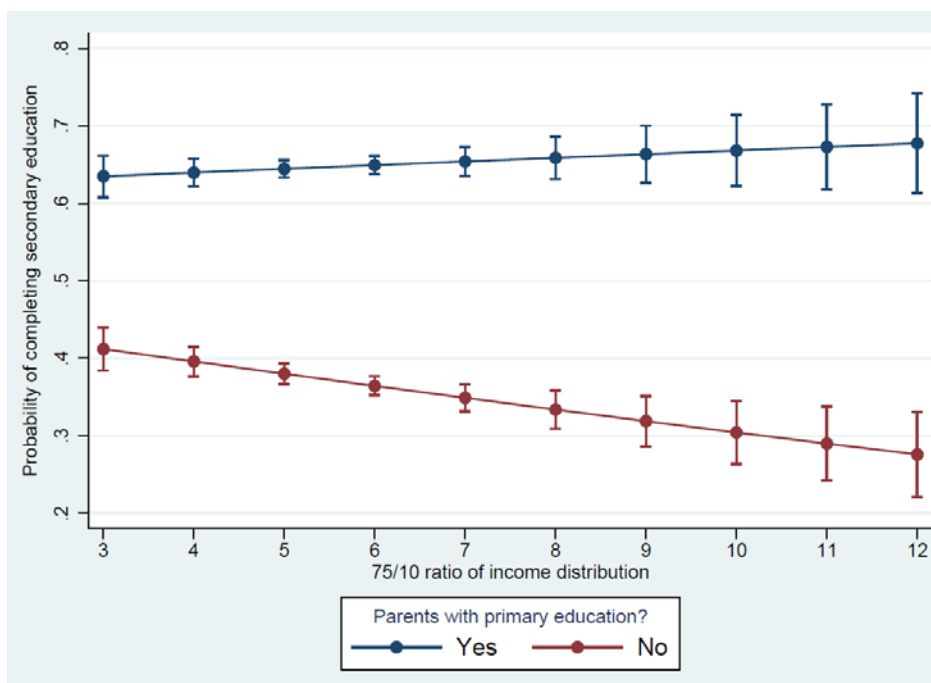
Notes: The 75/10 ratio represents the relation between the income earned by individuals in the 75th percentile and
the earnings of individuals in the 10th percentile. Estimations of income inequality based on the 75/10 ratio
of total income of the economically active population aged 15 or over and with earnings greater than zero.

Tables

Table 1. Weighted Descriptive Statistics (PNAD-2014)

