UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL, HUMAN AND MATHEMATICAL SCIENCES

Department of Economics

Educational Mismatch and Labour Market Outcomes in Brazil

by

Larissa da Silva Marioni

Thesis for the degree of Doctor of Philosophy

September 2019

To my parents Rogério and Valéria,
for their endless support and unconditional love.

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL, HUMAN AND MATHEMATICAL SCIENCES
Department of Economics

Doctor of Philosophy

EDUCATIONAL MISMATCH AND LABOUR MARKET OUTCOMES IN BRAZIL

by Larissa da Silva Marioni

In this thesis, I investigate in more depth important topics in labour economics in Brazil. In particular, in the first chapter, I examine the incidence of educational mismatch in the Brazilian formal labour market, and I estimate its effects on wages controlling for workers' heterogeneity. Furthermore, I test what would be the effect on aggregate wages if the mismatched workers were reallocated to a well-matched position. As a result, I find that half of the workers are mismatched, and overeducated (undereducated) workers earn less (more) than their co-workers with the same attained education, but who hold a well-matched position. Additionally, I find that reassigning workers to a well-matched occupation would increase aggregate wages while changing people's education to eliminate the mismatch would decrease aggregate wages.

In the second chapter, I use two measures of the minimum wage to investigate how changes in the minimum wage policy affect investments on education acquisition on-the-job, through changes in the wage distribution and skill premia. The findings show that changes in the absolute minimum wage are associated with compression of the skill premia, particularly for higher education, and has a negative impact on educational investments. On the contrary, the relative minimum wage is associated with an increase in the skill premia for higher education but has no significant effect on education acquisition on-the-job on average.

Finally, in the third chapter, I examine the effects of internal migration on job transitions by instrumenting migration with violent crime shocks at the origin state. In particular, I focus on young people entering the labour market. The results show that without controlling for self-selection, migrants are positively selected and earn more, on average, than non-migrants. In contrast, after the self-selection is taken into account, on average, migrants are negatively selected and earn less. The impact is stronger on men and low-educated workers. In addition, I find that migrants coming from richer states are positively selected, while those coming from poorer states are negatively selected. These findings may indicate forced migration only for poorer origin states.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Larissa da Silva Marioni , declare that the thesis entitled *Educational Mismatch and Labour Market Outcomes in Brazil* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission.

Signed:...

Date:...

Acknowledgements

First, I would like to thank my supervisors, Thomas Gall and Michael Vlassopoulos for their continuous support, patience and advice. Thomas' has been an extraordinary source of advice and his detailed comments and great knowledge helped me in all stages of my PhD. Michael has always been available and he has provided me with great input, especially in decisive moments. I could not have imagined having better supervisors, and this thesis would not have been possible without their guidance. I would like to express my gratitude to Corrado Giulietti, Jackline Wahba, Brendon McConnell and Hector Calvo Pardo for their comments and feedback along the way. I also thank the Economic and Social Research Council for the scholarship. I would like to thank my fellow PhD students, Abu Siddique, Armine Ghazaryan, Lunzheng Li, Marius Strittmatter and Xiaocheng Hu for the exciting discussions, for their help and for all fun we had during the last few years.

I would like to express my special appreciation and thanks to Bruno Soriano for his patience, encouragement and understanding. He has been a great supporter and his love was essential during all times.

Lastly, and very importantly, I would like to thank my family. I am especially grateful to my parents, Rogério and Valéria who always believed in me, gave me strength and courage and made great efforts to make this achievement possible. Thank you for inspiring me to follow my dreams. Thank you to my brother Matheus, for his support and friendship. Also, I would like to thank my grandma Therezinha and my godparents Arthur and Ivani for encouraging me in all of my pursuits.

Abbreviations

BRL	Brazilian Real
CBO	Brazilian Classification of Occupation
DATASUS	Public Health Care System Information
EAD	Distance Education Courses
EDUC	Attained education
EJA	Youth and Adult Education
FE	Fixed Effects estimation
FEIV	Fixed Effects estimation with Instrumental Variables
GDP	Gross Domestic Product
HCT	Human Capital Theory
IBGE	Brazilian Institute of Geography and Statistics
ICD-10	International Classification of Diseases 10th Edition
INPC	National Consumer Price Index
IPCA	Brazilian Consumer Price Index
IV	Instrumental Variable approach
LATE	Local Average Treatment Effect
LPM	Linear Probability Model
MTE	Brazilian Ministry of Labour and Employment
MW	Minimum wage
OECD	Organization for Economic Cooperation and Development
OLS/POLS	Ordinary Least Squares/Pooled Ordinary Least Squares
PROUNI	University for All Program
RAIS	Annual Social Information Report
REQ	Required education
REUNI	Support Program for Restructuring and Expansion Plans of Federal Universities
SIM	Brazil Mortality Information System
WHO	World Health Organization

Chapter 1

Introduction

In the first decade of the 21st century, Brazil experienced a positive performance in its economy and labour market. Inequality, extreme poverty, and unemployment rates declined, while labour earnings, formal employment and the educational level of the workforce increased (Barros et al., 2010; Corseuil and Foguel, 2016; Firpo and Pieri, 2018).

It has been argued that the fast growth in education was a determinant behind the sharp decline in income inequality in Brazil during the 2000s (Barros et al., 2010; Gasparini et al., 2011; Dedecca, 2015). The public expenditure in education as a percentage of the GDP increased by 2 percentage points over the first decade, from 3.5% in 2000 to 5.6% in 2010 (OECD, 2013). In order to understand the changes in education in Brazil, this is how its educational system is structured: primary education is compulsory and it is divided into two, basic education I, which goes from 1st to 5th grade and basic education II from 6th to 9th grade. After completing primary education, students start the non-compulsory secondary education or high school. At age 18, students can enter higher education. In contrast to the United States and the United Kingdom, public universities are the most prestigious ones and they are free of tuition, but access to them is very competitive.

The expansion in the educational system started in earnest in the 90s with the increase in the number of private universities, enrolments in higher education and growth in the distance education courses (*Educação à Distância* - *EAD*). During the 2000s, the government adopted policies that aimed to increase access to higher education for marginalised groups that previously did not have this privilege. For example, the University for All Program *(Programa Universidade Para Todos - PROUNI)*, which facilitates the access of low-income students to private universities by granting full or partial scholarships, and the Support Program for Restructuring and Expansion Plans of Federal Universities *(Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais - REUNI)*, which has promoted the expansion and internalization of federal universities through the growth in government expenditure on public higher education, attempting to build an inclusive higher education by doubling the number of vacancies and to reduce regional inequalities. In addition to the programs, the government also adopted affirmative action policies that prioritize low-income students, for instance by setting quotas at federal universities for students from public schools. These policies were directly responsible for the 44.6% increase in the number of enrolments in higher education in the period 2003-2013 (INEP, 2013a).

However, the educational boom was not limited to higher education. The universalisation of access to basic education and programs such as Youth and Adult Education *(Educação de Jovens e Adultos - EJA)*, which provides free education (mostly evening classes) for young people, adults and the elderly that could not complete fundamental and/or secondary education at the regular age, caused an increase in the average schooling of the workforce, and thus in the percentage of workers with at least secondary education (IBGE, 2014).

As previously discussed, education is crucial to economic development. However, if the increase in the supply of highly educated workers is not followed by the demand side, the educational mismatch may arise. The educational mismatch occurs if workers have more or less education than it is required by their jobs – i.e. overeducation and undereducation, respectively – and it may indicate inefficiencies in the labour market. Therefore, the educational mismatch can contribute to the misallocation of skills if workers become over or underemployed, raising questions about the relationship between education and the labour market (Groot and Van Den Brink, 2000).

The occurrence of educational mismatch, especially overeducation, highlights the misalignment between firms' needs and individuals' training. Firms look for workers with a high stock of human capital or with productive attributes to fulfil their needs. But the educational system may not focus on workers' productive development or may provide low-quality education to students. In this case, it would be better for firms to hire highly educated workers, which are easy to train, even if the occupations do not demand such education (Hartog, 2000).

Some theories argue that educational mismatch is a temporary phenomenon. In the matching theory framework, the educational mismatch could occur due to inefficiencies in the labour market related to imperfect information and mobility (Jovanovic, 1979). For the career mobility theory, the educational mismatch occurs once young individuals are willing to accept jobs for which they are overeducated because of their lack of experience (Sicherman and Galor, 1990). In both frameworks, workers can improve their matching either by changing jobs across firms or changing jobs within the same firm. In fact, some studies found that overeducation is related to higher mobility rates or within firm promotions (Alba-Ramirez, 1993; Sloane et al., 1999; Groot and Van Den Brink, 2000; Groeneveld and Hartog, 2004).

Additionally, the misallocation of resources in frictional labour markets leads to a decrease in economic aggregate productivity (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2013). Hence, the misallocation of talent is relevant for economists since one could increase the efficiency of human capital allocation using a different assignment of the resources. Furthermore, the misallocation problem is more severe in labour markets with more frictions, which is the case for developing countries.

In the literature, the effect of the educational mismatch on productivity is usually estimated indirectly from wages, job satisfaction and other features related to productivity (Hartog, 2000). Nonetheless, most of the studies are for developed countries, and only a few explore this aspect in developing countries (Quinn and Rubb, 2006; Mehta et al., 2011; Reis, 2017).

Hence, a potentially serious problem arises when educational mismatch occurs, particularly in Brazil where high-quality education is still a challenge. That is the focus of Chapter 2. The boost in the educational level of the population in Brazil allows me to check whether the labour market is absorbing or not the highly qualified workers and the effects of educational mismatch on wages, helping policy makers to improve efficiency and efficacy in use of public resources. To the best of my knowledge, the use of a rich employer-employee dataset to analyse the effect of educational mismatch on workers' wages and to assess the impact of reassigning workers to well-matched occupations on aggregate wages, is a pioneering approach to the Brazilian context.

In addition to the expansion in the educational system, the active minimum wage policy was another determinant of the decline in inequality in Brazil during the 2000s. From 2003 to 2013 the real minimum wage increased by more than 70% (Ipeadata, 2017). In the same period, there was also an increase on real earnings since the minimum wage acts as a reference for wage determination. That is, workers' wages are calculated as multiples of the minimum wage, and thus changes in the minimum wage affect not only the bottom part of the distribution. A usual concern is that minimum wage increases may encourage firms to hire workers without a formal contract, rising informality in the labour market. However, Brazil experienced an expansion on its formal labour market driven by the economic growth in the period. The total number of formal jobs increased by 65.7% in ten years according to the Brazilian Institute of Geography and Statistics (IBGE, 2013).

Most of the economic literature focuses on the effects of minimum wage policies on employment (see Brown et al., 1982; Card, 1992; Card and Krueger, 1994; Machin and Manning, 1994; Giuliano, 2013; Dolton et al., 2015). Another concern, the possible effects of minimum wages on education acquisition, has received less attention (see Mattila, 1978; Ehrenberg and Marcus, 1982; Neumark and Wascher, 1995a, 2003). In any event, little work has been done uncovering the underlying mechanisms.

It is well known that the labour market rewards investments in education, thus interventions that affect the labour market equilibrium also impact the individual's decisions to acquire education. Individuals' expectations about the future labour market will influence their educational decisions. Consider for instance firms hiring highly educated workers to take advantage of their abilities. If workers know that this is the case, acquiring more education may give some signals about their abilities to employers. Furthermore, changes in the minimum wage will impact the wage distribution and thus skill premia. On the one hand, if the minimum wage compresses the wage distribution, the returns to education may diminish, discouraging educational investments. On the other hand, if the minimum wage increases the returns to education, particularly higher education, this will encourage further education acquisition.

Hence, changes in the minimum wage policy may alter the skill structure of the workforce, especially if it excludes low-productive workers. For instance, raises in the minimum wage may incentivize higher-ability individuals to drop out of school and pursue a minimum wage job (Neumark and Wascher, 1995a, 2003). Moreover, a higher minimum wage could increase unemployment, especially for low-skilled workers. The reduction of low-skilled workers reduces the skill premium, and thus educational investments. Conversely, the increase in the minimum wage makes it more difficult to find a job, enhancing the role of education (Cahuc and Michel, 1996). Depending on which of these forces prevail, the educational attainment of the workforce may increase or decrease. For example, Bárány (2016) finds that a decrease in the minimum wage in the United States led to a reduction in the skill premium and thus, reshaped the share of low and high-skilled workers.

Differently from developed countries, minimum wages and education acquisition are not necessarily substitutes in Brazil. As discussed at the beginning of this introduction, due to the growth in the distance education courses and the evening classes offered, workers can hold a full-time job and study at the same time. Among the people attending educational courses after the regular age of attending school at the program Youth and Adult Education (EJA), more than half declared to be also working (INEP, 2013b). In Chapter 3, I explore that by analysing the effects of the minimum wage policy on educational investments on-the-job in Brazil.

Chapter 3 also relates to the literature that analyse the impact of the minimum wage on-the-job training. The human capital theory (Becker, 1964; Mincer, 1974) suggests that part of human capital is accumulated on the job and workers finance these investments through lower wages. Therefore, in a perfectly competitive labour market, the increase of the minimum wage reduces general training acquisition, since it will prevent workers to accept lower wages in order to pay for training costs (Rosen, 1972; Hashimoto, 1982). Additionally, in a competitive labour market, mobility is high, then firms have no incentive to invest in worker's general training since other firms might "poach" their workers. On the contrary, Acemoglu and Pischke (1999, 2003) suggest that in imperfect markets, raising the minimum wage compresses firms' wages, and encourages firms to provide training in general skills, since low-skilled workers will be receiving higher wages anyway. Thus, according to theory the effect of minimum wage may be positive or negative, and the empirical studies presented mixed results (Acemoglu and Pischke, 1999; Grossberg and Sicilian, 1999; Neumark and Wascher, 2001; Acemoglu and Pischke, 2003; Arulampalam et al., 2004). Chapter 3 adds new empirical evidence for the consequences of changes in the minimum wage on investments in general training, specifically on education.

Despite the good economic performance in the 2000s, Brazil also experienced an increase on its job turnover rate. The Brazilian labour market presents one of the highest rates of job turnover in the world, especially among young people entering the labour market (Ribeiro, 2010; Corseuil et al., 2014; Gonzaga and Cayres Pinto, 2014; Rocha et al., 2018). As pointed out by the human capital theory, labour productivity is associated with workers' human capital levels (education attainment and/or training accumulated on-the-job). Thus, high rates of turnover may discourage investments in human capital and then may prevent the growth of productivity. Furthermore, the investigation of job and worker flows helps us to understand the efficiency of human capital reallocation in the economy.[1]

Directly associated with workers and job flows, is the internal migration of workers across states, which is another feature of the Brazilian labour market. Due to its regional inequality, large territory, and heterogeneous states, Brazil has been long studied for its internal migration (Sahota, 1968; Yap, 1976; Fiess and Verner, 2003; Queiroz and Golgher, 2008; Ramalho and Queiroz, 2011). These studies show that worker mobility across regions is determined not only by regional earnings differences, but also by regional socioeconomics characteristics, such as infrastructure, and environmental features, e.g. climate amenities.

Historically, the urbanization process in Brazil started after 1950 and was driven by migration flows from rural areas to new industrial cities. Deterioration of agricultural prices contributed to the rural-urban migration until 1980, shaping a new profile for society. There was a huge population redistribution, particularly to metropolises and coastal cities, and the new demographic concentration increased the existing regional inequalities. Interstate migrants moved mainly from the North towards the Southern industrialised states and the capital Brasília. The share of migrants increased from 9% in 1940 to around 20% in 1980 (Graham, 1970; Martine, 1990; Schmertmann, 1992). In the 90s, market oriented reforms to promote local economic development, and trade reforms to integrate the country to the global economy accentuated the migration flow, and the interstate migration rate reached 40% in 1999, reinforcing income concentration in the richest regions (Fiess and Verner, 2003; Aguayo-Tellez et al., 2010).

[1]The literature on job and workers flows started with Blanchard et al. (1990) and Davis and Haltiwanger (1992).

The internal migration rate declined over the 2000s, reaching 14.5% in 2010, according to the Brazilian Census (IBGE, 2010). This decrease was followed by changes in the migration patterns, that became more complex to understand. For instance, the North-South migration slowed down. Even though the Northeast region still presents negative net migration, some states became a net recipient of migrants, e.g. the state of Ceará. On the other hand, states like São Paulo and Rio de Janeiro, which traditionally received a large number of immigrants, only displayed a modest positive net migration.

Hence, these recent transformations in the internal migration may contribute to reducing inequality across regions, since they are directly related to the dynamics of job and worker flows in the Brazilian labour market (Mendes et al., 2017). Additionally, migrants often face job changes decisions when moving to a different location, and it is unclear the impact of the internal migration on shifting to a better job. In Chapter 4, I analyse the effects of internal migration of young people entering the labour market on job transitions in Brazil. The focus of the literature lies on the effects of internal migration on wages. However, impacts on other job features received scant attention, particularly for developing countries.

However, migration decisions are associated with individual's preferences, and therefore are endogenous. To overcome this obstacle, I use shocks on homicide rates at the state of origin to get an exogenous variation on the individual's decision to move. In Brazil, violence is an important issue. According to a recent study from the Special Secretary for Strategic Affairs, the cost of violence varied from 4.4% to 5.6% of the Brazilian GDP in the period 1996-2015 (SAE, 2018). Additionally, the Brazilian Census reported that half of the population living in urban areas declared to feel unsafe about where they live (IBGE, 2010). Thus, the increase in violence decreases people's quality of life, and this sense of insecurity may trigger one's migration. Although there are several studies about internal migration in Brazil, Chapter 4 presents a novel approach by analysing the impacts of internal migration driven by violence shocks.

All chapters contribute to policy makers in reducing income inequality by focusing on: i) educational increase of the workforce, ii) minimum wage and iii) labour market integration. As previous mentioned, the rapid increase in the educational level of the workforce may have contributed to the decrease in labour earnings inequality by changing the distribution of labour earnings. Directly related to that is the minimum wage, which aims to boost earnings at the bottom of the wage distribution and also may change the distribution of labour earnings. Finally, labour market integration of migrants may affect inequality depending on labour market opportunities and wages. Although inequality has been decreasing in Brazil, it is still above the world's average level. As Barros et al. (2010) points out, most of the decrease in inequality over the last decade is due to social policies, which are still contradictory. Thus, inequality reductions remain limited as a consequence of inconsistent public policies and market failures, leaving room for improving policy design and to further decrease inequality and poverty.

In this thesis, I study educational mismatch, minimum wage policies and internal migration in Brazil. The following chapters can be summarised as follows. In Chapter 2, I examine the educational mismatch in the labour market and its impact on wages. In addition, I estimate the effects on aggregate wages of reassigning workers to a well-matched job or correcting their attained education in the current job. In Chapter 3, I analyse potential encouragement or discouragement of a minimum wage policy on education acquisition on-the-job, through changes in the wage distribution and skill premia. To identify the effects of the national minimum wage policy, I use variation of price levels across time and states. Lastly, in Chapter 4 I investigate the effects of internal migration on job transitions by instrumenting migration with violent crime shocks at the origin state. In particular, I focus on young people entering the labour market.

Chapter 2

Overeducation in the Labour Market: Evidence from Brazil

Abstract. The educational mismatch is a concern because it may signal large inefficiencies in the labour market. This paper analyses the impact of educational mismatch on wages in Brazil using data from a panel of all formal workers in the Brazilian labour market (Annual Social Information Report, RAIS). I find that half of the Brazilian formal labour market is mismatched, with similar proportions of over and undereducated individuals. After individual heterogeneity is controlled for, I find that overeducated (undereducated) workers earn significantly lower (higher) than their co-workers who hold a well-matched job; and the penalty for overeducation is equal the premium for undereducation. However, the overeducation penalty is about half of premium for going to university. I also find that given the symmetry of over/undereducation when I correct the educational mismatch, by changing workers' education or occupations, the effects on the aggregate wages are very small.[1]

2.1 Introduction

The occurrence of educational mismatch (when the individual has more/less education than required by the job, or simply over/undereducation) may suggest evidence for inefficiencies in the labour market, and it is a social problem if reassigning workers to jobs could increase the aggregate output. This phenomenon is relevant, especially because many employees are mismatched; taking into account the existing studies from Asia, Australia, Europe, Latin America, United States and Canada, on average 26%

[1]I would like to thank Thomas Gall, Michael Vlassopoulos, Corrado Giulietti, Brendon McConnell, Fernanda Estevan, Hector Calvo Pardo, as well as conference participants at the 2018 Royal Economic Society Annual Conference, 2017 Economics PhD Annual Workshop and University of Southampton internal seminars for valuable comments and suggestions on earlier drafts of this paper. I gratefully acknowledge financial support from the *ESRC*.

are undereducated and 30% are overeducated, as reported by Leuven and Oosterbeek (2011).[2] The concern with educational mismatch is associated with the search for more efficiency in allocations in the labour market. If education is a good measure of one's ability, then the mismatch is a sign for inefficiencies (e.g., Rumberger, 1987). However, if education is a poor measure, then educational mismatch is part of an efficient labour market (Hersch, 1991; Sicherman, 1991). Hence, the educational mismatch rise questions on the interpretation of the theories that relate the labour market and the educational system. The improvement in the educational level of the workforce has been followed by higher than average growth rates in jobs for more educated workers and, in some jobs, the necessary skills to perform them correctly have been upgraded, encouraging workers to acquire more education (Groot and Van Den Brink, 2000). If the growth in supply is bigger than the demand for highly qualified workers, then possibly there will be overeducation of the workforce. As a result, overeducation can lead to misallocation of skills among jobs, since high qualified workers may be underemployed, i.e., they may end up in jobs that required less than their actual abilities. Thus, the boost in the educational level of the population allows evaluating if the labour market is absorbing the qualifying offer, through the effect on wages, helping policy makers to improve efficiency and efficacy in use of public resources.

Educational mismatch and its effects on wage returns have been analysed in the literature, especially in developed countries (Groot and Van Den Brink, 2000; Hartog, 2000; Sloane, 2003; McGuinness, 2006; Leuven and Oosterbeek, 2011). Freeman (1976) was the first to study the overeducation phenomenon motivated by diminishing returns to education in the North American labour market.[3] Later, Duncan and Hoffman (1981) introduced a model in which wages are a function of education and divided the education into three parts - overeducation, undereducation and required education, all measured by years (ORU specification) - to analyse the effect of each one on wages through an extended version of the Mincerian wage equation. Another approach used by empirical studies on educational mismatch is the Verdugo and Verdugo (1989) specification. Unlike Duncan and Hoffman's approach, this specification takes into account the worker's attained education, a dummy variable for overeducation and a dummy variable for undereducation.[4] Therefore, the discussion on over/undereducation is led by stylized facts: (i) overeducated workers have

[2]The authors report the overall unweighted means for previous available studies. In particular, studies for the United States and Canada present on average larger shares of overeducation, while studies for Europe report on average larger shares of undereducation.

[3]Freeman (1976) argued that this occurred due to the increased labour supply of more educated workers relative to labour demand. Smith and Welch (1978) added two years to the beginning and end of the period used by Freeman. The authors found less pronounced results and show that the results are more consistent with an overcrowding labour market to new entrants because of larger size cohorts than with overeducation situation.

[4]Studies using Duncan and Hoffman's approach found that the returns on overeducation are positive, while returns on undereducation are negative (Sicherman, 1991; Hartog, 2000; Bauer, 2002; Rubb, 2003; Dolton and Silles, 2008; Tsai, 2010). Conversely, studies using Verdugo and Verdugo's approach found the opposite, overeducation has a negative effect on earnings and undereducation has a positive effect (Rubb, 2003; Di Pietro and Urwin, 2006; Green and McIntosh, 2007; Sánchez-Sánchez and McGuinness, 2015).

lower wages than those who have the same educational background, but instead hold jobs that require the exactly level of education they have. However, those overeducated workers earn more than their colleagues who are not overeducated; (ii) undereducated workers have higher wages than those who have the same educational level working in jobs that require the exact level they have. In spite of that, these undereducated workers receive lower salaries than their colleagues who have the required or higher level of education.

Although the educational mismatch has been extensively studied for developed countries, the literature is still limited for developing countries. Mehta et al. (2011) suggest that the main reason for the lack of studies in developing countries is due to the absence of proper data. Few exceptions are the studies by Quinn and Rubb (2006), which presents the results of educational mismatch on wages in Mexico using cross-sectional data, and by Mehta et al. (2011), which the authors propose a new method to identify overeducated workers and apply it to data from India, Mexico, the Philippines and Thailand.[5] Overeducation is, by definition, more prevalent in labour markets where the average educational level of the workers is high, which is a characteristic of developed countries. Then the overeducation phenomenon in developing countries, particularly in Brazil might represent a serious problem, considering the Brazilian labour market deficiency of skilled labour and good education (Santos, 2002). Since the 1990s, Brazil has been facing a major expansion in higher education, mostly because of the growth of private education. According to the Higher Education Census, Brazil has about 2400 higher education institutes, which more than 2000 are private, and the number of students starting undergraduate courses increased by 76.4% in the period 2003-2013 (INEP, 2013a).[6] Therefore, if the increase in the supply of college graduated workers is not absorbed by the labour market, the educational raise in the labour supply may increase the share of overeducation in Brazil. For instance, Manacorda et al. (2010) find that the demand for skilled workers has been stable in Brazil during the 1990s.[7] This result is corroborated by Pauli et al. (2012) using a different dataset.[8] Notwithstanding, this expansion may affect the quality of education. For instance, public universities are more selective (the entrance exam to university is very competitive) and because of

[5] The standard method determines which workers have more education than required to perform their jobs and then estimates the return of the educational surplus. Conversely, Mehta et al. (2011), determine first which jobs do not pay a reasonable return to education, and then detect how many workers those jobs employ.

[6] Moreover, according to the Brazilian Institute of Geography and Statistics (IBGE, 2014), the proportion of employed individuals with secondary and higher educational level increased from 2004 to 2013, the average schooling of the 25 years population or more increased from 6.4 to 7.7 years of study. Likewise, in the same period, the proportion of people in the age group from 25 to 34 years who have higher education almost doubled from 8.1% to 15.2%.

[7] Manacorda et al. (2010) assess the role of changes in demand and supply of skilled workers and the relation with returns to education in Latin America. The authors find that the demand for college educated workers has been increasing in Mexico and Colombia, but not in Argentina and Brazil.

[8] Different from Manacorda et al. (2010) that use the Brazilian National Household Sample Survey (PNAD), Pauli et al. (2012) use data for the Brazilian Annual Social Information from 1990 to 2007. The latter authors find that the demand side for qualified workers did not follow the expansion of the supply side.

that, they have been able to preserve educational quality. However, private institutions are more concerned about profit and they are not always concerned with the quality of education provided. Usually, those private institutions present low admission and approval requirements in their courses.[9] If the quality in higher education does not come with quality improvement then those new higher educational institutions bring little to society, since they do not help increasing individual's productivity. By attending a low-quality university, one may be overeducated, but may not be overqualified for the job. The extra education may be necessary to fulfil the right amount of human capital required to perform the job (Robst, 1995; Ordine and Rose, 2009). In this framework, the educational mismatch is not an inefficiency: over/undereducation reflects one's true ability, while formal education certificate may not.

In order to contribute to the literature on educational mismatch in the labour market, this paper analyses wage differentials for over/undereducated individuals in the Brazilian labour market using a panel data from 2006 to 2013 from the Annual Social Information Report (RAIS), which is a census of the formal Brazilian labour market. The purpose of this study is to identify the effects of over/undereducation on workers' hourly wages and offer a first pass on the general equilibrium effects on wages of educational mismatch. To do this, the paper proceeds in two parts. In the first part, I rely on an earning function specification controlling for individual's fixed effects to address the potential omitted-variable problem that may stem from individual unobserved heterogeneity. As far as I know, the use of the Annual Social Information Report (RAIS) as a panel dataset for the educational mismatch literature is a pioneering approach to Brazilian data. In the second part, I perform a postestimation analysis to assess the impact on aggregate wages if one changes individuals' attained education and jobs, reassigning them to well-matched positions.

Few studies in the literature on educational mismatch use panel dataset (Bauer, 2002; Frenette, 2004; Tsai, 2010; Carroll and Tani, 2013; Mavromaras et al., 2013; Reis, 2017), and they have found that over/undereducation effects decrease after controlling for individual's heterogeneity. In Brazil, the overeducation literature is being developed recently and it presents relevant empirical evidence (Santos, 2002; Diaz and Machado, 2008; Esteves, 2009; Reis, 2017). These studies show that educational mismatch is present in the Brazilian labour market, indicating the need to analyse it. However, the majority of these use cross-sectional data, which may present biased results because of the individuals' unobserved heterogeneity.[10] Therefore, this paper contributes to the scarce

[9]For more details on the expansion and quality of higher education in Brazil see Schwartzman (2004).

[10]As far as I investigate, the only exception is the study by Reis (2017), which uses panel data for six main Brazilian metropolitan areas. In his paper, Reis (2017) uses two observations for each individual from the Monthly Employment Survey (PME) and Continuous National Household Sample Survey (PNAD Contínua) datasets to investigate the impact of educational mismatch on earnings. He finds that an additional year of surplus schooling increases earnings, but this effect is smaller than that of an additional year of required schooling. Moreover, a year of deficit schooling is associated with a penalty on earnings, although undereducated workers still earn more than those well-matched with the same educational level.

literature on educational mismatch in developing countries, accounting for individual heterogeneity in the form of fixed effects.

I find that half of the Brazilian formal labour market is mismatched, with similar proportions of over and undereducated individuals. This high incidence of the educational mismatch is consistent with previous results for the Brazilian labour market (e.g., Reis, 2017). I find that after taking into account the individual heterogeneity, over- and undereducation have the same impact in magnitude on earnings, but in opposite directions. While overeducated workers suffer from a wage penalty, undereducated ones receive a premium, comparing them to their well-match counterparts with the same educational level. Nevertheless, the penalty for overeducation is half of the premium for going to college. Furthermore, I find that correcting the educational mismatch by reassigning workers into different jobs would increase aggregate wages while switching their education would decrease aggregate wages. However, these changes are very small, suggesting that given the symmetry of over and undereducation the mismatch is linear and then, overeducation and undereducation are cancelling each other.

The remainder of this paper is organised as follows. Section 2.2 reviews some theoretical pointers that link the educational mismatch to the labour market. Section 2.3 describes the dataset and key variables used for the empirical analysis. Section 2.4 explains the empirical methodology employed. Section 2.5 provides a discussion of the empirical results and section 2.6 concludes.

2.2 Theoretical pointers

There is no consensus in the literature about a unique theory to explain the educational mismatch. Instead, the phenomenon has been interpreted according to different labour market theories. To begin with, the human capital theory (HCT) asserts that higher productivity yields higher earnings, resulting in a positive relationship between wages and human capital (Becker, 1964). This relationship is represented by an equation known as the Mincerian wage equation, in which the attained education affects wages (Mincer, 1974). This means that in equilibrium should be no underutilisation of one's stock of human capital. Therefore, overeducation can only be explained by HCT if it is a temporary phenomenon. The HCT states that workers will always earn their marginal product because firms are ready to fully use workers skills by adjusting the production process to changes in the labour supply. In the presence of overeducation, wages rates will be below the individuals' marginal product and in the short-run, the market will not be in equilibrium. That is, it will last until workers find a more suitable job to their potential, or until the time firms take to update their production processes to use fully the available human capital. Hence, persistent overeducation is inconsistent with the HCT. However, another explanation that copes with the HCT is that additional education would help to compensate for the lack of other factors such as training, experience and abilities. (Dolton and Vignoles, 2000; McGuinness, 2006). The other way around is also true, i.e., experience and training may compensate for undereducation.

Contrary to the HCT, in the job competition model (Thurow, 1975) individuals compete for jobs based on their training costs. In this framework, the allocation of jobs is fixed, and individuals compete for job opportunities. On the other hand, employers consider education as an indicator of the cost of investing in workers' training to perform a specific job. Then more educated individuals are abler and need less training by the firm, and they may accept the job for which they are overeducated while competing for a job. Wages, in this environment, are determined only by the productivity characteristics of the job. Therefore, given the presence of overeducation, individuals have an incentive to invest in education to preserve their position, even though the job not necessarily requires such high levels of education.

Similarly to the job competition theory, in the signalling/screening model (Spence, 1973) education acts as a signal in labour market with imperfect information, although on the latter the individual's investment on education is limited based on the earnings and costs of education. Hence, if the cost of education is lower for higher ability workers then an exogenous decrease in the educational cost will incentive lower ability individuals to get more education. Therefore, it will increase the average educational level of the workforce forcing the firms to raise the qualification requirements to perform the same job in order to guarantee they will hire the workers with the specific ability. In this framework, the qualification mismatch arises, but not ability mismatch.

Another explanation can be found in the hedonic/assignment models (Sattinger, 1993). According to these models, changes in earnings are not explained by human capital or job features individually. Instead, wage's variations are determined within a hedonic price equation including demand and supply parameters. This means that assignment models focus on the fact that investments in human capital may depend on the match between the worker and the job – i.e., even if education increases worker's productivity, the actual productivity also depends on the nature of the job. These models argue that productivity reaches a maximum when workers are assigned according to their skills, meaning that most skilled workers are allocated in more complex jobs while less skilled workers are allocated in simpler jobs. This creates the possibility of balance between workers skills and abilities required by a specific job that show some kind of deviation. These deviations would be over and undereducation, explained by different proportions of complex/simple jobs and skilled/unskilled individuals. In this framework, the returns of education are limited if the jobs do not fully use the individual's schooling. Hence, overeducated workers earn less than their colleagues with the same educational level because of a job limit on productivity.

Alternatively, some theories propose that overeducation is transitory and occurs because of imperfect information or strategic behaviour, which means that individual preferences should be taken into account (McGuinness and Pouliakas, 2017). For instance, the career mobility theory (Sicherman and Galor, 1990; Sicherman, 1991) points out that workers with a given innate ability may prefer at the beginning of their career to be in a job that does not match their educational level if it there is a high probability of being promoted. Thus, they can acquire skills through on-the-job training and/or work experience in order to progress on the job quickly. The matching theory of job search (Jovanovic, 1979), on its turn, suggests that workers may accept jobs for which they are overeducated as a consequence of the costs to search for an appropriate job. These job search costs are more prevailing in developing countries. Over/undereducation may be indications of frictions, which is a more prevalent problem in developing countries. In this environment, Albrecht and Vroman (2002) developed a model in which workers are not allowed to search on-the-job. Therefore, two equilibria may stem depending on the productivity gap between high-skilled and low-skilled jobs, and on the proportion of high-skilled workers in the labour force. In one scenario, high educated workers match with high-skilled and low-skilled jobs; alternatively, those high educated workers decline low-skilled jobs. Conversely, Gautier (2002) proposes a model of on-the-job search and matching, which allows highly educated workers to hold a job for which they are over-skilled. Then, the firm's lack of information on the workers' skills drives, in the equilibrium, the mismatch in the labour market. On the other hand, McGuinness and Sloane (2011), suggest that the mismatch in the labour market may be due to the exchange of accepting a job for which one is overeducation for other job features (e.g., job security), for which they have stronger preferences.

The consumption value of education that maximises lifetime earnings determines the acquisition of schooling for each individual. Because people may have different taste for education, individuals may choose different levels of schooling depending on their preferences. Thus, some people may like education and then overinvest in it, while undereducation may simply reflect the presence of individuals who intensely dislike education. Leuven and Oosterbeek (2011) point out this preferences theory could be used to interpret over/undereducation, but it has been neglected in the literature. An exception is the study of Oosterbeek and Van Ophem (2000) in which they find that for all individuals their sample, the schooling parameter is positive. This means that even workers with a low taste for education benefit from a positive utility from it.

Although several theories explain the educational mismatch in the labour market, one could use some of those to explain the phenomenon in Brazil. For instance, according to the human capital theory, individuals could be acquiring more education to compensate for other factors. One of these factors could be low educational quality (Robst, 1995; Mehta et al., 2011), which is a plausible explanation for Brazil. With the increase of private institutions, the quality of education declined significantly. Thus, educated workers may end up on unskilled jobs due to their low-quality education. If there is an increase on the supply of workers with low-quality education in the labour force, then there would be an increase in the demand for high-return jobs, leading on its turn to an increases on the education returns. If that is the case, then overeducation arises not from the excess of educated workers, but from the lack of high-quality education. Moreover, low-quality education may be well linked to the cost of the mismatch. Therefore, public policies should aim to increase the quality of education offered, so workers will be more likely to find a job that fits their attained education.[11] Another reason could be that overeducation is related to asymmetric information. As pointed by Quinn and Rubb (2006), in developing countries like Mexico, information is not likely to be quickly available at employment time, and that could also be the case for Brazil.

2.3 Data and descriptive statistics

2.3.1 Data

The analysis is based on two datasets from the Brazilian Ministry of Labour and Employment (MTE), the Annual Social Information Report (RAIS) and the Brazilian Classification of Occupation (CBO). RAIS is a statistical database on the formal labour market in Brazil with socioeconomic information. It is designed to monitor labour activities in Brazil and to provide labour market information to governmental entities. Established in 1975, it is held annually and has national coverage. All legal organizations are obliged to report to it, which is likely to guarantee high-quality data. Moreover, it is possible to determine from RAIS the number of formal employees in Brazil, how many workers were laid off, or which sectors hired more. The data is stratified by municipality, economic class and occupation, age, education level, tenure and range of average income. RAIS is a reliable data source for the formal labour market in Brazil, especially because of its census nature, the extent of its information, and its geographic and temporal dimension (the information is collected every year). Moreover, RAIS allows longitudinal analysis, which enables the research of various problems related to the dynamics of the labour market and income distribution with a significant impact on public policies. The empirical analysis in this paper relies on a 1% representative random sample panel – stratified by federal states – of individuals selected in 2006 and followed until 2013. This paper follows 64,244 workers from 25 to 58 years old (in 2006) over eight years. The cut off age was picked in order to exclude workers who had a part-time job while attending university at the regular age and to make sure that retirees do not affect the education mechanisms.

[11] On subsection 2.5.2, I check what happen to the economy if all workers are on jobs for which their education is the same as the required education to perform the job.

The second dataset used is the Brazilian Classification of Occupation (CBO) from 2010. CBO is an enumerative and descriptive classification.[12] The enumerative classification codes jobs for statistical purposes, while the descriptive classification details the activities performed at the job, the requirements and experience needed to perform the job and the work conditions. It is worth noting that the CBO classification is based on occupations and it is time and state invariant. While the dataset from RAIS only uses the CBO enumerative function to classify workers occupations, this paper also uses the CBO descriptive function to determine the presence of under and overeducation in the dataset. The occupations are described by occupational families and each one of these families consists of a set of similar occupations.[13] In total, there are 607 occupational families. The MTE elaborates the CBO using the occupational analysis method. This methodology uses a committee of professionals working in each of the occupational families to describe what the worker has to do in the job. It is the same method used in the United States and Canada adapted, however, to the Brazilian labour market institutions (MTE, 2010).

2.3.2 Measuring overeducation

Three main approaches have been used in the literature to measure the required education for some occupation: job analysis (JA), worker self-assessment (WA) and realized matches (RM) (Hartog, 2000; Verhaest and Omey, 2006; Leuven and Oosterbeek, 2011). Job analysis means that professional job analysts establish the required level of education for each occupation. Worker self-assessment is to elicit the required education for the job from workers themselves. Realized matches mean using a measure of the effective observed education of workers (usually it is used the mean or mode of the distribution).

Hartog (2000) and Verhaest and Omey (2006) argue that the job analysis method is preferable.[14] First of all, job analysis measures depend on the technology of the job. Second, job analysis presents clear definitions and objectivity of overeducation. Moreover, measurement instructions are very precise, thus ensuring that there are no reasons to expect systematic bias. However, the use of the job analysis method will depend on the availability of data. One possible limitation of this method discussed by Verhaest and Omey (2006) is related to the low frequency of updates in the classifications of the

[12] Administrative records, population censuses and other household surveys use the enumerative classification for statistics purposes. Activities that need information about the content of the job uses the descriptive classification, such as educational activities in firms and unions, immigration services. The latter provides a detailed description of the activities performed at the job, the training requirements and professional experience, and working conditions.

[13] For example, the occupational family Economist includes economic analysts, agroindustrial economist, financial economist, industrial economist, public sector economist, environmental economist, regional and urban economist.

[14] The authors compared five ways of measure overeducation: job analysis, worker-assessment of the required level to do the job, worker-assessment of the required level to get the job, the mean educational level of realized matches and the modal educational level of realized matches. They concluded that for all variables analysed (wages, job satisfaction, mobility and training) job analysis method was the best among all.

occupations because of the high costs involved. Hence, those classifications can become out-of-date, since the labour market dynamics can cause changes in the educational requirement for the occupations and then, the long-term monitoring of overeducation might be flawed.

Following the argument above, I use the job analysis approach to determine the required level of education for each occupational family. Education in the sample is divided into the following categories: illiterate, incomplete basic education I, complete basic education I, incomplete basic education II, complete basic education II, incomplete secondary school, complete secondary school, incomplete higher education, complete higher education, master and doctorate. In this analysis, I excluded those occupational families whose required education is less than basic education II because it was compulsory in Brazil during the period analysed.[15] Besides, I do not consider those occupational families that consist of armed forces, police and military firefighter, senior members of the government, leaders of public interest organisations and companies, because one cannot determine the required education in those due to the heterogeneity of the occupations.

Finally, after determining the required level of education for each occupational family, I compute the educational mismatch for each worker comparing individual's attained level of education and the required level of education to perform her/his job. Thus, individuals who have more/less education than required by their jobs are classified as over/undereducated, while those who have the same educational level as required by the job are classified as well-matched.

2.3.3 Descriptive analysis

Based on the aforementioned information, I classify each worker as overeducated, well-matched or undereducated. Table A.1 (in appendix) shows that half of the Brazilian labour market is mismatched, about 24.67% of the sample is overeducated, while 25.67% is undereducated.[16] The results corroborate previous results found by Reis (2017) using another dataset, that one quarter of the Brazilian workforce is overeducated and one quarter is undereducated. Table A.2 shows the variation in the educational difference over the years. The levels are computed based on the difference between individuals' attained education and the required educational level to perform the job. Negative levels represent undereducated individuals and positive levels represent overeducated individuals. For instance, consider a work whose attained education is complete higher education, but who holds a job which requires complete secondary school. In that case, this worker presents two positive levels of difference. A difference of zero means that the individual is well-matched to her/his job. Additionally, Table A.2 shows that during

[15] Basic education I goes from 1st to 5th grade and basic education II from 6th to 9th grade.

[16] Table A.1 contains the descriptive statistics of the sample for the main variables used in this study: hourly wage (deflated by the Brazilian Consumer Price Index – IPCA using 2013 as the base year), hours worked, age, tenure, gender, education, working class, firm size and industry.

the period analysed the total of undereducated individuals decreased, while the total of overeducated and well-matched individuals increased. Therefore, workers are getting more educated and/or changing to jobs that require less education than the previous one.

Table 2.1: Conditional profile

Variables		**Undereducated**	**Exactly level**	**Overeducated**	**Total sample**
Gender	Male	0.67	0.53	0.62	0.58
		(0.004)	(0.003)	(0.003)	(0.002)
	Female	0.33	0.47	0.38	0.42
		(0.004)	(0.003)	(0.003)	(0.002)
Sector	Industry	0.39	0.26	0.36	0.31
		(0.004)	(0.002)	(0.003)	(0.002)
	Services	0.61	0.74	0.64	0.69
		(0.004)	(0.002)	(0.003)	(0.002)
Class of worker	Blue collar	0.36	0.14	0.30	0.23
		(0.004)	(0.002)	(0.003)	(0.002)
	White collar	0.64	0.86	0.70	0.77
		(0.004)	(0.002)	(0.003)	(0.002)
Region	North	0.05	0.06	0.04	0.05
		(0.002)	(0.001)	(0.001)	(0.001)
	Northeast	0.17	0.15	0.14	0.15
		(0.003)	(0.002)	(0.002)	(0.001)
	Midwest	0.07	0.07	0.08	0.07
		(0.002)	(0.001)	(0.002)	(0.001)
	Southeast	0.52	0.54	0.54	0.54
		(0.004)	(0.003)	(0.004)	(0.002)
	South	0.20	0.18	0.20	0.19
		(0.003)	(0.002)	(0.003)	(0.002)
Changed occupation	Yes	0.07	0.10	0.10	0.10
		(0.001)	(0.001)	(0.001)	(0.001)
	No	0.93	0.90	0.90	0.90
		(0.001)	(0.001)	(0.001)	(0.001)
Changed education	Yes	0.02	0.05	0.09	0.06
		(0.001)	(0.000)	(0.001)	(0.000)
	No	0.98	0.95	0.91	0.94
		(0.001)	(0.000)	(0.001)	(0.000)
Age group	[25-29]	0.05	0.07	0.08	0.07
		(0.001)	(0.001)	(0.001)	(0.001)
	[30-34]	0.13	0.18	0.20	0.17
		(0.002)	(0.001)	(0.002)	(0.001)
	[35-44]	0.36	0.37	0.37	0.37
		(0.003)	(0.002)	(0.003)	(0.001)
	[45-54]	0.33	0.29	0.27	0.29
		(0.003)	(0.002)	(0.003)	(0.001)
	[55-65]	0.12	0.09	0.07	0.09
		(0.002)	(0.001)	(0.002)	(0.001)

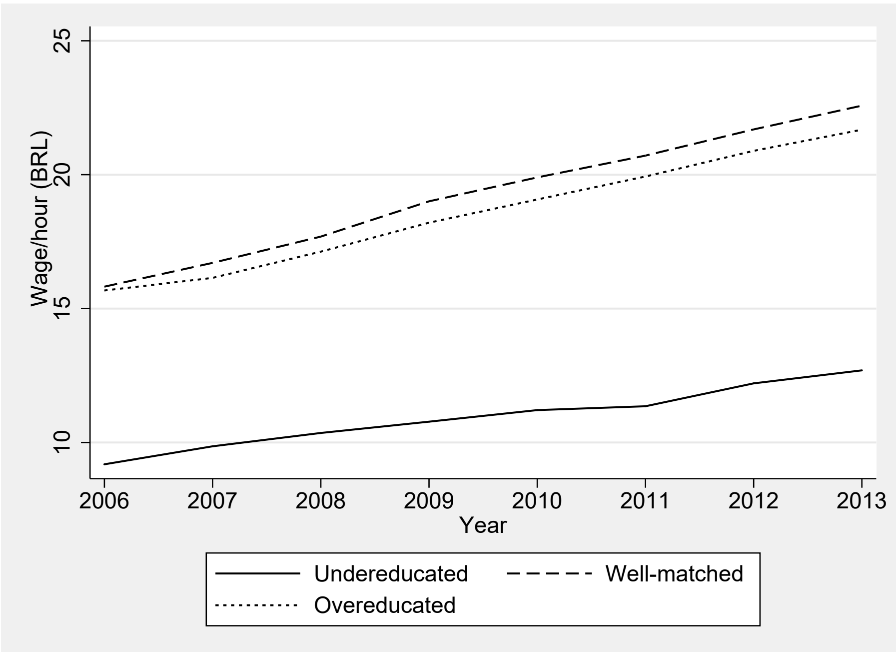
Source: RAIS, 2006-2013.
Notes: Standard deviations in parenthesis. Changed occupation and changed education are dummy variables that takes value 1 if the worker changed her/his occupation/education compared to the previous year, 0 otherwise.

In order to better understand who are those over/undereducated workers, Table 2.1

presents the conditional probability of various characteristics for over/undereducated workers and for well-matches ones. For example, the probability of being female is bigger for those who are overeducated than for those who are undereducated. An individual is also more likely to work on the services sector and be white collar work if they are overeducated rather than undereducated. Among the regions, conditional on the educational mismatch is more likely to be in the Southeast and South regions respectively. Those are the richest regions in the country. The probability of living in the Southeast region is slightly bigger for overeducated than for undereducated, however, the probability of living in the South is the same for over and undereducated individuals. It is worth mentioning that overeducated individuals are more likely to change jobs and education than undereducated workers. Finally, considering the age groups, the probability of being younger is greater for overeducated individuals than for undereducated ones.