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
State			Population		Income distribution (R$)			Net enrolment ratio in education			Average years of schooling				Standard deviation		
Name	Abbrev.	Total	Average Age	Ratio in rural areas	Bottom (10%)	Middle (50%)	Top (10%)	7–14	15–17	16–24	Observations	Children	Fathers	Mothers	Children	Fathers	Mothers
Rondônia	RO	1.748.531	30,65	0,2369	204	593	1.666	0,9921	0,7948	0,2972	541	8,4842	2,9488	3,2734	4,5993	3,5316	3,9711
Acre	AC	790.101	27,27	0,2591	128	400	1.500	0,9653	0,7616	0,3191	415	7,3954	2,9512	3,7874	5,3104	4,2085	4,7970
Amazonas	AM	3.873.743	28,49	0,1634	150	438	1.500	0,9776	0,8294	0,3482	1.045	8,2446	3,8986	4,1803	4,7067	4,2232	4,2799
Roraima	RR	496.936	28,37	0,1680	191	530	1.862	0,9865	0,7576	0,3238	160	10,0415	4,1797	4,5257	4,2826	4,2936	4,3761
Pará	PA	8.073.924	29,82	0,2992	133	399	1.185	0,9824	0,8454	0,3268	2.019	7,3653	3,3370	3,6478	4,4826	3,7415	4,0299
Amapá	AP	750.912	27,46	0,1039	200	499	1.860	0,9917	0,8541	0,3197	217	9,3247	4,5997	4,2800	4,8566	4,6337	4,4242
Tocantins	TO	1.496.880	31,59	0,2148	164	500	1.674	0,991	0,8205	0,3233	545	7,6599	2,3349	2,9175	4,7665	3,3784	3,9490
North		**17.231.027**	**29,50**	**0,2408**	**146**	**437**	**1.433**	**0,9826**	**0,829**	**0,3277**	**4.942**	**7,8259**	**3,3986**	**3,7343**	**4,6621**	**3,9186**	**4,1553**
Maranhão	MA	6.850.884	30,14	0,4083	89	333	1.015	0,9825	0,8508	0,2747	864	6,2833	1,9498	2,6083	4,9003	3,3379	3,8792
Piauí	PI	3.194.718	32,26	0,3247	117	400	1.114	0,9879	0,8552	0,3149	628	6,0786	1,7586	2,2864	5,1269	3,2388	3,8753
Ceará	CE	8.842.791	33,38	0,2648	120	400	1.134	0,9793	0,8346	0,2742	2.018	6,7593	2,3395	2,8007	4,9012	3,7040	3,9939
Rio Grande do Norte	RN	3.408.510	32,90	0,2352	150	434	1.314	0,9936	0,8235	0,2861	560	6,9774	2,3294	3,1026	4,8696	3,5646	4,0347
Paraíba	PB	3.943.885	33,08	0,1839	145	436	1.400	0,9752	0,7962	0,2943	672	6,7942	2,7729	3,0116	5,1071	4,1112	4,1259
Pernambuco	PE	9.277.727	33,50	0,1894	140	437	1.308	0,9831	0,8162	0,2627	2.541	7,4292	3,4088	3,5120	4,9493	4,3138	4,3208
Alagoas	AL	3.321.730	30,91	0,2833	95	348	1.005	0,9706	0,7744	0,2814	563	6,3830	2,9068	3,0360	5,1566	4,1516	4,2944
Sergipe	SE	2.219.574	32,09	0,2812	156	431	1.200	0,977	0,8352	0,3242	652	6,5579	1,9785	2,4805	4,8874	3,4009	3,5932
Bahia	BA	15.126.371	33,11	0,2488	139	431	1.400	0,985	0,8461	0,3171	3.359	7,1842	2,8426	2,9858	4,9479	3,9217	4,1399
North-east		**56.186.190**	**32,62**	**0,2632**	**126**	**403**	**1.250**	**0,9827**	**0,8315**	**0,2904**	**11.857**	**6,8927**	**2,6340**	**2,9858**	**4,9778**	**3,8801**	**4,1025**
Minas Gerais	MG	20.734.097	34,77	0,1544	236	706	1.933	0,9876	0,8674	0,2785	4.111	7,8235	3,4013	3,4944	4,7173	3,7202	3,8092
Espírito Santo	ES	3.885.049	34,25	0,1553	225	700	2.066	0,9728	0,8133	0,324	772	8,3036	3,3526	3,2428	4,5309	3,6945	3,6406
Rio de Janeiro	RJ	16.461.173	36,80	0,0268	266	750	2.566	0,9899	0,8736	0,3319	3.085	9,5006	6,0272	5,4156	4,3063	4,7137	4,2991
São Paulo	SP	44.035.304	35,29	0,0344	326	860	2.650	0,9938	0,8645	0,2907	4.339	9,9718	5,1388	4,7980	4,3038	4,4431	4,3142
South-east		**85.115.623**	**35,41**	**0,0676**	**277**	**766**	**2.405**	**0,9905**	**0,8645**	**0,2968**	**12.307**	**9,2358**	**4,7551**	**4,4993**	**4,5212**	**4,3904**	**4,2187**
Paraná	PR	11.081.692	34,52	0,1252	315	817	2.325	0,9881	0,8231	0,2923	2.301	8,3716	3,8072	3,6635	4,6968	3,8634	3,9054
Santa Catarina	SC	6.727.148	35,59	0,1589	390	1000	2.500	0,9914	0,8207	0,2948	1.117	8,7443	4,3552	4,1688	4,6119	3,7287	3,6662
Rio Grande do Sul	RS	11.207.274	36,90	0,1501	300	850	2.500	0,9869	0,837	0,3084	3.702	8,4965	4,0837	3,9188	4,4833	4,0926	3,9434
South		**29.016.114**	**35,69**	**0,1426**	**320**	**870**	**2.444**	**0,9884**	**0,8278**	**0,2989**	**7.120**	**8,5022**	**4,0360**	**3,8726**	**4,5912**	**3,9347**	**3,8773**
Mato Grosso do Sul	MS	2.619.657	33,19	0,1078	300	766	2.232	0,9849	0,7832	0,2918	615	8,5436	4,1849	4,0337	4,7883	4,5221	4,5129
Mato Grosso	MT	3.224.357	32,12	0,1720	285	733	2.000	0,9911	0,7919	0,284	627	8,7704	4,2925	4,7669	4,8419	4,6268	4,7849
Goiás	GO	6.523.222	33,35	0,0776	268	724	1.912	0,9923	0,8126	0,3153	1.361	8,0773	3,0697	3,4046	4,7121	3,7236	4,0061
Distrito Federal	DF	2.852.372	32,69	0,0442	301	1000	5.000	0,9931	0,8955	0,414	408	11,7179	7,1630	7,2125	3,8968	5,2154	5,3176
West Central		**15.219.608**	**32,94**	**0,0965**	**285**	**750**	**2.500**	**0,9909**	**0,8193**	**0,3237**	**3.011**	**8,7258**	**4,0041**	**4,2698**	**4,7976**	**4,4461**	**4,6084**
Brazil		**202.768.562**	**33,55**	**0,1494**	**200**	**662**	**2.000**	**0,987**	**0,8427**	**0,3002**	**39.237**	**8,3237**	**3,9598**	**3,9380**	**4,7947**	**4,2511**	**4,1976**

Notes: Column 1 refers to the IBGE estimation based on the PNAD-2014 data. Columns 2 to 9 are the author's own estimates based on all the observations from PNAD-2014. The values from columns 10 to 16 have been determined on the basis of the PNAD-2014 mobility supplement. The income distribution is based on monthly per capita domiciliary income. Bottom, middle, and top represent the poorest 10 per cent, the middle 50 per cent, and the richest 10 per cent, respectively, of the income distribution.