Figure 2.1 depicts the average hourly wages by educational mismatch over the period studied. As one can see, on average, workers who hold a well-matched job earned more than everyone else. Well-matched workers and overeducated ones started almost at the same hourly wage, and although the trend is similar – hourly wages rose for all workers – the gap among them increased. One possible explanation for that is the individual heterogeneity. It could be the case that well-matched workers are more skilled than the mismatched ones, and because of that, they earn more on average. Therefore, by acquiring more education workers could be compensating for the lack of abilities.

Figure 2.1: Evolution of hourly wages by educational mismatch



Source: RAIS, 2006-2013.
Note: Hourly wages were deflated by 2013 prices.

Table 2.2 shows that individual educational level generally increased over the period analysed. Comparing the first and the last year of the panel, the total number of individuals in lower levels decreased, while the number of individuals who completed secondary school, higher education, a master or a doctorate increased. For instance, the total number of individuals who completed higher education increased by 9.40 percentage points from 2006 to 2013. The increase in higher levels of education is consistent with the increase of private universities that Brazil faced in the past few years. It may, of course, also be the case that individuals are becoming more educated because they are taking evening courses at universities while working.

Considering the big increase on the percentage of workers who completed higher education (from 22.65% in 2006 to 32.05% in 2013 as reported in Table 2.2), I then present in Table 2.3 the profile of these highly-educated individuals in order to better understand this type of worker.

Table 2.2: Educational level: workforce shares in (%)

Educational level	**Years**		**Difference**
	2006	**2013**	**(p.p.)**
Illiterate	0.21	0.10	**-0.11**
Incomplete Fundamental Education I	2.72	1.59	**-1.13**
Complete Fundamental Education I	4.88	3.16	**-1.72**
Incomplete Fundamental Education II	7.99	5.48	**-2.51**
Complete Fundamental Education II	13.29	9.31	**-3.98**
Incomplete Secondary School	5.71	3.92	**-1.79**
Complete Secondary School	38.71	40.51	**1.80**
Incomplete Higher Education	3.68	2.97	**-0.71**
Complete Higher Education	22.65	32.05	**9.40**
Master	0.14	0.73	**0.59**
Doctorate	0.04	0.19	**0.15**

Source: RAIS, 2006-2013.

According to Table 2.3, among those workers who completed higher education, the majority of them are well-matched to the job. Moreover, the share of overeducated workers in this specific group (32.46%) is bigger than the share of overeducated individuals for the whole panel (24.67% as in Table A.2). On the other hand, the share of undereducated is very small, only 1.20%. Also, most of those higher educated workers are middle-aged employees, thus they may be getting training on the job. Additionally, the majority of the higher educated workers are females, white-collar workers, employed in the services sector, live on the Southeast region (the most developed region in Brazil) and did not change occupation.

Table 2.3: Profile of workers who completed higher education (%)

Variable	**Description**	**Frequency**
Educational mismatch	Undereducated	1.20
	Exactly level	66.34
	Overeducated	32.46
Gender	Male	40.15
	Female	59.85
Sector	Industry	15.50
	Services	84.49
Class of worker	Blue collar	2.16
	White collar	97.84
Region	North	8.07
	Northeast	20.16
	Midwest	14.94
	Southeast	42.21
	South	14.62
Changed occupation	Yes	11.84
	No	88.16
Age group	[25-29]	4.67
	[30-34]	15.4
	[35-44]	36.96
	[45-54]	31.76
	[55-65]	11.21

Source: RAIS, 2006-2013.
Note: This table does not include those workers who hold master or doctorate degrees.

Additionally, Figure A.1 depicts the educational mismatch over the educational levels for 2006 and 2013. Considering that the minimum education required is complete fundamental education II, then all individuals below this level are undereducated. For the educational levels above complete fundamental II, the proportion of overeducated workers increases while the proportion of undereducated decreases, comparing the first and the last year of the data. Furthermore, the expansion of workers holding a master/doctorate degree over the period (Table 2.2) suggests that the boost in education was not absorbed by the labour market in this educational level since the proportion of overeducated rose and the proportion of well-matched reduced.

Moreover, Table 2.4 presents the percentage of individuals who changed occupation over the period analysed. The occupational changes follow a fairly steady pattern. The changes are higher in 2007 and 2008 – 10.92% and 11.56% respectively – and then decreased gradually until 2013 reaching 8.46%. Considering that the total number of individuals changing jobs is decreasing and the total number of overeducated individuals is increasing (Table A.2 and Figure A.1), then I can confirm what I found in Table 2.2 that the individuals are getting more educated over the years.

Table 2.4: Changes of the occupational family (%)

Changed Occupation	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**Total**
	Year of observation							
Not changed	89.08	88.44	89.70	90.17	90.68	91.86	91.54	**90.21**
Changed	10.92	11.56	10.30	9.83	9.32	8.14	8.46	**9.79**
Total	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**

Source: RAIS, 2006-2013.

Lastly, Table 2.5 shows the relative frequency in changing education and occupation for each year. It seems that the share of workers changing both education and occupation is steady until 2009, and after that decrease. The same pattern is observed for those who change only one or another, i.e., the frequency of individuals changing shrink over time. The fact that these transitions are more frequent at the beginning of the period analysed could indicate that workers are more willing to change education and/or occupation when they are younger.

Table 2.5: Changes of occupational family and educational level

Year	**Changed Occupation**	**Changed Educational No**	**Changed Educational Yes**
2007	No	0.865	0.026
	Yes	0.063	0.046
2008	No	0.864	0.021
	Yes	0.069	0.047
2009	No	0.879	0.018
	Yes	0.058	0.046
2010	No	0.887	0.015
	Yes	0.063	0.036
2011	No	0.886	0.021
	Yes	0.050	0.043
2012	No	0.904	0.015
	Yes	0.047	0.034
2013	No	0.897	0.018
	Yes	0.047	0.038

Source: RAIS, 2006-2013.
Note: Relative frequency for each year.

2.4 Methodology

Following the literature of educational mismatch, I model the wage differential using a variation of the specification originally outlined by Verdugo and Verdugo (1989), who defined workers as overeducated if their attained level of education was at least one standard deviation over the mean level of education in the worker's occupation. One problem with the measurement adopted by Verdugo and Verdugo (1989) is that is defined by the actual attained education rather than the education required to perform the job. However, this is not a problem in this paper since overeducation is defined comparing the individual attained level of education with the required level of education to perform the job determined by the CBO (2010).

First, I estimate the pooled OLS (POLS) model:

$$log\, W_{it} = \sum_{k=1}^{10} \delta_{1(k)} Educ_{itk} + \delta_2 Over_{it} + \delta_3 Under_{it} + \beta_1 X'_{it} + \varepsilon_{it} \tag{2.1}$$

where $log\, W_{it}$ denotes the logarithm of the wage of individual i at time t, $Educ_{it}$ is a series of dummy variables measuring the individual attained schooling category, $Over_{it}$ is a dummy that takes value one if there is a surplus of individual schooling on required education by their job, $Under_{it}$ is a dummy that takes value one if there is a deficit of individual schooling on required education by their job, X'_{it} is the vector of observed characteristics, and ε_{it} is the error term.

An issue of greater concern is that of identification of the over/undereducation premia, which tends to be disregarded in the overeducation literature (Leuven and Oosterbeek, 2011). The identification is difficult[17], because even when controlling for observable characteristics, wage differentials may persist and stem from unobserved heterogeneity. As Ashenfelter et al. (1999) suggest, obtained education can be correlated to the error term. If that is the case, then required education, overeducation and undereducation will be correlated to the error term, so that the OLS estimators will be inconsistent.

One approach used to handle the endogeneity problem is to apply a fixed effects model (Bauer, 2002; Lindley and McIntosh, 2009; Korpi and Tåhlin, 2009; Tsai, 2010). In order to identify the fixed effects estimates, the same worker has to change his/her educational level, job level or both, which I verified happened to workers in the dataset at a rate of around 6% and 10% per year, respectively. Also, to produce unbiased estimates I need to assume that the unobserved characteristics do not vary over time.

[17]Identification of the econometric analysis requires sufficient variation of the education level and/or job for the same individual across time. Tables 2.2 and 2.4 show that in the sample there are workers changing educational level and/or jobs over time.

Then, to account for the individual's unobserved heterogeneity, I re-estimate equation (2.1) with a fixed effects estimator:

$$log\, W_{it} = \sum_{k=1}^{10} \delta_{1(k)} Educ_{itk} + \delta_2 Over_{it} + \delta_3 Under_{it} + \beta_1 X'_{it} + c_i + \varepsilon_{it} \tag{2.2}$$

where c_i is the term of unobserved characteristics.[18]

In addition, considering previous results in the overeducation literature highlighting the gender differences in the labour market (Groot and Van Den Brink, 2000; McGuinness and Bennett, 2007), I investigated the wage equation separately for females and males also using equation (2.2). Furthermore, earlier work suggests consistently that younger workers are more likely to be overeducated than elderly ones (Alba-Ramírez and Blázquez, 2004; Leuven and Oosterbeek, 2011), and the data also presents this feature (see Table 2.1), which motivated me to split the analysis into these two groups. Lastly, due to the increase in college graduates observed in the sample (see Table 2.2) I also estimate equation (2.2) for those individuals who have a higher education degree and below higher education degree.

Studies that used this specification found that overeducated workers pay penalties ($\delta_2 < 0$) and undereducated workers receive wage premiums ($\delta_3 > 0$)(Verdugo and Verdugo, 1989; Sicherman, 1991; Bauer, 2002; Rubb, 2003). It is worth mentioning that these studies compare overeducated workers with workers whose job requires a higher educational profile.

To provide a robustness check I will also estimate the wage equation using levels of educational mismatch instead of the over/undereducation binary indicator.

Like any study, I face some limitations. Although the fixed effects model is a good alternative to deal with the endogeneity problem, it may not address completely the issue. Changes in education and/or in occupations can be followed by other unobservable changes that may affect one's wage. If that is the case, fixed effects estimates may be biased. Another issue concerning this model is that does not control for potential selection. The identification relies on individuals changes in education and/or occupations, i.e., the variation is within individual, which may result in individual choice. I acknowledge these as possible limitations.

[18] I also took into account survey weights in this specification.

2.5 Results

Tables 2.6 and 2.7 report the results from the wage regressions using equation (2.2).[19] All specifications compare overeducated and undereducated workers to those who have the same attained educational level but hold a well-matched job. In this framework, the negative coefficient for overeducation implies that the overeducated individuals are in lower level jobs than those who are not overeducated.

Table 2.6 shows the results of four different specifications. Using equation (2.1) column (1) reports the POLS estimates and using equation (2.2) column (2) reports fixed-effects estimates. The inclusion of interactions terms aims to capture the differences for over and undereducated individuals between different groups, i.e. females and males, young and elderly workers, highly educated workers and less educated ones. Therefore, the interpretation of the baseline mismatch coefficient is for male, older than 40 years old and who hold a higher education degree. Column (3) presents fixed-effects estimates for the subsample of females, and column (4) presents fixed-effects estimates for the subsample of male.[20] For columns (3) and (4), the reference group for the mismatch coefficient is older than 40 years old and who hold a higher education degree.

Ignoring fixed effects, column (1) indicates that overeducated individuals earn significantly less and undereducated individuals earn significantly more than those who hold a well-matched job when controlling for tenure, age, gender, level of education, industry, states and year. The wage differential is about -29.2% for overeducated and 8.53% for undereducated. These results corroborate those found in previous studies, that educational mismatch explains wage differences among individuals. To account explicitly for gender differences, age differences and education differences the interactions between females and over/undereducation, age (less than 40 years old) and over/undereducation, and education (less than higher education) are also included. All interaction terms are significant. As expected female workers earn less than males, and the penalty is intensified for those female workers who are overeducated. On the other hand, the penalty for females is attenuated if they are undereducated. In addition, the returns of education are significant and positive for higher education as expected. Overall, the OLS estimates corroborate previous studies in the literature, i.e., comparing individuals with the same education, overeducated workers earn less and undereducated workers earn more than well-matched ones.

Further, when the individual fixed effects are added to the model (column 2), the wage differential coefficient drops substantially to -3.97% for overeducated and to 4.31% for undereducated, but still remains significant at the usual levels. These differences in

[19]Table A.3 shows the results without interaction terms and the results for each interaction term individually.

[20]Note that the category 'young' is fixed and consists of workers who are up to 40 years old in 2006. Similarly, 'elder' consists of workers who are older than 40 years old at the beginning of the sample.

over/undereducation coefficients from columns (1) to (2) show that individual heterogeneity plays a big part in educational mismatch context and these results are similar to previous findings for other countries (Bauer, 2002; Tsai, 2010; Carroll and Tani, 2013). The OLS model does not take into account that different workers have different unobserved characteristics (e.g. ability and motivation) as the fixed effects model does.[21] Therefore, the results in column (2) suggest that regardless of the individual heterogeneity, overeducated workers suffer a penalty while undereducated ones earn a premium. Since I control for workers heterogeneity, then the educational mismatch may arise due to frictions (Tsai, 2010), or it may be the case that overeducation also occurs due to individual choice.

To put the over/undereducation premia in context, note that despite the penalty for being overeducated, individuals who go to college (incomplete higher education) earn a premium that is more than double of the penalty for being overeducated. Also, over and undereducated females do not have an earnings disadvantage compared to well-matched females with the same educational level. This latter result is similar to those found by Bauer (2002) for Germany. Moreover, overeducated young continue to have an advantage compared to their well-matched counterpart. And for overeducated workers who have less than higher education degree the coefficient's signal changes – comparing columns (1) and (2) – meaning that they earn less than their well-matched colleagues when the unobservable characteristics are controlled for. This change in the signal suggests that those workers may have a higher ability which could contribute to their earnings, as one can see in OLS estimates. When this unobserved characteristic is controlled for in the fixed effects estimates, the high ability premium disappears and the overeducation penalty for those low-educated workers increases by 5.72%. Additionally, in column (2) some of the education dummies are not significant anymore. Those which remain significant confirm that returns to education increase as the educational level rise. The coefficients are smaller in magnitude, indicating that abilities bias have a big effect on returns to education. Finally, note that the overeducation penalty is almost the same as the undereducation premium after I control for the individual fixed effects.

Moving to columns (3) and (4) overeducated workers are paying a penalty for holding a mismatched job while undereducated workers are being rewarded for that. Both specifications present the same coefficient for overeducation, but different for undereducation. Even though the coefficients are larger for undereducated females than for undereducated males, I cannot conclude that there are significant differences between males and females, because the interactions in column (2) are not significant. According to evidence for Germany, overeducated men have a disadvantage of roughly 3%, however, there is no significant evidence for undereducated males nor for under/overeducated females (Bauer, 2002).

[21]The Hausman specification test rejects the null hypothesis that there is no correlation between individual error and explanatory variables. Therefore, the random effects model is rejected against the fixed effects model for this specification.

Table 2.6: Wage equations - full sample and by gender groups

Variables	(1) POLS	(2) FE	(3) FE Female	(4) FE Male
Overeducation	-0.292***	-0.0397***	-0.0393***	-0.0391***
	(0.00483)	(0.00717)	(0.00997)	(0.00800)
Undereducation	0.0853***	0.0431***	0.0695***	0.0271***
	(0.00499)	(0.00865)	(0.0120)	(0.00933)
Female	-0.373***	-	-	-
	(0.00336)	-	-	-
Overeducation*Female	-0.0166***	0.00190	-	-
	(0.00549)	(0.00855)	-	-
Undereducation*Female	0.0993***	-0.00312	-	-
	(0.00514)	(0.0109)	-	-
Overeducation*Young	0.0127***	0.0112**	0.00959	0.0100
	(0.00480)	(0.00489)	(0.00775)	(0.00625)
Undereducation*Young	0.0478***	0.00584	-0.0186**	0.0203***
	(0.00431)	(0.00523)	(0.00921)	(0.00602)
Overeducation*LHE	0.156***	-0.0572***	-0.0556***	-0.0481***
	(0.00627)	(0.0107)	(0.0148)	(0.0154)
Undereducation*LHE	-0.103***	0.00563	-0.00296	0.0190
	(0.0135)	(0.0209)	(0.0244)	(0.0391)
Tenure	0.00179***	0.000384***	0.000133	0.000586***
	(3.52e-05)	(4.83e-05)	(8.71e-05)	(5.72e-05)
Tenure^2	2.31e-07**	5.47e-08	8.25e-07***	-7.54e-07***
	(1.03e-07)	(1.69e-07)	(2.63e-07)	(2.20e-07)
Less than Fundamental Education II	-0.624***	-0.0114	-0.00480	-0.0247*
	(0.00480)	(0.0108)	(0.0178)	(0.0134)
Complete Fundamental Education II	-0.421***	0.00889	0.0272*	-0.00598
	(0.00359)	(0.00821)	(0.0143)	(0.0100)
Incomplete Secondary School	-0.238***	-0.0112	-0.0266	-0.00953
	(0.00456)	(0.00993)	(0.0195)	(0.0114)
Complete Secondary School	omitted	omitted	omitted	omitted
Incomplete Higher Education	0.388***	0.0963***	0.0940***	0.0864***
	(0.00849)	(0.0145)	(0.0205)	(0.0205)
Complete Higher Education	0.861***	0.213***	0.222***	0.190***
	(0.00380)	(0.01000)	(0.0129)	(0.0157)
Master/Doctorate	1.449***	0.376***	0.393***	0.329***
	(0.0165)	(0.0310)	(0.0348)	(0.0525)
Other controls	Yes	Yes	Yes	Yes
State dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Constant	1.449***	-0.871***	-0.992***	-0.774***
	(0.0234)	(0.0408)	(0.0713)	(0.0496)
Observations	507,612	507,612	212,032	295,580
Adjusted R-squared	0.551	0.928	0.921	0.934

Standard errors in parentheses clustered by individual; *** p<0.01, ** p<0.05, * p<0.1.

Notes: Other controls include age, firm size, sector and class of worker.

*Overeducation*Female* (*Undereducation*Female*) is a dummy that takes value one if the individual is overeducated (undereducated) and female; *Overeducation*Young* (*Undereducation*Young*) is a dummy that takes value one if the individual is overeducated (undereducated) and up to 40 years old; *Overeducation*LHE* (*Undereducation*LHE*) is a dummy that takes value one if the individual is overeducated (undereducated) and has less than higher education.

Tsai (2010) using data for the United States, reports significant results for overeducated females and undereducated males (-1.4% and 1%, respectively).

It is worth mentioning that as expected, tenure has a positive and significant coefficient in all specifications and the quadratic term is only negative and significant for men. According to the human capital theory, one can separate the human capital into general human capital and firm specific human capital. Accumulating general human capital allows workers to be more productive in all jobs, but accumulating firm specific human capital means that the worker will be more productive only at a specific firm. The positive relationship between tenure and earnings may be interpret as the result of the unobserved job match heterogeneity, that is, higher wages may reflect the high-quality employer-employee relationship, which may lead to longer tenure.[22]

Table 2.7 presents the results for age and education subsamples. Column (1) reports estimates for individuals up to 40 years old (in 2006), column (2) presents estimates for workers older than 40 years old (in 2006), column (3) shows estimates for individuals who have less than higher education, and finally, column (4) presents estimates for higher educated workers. For columns (1) and (2), the reference group for the mismatch coefficient is male who hold a higher education degree, and for columns (3) and (4) is male older than 40 years old.

First, looking at estimates in columns (1) and (2) results show the coefficients are conforming to previous expectations, i.e., overeducated individuals earn less than those who hold a well-matched job, while undereducated earn more than their well-matched counterparts. Overeducated younger workers earn -4.56% less and undereducated younger workers earn 3.67% more than well-matched workers. Even though the coefficient is negative, the coefficient is not different from zero for older workers. These results corroborate the career mobility theory that the wage penalty is larger for younger workers. However, undereducation remains significant and positive, so that undereducated older workers earn 4.38% more than their well-matched counterparts.

In columns (3) and (4), for those workers who are lower educated, the cost of being overeducated is earning -3.13% compared to their well-matched counterparts, and for those who are higher educated the cost is about -4.5%. On the other hand, the premium of being undereducated is about 4.73% for less-educated workers and it is not significant for more educated ones.

The results are consistent with those found previously in the literature, the educational mismatch is associated to wages differences in the labour market (Verdugo and Verdugo, 1989; Bauer, 2002; Rubb, 2003). Also, being overeducated results in wage penalty while being undereducated results in wage premiums.

[22]Table A.4 presents the results for interacting over/undereducation with tenure. After controlling for individual's unobserved characteristics, the interactions between overeducation and undereducation with tenure are not significant at the usual levels.

Table 2.7: Wage equations by age and education groups

Variables	(1) FE Young	(2) FE Elder	(3) FE Low-educ	(4) FE High-educ
Overeducation	-0.0456***	-0.00424	-0.0313***	-0.0446***
	(0.00788)	(0.0109)	(0.00781)	(0.0165)
Undereducation	0.0367***	0.0438***	0.0473***	0.0478
	(0.0109)	(0.0123)	(0.00850)	(0.0425)
Overeducation*Female	-0.000756	0.00211	-0.00255	-0.0248
	(0.0111)	(0.0135)	(0.00992)	(0.0197)
Undereducation*Female	0.0298*	-0.0361**	0.00333	-0.0285
	(0.0154)	(0.0158)	(0.0116)	(0.0429)
Overeducation*Young	-	-	-0.00455	0.0157*
	-	-	(0.00580)	(0.00882)
Undereducation*Young	-	-	0.000442	0.0532*
	-	-	(0.00508)	(0.0300)
Overeducation*LHE	-0.0616***	-0.0521***	-	-
	(0.0135)	(0.0171)	-	-
Undereducation*LHE	-0.0139	0.0117	-	-
	(0.0288)	(0.0303)	-	-
Tenure	0.000878***	0.000369***	0.000594***	-3.35e-05
	(8.34e-05)	(6.78e-05)	(5.00e-05)	(0.000110)
Tenure^2	-4.24e-06***	4.83e-07**	-3.70e-07**	1.03e-06***
	(4.27e-07)	(1.91e-07)	(1.73e-07)	(3.38e-07)
Less than Fundamental Education II	-0.0319**	0.0202	-0.0413***	-
	(0.0153)	(0.0158)	(0.0111)	-
Complete Fundamental Education II	0.00582	0.0199*	-0.0155*	-
	(0.0116)	(0.0120)	(0.00833)	-
Incomplete Secondary School	-0.0169	-0.00718	-0.0246**	-
	(0.0135)	(0.0149)	(0.00987)	-
Complete Secondary School	omitted	omitted	omitted	-
Incomplete Higher Education	0.146***	0.0197	-	-0.115***
	(0.0178)	(0.0263)	-	(0.0139)
Complete Higher Education	0.257***	0.134***	-	omitted
	(0.0136)	(0.0154)	-	
Master/Doctorate	0.427***	0.270***	-	0.137***
	(0.0650)	(0.0351)	-	(0.0266)
Other controls	Yes	Yes	Yes	Yes
State dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Constant	-1.075***	-0.335***	-0.868***	-0.449***
	(0.0927)	(0.116)	(0.0446)	(0.0967)
Observations	236,970	270,642	337,013	170,599
Adjusted R-squared	0.912	0.940	0.912	0.893

Standard errors in parentheses clustered by individual; *** p<0.01, ** p<0.05, * p<0.1.

Notes: Other controls include age, firm size, sector and class of worker.

*Overeducation*Female* (*Undereducation*Female*) is a dummy that takes value one if the individual is overeducated (undereducated) and female; *Overeducation*Young* (*Undereducation*Young*) is a dummy that takes value one if the individual is overeducated (undereducated) and up to 40 years old; *Overeducation*LHE* (*Undereducation*LHE*) is a dummy that takes value one if the individual is overeducated (undereducated) and has less than higher education.

Overall, the results highlight the importance of individual heterogeneity in overeducation context. After controlling for individual fixed effects, wage's penalty and premium are smaller, but still significant and at about 4%. These remained differences in earnings after controlling for unobserved heterogeneity suggests that wages may be related to job characteristics or the match between jobs and workers. The results discussed in this paper rely on individuals changing education and/or jobs. Another potential source of variation that allows me to identify the results is migration. Due to the nature of the dataset used in this study, if a worker move to a different city or state, she/he may also move to a different job with a different required education. The mismatch may be a reflection of a poor match because workers could not find proper jobs to their attained education or it may be due to an individual decision.

2.5.1 Robustness check

2.5.1.1 Levels of education

In order to provide robustness, I also estimate the wage equation using levels of educational mismatch instead of the over/undereducation indicator:

$$log\, W_{it} = \sum_{k=1}^{10} \delta_{1k} Educ_{itk} + \sum_{n=1}^{13} \delta_{2n} LevelsEducMismatch_{itn} + \beta_1 X'_{it} + c_i + \varepsilon_{it} \quad (2.3)$$

where $LevelsEducMismatch_{it}$ is a set of dummies for each level of education mismatch (13 dummies). As explained before, the levels were calculated based on the difference between individuals educational level and the required educational level to perform the job. In addition, I controlled for the same variables as in Table 2.6. The results are presented in Table A.5 in Appendix. Moreover, for the sake of comparison, specification (1) does not consider any fixed effect, while in the specification (2) adds individual fixed effects.

As expected, the magnitude of the coefficients increases as the educational mismatch raises, i.e., the penalty (premium) grows if the degree of overeducation (undereducation) becomes larger. Furthermore, after controlling for fixed effects the coefficients are smaller in magnitude, showing again that the unobserved characteristics must be taken into account or the estimation will be overestimated. An interesting point is that the penalty for workers slightly overeducated (one educational level of difference) is not significantly different from zero considering the individual fixed effects. This suggests that slightly overeducated individuals may have lower ability and because of that, they accept lower jobs in the first place.

2.5.1.2 Coarser education categories

The construction of the educational mismatch using eleven categories of education may give rise to a lot of small mismatch. Therefore, to check whether that is the case, first I use coarser education categories to construct the over- and undereducation dummies. Workers' attained education are now divided into 4 categories: less than complete basic education II, complete basic education II, complete secondary school and complete higher education. Then, as before, to compute the educational mismatch, I compared individual's attained education and the level of education required by the job. I find that 57.03% of the sample is well-matched, 22.60% is undereducated and 20.37% is overeducated. As expected, compared to the previous descriptives (Table A.1), part of the mismatch disappears as one uses coarser educational levels.

I also re-estimate equation (2.2) using the new education categories. The results are presented in Table A.6 in Appendix. Looking at the fixed effects estimation (column 2), the wage differential is -4.76% for overeducated and 4.28% for undereducated, compared to their well-matched counterpart. These results are similar to those found in Table 2.6.

2.5.2 What would happen if there was no overeducation?

As McGuinness (2006) points out overeducation may be costly to the economy, for example, by spending taxes to finance individuals to get inefficient levels of education. Moreover, it may be costly for firms if overeducated individuals are less productive than their well-matched colleagues. Following the human capital theory approach, in which wages reflect productivity, and considering the specification adopted in equation (2.2), the previous results indicated that overeducated (undereducated) workers are less (more) productive than those who are well-matched to their job.[23]

In the case where mismatched workers are compared to their counterparts with the same educational level, the over/undereducation wage premium is determined by the required education to perform a job. As one has seen, the individual's wage is a function of the attained education (e) and the required education to perform the job (j):

$$
\begin{aligned}
w(e,j) \qquad \text{then if} \qquad & e<j, \text{undereducated;}\\
& e>j, \text{overeducated;}\\
& e=j, \text{well-matched.}
\end{aligned}
$$

[23]The human capital theory considers that human capital and earnings are proportional to workers' productivity on the job (Rumberger, 1987).

Therefore, if an individual's attained level of education is equal to 2 and the required level to perform the job is equal to 1, for example, then the overeducation premium is given by $w(2,1) - w(2,2)$. Similarly, if an individual's attained level of education is equal to 1, but the required level to perform the job is equal to 2, then the undereducation premium is $w(1,2) - w(1,1)$.

Considering the potential cost of the educational mismatch, the results can be used to compute the following, simple counterfactual results: i) what would happen if all workers in the labour market were well-matched to their jobs? ii) What would be the impact on aggregate wages? To answer those questions I perform two exercises. In the first one, I change individuals' education in order to match the required education to perform the actual job. In the second exercise, I am reassigning the individual to a different job only by changing the required education to perform that job. In both exercises the over/undereducation will vanish. Table 2.8 shows an example of each framework. For instance, consider an individual in the sample, represented by column "Actual". This individual has a higher education degree but holds a job that requires only secondary school, i.e., this worker is overeducated. In the first exercise, illustrated in column "Exercise 1", I consider the case that the attained education is the same as the required education to perform the actual job, i.e., the attained education is secondary school. Now the overeducated worker is well-matched to his/her job. In column "Exercise 2" I move this worker to a job that requires her/his attained education, i.e., the required education is higher education. Again, the previous overeducated individual is now well-matched to her/his job. Note that in this case, I am assuming that there are no costs of changing those individuals jobs and education.

Table 2.8: Simulation exercises

Example	Actual	Exercise 1	Exercise 2
Attained education	Higher education	Secondary school	Higher education
Required education	Secondary school	Secondary school	Higher education

To compute the differences I first construct the counterfactual by predicting workers' wage if they were well-matched to their jobs and then I compare the average wages with the true average outcome.

$$log\,\widetilde{W_{it}} = \boldsymbol{REQ_{it}}\widehat{\boldsymbol{\delta}} + \boldsymbol{X_{it}}\widehat{\boldsymbol{\beta}} + \widehat{c_i} \tag{2.4}$$

$$log\,\widetilde{W_{it}} = \boldsymbol{EDUC_{it}}\widehat{\boldsymbol{\delta}} + \boldsymbol{X_{it}}\widehat{\boldsymbol{\beta}} + \widehat{c_i} \tag{2.5}$$

where $log\widetilde{W_{it}}$ are the logarithms of the predicted wages if workers were well-matched. The coefficients are estimated using equation (2.2), where $\widehat{\boldsymbol{\delta}}$ are the estimated coefficients for attained education, and $\widehat{\boldsymbol{\beta}}$ are the estimated coefficients for other observed characteristics. REQ_{it} is the required education to perform the job, $EDUC_{it}$ is the attained education, $\boldsymbol{X_{it}}$ is the vector of observed characteristics, and $\widehat{c_i}$ is the estimated fixed effect.

Table 2.9 presents the differences in aggregate wages by different subgroups. Considering the whole sample, if one changes workers' education the aggregate wage decreases by 0.32%, and if one switches his/her jobs it increases by 0.25%.[24] Note that depending on the change adopted to eliminate the educational mismatch the impact follows opposite directions, but it is very small.[25] In exercise 1, the counterfactual is a mere redistribution of education, i.e., the aggregate level of education remains the same after I have changed the individual's education.

Table 2.9: Differences in aggregate wages (%)

Switching	(1) Educ	(2) Occup		(1) Educ	(2) Occup
General	-0.32	0.25	Overeducated	-1.18	2.61
Less than Fundamental Education II	-1.37	-2.38	Undereductaed	0.02	-2.17
Complete Fundamental Education II	-0.96	-0.99	Female	-0.20	0.35
Incomplete Secondary School	0.65	-0.43	Male	-0.42	0.15
Complete Secondary School	0.97	0.20	Industry	-0.32	0.10
Incomplete Higher Education	0.30	2.44	Services	-0.33	0.29
Complete Higher Education	-1.29	0.96	Young	-0.42	0.18
Master/Doctorate	-1.33	2.29	Elder	-0.26	0.27

Source: RAIS, 2006-2013.

The results for the educational levels suggest that switching workers' education has a negative effect on lower and higher educational levels. Workers in lower levels are either undereducated or well-matched (Figure A.1), then in exercise 1 their education increase and it may be the case that the premium for the new educational level is not compensating for the loss of the undereducation premium. On the other hand, workers in higher educational levels are mostly overeducated and by switching their education I resign them to a lower level. In this case, the loss in their education reward may be greater than the loss of overeducation penalty, which justifies the negative effect. Unlikely, changing workers' occupation (exercise 2) yields a negative impact for lower educational levels and positive for higher educational ones. These results imply that low educated workers are in lower level jobs than before, while highly educated workers are in higher level occupations.

[24] Using the coarser education categories defined in 2.5.1.2 the results are similar, changing workers' education decreases aggregate wages by -0.29% and switching workers' to a well-matched job increases aggregate wages by 0.27%.

[25] Hsieh et al. (2013) using a Roy model of occupational choice for the United States from 1960 to 2008 find that reductions in barriers to occupational choice (for minorities) can explain 15 to 20% of the increase in aggregate output per worker.

The impact on over and undereducated workers follows opposite directions in both exercises. First, consider the impact of changing the individual's education. In this case, results indicate that overeducated workers would be worse off, suggesting that their loss in the education premium is greater than the loss of overeducation penalty. On the contrary, the loss in undereducated premium would be compensated by the positive return of education. In the second exercise, if I change workers' occupations, results indicate that overeducated hold higher level jobs than before, while undereducated workers hold lower level jobs. Thus, by correcting workers occupation, aggregate wages would increase.

The small results are not surprising given the symmetry of over and undereducation – see Table 2.2 and Table 2.6, column (2). Because the mismatch is linear, then overeducation and undereducation are cancelling each other when one corrects the educational difference. Note that the driver of these results is the symmetry in both over- and undereducation premia and incidence of the mismatch. However, one could also perform a counterfactual exercise that would rely less on averages. One approach would be testing the mismatch in a different way. For example, Mehta et al. (2011) test the incidence of overeducation in developing countries by checking which jobs offer low returns to education and how many educated workers are employed in these jobs. This way, under- and overeducation may have different weights in an unskilled labour market and in a skilled one, which may have a different impact on aggregate wages. However, further analysis is necessary to understand the effects of the mismatch on the two segmented labour markets.

2.6 Conclusion

This paper provides an analysis of the impact of educational mismatch in the Brazilian labour market. Using a panel dataset for the formal labour market in Brazil (RAIS), I examine the incidence of educational mismatch and estimate the wage effects of over and undereducation controlling for workers' heterogeneity. The main specification compares over/undereducated workers with their equally educated colleagues who hold a well-matched job. The use of the panel data allows the individual's attained education to vary over time.

The main findings show that one quarter of the Brazilian formal labour market is overeducated, and one quarter is undereducated. Those overeducated (undereducated) workers earn significantly lower (more) than their co-workers who hold a well-matched job after individual's unobserved abilities are controlled for. These results suggest that the remained differences may be due to job characteristics and to the match between worker-occupation. Despite the penalty for overeducation, the premium for going to college is more than double of that penalty suggesting that for the individual perspective, it worth acquiring higher education. Therefore, the overeducation may be arising due to an individual decision. As expected, the results also indicate that the individual heterogeneity plays an important role in the educational context.

In addition, I asked what would happen if there was no overeducation. The analysis shows that eliminating the educational mismatch by changing workers' education would result in a decrease of aggregate wages while eliminating the mismatch by changing workers' jobs would result in an increase of the aggregate wages. However, these results are very small indicating that given the symmetry of the educational mismatch observed in the dataset, over and undereducation effects are cancelling each other. Moreover, when I changed education for overeducated workers results imply that the loss of the education premium overcome the loss of the overeducation penalty, while for undereducated the educational premium compensates the loss in the undereducation premium. On the contrary, when I moved workers to a well-matched occupation I reassign overeducated (undereducated) workers to high (low) level jobs, hence the aggregate wages increase (decrease).

An alternative explanation for educational mismatch in Brazil is that the increase in the educational level of the population has been followed by a decrease in educational quality. Highly educated workers are not qualified to perform jobs for which they studied and then they end up in unskilled occupations, explaining the incidence of overeducation. For undereducated workers, the problem would be related to the lack of education, which may constrain their productivities in their jobs. Therefore, investments on-the-job training could be an alternative in which workers can acquire the necessary skills to perform their jobs.

From the policy perspective, the presence of educational mismatch is important to review policies that expand the access to education. Moreover, policy-makers should also concern with the quality of education provided, particularly for workers who are more likely to end up in an unskilled job, and with the creation of more education-intensive jobs. Thus, policies should focus not only on increasing the educational level of the labour force, but they also should focus on the increase of the educational quality provided by the institutions in order to raise the workers' productivity. However, it is important to highlight that one should not only focus on the effects of education on workers' productivity. Individuals could benefit from more schooling on other aspects such as better health outcomes and empowerment. Finally, it is important to highlight that the results hold for the Brazilian formal labour market only.

Chapter 3

The effects of minimum wage on education acquisition in Brazil

Abstract. Changes in the minimum wage will affect the wage distribution and possibly skill premia. In particular, the latter will affect education investment decisions. This paper examines potential effects of a minimum wage policy on education investment on-the-job. To do so I proceed in two steps: first, I examine the impact on the skill premium of changes in two different measures of the minimum wage policy: absolute and relative level of the minimum wage. To identify the effects of the national minimum wage policy, I exploit variation of price levels across time and states. Second, I measure the effect of the two measures on education investments. I find that an increase in the absolute minimum wage decreases the skill premium and has, on average, a negative effect on educational acquisition. The impact is more pronounced on the bottom of the wage distribution (10th and 25th percentiles). By contrast, an increase in the relative minimum wage increases the skill premium for top levels of education and, on average, has no significant effect on education investments. However, there is some evidence for heterogeneous effects as investments for the 25th, 50th and 75th percentiles of the wage distribution are positively affected.[1]

3.1 Introduction

Minimum wages are frequently used as a labour market policy instrument around the world, often with an aim to reduce poverty and inequality by increasing earnings in the bottom of the income distribution. Changes in the minimum wage will alter the distribution of wages and also skill premia. The latter will change educational investment

[1] I would like to thank Thomas Gall, Michael Vlassopoulos, Corrado Giulietti, Hector Calvo Pardo, Mirko Draca, Christian Schluter, as well as conference participants at the 2018 Economics PhD Annual Workshop and University of Southampton internal seminars for valuable comments and suggestions on earlier drafts of this paper. I gratefully acknowledge financial support from the *ESRC*.

incentives. Where the minimum wage compresses the wage distribution, depressing wage premia for intermediary and higher education the returns to education will diminish, and thus education acquisition may decrease, too. Conversely, if the minimum wage increases the wage premia for higher education the returns to education will increase, and thus education acquisition may increase. Hence, whether the minimum wage is going to encourage or discourage education acquisition is ambiguous.

The focus of the debate among policy-makers, and indeed much of the economic literature, lies on possible employment effects of minimum wage policies.[2] Potential effects of such policies on education acquisition, through changes in the wage distribution and skill premia, have, however, received scant attention. This paper examines potential encouragement or discouragement effects of changes in the minimum wage policy on educational investment. Investigating this question faces two formidable obstacles. Firstly, changes in the minimum wage policy may well be endogenous to labour market outcomes. Secondly, the impact of a change in the minimum wage policy on skill premia and incentives for education acquisition is very ambiguous and will depend on the precise effect on different quantiles of the wage distribution.

To deal with the first obstacle I use plausibly exogenous variation in the price levels over time and across regions as a source of variation of the nominal minimum wage. Using a rich panel dataset from Brazil allows me to exploit the vast regional differences in labour markets and business cycles across regions. I address the second obstacle by employing two different proxy variables of the real minimum wage that can be expected to be associated with different effects on the wage distribution and skill premia: an absolute and a relative minimum wage measure. The relative minimum wage captures the distance of the lower bound of the wage distribution to its mean, and thus reflects squeezes in the lower half of the wage distribution. The absolute minimum wage measure captures instead an aggregate effect, i.e. increases in the lower bound of the support unconditional on the mean or other moments of the wage distribution. Hence, I expect the relative minimum wage to be associated with a relative decrease in the wage premium for intermediate education and the absolute minimum wage with a relative decrease of the support of the wage and skill premium distribution.

The analysis proceeds in two steps. First, I estimate the impact of the two measures, absolute and relative minimum wage on the skill premium. Then I turn to estimate the effects of changes in the minimum wage policy on education acquisition on the job, using the two measures of changes in the minimum wage to disentangle the different distributional effects. The results of the first step show that the two minimum wage measures do indeed capture different effects on the skill premium: while the absolute

[2]Evidence for the US found initially negative effect of minimum wages on employment, while recent studies found negative and slightly positive effects (Brown et al., 1982; Card, 1992; Card and Krueger, 1994; Giuliano, 2013). In the UK, these effects are zero or marginally positive (Machin and Manning, 1994; Stewart, 2004; Dolton et al., 2015). In Latin America, studies found that wage compression and effects on employment are stronger than those found in developed countries (Maloney and Mendez, 2004).

minimum wage is associated with a reduction in the support of the skill premium distribution and a compression of skill premia in particular for higher education, the opposite is true for the relative minimum wage. Lower and intermediate education premia seem unaffected, but premia for higher education increase significantly. This is at first glance surprising, since an increase in the relative wage suggests a compression of the wage distribution, but only in its lower half. The relative minimum wage thus appears to capture a catching up of lower education with intermediate education wage premia, but not with the wage premium for higher education.

In the second step, I use the two measures, associated with opposing effects on the skill premia particularly for higher education, to estimate the effect of changes in the minimum wage policy accounting for the distributional channel. As expected the absolute measure of the minimum wage, capturing a degree of wage compression, has on average a negative effect on education acquisition. The impact is more pronounced at the bottom of the wage distribution (10th and 25th percentiles). By contrast, the relative minimum wage has no significant effect on educational investments on average. The zero aggregate effect may mask considerable heterogeneity, however: the effect on education acquisition is positive at the 25th, 50th and 75th percentiles of the wage distribution. These results confirm basic economic intuition: minimum wage policies that erode the skill premia for education, especially for secondary education and above, will dampen incentives for education acquisition, indicating that the use of minimum wage policies requires careful evaluation of likely effect on the wage distribution and skill premia.

Brazil is an ideal test case for several reasons. First, because over the past years, it has been experiencing a large increase in the schooling of its workforce due to several important reforms in the educational system.[3] However, the increase in the share of highly educated workers has been followed by changes in the wage premium. For instance, the changes in the real wage were positive for tertiary and primary education. Particularly, there had been a considerable increase in real wages for workers that have less than complete primary education (Bruns et al., 2011). Some economists associate this increase in real wages of low-educated workers to active minimum wage laws and the expansion of cash transfers programs over this period (Barros et al., 2010; Bruns et al., 2011). Second, there is ample variation of real minimum wage across states, which are very different in terms of income, so the national minimum wage will affect differently distinct states. Third, I rely on a unique dataset – Annual Social Information Report (RAIS), which allows me to analyse education acquisition on-the-job. According to the Educational Census (2013), 32.2% of the population between 15 and 29 years old work and study at the same time. This number is even higher for people who take courses to complete primary and secondary education after the regular age, around 55% declared that work and study simultaneously. At the university level, 70% of the students declared that they also have a job (INEP, 2013a,b).

[3] For more information about these reforms, see Bruns et al. (2011).

This paper contributes to our understanding of the impact of minimum wage policies on educational investments. To the best of my knowledge, this is the first paper to study the distributional effect of minimum wage changes and its impact on education acquisition on-the-job. This paper contributes to the existing literature in different ways. First, it relates to the literature on the redistributive effects of the minimum wage on the wage distribution (DiNardo et al., 1996; Lee, 1999; Teulings, 2003; Neumark et al., 2007; Butcher et al., 2012; Dolton et al., 2012). The main idea of these papers is that the minimum wage truncates the wage distribution and affects particularly those workers for whom the minimum wage is binding. In this paper, I provide evidence that while increases in the absolute minimum wage will affect the wage distribution by decreasing the skill premium. Conversely, increases in the relative minimum wage capture raises in the skill premium on top levels suggesting that the compression of the wage distribution is only on its lower half. Moreover, the impact is not only for workers for whom the minimum wage is binding, but it affects differently distinct parts of the wage distribution. By truncating the wage distribution, the minimum wage changes skill prices and thus affects people's decision of educational investments. Workers choose their education optimally, as they are heterogeneous on their abilities and face different time-costs of education. Higher ability workers are more productive, and education acquisition enhances their productivity, allowing them to entry the high-skilled sector, where in equilibrium the productivity per ability is greater (Bárány, 2016). Hence, if an increase in the minimum wage excludes low-productivity workers and education increases one's productivity, changes in the minimum wage modifies individual's optimal educational choices, which on its turn may change the skill composition of the workforce. An increase in the minimum wage leads to an outflow of low-skilled workers in the labour market (Card and Krueger, 2000). This outflow of low-skilled workers slowly changes the composition of the labour force and reduces the skill premium, which decreases the incentives of acquiring education. Contrary, the increase in the minimum wage makes more difficult to find a job, rising education importance, particularly for low-skilled workers. Moreover, as the workforce becomes more skilled and the decisions to acquire education also changed, the average skill in both low-skilled and high-skilled sectors increases. Because high-skilled workers are entering low-skilled sector, the skill premium may decrease even more.[4] Thus, depending on which force prevails, the effect of changes in the minimum wage police may increase or decrease the investments in education, depending whether raise or reduce the skill premium.

Second, it relates to the literature that investigates the effects of the minimum wage on additional education acquisition. Bárány (2016) develops a general equilibrium model to assess the permanent effects of a decrease in the minimum wage in the United States in the 1980s on inequality. The model includes minimum wages, endogenous educational choices and endogenous technological progress. The author finds that a decrease in the

[4] Autor et al. (2005) demonstrate that changes in the education composition of the workforce plays a significant role in reducing inequality in the lower-tail of the wage distribution.

minimum wage alters the share of low-skilled and high-skilled workers on the entire wage distribution, and reduces the skill premium. Much of the empirical studies analyse jointly how minimum wage impacts employment and educational enrolment (Mattila, 1978; Ehrenberg and Marcus, 1982; Neumark and Wascher, 1995a,b, 2003; Turner and Demiralp, 2001).[5] While higher minimum wage makes it more difficult for low-skilled workers to find employment, it also increases the attractiveness of minimum wage jobs. Overall, these studies indicate that as minimum wages rise, some higher-ability individuals may leave school to pursue a minimum wage job. Conversely, low-ability individuals are more likely to drop out of school and stay in unemployment. By contrast, this paper adds to the analysis of the effects of the minimum wage on education by investigating the educational investments on the job. Contrary to developed countries, where minimum wages and education are substitutes, in Brazil that is not necessarily the case. Around 30% of the enrolments in secondary and tertiary education are for night classes (INEP, 2013a), which allows people to work during the day and take classes to complete their education at night.

Third, this paper also relates to the literature that evaluates the effects of the minimum wage on-the-job training. Acemoglu and Pischke (1999, 2003) develop a hybrid model in which the minimum wage may increase or decrease the training depending on the degree of the labour market competition. The authors find that in imperfect markets because workers are not paid their marginal product when they move to a different firm, their general skills become firm-specific skills. Also, the increase in minimum wage compresses firms' wage structure and if firms are forced to pay higher wages to unskilled workers, then they will have more incentive to increase the productivity of their workers by increasing training in general skills. Several empirical studies document these effects of the minimum wage on training. Arulampalam et al. (2004) find an increase in training opportunities after the introduction of the national minimum wage in the United Kingdom in 1999. Lechthaler and Snower (2008) expand the Acemoglu and Pischke model by adding endogenous job separation and find that increases in the minimum wage may have negative or positive effects on training. The net effect depends on the distribution of workers' ability.[6] Neumark and Wascher (2001) explore increases in federal and state minimum wages in the United States and find that an increase in the minimum wage decreases formal training on the current job. Hara (2017) analyses the formal and informal firm-provided training and worker-initiated training for Japan. The author finds that increases in the minimum wage lead to a decrease in formal training, but there is no effect on informal and worker-initiated training.

Thus, the theoretical predictions are ambiguous and the empirical literature has no consensus about the effects of the minimum wage on training. Some studies have found

[5] Ravn and Sørensen (1999) find that an increase in the minimum wage reduces firm-provided on-the-job training and encourages schooling prior to entering the labour market.