Table 2. Correlation Coefficients by Birth Cohort

State		Cohort: 1940–1989		Cohort: 1940–1949		Cohort: 1950–1959		Cohort: 1960–1969		Cohort: 1970–1979		Cohort: 1980–1989	
Name	Abbrev.	Obs.	Correlation	Obs.	Correlation	Obs.	Correlation	Obs.	Correlation	Obs.	Correlation	Obs.	Correlation
Rondônia	RO	669	0.379***	56	0.115	84	0.611***	121	0.260**	191	0.304***	217	0.512***
Acre	AC	325	0.502***	19	0.254	39	0.655***	45	0.491**	92	0.524***	130	0.526***
Amazonas	AM	915	0.419***	64	0.822***	105	0.463***	152	0.395***	264	0.361***	330	0.424***
Roraima	RR	190	0.253***	7	0.118	32	-0.0118	21	0.373	55	0.232	75	0.401***
Pará	PA	1.673	0.439***	139	0.518***	230	0.561***	315	0.465***	430	0.442***	559	0.428***
Amapá	AP	198	0.351***	12	0.441	25	0.223	35	0.321	48	0.511***	78	0.361**
Tocantins	TO	484	0.370***	56	0.301*	73	0.354**	98	0.213	121	0.476***	136	0.414***
North		**4.454**	**0.425*****	**353**	**0.511*****	**588**	**0.503*****	**787**	**0.401*****	**1.201**	**0.419*****	**1.525**	**0.460*****
Maranhão	MA	620	0.377***	60	0.176	87	0.321**	111	0.384***	156	0.336***	206	0.462***
Piauí	PI	562	0.480***	63	0.696***	86	0.552***	89	0.458***	163	0.426***	161	0.529***
Ceará	CE	1.464	0.440***	147	0.458***	212	0.456***	305	0.469***	344	0.493***	456	0.469***
Rio Grande do Norte	RN	486	0.410***	51	0.631***	51	0.491***	122	0.437***	124	0.369***	138	0.468***
Paraíba	PB	598	0.461***	56	0.400**	74	0.594***	137	0.481***	153	0.559***	178	0.388***
Pernambuco	PE	1.965	0.472***	228	0.437***	291	0.500***	409	0.531***	483	0.545***	554	0.419***
Alagoas	AL	371	0.497***	37	0.710***	59	0.649***	67	0.358**	93	0.500***	115	0.518***
Sergipe	SE	530	0.471***	58	0.634***	70	0.449***	94	0.549***	143	0.484***	165	0.485***
Bahia	BA	2.744	0.488***	263	0.628***	378	0.529***	602	0.469***	644	0.501***	857	0.533***
North-east		**9.340**	**0.466*****	**963**	**0.517*****	**1.308**	**0.519*****	**1.936**	**0.474*****	**2.303**	**0.484*****	**2.830**	**0.493*****
Minas Gerais	MG	3.746	0.454***	415	0.637***	608	0.433***	759	0.459***	945	0.450***	1.019	0.491***
Espírito Santo	ES	733	0.451***	56	0.529***	118	0.523***	152	0.310***	193	0.572***	214	0.457***
Rio de Janeiro	RJ	2.813	0.510***	359	0.529***	527	0.591***	536	0.418***	668	0.481***	723	0.575***
São Paulo	SP	4.565	0.449***	492	0.524***	766	0.495***	906	0.449***	1.161	0.496***	1.240	0.448***
South-east		**11.857**	**0.472*****	**1.322**	**0.556*****	**2.019**	**0.511*****	**2.353**	**0.452*****	**2.967**	**0.488*****	**3.196**	**0.499*****
Paraná	PR	2.237	0.409***	215	0.456***	375	0.449***	507	0.363***	528	0.424***	612	0.480***
Santa Catarina	SC	1.083	0.407***	93	0.537***	182	0.456***	256	0.318***	275	0.439***	277	0.515***
Rio Grande do Sul	RS	3.120	0.439***	371	0.454***	586	0.489***	671	0.444***	693	0.423***	799	0.501***
South		**6.440**	**0.421*****	**679**	**0.480*****	**1.143**	**0.467*****	**1.434**	**0.386*****	**1.496**	**0.427*****	**1.688**	**0.500*****
Mato Grosso do Sul	MS	674	0.476***	59	0.809***	104	0.464***	131	0.534***	175	0.546***	205	0.399***
Mato Grosso	MT	695	0.419***	50	0.368*	100	0.444***	143	0.421***	187	0.407***	215	0.461***
Goiás	GO	1.375	0.356***	129	0.271**	217	0.454***	267	0.319***	349	0.399***	413	0.408***
Distrito Federal	DF	922	0.492***	85	0.600***	107	0.590***	171	0.441***	269	0.475***	290	0.452***
West Central		**3.666**	**0.445*****	**323**	**0.482*****	**528**	**0.501*****	**712**	**0.428*****	**980**	**0.458*****	**1.123**	**0.473*****
Brazil		**35.757**	**0.475*****	**3.640**	**0.536*****	**5.586**	**0.517*****	**7.222**	**0.454*****	**8.947**	**0.492*****	**10.362**	**0.533*****

Notes: Estimations based on OLS regressions using years of schooling of children and their (better-educated) parent. Results are controlled by the variation over time in standard deviation in education. The lower the correlation coefficients, the lower the persistence in education across generations (or the higher the level of mobility). Statistically significant: *p<0.05, **p<0.01, ***<0.001.

Source: PNAD-2014, author's own estimates.

Table 3. The Impact of Inequality on Educational Attainment

Birth cohort	All	1965	1970	1975	1980	1985
Birth years	(1965-1989)	(1965-1969)	(1970-1974)	(1975-1979)	(1980-1984)	(1985-1989)
Socio-economic Marginalisation # Inequality	-0.0542***	-0.0486	-0.0353	-0.00346	-0.0982**	-0.0519
	(0.0186)	(0.0644)	(0.0473)	(0.0335)	(0.0486)	(0.0462)
Socio-economic Marginalisation	-0.405***	-0.535	-0.492*	-0.705***	-0.261	-0.270
	(0.104)	(0.350)	(0.287)	(0.208)	(0.251)	(0.237)
Inequality	0.0129	-0.0256	0.0387	-0.0164	0.0929***	-0.0170
	(0.0138)	(0.0503)	(0.0377)	(0.0242)	(0.0327)	(0.0320)
Male	-0.203***	-0.0988*	-0.114**	-0.217***	-0.291***	-0.257***
	(0.0208)	(0.0531)	(0.0487)	(0.0455)	(0.0425)	(0.0445)
Rural	-0.650***	-0.575***	-0.812***	-0.784***	-0.687***	-0.490***
	(0.0333)	(0.0854)	(0.0831)	(0.0723)	(0.0674)	(0.0692)
Living with both parent	0.0830***	-0.0687	0.0312	0.0614	0.0505	0.256***
	(0.0248)	(0.0683)	(0.0601)	(0.0537)	(0.0497)	(0.0494)
Birth year	0.0159***	0.00722	-0.00521	0.0157	0.0116	-0.0373**
	(0.00156)	(0.0184)	(0.0171)	(0.0161)	(0.0155)	(0.0158)
White (reference)	-	-	-	-	-	-
Black	-0.160***	-0.189**	-0.217**	-0.0887	-0.184**	-0.152*
	(0.0366)	(0.0927)	(0.0872)	(0.0801)	(0.0730)	(0.0787)
Mixed (white/black)	-0.271***	-0.374***	-0.305***	-0.302***	-0.256***	-0.166***
	(0.0222)	(0.0568)	(0.0521)	(0.0489)	(0.0455)	(0.0476)
Asian	0.296*	0.635	0.938***	0.356	-0.413	0.130
	(0.159)	(0.387)	(0.353)	(0.373)	(0.282)	(0.355)
Indigenous	-0.346*	-0.653	-0.555	-0.209	-0.147	-0.381
	(0.191)	(0.619)	(0.457)	(0.396)	(0.357)	(0.362)
Constant	-30.86***	-13.43	10.54	-30.38	-22.57	74.74**
	(3.109)	(36.29)	(33.80)	(31.78)	(30.73)	(31.47)
Observations	23,008	3,699	4,223	4,724	5,387	4,975

Notes: Statistically significant: *p<0.05, **p<0.01, ***<0.001. Standard errors in parentheses.
Source: PNAD-2014, author's own estimates.

Table 4. Robustness Checks

Model	Main Model	Alternative Econometric Approaches			Alternative Model Specifications			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Model	Probit	LPM	LPM	Bivariate Probit	Probit	Probit	Probit	Probit
Estimation method	MLE	OLS	2SLS	MLE	MLE	MLE	MLE	MLE
Changes to Specidication (1)	-	No	No	No	No Migrants	Ratio 90/10	Illiterate Parents	IlliteP & NoEduP
Coefficient of MSB # Inequality	-0.0540***	-0.0179***	-0.0175***	-0.0487***	-0.0695*	-0.0199**	-0.0318	-0.0307
	(0.0186)	(0.00636)	(0.00636)	(0.0178)	(0.0418)	(0.00796)	(0.0227)	(0.0216)
APRs for MSB and Inequality								
1bn._at	-0.223***	-0.220***	-0.223***	-0.175***	-0.210***	-0.209***	-0.236***	-0.201***
	(0.0196)	(0.0179)	(0.0179)	(0.0212)	(0.0430)	(0.0271)	(0.0257)	(0.0227)
2._at	-0.244***	-0.238***	-0.241***	-0.190***	-0.237***	-0.217***	-0.247***	-0.213***
	(0.0134)	(0.0125)	(0.0125)	(0.0210)	(0.0288)	(0.0242)	(0.0184)	(0.0154)
3._at	-0.264***	-0.256***	-0.258***	-0.205***	-0.263***	-0.224***	-0.258***	-0.225***
	(0.00884)	(0.00863)	(0.00861)	(0.0210)	(0.0181)	(0.0213)	(0.0130)	(0.0101)
4._at	-0.284***	-0.274***	-0.276***	-0.220***	-0.289***	-0.232***	-0.268***	-0.237***
	(0.00876)	(0.00853)	(0.00851)	(0.0215)	(0.0183)	(0.0185)	(0.0115)	(0.0102)
5._at	-0.305***	-0.291***	-0.293***	-0.236***	-0.314***	-0.240***	-0.278***	-0.248***
	(0.0130)	(0.0123)	(0.0123)	(0.0223)	(0.0286)	(0.0159)	(0.0149)	(0.0155)
6._at	-0.324***	-0.309***	-0.311***	-0.253***	-0.338***	-0.248***	-0.288***	-0.260***
	(0.0187)	(0.0177)	(0.0177)	(0.0235)	(0.0417)	(0.0133)	(0.0205)	(0.0225)
7._at	-0.344***	-0.327***	-0.328***	-0.270***	-0.361***	-0.255***	-0.298***	-0.272***
	(0.0248)	(0.0235)	(0.0235)	(0.0251)	(0.0553)	(0.0111)	(0.0267)	(0.0300)
8._at	-0.363***	-0.345***	-0.346***	-0.287***	-0.383***	-0.263***	-0.307***	-0.283***
	(0.0309)	(0.0296)	(0.0296)	(0.0271)	(0.0687)	(0.00927)	(0.0330)	(0.0376)
9._at	-0.382***	-0.363***	-0.363***	-0.305***	-0.404***	-0.271***	-0.315***	-0.295***
	(0.0369)	(0.0358)	(0.0358)	(0.0294)	(0.0819)	(0.00822)	(0.0392)	(0.0452)
10._at	-0.401***	-0.381***	-0.381***	-0.323***	-0.423***	-0.278***	-0.324***	-0.306***
	(0.0427)	(0.0420)	(0.0420)	(0.0320)	(0.0947)	(0.00819)	(0.0452)	(0.0528)
Observations	23,008	23,008	23,008	23,008	5,340	23,008	24,842	21,668