[6] The authors argue that for low-ability workers the impact is negative since it reduces the profitability of the worker and the firm is more likely to lay them off, while for high-ability workers the effect is positive due to wage compression.

that minimum wage increase has a negative effect on training (Mincer and Leighton, 1981; Hashimoto, 1982; Neumark and Wascher, 2001; Hara, 2017; Schumann, 2017). Others have found positive results (Arulampalam et al., 2004) and mixed results or no effects (Grossberg and Sicilian, 1999; Acemoglu and Pischke, 2003; Fairris and Pedace, 2004; Lechthaler and Snower, 2008). This paper provides new empirical evidence for the effects of the minimum wage on general training, i.e. educational investments.

Finally, this paper contributes to the literature on determinants of decreasing inequality and changes in the wage premia for workers with different educational levels in Brazil (Barros et al., 2010; Bruns et al., 2011; Gasparini et al., 2011; Barbosa et al., 2015; Dedecca, 2015; Engbom and Moser, 2017).

The rest of this paper is organised as follows: Section 3.2 discusses the minimum wage in Brazil and describes the two measures used in this study; Section 3.3 describes the data and descriptives; Section 3.4 contains the empirical approach; Section 3.5 presents the main results; Section 3.6 shows the results for subgroup analysis; Section 3.7 presents robustness checks, and Section 3.8 concludes.

3.2 Minimum wage policy in Brazil

3.2.1 Background

In Brazil, a minimum wage has been in place since 1940. Since 1984 the minimum wage has been determined by federal law at a national level.[7] From 2011 on, by law the minimum wage has to increase at the rate of inflation (National Consumer Price Index – INPC) during the last 12 months plus the real rate of GDP growth two years prior:

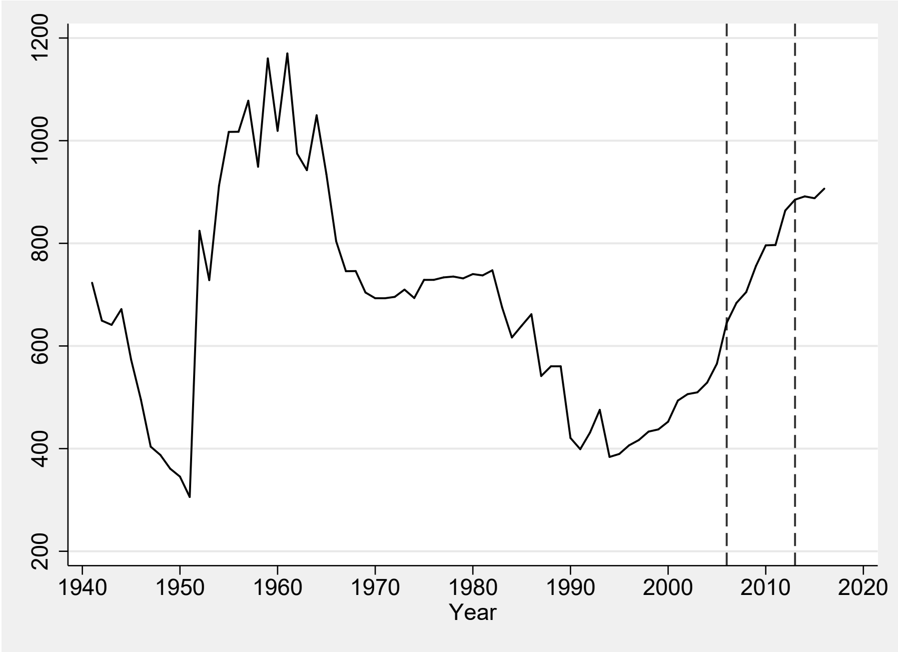
$$\Delta MinWage_t = \Delta INPC_{t-1} + \Delta GDP_{t-2} \tag{3.1}$$

Although the present law was established in 2011, the government has been using these rules to adjust the minimum wage officially since 2005, but it was a non-legislated policy.

The minimum wage is enforced by the Labour Justice system, which charges fines on noncompliant employers. Additionally, the enforcement of the minimum wage law is fiscalised by unions, which increased their power over the last years accentuating their influence on labour rights (Simão, 2009).

[7]Law No. 12,382 of February, 2011.

Figure 3.1: Brazilian real minimum wage 1940-2016 (in 2016 prices, BRL)



Source: Ipeadata (2017).
Note: The dashed lines highlight the period analysed in this paper.

Figure 3.1 shows the evolution of the real minimum wage in Brazil since its creation in 1940. As one sees, the national real minimum wage presented high variation over time. Originally, the minimum wage had a term of validity of three years. After the adjustment in 1943, the nominal minimum wage was not adjusted until 1951, losing its real value due to inflation. During the 50s, the real minimum wage was adjusted often and increased due to high productivity and the populist government. Then it decreased again because of the acceleration of inflation. In 1964, the dictatorship government changed the way that the minimum wage was adjusted because of the high inflation. The under-indexation of the nominal minimum wage transformed it in a stabilization policy (Camargo, 1984). At the end of the 70s, the nominal minimum wage was adjusted monthly by 10% above the inflation rate, which was responsible for a small recovery in its real value. During the 80s and early 90s, Brazil witnessed several plans to control high inflation. The nominal minimum wage was increasing, however, it was eroded by the increasing inflation (Lemos, 2004). After 1994, the real minimum wage increased again after the economic stabilization and the growth was particularly high after 2005.

It is widely held that the active minimum wage policy was one of the main factors that led to a decrease of inequality in Brazil in the first decade of the 21^{st} century (Barbosa et al., 2015). By linking the minimum wage to GDP, real minimum wages increased faster than average wages indicating that the policy led to some wage compression during the period.[8]

[8]The increase in the minimum wage jointly with formal job creation led to a wage level convergence within and between sectors, regions and social groups in Brazil (Dedecca, 2015).

The increase in the minimum wage is associated with the shrinking of the formal sector and an increase of low wages in the informal sector, though.[9] This was not the case for Brazil, which faced an expansion on employment and wages in the formal labour market during the period of active increase of the minimum wage (Barbosa et al., 2015).

Minimum wage policy still faces some criticism about the fact that few people are impacted directly by it.[10] Despite one should expect bigger effects for workers for whom the minimum wage is binding, some theories suggest that the indirect effects of the minimum wage may impact the higher part of the wage distribution by changing the general equilibrium in the labour market (Burdett and Mortensen, 1998). The magnitude of those spillover effects of the minimum wage is an open question though. Moreover, there is evidence that the minimum wage has been used as a numeraire in Brazil, acting as a reference for wage determination.[11] Therefore, because of the numeraire effects not only low-wage workers should be impacted by changes in the minimum wage, but also workers earnings above the minimum. Hence, a minimum wage policy should affect different parts of the wage distribution.[12]

Lastly, Figure 3.2 shows that the earnings distribution for the period analysed in this paper (2006-2013). It can be seen that in the first year of the analysis the wages are more concentrated in the left tail than in the last year, i.e. there is a movement to the right from 2006 to 2013. This movement may be related to the increase in the minimum wage over this period, as shown in Figure 3.1.[13]

[9] The formal sector minimum wage is used as a reference in the informal sector, which is called "lighthouse effect". The lighthouse effect or *Teoria do Farol* it was first introduced by Souza and Baltar (1979).

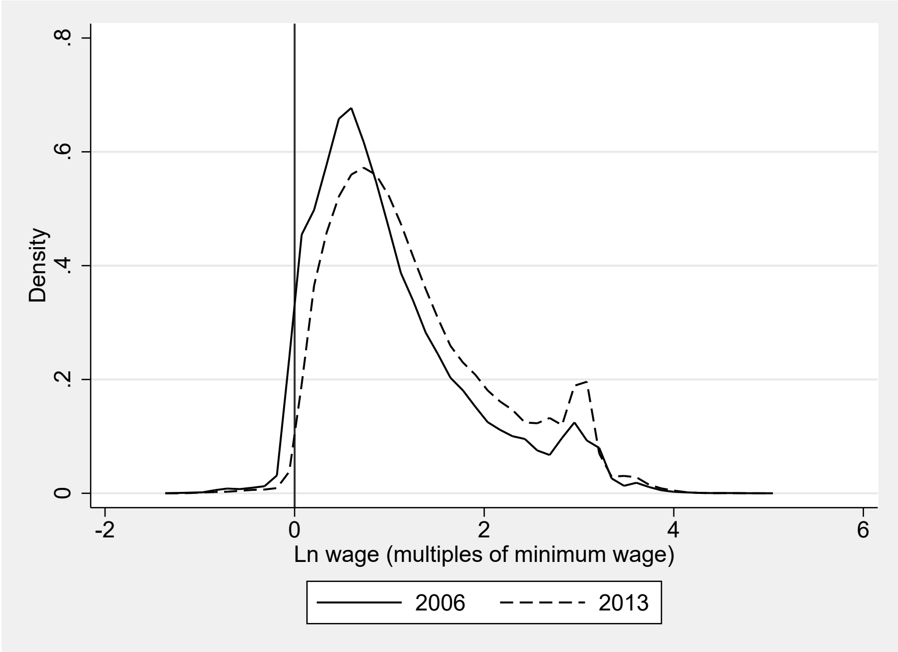
[10] Figure B.4 in Appendix shows the share of workers earning exactly the minimum wage, the share below the minimum wage, and the share up to 10% above the minimum wage. At the beginning of the analysis, at least 4% of workers are binding at the minimum wage, and 4% earn up to 10% above the minimum wage.

[11] Neri et al. (2000) find evidence of numeraire effects all over the wage distribution in Brazil. They find that in the formal labour market, 6% of workers received multiples of the minimum wage (0.5, 1, 1.5, 2 or 3).

[12] Empirical evidence shows that the minimum wage affects the earnings for workers with higher wages, and not only for those earning the minimum wage (Lee, 1999; DiNardo et al., 1996; Acemoglu and Pischke, 2003). Lechthaler and Snower (2008) provide evidence of the spillover effects of the minimum wage on training.

[13] Figures B.1, B.2 and B.3 in Appendix present the earnings distribution by year for all workers, male and female, respectively. From these figures, one sees that the left tail of the distribution is more compressed for women than for men.

Figure 3.2: Histogram of the earnings distribution



Source: RAIS, 2006-2013.

3.2.2 The minimum wage measures

The nominal minimum wage is the same for the entire country and does not capture regional differences. However, the effects of the minimum wage vary from low wages states to high wages states, which is not captured by the national minimum wage. Therefore, using data from the Brazilian Ministry of Labour and Employment (MTE) I construct two different measures of the minimum wage to proxy for distributional effects.

The first one, which is more straightforward, is the *absolute minimum wage* for each state. In order to get variation of the national minimum wage across states I deflate it by the regional Brazilian consumer price index, IPCA.[14] The IPCA price index covers 11 metropolitan areas in Brazil: Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba, Porto Alegre, Brasília and Goiânia. However, there are 26 states in Brazil plus the Federal District, then some states use the same deflator index.[15]

The first measure captures the absolute level of the minimum wage in real terms. By contrast, the second measure aims to capture the size of the minimum wage relative to average wages. The *relative minimum wage* is defined as $MW_t^N/\overline{W}_{st}$, where MW_t^N is the national minimum wage at year t and $\overline{W}_{st}$ is the average wage at state s at year t.

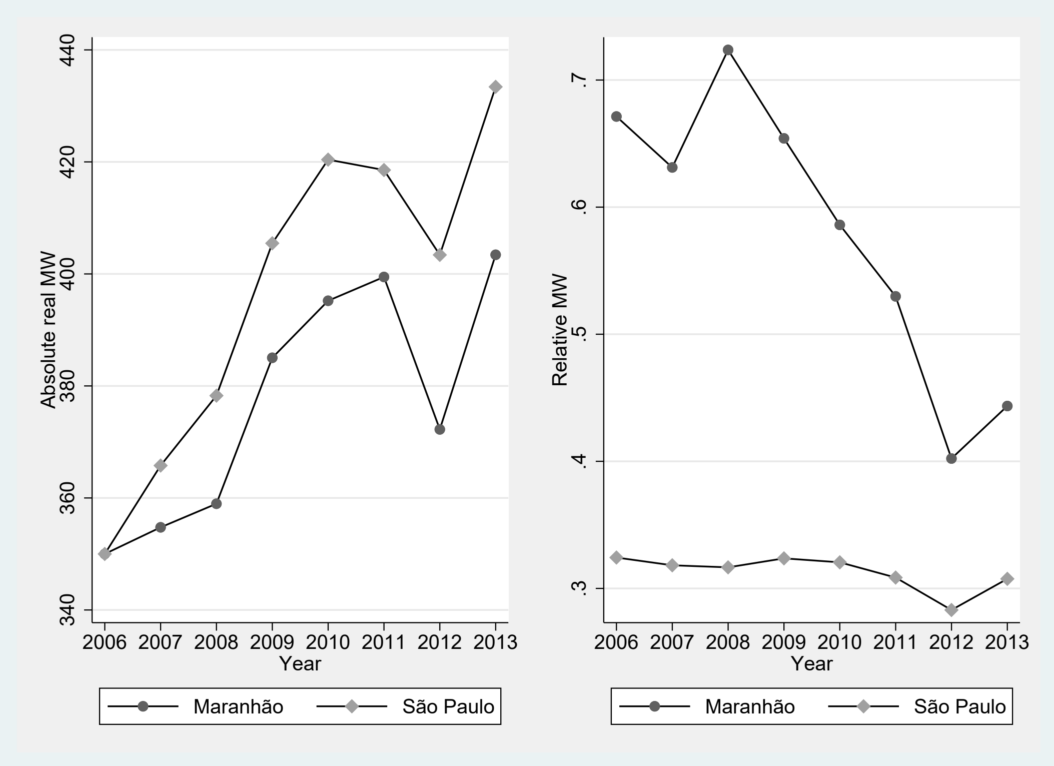
Thus, both the absolute minimum wage and the relative minimum wage are exogenous, because they vary with the price levels over time and states. That is, changes in my

[14] The IPCA is the official price index in Brazil and it is used by the Brazilian Central Bank to determine inflation targets.

[15] I selected which deflator index to use based on the proximity and similarity of the states.

measures of the minimum wages do not depend on the nominal level set by policy-makers. Their variation stems from differences in business cycles between different states. However, while the absolute minimum wage captures both shifts and compression, the relative minimum wage captures the compression in the lower half of the wage distribution.

Figure 3.3: Absolute and relative minimum wage evolution over time



Source: RAIS, 2006-2013.

Note: The absolute real minimum wage is in BRL and it was deflated by 2006 prices.

Brazil is a heterogeneous country and the wage gap between states is large. According to the Brazilian Institute of Geography and Statistics (IBGE), the wage gap between the poorest and richest state is almost 250% (IBGE, 2016).[16] Consider for example two extremes, on one hand, is Maranhão, one of the poorest states in the country, and on the other hand is São Paulo, one of the richest states. Figure 3.3 shows the evolution over time of the absolute minimum wage (left) and the relative minimum wage (right) for the two states. First, on the left side, one sees that as the national minimum wage was deflated by 2006 prices, the absolute minimum wage is the same for both states. However, as price levels evolve differently over time in these two states, the absolute minimum wage is smaller in Maranhão. Note that the steep decrease of the absolute real minimum wage in 2012 is because the adjustment in the nominal minimum wage in that year was very small. On the right side, one sees different patterns for the two states. These differences are due to distinct movements on the state's average wage. Thus, due

[16] The wage gap was calculated by the IBGE using data from the National Household Survey (PNAD) for 2016.

to this heterogeneity in the states, the minimum wage policy may have different effects in different states.

In order to make the two proxies comparable, I normalised them between zero and one to create an index. Both measures vary within the country, by state or by metropolitan region:

$$MWindex_{st} = \frac{MW_s - min(MW_s)}{max(MW_s) - min(MW_s)} \tag{3.2}$$

where $MWindex_{st}$ is the index of the minimum wage in state s at time t, MW_s is either the absolute minimum wage or relative minimum wage at state s, $min(MW_s)$ is the minimum value of the minimum wage among all states at a particular time and $max(MW_s)$ is the maximum value of the minimum wage among all states at a particular time.

3.3 Data and descriptive statistics

3.3.1 Data

The dataset used in this paper stems from the Annual Social Information Report (RAIS), and covers the years 2006 to 2013. RAIS is an administrative dataset that contains socioeconomic information of all firms and their employees in Brazil. Since all tax registered firms have to report their information by law, RAIS can be viewed as a census of the Brazilian formal labour market. Hence, any results generalise immediately to the entire formal sector in Brazil in 2006-2013. The data used for the empirical analysis is a 0.5% representative random sample panel of firms selected in 2006 and followed until 2013. All workers from those selected firms in 2006 were followed until 2013. In addition, those individuals who entered the selected firms after 2006 were followed back in time in order to get their prior information.[17] This paper follows 162,744 workers from 25 to 58 years old (in 2006).[18]

The investments in education on-the-job variable used in this study is a binary dummy variable that takes the value 1 if the worker changed her/his attained education compared to the previous year and 0 otherwise. Attained education is a categorical variable in the dataset and it is separated into 11 categories: illiterate, incomplete fundamental education I, complete fundamental education I, incomplete fundamental education II, complete fundamental education II, incomplete secondary school, complete secondary school, incomplete higher education, complete higher education, master and doctorate.

[17] Initially I selected 8,355 firms, but due to the incorporation of previous information of those workers who entered the selected firms later, the sample has 79,072 firms in total.

[18] The cut off age was picked in order to exclude workers attending university at the regular age and retirees.

Reassuringly, all workers only ever increase their education, so the scope for misreporting is rather limited.

Additionally, the dataset includes information on tenure, age, gender, educational mismatch (if the worker has a different education than required by the job she/he performs), class of worker, firm size, sector, state and year. Table B.1 presents the descriptive statistics for the sample.

3.3.2 Descriptives

As discussed in the introduction, education impacts earnings distribution through changes in education distribution and its price effects.[19] Thus, I present some descriptives in order to better understand the returns to education and schooling acquisition before estimating the impact of the minimum wage on investments in education.

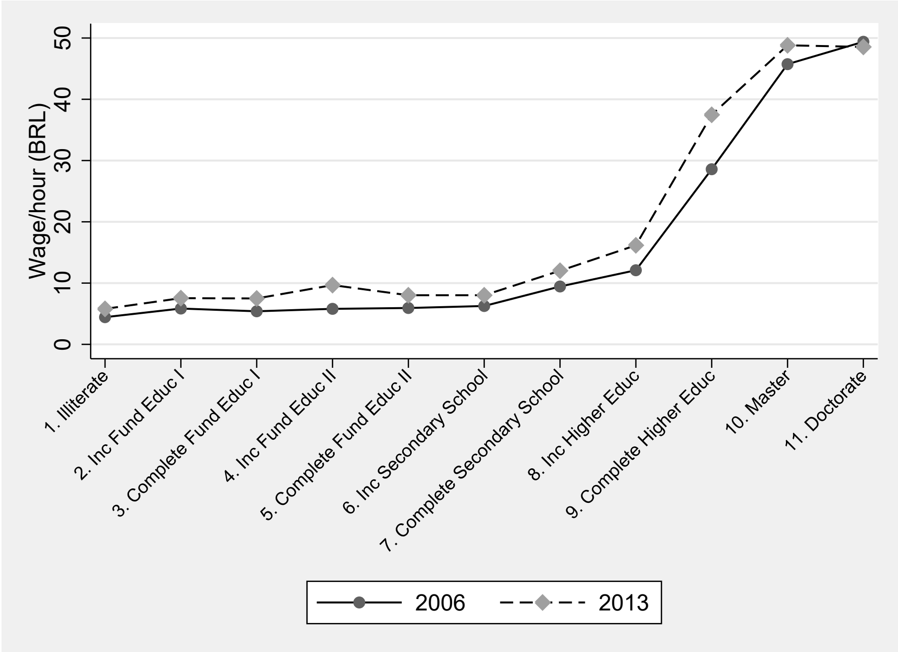
Figure 3.4 depicts the average earning by educational levels for 2006 and 2013. In both years, the impact of schooling on earnings becomes steeper for the last degrees (higher education). Indeed, the returns are higher in 2013 than in 2006 for almost all levels of education with one exception: achieving a doctorate is associated with a drop in earnings compared to a Master. One possible explanation is that the demand for doctors in Brazil is not as high as the supply. The number of people who hold a doctorate degree increased by 486% from 1996 to 2014 (CGEE, 2016). Moreover, unemployment is increasing among doctors because of monetary cuts in science that are directly affecting universities, which are the main market for postgraduates in Brazil.

The labour earnings in Brazil have risen over time for all levels as shown in Figure 3.5. This increase is higher for tertiary education, in particular after 2008. Also, it highlights the difference in earnings for the different educational levels. The gap between tertiary and secondary education is greater than the gap between secondary and primary education[20].

[19] Becker and Chiswick (1966); Sattinger (1993); Barros et al. (2010).

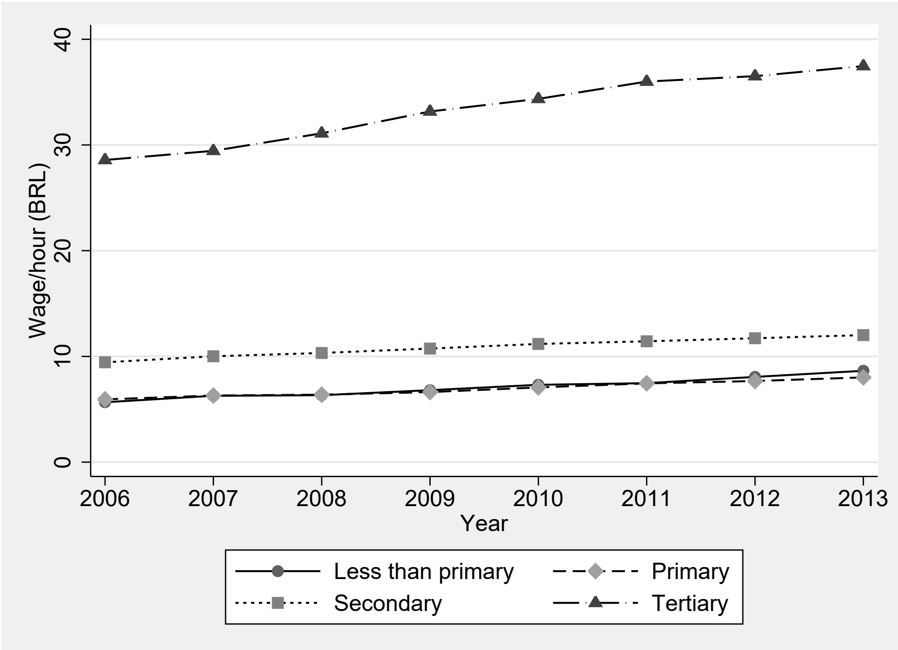
[20] Less than primary education includes illiterate, incomplete and complete fundamental education I and incomplete fundamental education II. Primary consists of complete fundamental education II. Secondary comprehends only complete secondary school, and tertiary consist of only complete higher education.

Figure 3.4: Education and labour earnings



Source: RAIS, 2006-2013.

Figure 3.5: Evolution of labour earnings by education



Source: RAIS, 2006-2013.

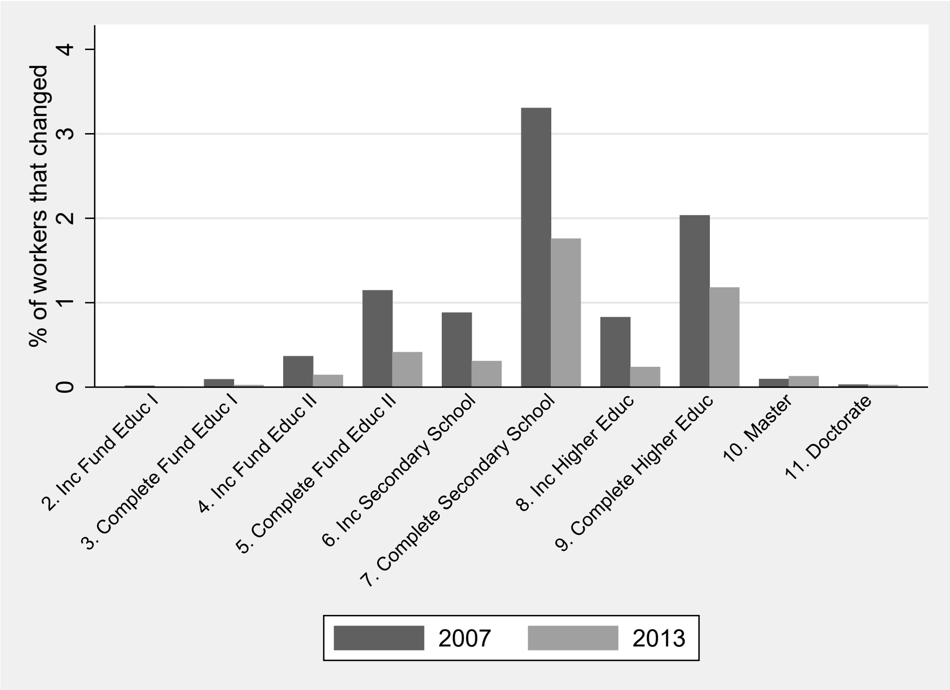
Considering the wage ratio among educational levels, one sees in Table 3.1 that it decreases as education increases. Comparing 2006 and 2013, the wage ration raised for primary and secondary, but it declined for tertiary education. The gap between tertiary and secondary education expanded over time due to the higher increase in wages for the higher level (as depicted in Figure 3.5).