Notes: The coefficients of the interaction between socio-economic marginalisation (MSB) and the inequality level show how the effects of having (un)educated parents on the children's schooling level vary with different values of inequality. The adjusted predictions at representative values (APRs) fix the covariate "75/10 ratio" to each of the 10 deciles of the inequality distribution, showing the gap in the chances of achieving a secondary school certificate between the two populations investigated: children from parents with and without (primary) education. For the LPMs, the standard errors are robust to arbitrary heteroskedasticy. Statistically significant: *p<0.05, **p<0.01, ***p<0.001. Standard errors in parentheses. All predictors at their mean value.

Source: PNAD-2014, author's own estimates.

Appendix

A   Model of the Intergenerational Transmission of Income Inequality

The Solon's (2004) model, which is based on the theoretical approach of Becker and Tomes (1979, 1986), has been used in the economic literature as the starting point to understand the correlation between income inequality and intergenerational mobility.[39]

In this model, the family i is composed of a parent from generation $t-1$ and a child from generation t and it is assumed that an intergenerational decision maker should allocate the lifetime earnings gained from parent $y_{i,t-1}$ into only two goods: the parent's own consumption $C_{i,t-1}$ and investment in the child's human capital $I_{i,t-1}$, subject to the budget constraint:[40]

$$y_{i,t-1} = C_{i,t-1} + I_{i,t-1} \qquad (A.1)$$

The investment $I_{i,t-1}$ presents diminishing marginal returns on education and will form the child's stock of human capital h_{it} used in the future to produce economic value in the labour market

$$h_{it} = \theta \, log \, I_{i,t-1} + e_{it} \qquad (A.2)$$

The diminishing marginal returns on education in equation 2 refer to the fact that (more) investments produce even a positive marginal product for human capital stock $\theta > 0$, but of a lesser and less additional value, due to the semi-log form of the function. The error term e_{it} accounts for the variation in the child's human capital endowment that cannot be explained by the investment of parents, referring mainly to genetic endowment and personality traits which are also transmitted in the family environment and play an important role in human capital accumulation.

The independent human capital endowment e_{it} follows a first-order autoregressive process as:

$$e_{it} = \delta + \lambda \, e_{i,t-1} + v_{it} \qquad (A.3)$$

in which $v_{i,t}$ represents a white-noise error term and the parameter λ is a heritability coefficient with $\lambda \in [0,1]$.

Solon (2004) assumed in his model that the lifetime income of child y_{it} can be regarded as a semi-log earnings function, where p represents the earnings return on human capital.

$$log \, y_{it} = \mu + \rho \, h_{it} \qquad (A.4)$$

Substituting equation (A.2) into equation (A.4), we have:

$$log \, y_{it} = \mu + \rho \, \theta \, log \, I_{i,t-1} + \rho \, e_{it} \qquad (A.5)$$

---

39 The description of Solon's (2004) theoretical model in this appendix refers to the simplified version of the model presented by Solon (2014).

40 For purposes of simplification, Solon's (2004) model as presented in this section does not take into account variables such as taxation, government investment in children, and the borrowing and bequeathing of financial assets. See Becker and Tomes (1986) for a more complete version of the model.

From equation (A.5) the value $p\,\theta$ can be interpreted as the elasticity of the child's income in relation to the human capital investment $I_{i,t-1}$, representing in this way the earnings return on human capital investment, in the following briefly termed γ.

$$log\ y_{it} = \mu + \gamma\ log\ I_{i,t-1} + \rho\ e_{it} \qquad (A.6)$$

To make the optimal decisions concerning the investment in the child's human capital, the family considers a two-good world, in which the parents' lifetime income $y_{i,t-1}$ needs to be allocated between their own consumption $C_{i,t-1}$ and investment $I_{i,t-1}$ in the child's human capital. The family wishes to maximise the utility, denoted $U_{i,}(C_{i,t-1}, I_{i,t-1})$, subject to the budget constraint in (A.7):

$$U_i = (1-\alpha)\ log\ C_{i,t-1} + \alpha\ log\ y_{it} \qquad (A.7)$$

with $\alpha \in [0,1]$ indicating the degree of altruism of parents for child's income y_{it} in relation to their own consumption $C_{i,t-1}$. Plugging (A.1) and (A.6) into equation (A.7):

$$U_i = (1-\alpha)\ log\ (y_{i,t-1} - I_{i,t-1}) + \alpha\ (\mu + \gamma\ log\ I_{i,t-1} + \rho\ e_{it}) \qquad (A.8)$$

And rewriting it:

$$U_i = (1-\alpha)\ log\ (y_{i,t-1} - I_{i,t-1}) + \alpha\ \mu\ + \alpha\ \gamma\ log\ I_{i,t-1} + \alpha\ \rho\ e_{it}) \qquad (A.9)$$

In order to solve the problem, the main condition for the maximisation of the utility function is that

$$\frac{\partial U_i}{\partial I_{i,t-1}} = \frac{-(1-\alpha)}{y_{i,t-1} - I_{i,t-1}} + \frac{\alpha\gamma}{I_{i,t-1}} = 0 \qquad (A.10)$$

Solving for the optimal choice of $I_{i,t-1}$, we can rewrite the first-order condition as

$$I_{i,t-1} = \left\{ \frac{\alpha\gamma}{1-\alpha(1-\gamma)} \right\} y_{i,t-1} \qquad (A.10)$$

Note that the investment in the child's human capital $I_{i,t-1}$ increases by increasing parents' income $y_{i,t-1}$, altruism α, and earnings return on human capital investment γ. From these results, we can deduce the two most important conclusions from Solon's (2004) model: Firstly, parents with higher income have a higher financial capacity to invest in the human capital of their children, and secondly, they also have a greater incentive to make this investment if the return on investment in human capital increases over time.

B Stylised Model of the Decision to Drop Out of the Education System

Kearney and Levine (2014) presented the theoretical model used in this paper to explain the causal relationship between higher income inequality and the higher probability that children from socially disadvantaged families will drop out of school.

Let us assume that the child (student) tends to maximise an intra-generational utility function between utility in the current (t) and future period ($t + 1$). If the student drops out of the education system in (t), he or she will achieve the current-period utility u^d and the present discounted sum of future period V^d. Otherwise, the student has u^e and V^e from the decision

to remain enrolled in the school. The generalisation of the individual decision to drop out can be written as:

$$u^d + E(V^d) > u^e + E(V^e) \qquad (B.1)$$

Given the positive returns on education, we assume that dropping out of school has a negative effect on the utility in period $t + 1$, due to the reduction in the level of future consumption, such that $E(V^e) > E(V^d)$.

The decision to drop out of education will be never optimal so long as $u^d \leq u^e$. However, if $u^d > u^e$ in the case that the student's participation in the school system is associated with substantial utility costs, such as psychic costs, then dropping out of school can be considered an alternative.

Suppose that the child's utility in the future can achieve U^{high} or U^{low} – i.e. a high or low value, respectively – and U^{low} represents the utility level in the case that the child drops out of school. If the student remains enrolled in the education system, he or she will have the probability $p \in [0,1]$ of attaining the high-utility position. Assuming V^{low} as the deterministic present discounted value of the utility, we can rewrite equation (B.1) as:

$$u^d + V^{low} > u^e + pV^{high} + (1 - p)V^{low} \qquad (B.2)$$

By rearranging the terms in equation (B.2), the condition for remaining in school yields:

$$\left[pV^{high} + (1 - p)V^{low} \right] - V^{low} > u^d - u^e \qquad (B.3)$$

Thus, the student will continue studying as long as the likelihood of attaining a high utility in the future is greater than the current loss of utility caused by school attendance and his/her consequent sacrifice of leisure. Given the uncertainty associated with the future, the child cannot determine p with the best possible accuracy in the period t, working in this way with its individual subjective perception of success q.

Let's assume q as a function of p and x, such that $q = q(p, x)$ in which x represents external factors affecting the individual's perception of returns on schooling. Kearney and Levine (2014) have pointed out that these external factors can be influenced strongly by the lived experience during childhood and adolescence. Children who grow up in poverty have restricted contact to highly qualified individuals and may assume that a college degree is an objective very far from their reality, leading to an underestimation of the probability p. As a result, at the same level of p, students from different socio-economic backgrounds (SES) will present different individual perceptions of q.

This means that income inequality will affect the perceived returns on education p in two ways: Firstly, it affects x given that the higher the inequality, the higher the perception of social exclusion for poor children. Secondly, higher income inequality will increase the current return on investment in schooling, leading to a rise in the individual perception of return p. Then the condition for the student to continue studying follows:

$$\left[qV^{high} + (1 - q)V^{low} \right] > V^{low} + (u^d - u^e) \qquad (B.4)$$

From equation (B.4) it becomes evident that the chance of remaining enrolled rises with increasing q. Therefore, the student will invest more time in schooling if he or she notes that this investment will increase the chance of achieving V^{high}. However, if children are right in assuming that independently of their educational attainment they will never leave the situation of social exclusion, i.e. if q is very low, this increases the incentive to drop out of school.

Solving the equation (B.4) for q, we can define the reservation subjective probability q^r required for students' continuation of schooling.

$$q \geq q^r = \left\{ \frac{u^d - u^e}{V^{high} - V^{low}} \right\} \qquad (B.5)$$

The derivative from (B.5) to the socio-economic backgrounds (SES) represented an increasing function at point q, indicating that the higher the SES, the greater the perception of success as a consequence of educational attainment, such that:

$$\frac{\partial q}{\partial (SES)} > 0 \qquad (B.6)$$

Kearney and Levine (2014) propose that the perception of success q can also be described as a function of SES and income inequality in the society. For children from socially weaker families, the increase in the gap between the bottom and middle of the income distribution might lead to a reduction of the subjective perception p.

$$\frac{\partial q}{\partial (SocIneq)} < 0 \qquad (B.7)$$

In practice, it means that the farther away poor children's experiences are from the experiences of the middle class, the greater their perception of "social exclusion," strengthening in this way the individual view that "it is not for people like me."