Table 3.1: Wages ratio, 2006/2013

Education	Wages ratio 2006	Wages ratio 2013	% Δ Wage ratio	Ratio
Primary	0.9544	1.0780	12.95	incomplete/complete
Secondary	0.6286	0.6671	6.13	primary/secondary
Tertiary	0.3304	0.3207	-2.93	secondary/tertiary

Source: RAIS, 2006-2013.

Figure 3.6: Changes in attained education compared to the previous year



Source: RAIS, 2006-2013.

The investments in education on-the-job are identified by any worker changing their stated attained education level compared to the previous year. Figure 3.6 presents the share of workers who increased their education in each year during the period 2007 to 2013.[21] As shown in the graph, the changes in education are greater at the beginning of the analysis than at the end. Since I am following the same individuals over time, this difference indicates that workers are more prone to acquire education while younger. Moreover, the figure shows that in both years the majority of these investments in education are in order to complete secondary school and higher education, respectively. Note that in Brazil, evening classes are common for secondary and tertiary education. According to the Educational Census (2013), around 30% of the enrolments in secondary school are for night classes. For tertiary education the numbers are similar for public universities, 30% and 40% of students are enrolled at evening classes in Federal universities and state

[21] In total, 6.79% of workers in the sample changed their attained education.

universities, respectively. However, in private universities, the number of night classes enrolment predominates, around 70% (INEP, 2013a,b).

Lastly, is important to note that while in developed countries minimum wage and education are substitutes, the former is procyclical and the latter is countercyclical, it seems not the case for Brazil. As presented previously in this paper, at the same period that the minimum wage increased, there was a great expansion in the educational system as a result of several reforms implemented, increasing the schooling of the country workforce. For example, the Youth and Adult Education program *(Educação de Jovens e Adultos - EJA)*[22] focuses on youth and adult education for those who could not complete fundamental education or secondary education at the regular age. According to the program, the public education system shall ensure free educational opportunities for young people and adults, who are unable to undertake studies at the appropriate age, taking into account the characteristics of the student, their interests, living and working conditions. Another example is the University for All Program *(Programa Universidade Para Todos - PROUNI)*[23], which grants full and partial scholarships in undergraduate courses in private higher education institutions.

3.4 Empirical Approach

In order to check whether the minimum wage impacts education investments on the job, I proceed in two steps. First, I analyse the impact of the absolute and relative minimum wages on the wage distribution by looking at their effect on the skill premiums. Then, I look at how the distributional effects, proxied by the absolute and relative minimum wage, affect the education acquisition.

To measure the impact of the minimum wage on the skill premium for different education levels I estimated the following augmented Mincerian wage equation:

$$log\, W_{it} = \delta_1 MWindex_{st} + \sum_{k=1}^{10} \delta_{2k} Educ_{itk} + \sum_{k=1}^{10} \delta_{3k} Educ_{itk} \times MWindex_{st} + \beta_1 X'_{it} + \varepsilon_{it} \tag{3.3}$$

where $log\, W_{it}$ denotes the logarithm of the wage of individual i at time t, MW_{st} is the minimum wage index, either for the absolute minimum wage or relative minimum wage, $Educ_{it}$ is a series of dummy variables measuring the individual attained schooling category, X'_{it} is the vector of observed characteristics, and ε_{it} is the error term.

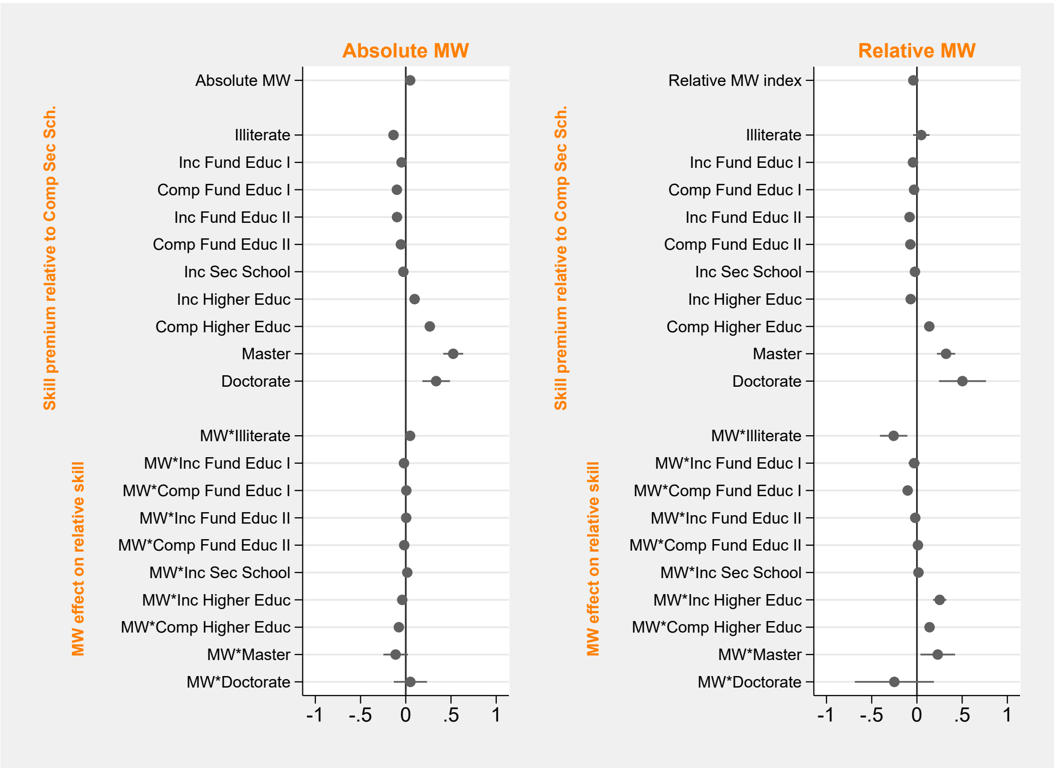
[22] Law No. 9,394/96.
[23] Law No. 11,096/05

To take into account the individual heterogeneity, I re-estimate equation (3.5) including a workers' fixed effect term, c_i:

$$log\, W_{it} = \delta_1 MWindex_{st} + \sum_{k=1}^{10} \delta_{2k} Educ_{itk} + \sum_{k=1}^{10} \delta_{3k} Educ_{itk} \times MWindex_{st} + \beta_1 X'_{it} + c_i + \varepsilon_{it} \tag{3.4}$$

The results are presented in Tables B.2 and B.3. In order to help the comparison of the results for the two measures of the minimum wage, the coefficients of interest are presented in Figure 3.7.[24]

Figure 3.7: Skill premium and minimum wage



Note: Coefficients are the same as in column (2) in Tables B.2 and B.3 in Appendix. Omitted education variable: complete secondary school.

The skill premium associated to complete secondary school has a similar pattern in both regressions. While the returns to education associated with educational levels below complete secondary school (omitted category) are negative, the levels above are positive, consistent with previous expectations. However, when one looks at the effect of the minimum wage on skill, the two measures of minimum wages present divergent results. First, consider the effect of the absolute minimum wage (left column of Figure 3.7), notice that it decreases the wage premium for higher education and master and it increases for illiterate. Conversely, the relative minimum wage (right column of Figure

[24]Figure 3.7 presents only the coefficients for column (2) of tables B.2 and B.3.

3.7) decreases wages for lower levels of education, e.g. illiterate, incomplete and complete fundamental education I, and raises for higher levels, such as incomplete and complete higher education and master.

Overall, the results show that a change in the two proxies is accompanied by two different changes in the wage distribution. While the absolute minimum wage is a proxy for skill premium reduction, the relative minimum wage is a proxy for skill premium increase. Hence, I can interpret the two measures as capturing very distinct changes in wage distribution related to a change in the minimum wage policy. This allows me to disentangle two different distributional shifts from wage compressions.

To account for the effects of minimum wage on investments in education on-the-job, I estimate the following linear probability model (LPM):

$$CEduc_{ist} = \alpha + \delta_1 MWindex_{st} + \delta_2 Mismatch_{it-1} + \delta_3 State_s + \delta_4 Year_t + \beta_1 X'_{ist} + \varepsilon_{ist} \tag{3.5}$$

where $CEduc_{ist}$ is the increase in education dummy variable for individual i at state s in year t.[25] $MWindex_{st}$ is either the index for the absolute minimum wage or relative minimum wage. $Mismatch_{(t-1)}$ refers to the educational mismatch or overeducation/undereducation (more/less education than required by the job, respectively) at the previous period. Dummy variables for each state, $State_s$, control for differences in acquiring more education between states. Additionally, dummy variables for each year, $Year_t$, captures national trends in getting more education during the period analysed. Lastly, X'_{ist} is the vector of observed characteristics, and ε_{ist} is the idiosyncratic error term. The analysis uses robust standard errors clustered by workers.

Note that the main specification includes an educational mismatch variable. I expect mismatched workers to have different incentives to acquire education on-the-job. That is because of the stylized facts in the literature: i) overeducated workers earn less than those with the same educational level, but more than their colleagues that are not overeducated and ii) undereducated workers earn more than those with the same education, but less than their co-workers (Hartog, 2000; McGuinness, 2006; Leuven and Oosterbeek, 2011). Hence, undereducated workers should be more likely to invest in education, while overeducated ones should be less likely.

The decision to acquire education on-the-job may be endogenous. Therefore, to address this problem, I will re-estimate equation (3.5) including a fixed effects estimator to account for workers' unobserved heterogeneity, c_i:

$$CEduc_{ist} = \alpha + \delta_1 MWindex_{st} + \delta_2 Mismatch_{it-1} + \delta_3 State_s + \delta_4 Year_t + \beta_1 X'_{ist} + c_i + \varepsilon_{ist} \tag{3.6}$$

[25] Again, note that workers can only upgrade their educational level.

Note that, empirical results need to be interpreted carefully due to limitations, particularly related to policy enforcement and informal labour market. Although the enforcement of the minimum wage in Brazil is highly supported by the Labour Justice system, which imposes high fines on noncompliants employers and demands compensations to workers based on the rules of the labour code, informality plays a big part in the Brazilian labour market. Because the informal labour market uses the minimum wage as a numeraire in Brazil, the formalization of the labour market that occurred during the 2000s made the minimum wage even more important for the informal sector. As Barbosa et al. (2015) show, the active minimum wage policy increased not only the employment in the formal labour market, but also raised wage levels in the informal market, even for low-productivity sectors. In the informal labour market the minimum wage is closer to the average wages than in the formal labour market, particularly in the agriculture sector.[26] Thus, one should look at how the minimum wage and average wages change within sectors. This paper only uses data on the formal labour market and the empirical results should be interpreted for the formal labour market only. I acknowledge this as a possible limitation.

3.5 Results

Table 3.2 presents the results of the linear probability regression using equations (3.5) and (3.6) for the absolute minimum wage (columns 1, 2 and 3) and the relative minimum wage (columns 4, 5 and 6).

First, focusing on the absolute minimum wage, column (1) shows a positive and significant effect on education acquisition (1.99%) moving from the smallest to the highest index. When the educational mismatch is introduced (column 2), the coefficient remains significant and positive but is now slightly smaller (1.76%). On the other hand, when workers' fixed effects are controlled for (column 3), an increase in the absolute minimum wage decreases on average the probability of acquiring more schooling (-0.56%). An increase in the absolute minimum wage is associated with a skill premium decrease, which discourages educational investments. Thus, the discouragement is bigger for states with higher real minimum wages. For example, the absolute minimum wage in São Paulo is 7.43% higher than in Maranhão in 2013. Therefore, someone living in São Paulo has a probability of acquiring education there 0.5 percentage points lower than someone living in Maranhão.[27]

Moving to the effect of the relative minimum wage, results in column (4) show that an increase in the relative minimum wage increases on average the likelihood of education

[26] As previous mentioned, agriculture was excluded of the analysis.

[27] To calculate the difference between the two states, first I checked how the changes in the minimum wage impact the index for each state using equation (3.2). Second, I calculated the impact of each state on the educational investment by multiplying the index by δ_1. Lastly, I compared both impacts.

acquisition (4.43%). Although, this result holds when the educational mismatch is taken into account (column 5), the impact is smaller, 3.06%. Following the logic of Bárány (2016), an increase of the minimum wage makes more difficult to find a job for lower ability workers and in order to avoid unemployment, the role of education rises. Therefore, in the long run, the average educational level of the workforce will increase. However, when workers' fixed effects are controlled for, the impact of the relative minimum wage is no longer significant, although the coefficient is still positive. The relative minimum wage is associated with an increase in the skill premia for higher education levels, therefore investments in education rise. For instance, consider the years 2012 and 2013 in São Paulo. The relative minimum wage is 8.65% higher in 2013 than in 2012, which means that the probability of acquiring education increases by 0.04 percentage points compared to the previous year for a worker with completed secondary education level.

Table 3.2: Effects of the minimum wage on investments in education on-the-job

Variables	**(1) OLS**	**(2) OLS**	**(3) FE**	**(4) OLS**	**(5) OLS**	**(6) FE**
Absolute minimum wage	0.0199***	0.0176***	-0.00564*	-	-	-
	(0.00345)	(0.00321)	(0.00327)	-	-	-
Relative minimum wage	-	-	-	0.0443***	0.0306***	0.00442
	-	-	-	(0.00890)	(0.00833)	(0.00827)
Educational mismatch	No	Yes	Yes	No	Yes	Yes
Individual fixed effects	No	No	Yes	No	No	Yes
Adjusted R-squared	0.061	0.176	0.463	0.061	0.176	0.463
Observations	675,401	675,401	675,401	675,401	675,401	675,401

Dependent variable: changed education.

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.

Note: All regressions include tenure, age, attained education, firm size, sector, class of worker, state and year. OLS regressions also include gender.

Figures B.5 and B.6 present the estimated impact on the 2006-2013 education acquisition on-the-job of changes in the absolute e relative minimum wages that occurred over the period. Rising the absolute minimum wage is estimated to have increased the likelihood of investing in more education by 0.27% in Piauí, Ceará and Rio Grande do Norte. Conversely, in Bahia and Sergipe there was a decrease of -0.72% on the probability of acquiring more education over the period. On the other hand, rising the relative minimum wage is estimated to have increased the probability of investing in education on-the-job by 0.72% in the Distrito Federal, while a decrease of -0.31% is estimated in Acre. Note that although the estimated coefficient is negative (positive) for the absolute (relative) minimum wage (columns 3 and 6 on Table 3.2, respectively) the overall effect depends of the variation of the state's index.[28]

[28]To calculate the implied effect for each state over the period, I first checked the effect of changes in the minimum wage on the index for each state and then I computed the effect of the variation on education acquisition using the measured coefficient of equation 3.6.

Overall, it emerges that controlling for workers' fixed effects plays an important role in the effect of the minimum wage on education acquisition on-the-job. This is quite expected, as a rise in the absolute minimum wage is likely to decrease the incentives to acquire education on the job, because the absolute minimum wage is a proxy for skill premium shrinking. On the other hand, an increase in the relative minimum wage is likely to increase the incentives to acquire education, because the relative minimum wage is a proxy for skill premium raise, in particular on top educational levels. Although the results suggest that the relative minimum wage has no significant effect on education acquisition, the coefficient is positive.

These results reflect the average effects of the two measures of the minimum wage, which may mask considerable heterogeneity across subgroups. In the next section, I try to shed some light of the effects of minimum wage on education acquisition on-the-job for different subgroups.

3.6 Subgroup analysis

3.6.1 Percentiles

So far, I have focused the analysis on the entire wage distribution. I now focus the investigation on different percentiles of the earnings distribution. I re-estimated equation (3.6) for various percentiles. Particularly, I focus on the 10th, 25th, 50th, 75th and 90th percentiles. The intuition is that workers in the lower tail of the wage distribution are most impacted by the minimum wage than workers at the upper tail, i.e. the effect would be more pronounced for whom the minimum wage is binding.

Table 3.3: Effects of the absolute minimum wage on investments in education by earnings percentiles

Variables	**(1) 10th percentile**	**(2) 25th percentile**	**(3) 50th percentile**	**(4) 75th percentile**	**(5) 90th percentile**
Absolute minimum wage	-0.0229**	-0.0278***	0.00104	0.00532	-0.00965
	(0.0105)	(0.00678)	(0.00720)	(0.00670)	(0.00820)
Educational mismatch	Yes	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.539	0.512	0.487	0.447	0.375
Observations	56,907	150,854	164,309	175,048	76,553

Dependent variable: changed education.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parenthesis clustered by individual.
Note: All regressions include tenure, age, attained education, firm size, sector, class of worker, state and year.

Table 3.3 present the effect of the absolute minimum wage in obtaining more education on-the-job for various percentiles of the wage distribution. An expected increase in the absolute minimum wage reduces the likelihood of getting more schooling on-the-job for workers at the 10th and 25th percentiles, -2.29% and -2.78% respectively. This is consistent with expectations. If the wage distribution shifts to the right, but the skill premium shrinks, then investments in education decrease. Another explanation for this result may be credit constraints. Due to serious imperfections in the credit market in Brazil, poor families are less likely to invest in human capital and are more likely to perpetuate poverty (Barros et al., 2001). Moving along the percentiles, one observes that the absolute minimum wage has no significant impact on the rest of the wage distribution. As expected, the effects of the absolute minimum wage are more pronounced at the bottom of the wage distribution.

Table 3.4: Effects of the relative minimum wage on investments in education by earnings percentiles

Variables	**(1) 10th percentile**	**(2) 25th percentile**	**(3) 50th percentile**	**(4) 75th percentile**	**(5) 90th percentile**
Relative minimum wage	-0.0139	0.0352*	0.0479**	0.0441**	-0.0478***
	(0.0307)	(0.0208)	(0.0209)	(0.0187)	(0.0174)
Educational mismatch	Yes	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.538	0.512	0.487	0.447	0.390
Observations	56,907	150,854	164,309	175,048	76,553

Dependent variable: changed education.

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.

Note: All regressions include tenure, age, attained education, firm size, sector, class of worker, state and year.

Table 3.4 shows the effects of the relative minimum wage on education acquisition on-the-job for different percentiles of the wage distribution. Column (1) indicates that an increase in the relative minimum wage does not affect the likelihood of acquiring education for workers at the very bottom of the earnings distribution. However, the effect is positive and significant (3.52%) for the 25th percentile (column 2). Intuitively, as the minimum wage increases, finding a job becomes more difficult for low-ability workers and the educational role enhances. Because of the increase in skill premia, workers have more incentive to acquire education. The effect is also positive and significant for the 50th and 75th percentiles, 4.79% and 4.41%, respectively. Contrary, at 90th percentile, the effect of an increase in the relative minimum wage is negative and significant (-4.78%). This can be explained by the fact that at the top end of the wage distribution, workers already hold a higher education or postgraduate degrees.

3.6.2 Educational mismatch

As shown above, controlling for the educational mismatch changes the effect of the minimum wage measures. Due to these stylized facts in the literature, undereducated workers should be more likely to invest in education, while overeducated ones should be less likely. That is because overeducated workers earn less than those with the same educational level, but more than well-matched co-workers. Conversely, undereducated workers earn more than workers with the same educational level but less than well-matched co-workers. Thus, Table 3.5 shows the effect of the minimum wage controlling for the educational mismatch and the interactions of educational mismatch and minimum wage. Column (1) presents the results for the absolute minimum wage and column (2) reports the results for the relative minimum wage.

Table 3.5: Effects of the minimum wage on investments by educational mismatch

Variables	**(1) Absolute MW**	**(2) Relative MW**
Minimum wage	0.00908**	-0.0455***
	(0.00425)	(0.00923)
Overeducation$_{(t-1)}$	-0.201***	-0.291***
	(0.00505)	(0.00633)
Undereducation$_{(t-1)}$	0.488***	0.434***
	(0.00588)	(0.00768)
Overeducation$_{(t-1)}$ × Minimum wage	-0.0484***	0.103***
	(0.00629)	(0.0106)
Undereducation$_{(t-1)}$ × Minimum wage	-0.00734	0.0889***
	(0.00698)	(0.0123)
Individual fixed effects	Yes	Yes
Adjusted R-squared	0.463	0.463
Observations	675,401	675,401

Dependent variable: changed education.
*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.
Note: All regressions include tenure, age, attained education, firm size, sector, class of worker, state and year.

Table 3.5 shows the results, which confirm expectations: workers who were overeducated at the previous year are less likely to acquire more education, -20.1% and -29.1% in columns (1) and (2) respectively. On the other hand, those who were previously undereducated are more likely to acquire more education, 48.8% in column (1) and 43.4% in column (2). Interacting the absolute minimum wage and overeducation indicates that an increase in the absolute minimum wage decreases, even more, the probability to acquire more schooling for those workers who were overeducated. For previously undereducated workers, although the interaction is negative, the coefficient is not significant at the usual levels. These results confirm the results above: as the skill premia decline (higher absolute minimum wage), investments in education falls. On the other hand, interacting the relative minimum wage and educational mismatch shows that an increase in the relative minimum wage also increases the likelihood for previously overeducated workers by 10.3%. Although the interaction is positive the overall effect on overeducated workers remains negative. Lastly, an increase in the relative minimum wage boosts, even more, the likelihood of investing in more education for workers who were undereducated (8.89%). These results also are in accordance with previous findings: as the skill premium increases (higher relative minimum wage) investments in education rise.

3.6.3 Gender differences

The next analysis consists in separating the sample by gender. Considering that the goal of the minimum wage is to redistribute income to the bottom of the wage distribution, one should be concerned with the impact of the minimum wage according to different characteristics, such as gender.[29] Female workers are more likely to be low wages workers, and as shown in Figures B.3 and B.2, wages are more concentrated at the lower tail of the distribution for them. Therefore, one should expect female workers to be more affected by changes in the minimum wage. Due to the difference in the wage distribution for males and females, I re-estimated equation (3.6) for females and males separately. Results are presented in Table 3.6.

An increase in the absolute minimum wage has a negative and significant effect on female workers, -1.62% (column 1), but it has no significant effect on male ones (column 2). As expected, the shrink in the skill premium (higher real minimum wage) has a greater impact on female workers, who are more likely to be at the bottom part of the earnings distribution. Moreover, women are already more educated than men (Table B.4), which may explain why they less likely to acquire education when there is a compression on the skill premium.

Conversely, an increase in the relative minimum wage does not affect either females or males workers. Although the effect is not significant, the coefficient is negative for female and positive for males. It seems driven by different parts of the wage distribution (Table 3.4), in which female workers predominate at the very bottom, and male ones prevail at the rest of the distribution.

[29] Abowd et al. (1999) analyse the effect of the minimum wage on employment by age and gender.

Table 3.6: Effects of the minimum wage on investments by gender

Variables	(1) Female	(2) Male	(3) Female	(4) Male
Absolute real minimum wage	-0.0162***	0.00180	-	-
	(0.00498)	(0.00434)	-	-
Relative minimum wage	-	-	-0.00727	0.0159
	-	-	(0.0123)	(0.0112)
Educational mismatch	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes
Adjusted R-squared	0.429	0.483	0.429	0.483
Observations	270,739	404,662	270,739	404,662

Dependent variable: changed education.
*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.
Note: All regressions include tenure, age, attained education, firm size, sector, class of worker, state and year.