C   Structure of the Brazilian Educational System

The current Brazilian educational system is anchored in the 1988 Constitution, which recognises education as a right for the population and an obligation of the government.

The same legislation distributes the responsibility for education between all three administrative levels of the federation: the federal, state, and municipal governments. Thus, the municipalities are responsible for providing and regulating pre-school education, while the states are involved with the same tasks for the primary and secondary education. The federal government plays only a secondary role in this context, providing financial and technical assistance to the states and municipalities in order to promote equality of opportunity and minimum quality standards.

The main responsibility of the federal government lies in providing education in its institutions – the vast majority of them related to tertiary education – and in regulating the private sector, which is free to operate within all three educational levels.[41] After 1996, which saw the

---

41 The author's own calculations on the basis of the National Household Sample Survey (PNAD) from 2014 indicate that the share of students enrolled in private institutions in Brazil reached 14 per cent in primary education, 13 per cent in secondary education, and 75 per cent in tertiary education.

publication of the Law of Directives and Bases of National Education (Lei de Diretrizes e Bases da Educação) or LDB, the central government also became responsible for defining a common national basis for curriculum in primary and secondary education, which needs to be used by states and municipalities as the basis for the development of their own curriculums.

Since the 1934 Constitution, there has been compulsory education in Brazil. However, in the beginning only children aged between 7 and 10 years were obliged to undertake full-time education. Over the years the obligatory period of schooling has grown steadily, so that in 1971 compulsory education ended at the age of 14, and in 2010 at 17.[42] Table A1 provides an overview of the Brazilian educational system and the changes made to it over the last six decades.

Table A1. Structure of Brazilian Educational System

Year	Level	Duration (in years)	Age group	Compulsory
Until 1971	Pré-escola (Pre-school)	3	4 to 6	No
	Escola primária (Primary school)	4	7 to 10	Yes
	Ginásio (Lower high school)	4	11 to 14	No
	Colégio (High school)	3	15 to 17	No
	Ensino superior (College)	variable	>=18	No
1971 to 1995	Pré-escola (Pre-school)	3	4 to 6	No
	1º grau (1st Degree)	8	7 to 14	Yes
	2º grau (2nd Degree)	3	15 to 17	No
	Ensino superior (College)	variable	>=17	No
1996 to 2009	Educação infantil (Early childhood education)	7	0 to 6	No
	Ensino fundamental (Primary education)	8	7 to 14	Yes
	Ensino médio (Secondary education)	3	15 to 17	No
	Ensino superior (College)	variable	>=17	No
Since 2010	Educação infantil (Early childhood education)	4	0 to 3	No
	Educação pré-fundamental (Pre-primary education)	2	4 to 5	Yes
	Ensino fundamental (Primary education)	9	6 to 14	Yes
	Ensino médio (Secondary education)	3	15 to 17	Yes
	Ensino superior (College)	variable	>=17	No

Sources: Law 4,024 of 20 December 1961, Law 5,540 of 28 November 1968, Law 5,692 of 11 August 1971, and Constitutional Amendment 59 of 11 November 2009.

Currently, the Brazilian educational system is composed of five distinct levels: early childhood, pre-primary, primary, secondary, and tertiary education. Individuals aged between 4 and 17 years are obliged to attend school. Children under four may attend the optional early childhood education. Attendance at the pre-primary educational level, usually at the age of 4, is the first phase of compulsory education. This is followed by the primary educational level, which comprises nine years of schooling. The third level of the educational system in Brazil is known as secondary education and lasts for a period of three years. Students who complete this level have the right to attend vocational training, or to start pursuing higher education qualifications: a bachelor's degree, for example, usually takes four years. Individuals who hold a university degree are eligible to undertake graduate studies, which consist of a master's degree followed, potentially, by a doctoral degree.

---

42 See Wjuniski (2013) for a detailed description of the changes over time in the legal framework for the educational system in Brazil.

The current requirement that children complete 14 years of compulsory education in Brazil was stipulated by constitutional amendment 59 of 11 November 2009, which created an obligatory 2+9+3 pattern in the education system. This was an increase from the previous system (valid until 2009), where students were required to remain in school only for nine years.[43]

D Codification of Years of Schooling

Based on the PNAD sample, this paper used two main variables related to education for the investigation of intergenerational mobility: the number of completed years of education (years of schooling) and the (highest) educational level achieved. The PNAD already provides both variables for the children's generation, but for the parents the information on years of schooling is missing.

Given this limitation, I calculated the parents' years of schooling according to their educational levels. Table A2 presents the matching procedure used for the codification.[44]

Table A2. Codification of Parents' Years of Schooling

	00	if	Only pre-primary education
	00	if	Went to school, but never completed 1st grade
	02	if	Completed 1st grade but didn't complete all grades up to 4th grade (before 1971)
	02	if	Uncompleted literacy classes (young people and adults)
	03	if	Attended literacy classes (young people and adults), but do not know if they were completed
	03	if	Attended primary school, but do not know if all grades up to 4th grade were completed (before 1971)
	04	if	Completed up to 4th grade
	05	if	Completed literacy classes (young people and adults)
	05	if	Completed 1st grade but didn't complete all grades up to 8th grade (after 1971)
	06	if	Completed 5th grade but didn't complete all grades up to 8th grade (before 1971)
	07	if	Attended 1st degree, but do not know if all grades up to 8th grade were completed (after 1971)
Years of education	07	if	Attended lower high school, but do not know if all grades up to 8th grade were completed (before 1971)
	08	if	Completed up to 8th grade
	09	if	Completed 9th grade but didn't complete all grades up to 11th grade
	10	if	Attended 2nd degree, but do not know if all grades up to 11th grade were completed (after 1971)
	11	if	Completed up to 11th grade
	13	if	Completed 1st year in college/university, but didn't graduate
	14	if	Attended college/university, but do not know if graduated
	15	if	Graduated college/university
	16	if	Incomplete master's or doctorate
	17	if	Attended master's or doctoral studies, but do not know if they were completed
	19	if	Completed master's or doctorate