As figures B.2 and B.3 show, male earnings move quicker than female earnings. These differences may stem from the fact that women are more concentrated in the services sector than men (Table B.4). Moreover, women tend to be less mobile in the labour market compared to men: the share of women that changed firm over the period is 2 percentage points smaller than the share of men (15.95% and 17.94%, respectively) and women stay on average 76 months on the same firm, while men stay 67 months. Thus, the difference by gender may be due to females selecting jobs at the lower part of the wage distribution. Higher minimum wage tend to affect more the lower part of the wage distribution, where women are more concentrated. If the minimum wage compresses the skill premium, then women will be less likely to invest in more education. The increase in the minimum wage may also lead to a decrease in education acquisition due to higher opportunity costs. For instance, according to Hara (2017) if the worker can only participate in training by forgoing work, then the investments on general training would decrease.

3.7 Robustness

In this section, I perform robustness checks in order to test the sensitivity of the results to changes in the specification.

3.7.1 Smaller sample

First, I test whether an alternative sample affects the results. To do so, I repeat the estimation using only those workers who are observed all the years, that is, 8 times.

Table 3.7 shows that the results for the restricted sample are similar to those presented in Table 3.2. The coefficients are slightly larger in the smaller sample compared to the full sample. The bigger effect size may be because in the balanced panel, one looks at the effects of changes in the minimum wage (absolute and relative) for workers who do not leave the formal labour market during the period. Moreover, the individuals who left the panel are on average less educated than those who were observed all years. So, although the direction and significance of the coefficients are similar in both samples, it is likely that the magnitude differs due to selection bias and more investigation is needed to properly compare both samples.

Table 3.7: Effects of the minimum wage on investments in education on-the-job - Smaller sample

Variables	**(1) OLS**	**(2) OLS**	**(3) FE**	**(4) OLS**	**(5) OLS**	**(6) FE**
Absolute minimum wage	0.0240***	0.0205***	-0.00742*	-	-	-
	(0.00426)	(0.00400)	(0.00394)	-	-	-
Relative minimum wage	-	-	-	0.0749***	0.0684***	0.00962
	-	-	-	(0.0100)	(0.00944)	(0.00923)
Educational mismatch	No	Yes	Yes	No	Yes	Yes
Individual fixed effects	No	No	Yes	No	No	Yes
Adjusted R-squared	0.043	0.122	0.362	0.043	0.122	0.362
Observations	314,966	314,966	314,966	314,966	314,966	314,966

Dependent variable: changed education.
*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.
Note: All regressions include tenure, age, attained education, firm size, sector, class of worker, state and year. OLS regressions also include gender.

3.7.2 Minimum wage lagged

So far, the analysis has used the minimum wage in time t. Table 3.8 presents the results using the lag of the minimum wage indexes. Overall, the results are in line with those presented in Table 3.2.

Starting with the absolute minimum wage, without controlling for workers' fixed effects, the coefficient is positive (columns 1 and 2), although they are less significant than using the contemporaneous absolute minimum wage (Table 3.2). When workers' fixed effects are controlled for, the coefficient becomes negative and significant. As previously discussed, the negative coefficient suggests that an increase in the absolute minimum wage decreases the investments in education on-the-job, as the absolute minimum wage is associated with decreases in the skill premium.

Table 3.8: Lag effects of the minimum wage on investments in education on-the-job

Variables	**(1) OLS**	**(2) OLS**	**(3) FE**	**(4) OLS**	**(5) OLS**	**(6) FE**
Absolute minimum wage lagged	0.00610	0.00607*	-0.0145***	-	-	-
	(0.00395)	(0.00362)	(0.00343)	-	-	-
Relative minimum wage lagged	-	-	-	0.0474***	0.0300***	0.00308
	-	-	-	(0.0100)	(0.00917)	(0.00881)
Educational mismatch	No	Yes	Yes	No	Yes	Yes
Individual fixed effects	No	No	Yes	No	No	Yes
Adjusted R-squared	0.057	0.165	0.472	0.057	0.165	0.472
Observations	552,220	552,220	552,220	552,220	552,220	552,220

Dependent variable: changed education.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parenthesis clustered by individual.
Note: All regressions include tenure, age, attained education, firm size, sector, class of worker, state and year. OLS regressions also include gender.

The results for the relative minimum wage are similar to those presented in Table 3.2. Without controlling for workers' fixed effects the relative minimum wage presents a positive and significant coefficient (columns 1 and 2). But, when individuals' fixed effects are controlled for, the impact of the relative minimum wage is not significant, although remains positive. In addition, the magnitude of the coefficients does not change compared to the results using the contemporaneous relative minimum wage (Table 3.2).

3.8 Conclusion

This paper provides empirical evidence on the effects of minimum wage changes on the skill premium and hence on investments in education on the job. Using a unique dataset from Brazil I proceed in two steps: first I analyse how changes in the absolute and relative minimum wages affect the skill premium and therefore the wage distribution. Second, I measure the effects of the two variables on educational investment decisions.

The results show that the absolute minimum wage is associated with compression of skill premia in particular for higher education. Conversely, the relative minimum wage is associated with an increase in the skill premia for higher education, but lower and intermediate education premia seem unaffected. Moreover, the absolute minimum wage has a negative effect on educational investments. The impact is more pronounced at the very bottom of the wage distribution (10th and 25th percentiles). By contrast, the relative minimum wage has no significant effect on educational investments on average. Nonetheless, the effect is positive at the 25th, 50th and 75th percentiles of the wage distribution. Additionally, the higher absolute minimum wage has a stronger effect on overeducated and female workers, which are less likely to invest in education acquisition.

Conversely, higher relative minimum wage impacts positively over and undereducated workers, but it is not significant for females nor males.

From the policy perspective, the effects of the minimum wage on the skill premium and the education acquisition are relevant for policy makers in order to reduce inequality. The analysis reveals that an increase of the minimum wage affect inequality though changes in the skill premium. Workers at the bottom of the wage distribution are negatively affected through education incentives. Thus, minimum wage policies that erode the skill premia, in particular for secondary education and above, will decrease incentives for education acquisition. Contrary, minimum wage policies that increase the skill premium, will raise the incentives to invest in education. Therefore, if the absolute minimum wage increases less than the relative minimum wage, then the results suggest that the skill premium increases for higher education levels, which impacts positively investments in education on-the-job. Thus, policies aiming to increase the relative minimum wage over the absolute minimum wage will maximise the impact of the minimum wage. The minimum wage policy also affects inequality the top end of the wage distribution, in a smaller extent, due changes in the skill premium and the possible change in the workforce composition. However, Brazilian states are heterogeneous and thus a decentralised minimum wage policy may be more effective to reduce the gap among them. Moreover, the active minimum policy alone may not be sufficient to reduce inequality, particularly for those workers at the bottom of the distribution who are more affected, and it is important that an active minimum wage and the expansion of the educational system are followed by increases in job opportunities, so workers can fully explore their productivity on the job (acquired through formal education). Otherwise, the misallocation of resources leads to a decrease in economic development.

Chapter 4

Internal migration driven by violence shocks and its effects on job transitions

Abstract. This paper studies the labour market impact of internal migration in Brazil by instrumenting migration with violent crime shocks at the origin state. I find that people are more likely to move away from their original state after positive shocks in violent crimes. Moreover, without controlling for self-selection, migrants are positively selected, and earn more than non-migrants. Conversely, after controlling for self-selection, I find that on average, migrants are negatively selected, and earn less than non-migrants. The effect is more pronounced on men and low-educated workers. In addition, I find that migrants who came from richer states are more positively selected than migrants who came from poorer states. Thus, these results may indicate forced migration only for poorer origin states.[1]

4.1 Introduction

Individuals decision to move to another area is driven by different reasons, such as economics, social and environmental factors. Moreover, the migration decision is followed by the decision to change jobs. However, it is not clear what is the effect of internal migration on moving to a better job, i.e. high-wage jobs, high-status jobs, and education-job match. Or to what extent other characteristics such as individuals' abilities or regions characteristics affect returns to migration. Most of the literature focuses on measuring

[1] I would like to thank Thomas Gall, Jackline Wahba and Michael Vlassopoulos for valuable comments and suggestions on earlier drafts of this paper. I gratefully acknowledge financial support from the *ESRC*.

the impact of migration on wages, but only a few studies analyse other job quality features (e.g. Venhorst and Cörvers, 2018).

However, the migration decision is not always voluntary. Involuntary migration takes place due to factors that are outside the control of the migrants, such as wars and natural disasters. This kind of migration has been receiving more attention due to recent conflicts, e.g. the Syrian civil war and the refugee crisis. Potential effects of forced migration may differ from voluntary migration at the region level, and at the individual level. Nevertheless, the difference between voluntary and involuntary migration is not always straightforward. Even during periods of war, people may choose not to move, even if the risk of staying is too high, e.g. death (Dustmann et al., 2017; Becker and Ferrara, 2019).

This paper examines the effects of internal migration on job match transitions, focusing on young people who are entering the labour market. However, measuring the effects of internal migration faces an identification obstacle. The main concern is that the migration decision is endogenous, therefore OLS estimates are likely to be biased and do not provide causal inference. To address this issue I use an instrumental variable approach (IV). Using a matched employer-employee dataset from the Brazilian formal labour market, I document the migration decision of individuals in Brazil for 8 years. Then to generate exogenous variation I use violent crime shocks at the origin state. Intuitively, the migration decision is endogenous due to self-selection. In turn, higher criminality rates at origin are correlated to one's decision to move and are not associated with the job transitions outcomes at the destination, making it a good candidate for instrument. The identification assumption underlying the estimation is that the criminality in the origin state does not affect the labour market outcomes at the destination other than through migration.

The results show a modest positive wage effect of migration when using the naïve model. Moreover, migrants are more likely to move to an education well-matched position, but also they are more likely to move down on the rank of occupations, that is, the occupation status. On the contrary, when the endogeneity is taken into account, the IV estimates show a negative wage effect of internal migration, and changing to a higher/lower rank and to an education well-matched occupation is no longer significant. However, after splitting the sample I find heterogeneous results, with a stronger effect on men and low-educated workers. Additionally, I find that migrants who came from richer states are more positively selected than migrants who came from poorer states. Thus, these results may suggest forced migration only for migrants who came from poorer states. Moreover, migrants induced by violence shocks may be different from other migrants. For instance, I find that LATE-complying migrants on average, earn less, are less likely to be highly educated and more likely to come from poorer regions.

Brazil is an ideal test case for several reasons. First, the Brazilian labour market is defined by high flows of jobs and workers, and it is among the highest rates of job turnover in the world (Gonzaga et al., 2003; Rocha et al., 2018). Moreover, Brazil is a diverse country characterized by great income inequality and these differences among regions drive internal migration (Santos Júnior et al., 2005; Ramalho and Queiroz, 2011). For instance, 14.5% of the population lives in a place other than their birthplace (IBGE, 2010), and around 5% of the formal workforce has moved to a different state (Aguayo-Tellez et al., 2010). Thus, internal migration could help to eliminate regional inequalities, through changes in the labour market dynamics (Ribeiro, 2010; Mendes et al., 2017). Second, Brazil is one of the most violent countries in the world, with one the highest homicide rates, according to the United Nations Office on Drugs and Crime (UNODC, 2013). Additionally, there is ample variation of the homicide rate across states. Over the last decade, the homicide rate in the state of São Paulo decreased by 56.7%, meanwhile, it increased by 256.9% in the state of Rio Grande do Norte (IPEA, 2018).[2] Moreover, homicides are concentrated in poor areas and are the main cause of youth mortality, especially for men. It accounts for 50.3% of the death of men between 15 and 29 years old (IPEA, 2018).

To the best of my knowledge, this is the first paper to study migration driven by violent crimes and its effects on job match quality, other than wages, in a developing country. This paper contributes to the literature in different ways. First, I provide credible estimates of the effect of migration on job match transitions. In contrast with the traditional literature, discussed later in this paper, I find evidence of negative selection on migrants, suggesting forced migration. Second, there is little evidence on the effect of crime on migration decision in developing countries with high homicide rate (e.g. Calderón-Mejía and Ibáñez, 2016). Investigating crime is relevant because high criminality is associated with socioeconomic costs, such as a negative impact on health and human capital (Monteiro and Rocha, 2017).

The rest of this paper is organised as follows: Section 4.2 provides some background; Section 4.3 describes the data and descriptives; Section 4.4 presents the empirical strategy; Section 4.5 shows the results; Section 4.6 presents robustness checks, and Section 4.7 concludes.

4.2 Background

4.2.1 Returns to migration

As the traditional literature points out, migrants move to regions with higher income levels (Sjaastad, 1962; Chiswick, 1978; Borjas, 1985; Borjas et al., 1992). According

[2] Figure C.1 presents a geopolitical map of Brazil. It contains 27 states divided into 5 regions.

to the wage differentials theory, people look for compensation for the investment in migration by comparing expected returns at the destination and expected returns in case they do not move. Hence, the costs of migration are inversely proportional to the decision to move. However, income is not the only rewarding outcome in the migration literature. In addition, returns to migration can be non-monetary and are related to the individual's preferences. For instance, Huttunen et al. (2018) find that workers are less likely to move away from where their family live, and migration is also related to family-forming decisions, such as divorce. Enrico (2011) findings suggest that individuals' decision to move is related to local amenities, such as cultural events or clean air. Also, migrants tend to avoid regions with high unemployment rates (Herzog Jr et al., 1993; Huttunen et al., 2018). Similarly, people avoid regions with higher rates of criminality (Morrison, 1993; Cullen and Levitt, 1999; Sousa, 2014; Chairassamee, 2018). Moreover, migration can help workers finding an education job-match as decreases the probability of overeducation (Büchel and Van Ham, 2003; Ramos and Sanromá, 2013; Iammarino and Marinelli, 2015; Waldorf and Yun, 2016).

Selection of migrants have been largely studied in the literature since Borjas (1987) to test empirically how migrants are different from non-migrants. Most of them rely on the Roy model of self-selection among countries. The model predicts that as long the skills are transferable across countries, the sorting is determined by international differences in the return to skills. Chiquiar and Hanson (2005) pioneer analysis show that skill-varying migration costs describes better the flow of migrants from Mexico to the United States. Moraga (2011), Borjas (2014) and Kaestner and Malamud (2014) conclude that negative selection can be due to unobserved characteristics that determine migrants' earnings. Contrary to most of the literature, that studies migration from poor to rich countries, Borjas et al. (2018) analyse emigration from a wealthy country, Denmark, to other countries. Other papers focus on the Roy model applied to internal migration (Borjas et al., 1992; Abramitzky, 2009; Bartolucci et al., 2018). For instance, Abramitzky (2009) analysing the internal migration in Israel find that migrants' self-selection depends on returns to skills in origin and destination.

The human capital theory predicts that workers will move if the returns on migration exceed the costs. The costs depend on the labour market conditions, family size among others. Following the standard framework, more able workers are more likely to move as one expects higher returns for them on the job searching. Besides, if more able workers have an advantage before moving, this could help them decreasing the costs of migration and increasing the likelihood of higher income. Therefore, if one expects migrants to be more able than non-migrants, failing to control for unobservable characteristics may lead to biased estimation of the returns on migration (Gabriel and Schmitz, 1995).

Another strand in the literature, related to the demand side, argues that if employers have information advantage on one's skills, they will be able to get the best local workers in the labour market, who are less expensive to train (Thurow, 1975). Therefore, the residual workers, who are less skilled, are compelled to leave the region, which may result in lower wages rather than higher wages for movers. These groups of migrants are known in the literature as "forced migrants" (Smits, 2001).

Empirical studies control for self-selection among migrants in several ways. Borjas et al. (1992) analyse the internal migration in the United States and find that migrants choose their destination based on the potential reward for their skills; i.e. high-skilled migrants tend to move to places where the skill premium is high compared to the local average wage, whereas low-skilled migrants tend to move to regions where the skill premium is low. Abramitzky et al. (2012) also show that selection is related to returns to skills, and migrants move to local labour markets that value their skills. That is, local labour markets with lower returns to skills attract low-skilled workers, while high-skilled workers tend to move to local labour markets with higher returns to their skills.

Some studies follow the Heckman (1979) selection model to estimate which part of the wage difference between migrants and non-migrants is due to migration itself. For example, Smits (2001) finds positive returns to migration for married men and women in the Netherlands without controlling for self-selection. However, after controlling for self-selection, the returns are negative for both groups. The author suggests that before moving migrants were in a relatively less favourable labour market situation than the non-migrants indicating that forced migration could play a role on returns. Yankow (2003) studies the return of migration for workers who changed jobs and finds that the returns for skilled migrants became positive after two years. On the other hand, the return is immediate for low-skilled migrants who also changed jobs. Détang-Dessendre et al. (2004) analyse young migrants in France and find no selection for loweducated migrants and positive selection for highly educated migrants. The effect is bigger for those who migrate from small labour market areas to bigger labour markets.

Another approach is the use of treatment effects models to estimate the effects of migration on wages controlling for self-selection. For instance, Nakosteen et al. (2008) use this method to separate observable and unobservable migrant characteristics. Their findings suggest that there is self-selection based on unobservable traits for men and women in Sweden. Additionally, they find that for women there is also evidence of selection based on measured income before moving, i.e. women with higher earnings are less likely to move than those with lower earnings. Using an instrumental variable approach, Venhorst and Cörvers (2018) calculate the impact of internal migration on job quality for university and college graduates in the Netherlands. The authors find a positive wage return, but it disappears after controlling for self-selection. Also, when the self-selection is controlled for, the impact on job quality decreases or is even negative, suggesting forced migration.

As one notices, the returns on mobility can be positive or negative, depending on the self-selection of the migrants. Moreover, it is important to correct for self-selection in order to get unbiased estimates. Considering that in the empirical analyses, I focus on workers who are at the beginning of their career and then they are more likely to move.

4.2.2 Involuntary migration

Contrary to voluntary migration, involuntary migration occurs when people have a very limited choice other than move. Particularly, two main factors are driving this kind of migration: large conflicts and natural disasters (Becker and Ferrara, 2019).

As previously discussed, the main concern in the migration literature is related to self-selection. In this context, usually forced migration is considered exogenous and its impacts are interpreted as causal effects. Some studies use an instrumental variable approach to analyse the impact of involuntary migration. For instance, Akgündüz et al. (2018) use the distance to the closest border crossing as an IV to capture the effects of Syrian refugees inflows into Turkey. Regarding internal migration, Morales (2018) uses an enclave IV exploiting migration driven by violence and past waves of migrants to estimate the effect of internal displacement on wages of local residents at the destination.

The literature is particularly concern with the effects of involuntary migration in developing countries, especially because of civil wars. Ruiz and Vargas-Silva (2013) review the literature on the impacts of forced migration and highlight that in developing countries, the impacts for those people forced to move are more severe, such as worse labour market outcomes and less consumption smoothing. Thus, forced migrants may end up in a suboptimal location, which may have consequences on their outcomes.

However, not necessarily all individuals move. In the case of people leaving their home region due to violence, the literature considers the threat to one's life an additional push factor. For instance, individuals may have a different perception of the threat, and therefore some people may choose not to move, even if staying is too risky (see Engel and Ibáñez, 2007; Ibáñez and Vélez, 2008). Shrestha (2017) analyses push and pull factors using data from Nepal. The author finds that an increase in the death rate increases the emigration rate.

As indicated by Ruiz and Vargas-Silva (2018), it is important to differentiate between the effects of voluntary and involuntary migration, as migrants in these groups may be very different. Although the literature on forced migration has been developing quickly over the last few years, it is still in early stages, particularly for developing countries due to the lack of data. Moreover, as pointed by Becker and Ferrara (2019) future research is needed to understand the factors driving migration that is not entirely voluntary.[3]

In this context, I analyse the effects of positive shocks on violence on individuals' decision to move and its effects on labour market transitions.

[3]See Becker and Ferrara (2019) for a complete survey of the consequences of forced migration over the last 10 years.

4.3 Data and descriptives

4.3.1 Labour market data

The sample stems from the Annual Social Information Report (RAIS), which an administrative dataset from the Brazilian Ministry of Labour and Employment (MTE). It covers the Brazilian labour market since all tax registered firms have to report their information by law. The data contains a unique time-invariant worker identifier, and firm identifier, which allows following employers and employees over time. The empirical analysis uses a 1% representative random sample panel of workers selected in 2006 and followed until 2013. All workers selected were starting their first job in 2006. Overall this paper follows 215,572 workers from 18 to 27 years old (in 2006).[4]

To determine if a worker is a migrant I use the information on the state one works. The most used definition of migration in the literature defines a migrant as someone who lives at a different place than his/her birthplace (UNDP, 2009). Unfortunately, I do not have information on workers birthplace, so I consider that their home state is the state where they get their first job.[5] Hence, a migrant is defined as a person whose state in which she/he works in period t differs from her/his state in $t-1$. Once the individual changes from the state of origin, she/he is considered a migrant.[6] Overall, 9.14% of the sample moved states over the period.[7] Table 4.1 provides summary statistics of this dataset.

The main labour market variables in this study are income and job match transition measures: (i) changed occupation, (ii) changed the rank of occupations, and (iii) changed to a well-matched (education) occupation. Income is recorded as the logarithm of the hourly wage in Brazilian Real (BRL) and is deflated using the Brazilian consumer price index (IPCA) with the base year 2006. Changed occupation is a dummy variable indicating whether or not the worker changed occupation compared to the previous period.

In ranking occupational families I use the average real wage for each occupational family for each state.[8] Then I calculate the quintiles of the occupational families in each state. Using the quintiles, I define high-rank occupations as those in the top quintile, and low-rank occupations as those at the bottom quintile. Finally, I construct a dummy variable, namely changed to higher (lower) rank, that takes value 1 if a worker moved to a higher (lower) quintile compared to the previous period, 0 otherwise.[9]

[4] The cut off age for men is 25 years old in 2006 and for women is 27 years old.

[5] Note that one's birth state is not necessarily the same state where she/he got his first job. I acknowledge this as a possible limitation.

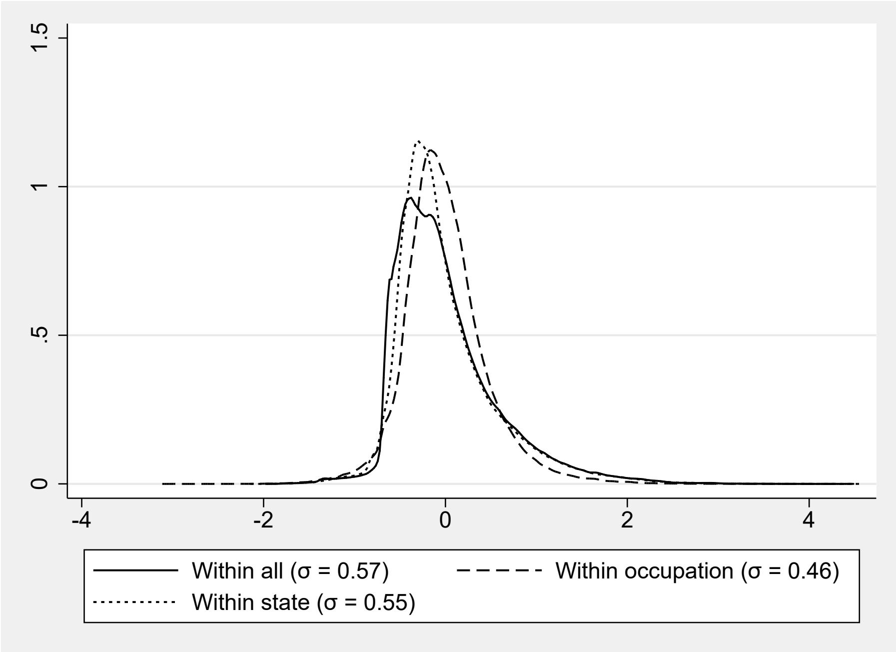
[6] Table C.10 presents the frequency of migrants at origin and destination states.

[7] Note that the majority of migrants moved only once in the period and around 0.65% of migrants moved twice.

[8] As previously explained in Chapter 1, an occupational family consists of a set of similar occupations.

[9] I follow a similar approach used by Haltiwanger et al. (2018) in ranking firms by wage.