---

43 Although the increasing of the compulsory education level via the constitutional amendment had already been established in September 2009, the states and municipalities had until 2016 to achieve its full implementation.

44 See Table 5 for an overview of the different educational levels used in the codification.

It is important to note that the information concerning the parents' educational level is based on the self-declaration of their children – i.e. the individuals who were interviewed by PNAD – and refers to educational attainment of parents when the children were 15 years old.[45] Thereby, three variables from PNAD have been used for the codification of parents' years of schooling: (a) the highest level of education attended, (b) whether the first year (grade) of this attended level was completed, and (c) whether the attended level was also completed.

For the first variable, 10 different educational levels were permitted: kindergarten, literacy classes for six-year-olds, literacy classes for young people and adults, primary school, lower high school, high school, 1st degree/primary education, 2nd degree/secondary education, college, and master's or doctorate. For the second and third variables only three answers were possible: yes, no, or unknown.

E   Data Harmonisation

The empirical investigations in this paper are based on the Brazilian National Household Sample Survey (PNAD). This nationally representative survey has been conducted annually since 1981 by the Brazilian Institute of Geography and Statistics (IBGE) and gathers information about household composition, educational attainment, labour market status, income, and a set of demographic variables (age, gender, location, race, etc.).[46]

In principle, it is possible to observe a relative consistency between the different sets of PNAD microdata over time. However, through the years the PNAD has undergone some restructuring in methodological terms, and for this reason some variables are not available for all the years and/or may not have been collected in the same way.

For this paper, the particularly relevant change was the reformulation of the definition of labour activities that occurred in 1992. The new formulation aimed to integrate some subsamples of the population involved in economic activities that were previously not included in the occupied population. Particularly noteworthy was the establishment of three additional categories of workers: those involved in production for self-consumption, construction for their own personal use, and paid domestic work. For this reason, it was necessary to harmonise all the PNAD's microdata to make the information about income inequality used in the investigations of the Great Gatsby curve and the decision to drop out of school compatible.

---

45 Because the investigation of intergenerational mobility in this paper is based on children born between 1940 and 1989, the reform of the education system through constitutional amendment 59 of 11 November 2009 had no consequences for the codification of parents' education.

46 Until 2003, the rural areas of Rondônia, Acre, Amazonas, Roraima, Pará e Amapá were not part of the PNAD. These six states compose Brazil's northern region and their rural population constitutes approximately 3 per cent of the total Brazilian population.

For the standardisation process I took the survey from 1981 as the initial base and made the subsequent PNADs compatible with it. This required that only those variables which already existed in the 1981 sample be maintained for the investigation.[47]

For the measures of Gini coefficient and 75/10 ratio, I followed the theoretical approach of Hoffmann (2006) and calculated the income inequality based on (positive) monthly personal income for the economically active population aged 15 or over. In the integrated data there are 12 variables related to income that are common to all the samples. For the investigation, I used the variable (personal) monthly income from all sources, which is derived from the sum of all job income, retirement, pension, rent, allowances and other sources. Subsequently, the variable related to income was deflated to the year 2012 with help of an income deflation based on the National Consumer Price Index (INPC/IBGE). Not least, I omitted the observations with income equalling zero to exclude those individuals performing unpaid work (care work, voluntary work, etc.) from the analysis.

In this paper, the economically active population consists of those individuals who were either employed or actively seeking employment in the PNAD reference week. Finally, because the state of Tocantins was created only in year 1989, I aggregated its data with Goiás for those years in which the separation had already occurred.[48]

The investigation of mobility conducted in this study is based on an intertemporal choice about (more) educational attainment that occurred 14 years after the birth of the children. Therefore, the first PNAD sample (1981) was used to investigate the educational choices of the individuals born in 1967, and the PNAD from 2003 for the investigation of children born in 1989. Because there are no nationally representative databases for the period prior to 1981 that could be harmonised in a reliable way with the PNADs generated after 1981, this paper limited the estimations in sections 5.2 and 5.3 to the individuals born from the year 1965 onwards.[49]

---

47 This standardisation process was undertaken using the "datazoom-pnad" package developed by the Department of Economics at Pontifical Catholic University of Rio de Janeiro (PUC-Rio), which aimed to compile all the variables from the last four decades that could be obtained and organised in a conceptually consistent way.

48 In the 1988 Brazilian Constitution, the state of Tocantins was officially created from the northern two-fifths of Goiás and admitted as a new state.

49 For the individuals born in 1965 and 1966, I used the inequality level from 1981 as a proxy. In the years 1991, 1994, and 2000 the PNAD was not carried out. For that purpose, I used the inequality levels for the respective following years (1992, 1995, and 2001) via the investigation of individuals born in 1977, 1980, and 1986, respectively.

GIGA Working Papers

Recent Issues