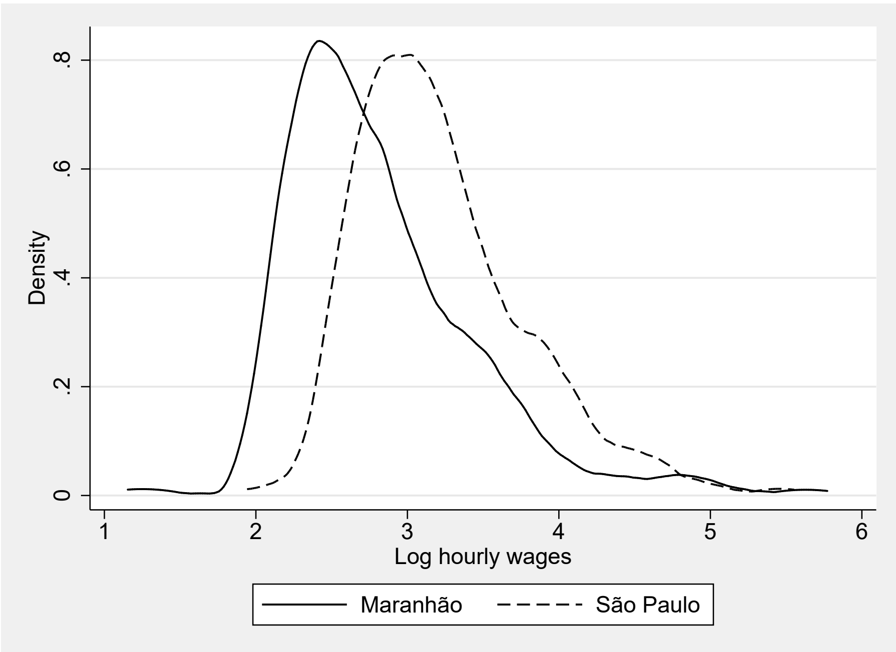
Figure 4.1: Wage density in Brazil, 2006-2013



Source: RAIS, 2006-2013.

Note: Kernel density for within-group wage dispersion of logarithm mean hourly wages. Standard deviations in parenthesis.

Figure 4.2: Earnings distribution - Occupational families

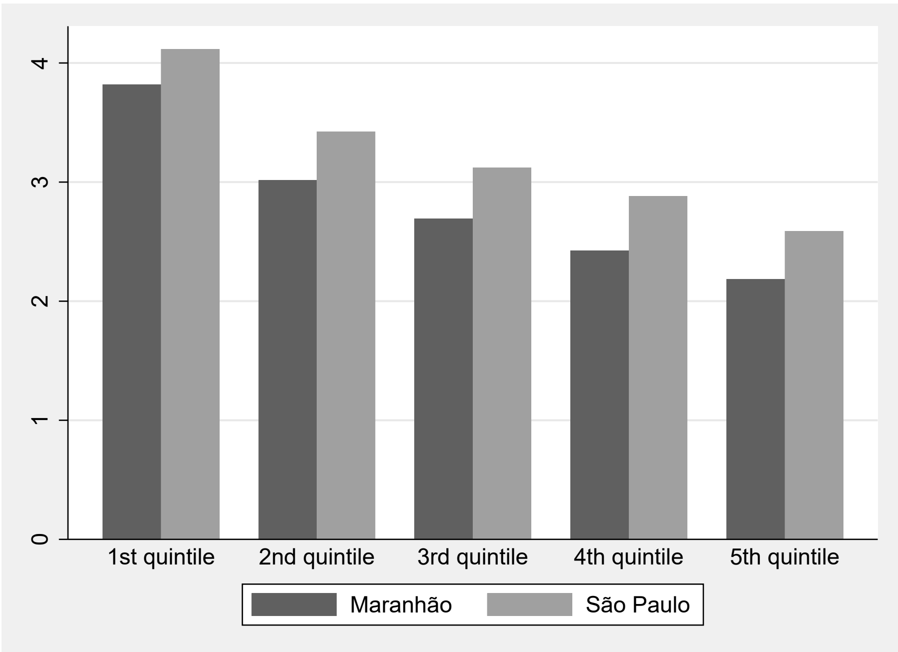


Source: RAIS, 2006-2013.

Figure 4.1 shows the kernel density for the wage dispersion within all workers across states and occupational families (solid line), within occupational families across states (dashed line), and within states across occupational families (dotted line). Note that the wage dispersion conditional on the occupational family is more compressed than the other two distributions. According to the standard deviation, the wage difference conditional on the occupational families is 19% lower compared to the unconditional wage distribution. Thus, part of the wage difference disappears conditional on the occupation, which suggests the relevance of occupations in the Brazilian labour market.

The wage gap between states in Brazil is large, and this heterogeneity in the states may reflect in the ranks. For instance, consider two extremes states: Maranhão, one of the poorest states in Brazil, and São Paulo, one of the richest. Figure 4.2 depicts the earnings distribution of the occupational families in these two states. The wage distribution is more concentrated in the left tail in the state of Maranhão. Moreover, in the state of São Paulo, the average wages are larger conditional on the occupational families, as the wage distribution is shifted to the right compared to the distribution of the state of Maranhão. Hence, the state of São Paulo presents higher average wages by quintile (see Figure 4.3). Therefore, it may be the case that a worker moving from Maranhão to São Paulo may earn more, but move to a lower rank occupation, and vice versa.[10]

Figure 4.3: Average wage by quintile



Source: RAIS, 2006-2013.

Lastly, changed to a well-matched (education) occupation is a dummy variable indicating whether or not the worker changed to an occupation at his/her corresponding level of education compared to the previous period. In addition, I observe one's age, tenure, gender, education, firm size, sector, occupational family, state and year.

Comparing migrants and non-migrants in Table 4.1 shows that internal migrants in Brazil earn around 32% more than non-migrants. Also, migrants are 18 percentage points more

[10]Tables C.1 and C.2 illustrate the top ranked and worst ranked occupations in the state of Maranhão and São Paulo.

likely than non-migrants to change occupation (49% and 31%, respectively). The share of female among migrants is 9.55 percentage points lower than non-migrants (39.07% and 48.62%, respectively), and the share of low-educated (below complete secondary school) is 7.59 percentage points higher among non-migrants.

Table 4.1: Summary statistics

Variable	All	Migrant	Non-migrant	t-statistics of mean comparison test	p-value
Wage/hour (BRL)	4.07	5.22	3.96	-78.58	0.0000
	(4.46)	(6.71)	(4.14)		
Tenure (months)	22.35	14.00	23.19	124.58	0.0000
	(20.51)	(15.49)	(20.76)		
Age	24.95	25.93	24.85	-93.09	0.0000
	(3.20)	(2.97)	(3.21)		
Share of female (%)	47.75	39.07	48.62	52.79	0.0000
Share that changed occupation (%)	32.78	49.15	31.14	110.00	0.0000
Education (%)					
Less than Fund Educ II	5.24	3.58	5.41	22.70	0.0000
Complete Fund Educ II	6.92	4.99	7.11	23.12	0.0000
Inc Secondary School	7.63	4.33	7.97	37.76	0.0000
Complete Secondary School	60.49	63.02	60.23	-15.72	0.0000
Inc Higher Educ	6.96	7.55	6.90	-7.08	0.0000
Complete Higher Educ	12.56	16.26	12.19	-33.89	0.0000
Master/Doctorate	0.19	0.27	0.18	-5.38	0.0000
Educational mismatch (%)					
Undereducated	8.68	5.17	9.04	37.95	0.0000
Well-matched	35.55	36.47	35.47	-5.67	0.0000
Overeducated	55.73	58.37	55.49	-15.90	0.0000
Region (%)					
North	6.71	6.44	6.74	3.31	0.0009
Northeast	18.48	19.72	18.36	-9.69	0.0000
Midwest	8.09	13.43	7.55	-59.51	0.0000
Southeast	52.05	45.26	52.73	41.25	0.0000
South	14.67	15.15	14.62	-4.12	0.0000
Firm size (%)					
Up to 19 employees	31.61	22.88	32.50	57.11	0.0000
20-99 employees	22.33	21.27	22.44	7.74	0.0000
100-499 employees	19.50	22.98	19.16	-26.64	0.0000
500-more employees	26.55	32.87	25.91	-43.48	0.0000
Sector (%)					
Industry	53.91	49.54	54.35	26.66	0.0000
Services	46.09	50.46	45.65	-26.65	0.0000

Source: RAIS, 2006-2013.
Notes: Mean is shown with standard deviation in parenthesis. Alternatively, frequency is shown when indicated. Wage/hour is measured in Brazilian Real (BRL).

4.3.2 Homicide data

Homicides rates are strongly correlated with other crimes in Brazil and can be considered as a lower bound estimate of overall crime, as demonstrated by Dix-Carneiro et al. (2018).

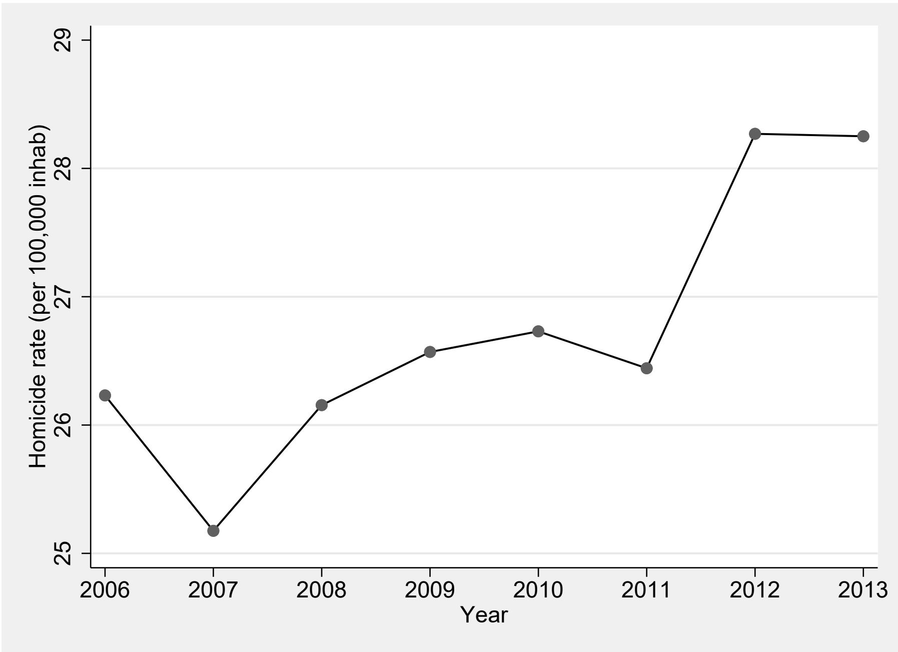
The homicide dataset comes from the Brazil Mortality Information System (SIM) from the Brazilian Ministry of Health. The data were extracted from the Department of the Public Health Care System Information (DATASUS). Homicides are defined as death registered with the codes X85 to Y09 in the International Classification of Diseases 10th Edition (ICD-10) from the World Health Organization (WHO). This corresponds to the coding of violent deaths involving aggression. This is the most reliable source of information on homicides in Brazil, despite the unregistered deaths are still problematic (Cerqueira, 2012; Murray et al., 2013).

The homicide rate H_{st} is the number of homicides per 100,000 inhabitants living in a state s at time t:

$$H_{st} = Homicides_{st} \times 100,000/Population_{st} \tag{4.1}$$

I use the homicide rate relative to population and not the absolute number of homicides because I expect the effect of criminality to be different if 1,000 people are murdered in a state of 1,000,000 people compared to one with 10,000,000 people.

Figure 4.4: Evolution of homicide rate in Brazil, 2006-2013



Source: DATASUS, 2006-2013.

Figure 4.4 shows the homicide rate in Brazil over the years. Although there are decreases in the homicide rate in 2007, 2011 and 2013, overall the period there is an increase. Brazil is among the countries with the highest homicide rates in the world. Its rate is almost 5 times the global average homicide rate, which is 6.2 in 2012 (UNODC, 2013). Despite the high level, the national homicide rate has changed little over the years. However, it presents significant variation between and within states. Figures C.2 and C.3 show the variation I will be exploring in the paper. As one can see the difference between the lowest and the biggest homicide rate is around 41 in 2006 and 54 in 2013. Within the states, there have been significant changes. For example, while the states of São Paulo (SP) and Rio de Janeiro (RJ) exhibited a decrease in their homicide rate from 2006 to 2013, the states of Ceará (CE) and Goiás (GO) presented an increase at the same period. Moreover, note that the states with the highest homicide rates are mostly located in the north or in the northeast regions, which are the poorest regions in the country.

4.4 Empirical strategy

To test the effect of internal migration on job match transitions, consider the following model:

$$Y_{ist} = \alpha + \delta Migrant_{ist} + X_{ist}\beta + \mu_i + \gamma_s + \theta_t + \varepsilon_{ist} \tag{4.2}$$

where Y_{ist} is the outcome of interest of individual i, in state s at time t. I use this specification to model the following scenarios: (i) the logarithm of the hourly wage (Mincerian wage equation); (ii) changed occupation; (iii) changed to a higher rank; (iv) changed to a lower rank; (v) changed to a well-matched (education) occupation. $Migrant_{ist}$ is a dummy variable indicating whether the worker is a migrant, thus δ is the coefficient of interest in this model. Vector X contains the control variables, μ_i, γ_s and θ_t are worker, state and year fixed effects. ε_{ist} is an idiosyncratic error term. The analysis uses robust standard errors clustered at the individual level.[11]

For the sake of comparison, I implement a naïve model using equation (4.2). However, this estimation does not provide causal evidence of the effect of migration on job match transitions. In particular, these estimates are likely to be biased due to an endogeneity problem. This is because the migration process is not random, i.e. workers who are more (less) likely to take advantage of moving may be more (less) likely to move. Thus, to deal with this problem I use an IV approach.

In order to be valid, the instrument has to be strongly correlated with the endogenous variable, but be uncorrelated with the error term, and therefore uncorrelated with the dependent variable (Angrist and Pischke, 2009). In other words, the IV should influence the outcome only through its effect on migration. In this paper, I use the deviation of the homicide rate at the state of origin from the average homicide rate of the 3 most violent states to create exogenous variation in the migration decision.

[11] I follow Abadie et al. (2017) to cluster the standard errors. The authors state that in a fixed effects regression (with fixed effects at the level of the relevant clusters), one should cluster if there is heterogeneity effects.

Therefore, to obtain an unbiased effect of migration, the two-stage regression with the IV and fixed effects is implemented. The intuition behind the first stage is the following: the higher the crime rates are in the state of origin, I expect people to move away from violence. The migration choice for individual i at time t is determined as:

$$Migrant_{ist} = \pi \Delta H_{ot-1} + X_{ist}\varphi + \mu_i + \gamma_s + \theta_t + \upsilon_{ist} \tag{4.3}$$

where ΔH_{ot-1} is the deviation of the homicide rate in the state of origin from the average homicide rate of the 3 most violent states at $t-1$. The instrument is constructed in this way to highlight the perception of violence.[12] That is, workers may perceive a state as unsafe if there is a shock in its homicide rate and it is greater than the average of the 3 most violent states. Formally:

$$\Delta H_{ot-1} = lnH_{ot-1} - lnH_{mt-1} \tag{4.4}$$

where lnH_{ot-1} is the logarithm of the homicide rate in the state of origin o at $t-1$ and lnH_{mt-1} is the logarithm of the average homicide rate in the 3 most violent states at $t-1$.

Finally, I use a linear approach to model all scenarios. Although the dependent variable is a binary variable in scenarios (ii)-(v), the linear model is used because the non-linear models do not allow the inclusion of fixed effects on an instrument variable estimation (Antman, 2011). Besides, there are other advantages in using a linear model: (i) coefficients are directly interpreted as marginal effects; (ii) it is easy to estimate robust clustered standard errors to deal with the heteroskedasticity issue.

4.4.1 Validity of the instrument

The endogeneity problem arising in equation (4.2) is due to self-selection. Therefore, to properly identify the effect of migration I need an instrument that is uncorrelated to the job match transitions but is highly correlated to the migration decision. I argue that the homicide rate in the state of origin fits both conditions.

First, I focus on the exclusion hypothesis. It is straightforward that the deviation in the homicide rate in the state of origin does not impact the outcomes in the destination, and thus the exclusion condition holds. Moreover, this paper uses homicide rates in the

[12] Engel and Ibáñez (2007); Ibáñez and Vélez (2008) model the migration choice taking into account individuals' perception of the threat.

previous period. That is because potential migrants can observe the homicides in $t-1$. To reinforce the exogeneity of the instrument, I checked the correlation between crime rate and labour market outcomes. As shown in Table C.11, the crime rate instrument is low-correlated with the labour market outcomes. This is enough for identifying the migration decision. Additionally, the difference between the homicide rate in the origin and the average homicide rate of the 3 most violent states adds more variation to the IV. By doing that, I look at a “surprise” component of the crime rate, that is, unpredicted by spatial and time trends, and thus makes it more exogenous to contemporaneous outcomes.[13]

Then, I turn my attention to the correlation between the instrument and the endogenous variable. Individual’s decision to move to another area take into account urban externalities, such as crime. As previous studies have shown, regions with higher rates of criminality tend to expel inhabitants to other regions, and this effect is higher for developing countries (Morrison, 1993; Cullen and Levitt, 1999; Engel and Ibáñez, 2007; Ibáñez and Vélez, 2008; Sousa, 2014; Chairassamee, 2018). Moreover, Brazil presents one of the highest crime rates in the world and one should consider criminality when explaining its internal migration. Regarding the orthogonality of the instrument, I argue that one’s decision to move away does not impact criminality at the origin. That is because this paper focuses on workers in the formal labour market. This kind of worker is less likely to be engaged in criminal activities (Baker, 2015; Blattman and Annan, 2016), which emphasizes the exogeneity of the instrument. Table 4.3 present the first stage estimation results, which provide evidence of the correlation between migration and the instrument. Additionally, the F-statistics on the excluded instrument show values above 10, which suggests that the instrument applied in this paper is a good instrument.

4.5 Results

4.5.1 Main results

For the sake of comparison, I estimate a naïve model without controlling for migrants self-selection. Table 4.2 presents the results. All regressions include socio-economic control variables, state and year dummies, and workers and occupation fixed effects. Standard errors are clustered at the individual level.

First, column (1) shows a positive and significant coefficient for migrants. That is, migrants earn 6.21% more than non-migrants on average. Moreover, migrants are on average 8.7% more likely to change occupation (column 2). However, regarding moving up or down the rank of occupations (columns 3 and 4), one sees that migrants are less

[13] I also computed the Moran’s I for the crime rate instrument yearly (Table C.12). According to the test, one cannot reject the null hypothesis for the period 2006-2011, meaning that there is no spacial correlation. Note that in 2012, however, H_0 is rejected at the usual levels of significance.

likely to move to a higher rank (-1.95%) and more likely to move to a lower rank (1.52%). This confirms the hypothesis that a worker may move to higher pay occupation, but with a lower rank. Lastly, migrants are also 2.05% more likely to move to an education well-matched occupation (column 5), which indicate a better job-match. Overall, these results suggest that migrants have an advantage compare to non-migrants, which corroborates the summary statistics presented in Table 4.1, and are in accordance with previous expectations.

Table 4.2: Naïve estimation

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	FE	FE	FE	FE	FE
Migrant	0.0621***	0.0870***	-0.0195***	0.0152***	0.0205***
	(0.00370)	(0.00313)	(0.00242)	(0.00218)	(0.00258)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	yes	yes	yes	yes	yes
R-squared	0.853	0.567	0.619	0.497	0.525
Observations	909,616	909,616	909,616	909,616	909,616

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.
Notes: Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Table 4.3: FEIV estimation

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	FEIV	FEIV	FEIV	FEIV	FEIV
First stage					
Ln homicide rate origin t-1 (deviation)	0.015***	0.015***	0.015***	0.015***	0.015***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
First stage F statistics	69.16	68.90	68.04	68.04	68.97
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000
Second stage					
Migrant	-0.421**	0.557***	-0.236	-0.0976	-0.0438
	(0.168)	(0.191)	(0.148)	(0.132)	(0.138)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	yes	yes	yes	yes	yes
R-squared	0.272	0.323	0.404	0.261	0.218
Observations	909,616	909,616	909,616	909,616	909,616

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.
Notes: All specifications use IV-2SLS with the deviation of the homicide rate in the state of origin in t-1 as an instrument for migrant. Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Table 4.3 presents the results for the IV model, which takes into account the endogeneity. The first-stage analysis indicates whether people are more likely to leave the place they live after violence shocks. The results show a positive and significant coefficient on origin state homicide measure, suggesting that people are 0.015% more likely to move from their original state in response to a 1% increase in deviation of the homicide rate. The F-statistic is sufficiently high, greater than 10 for all specifications.[14]

The second-stage analysis investigates the effects of migration on job-match transitions after using an exogenous variation on individual decision to move caused by shocks in the homicide rate in the origin states. The first column shows that after controlling for self-selection, the coefficient is negative and significant for migrants. That is, migrants earn less than non-migrants on average (-42.1%). Contrary to most of the literature that find positive returns to migrants, this result is in line with previous results find by Smits (2001), suggesting that before moving migrants were in a less favourable situation in the labour market than non-migrants and thus their returns could be negative if they are forced to move. Column (2) shows that migrants are 55.7% more likely to change occupation than non-migrants, but the magnitude is greater compared to the naïve estimates in Table 4.2. Column (3) shows that the coefficient is not significant, but it remains negative, and it is greater in magnitude than the naïve model (-0.236). The last two columns of Table 4.3 show that the coefficients are not significant at the usual levels, however, both coefficients are negative, contrary to the naïve estimates.

Overall, there is a large difference between the naïve and IV estimates. Comparing the results in tables 4.2 and 4.3 suggests that migrants are positively selected in the naïve specification. However, when the endogeneity is accounted for, the effect of migration is negative indicating that movers are being forced to leave. These results are in accordance with previous expectations. Considering that homicides are greater in poorer states in Brazil, then the type of migrant that responds to shocks in the homicide rate, and thus captured in the IV estimates, may be different from the average migrant.

To shed some light on the role of self-selection in this case, I also estimated the OLS model and IV without workers' fixed effects. The results are presented in Tables C.3 and C.4, respectively. Looking initially at the OLS estimation, all coefficients are positive and significant indicating that migrants are positively selected. That is, migrants earn 5.51% than non-migrants on average, are 4.17% more likely to change occupation, 2.31% more likely to move to a higher rank compared to the same rank level, but also 4.35% more likely to move to a lower rank and 1.5% more likely to move to migrate to a well-matched job. After including individuals fixed effects (Table 4.2), as discussed at the beginning of this section, migrants are also positively selected. Note however, that the coefficients for the latter are slightly larger. Another difference when workers' heterogeneity are controlled for is that the coefficient for moving to a higher rank occupation becomes negative (column 3). Moving to the IV analysis without individuals' fixed effects, Table

[14]The F-statistic is greater than the Stock and Yogo (2005) critical value for a test of maximal size 0.1.

C.4 shows that the first stage is positive and significant, which means that people are 0.49% more likely to move from their home state in response to a 1% increase in the deviation of the homicide rate.[15] Looking at the second-stage, one sees that migrants are earn on average -36.8% less than non-migrants, are 26.7% more likely to change occupations, -20% less likely to move to a higher rank and 32.6% more likely to move to a lower rank. Change to a well-matched job is not significant at the usual levels, although the coefficient is positive. However, when workers fixed effects are controlled for only columns 1 and 2 are significant (Table 4.3). Once again, the coefficients are slightly larger in the latter.

4.5.2 Empirical estimates by gender

The next analysis consists in separating the sample by gender. First of all, the literature on labour market participation indicates that labour supply decisions are different between males and females workers. Second, the analysis of violent deaths in Brazil shows that the homicide rate per 100,000 inhabitants is 142.7 for young people (15 to 29 years old), and considering only the subgroup of young males, the rate is 280.6. That is, among all the young people murdered, 94.6% were men (IPEA, 2018). Moreover, labour market performance is different for men and women, e.g. wages and occupational distribution.[16] Figures C.4 and C.5 depict the share of workers and the average wage by gender in each quintile of the occupational rank, respectively. Among young workers, the share of women is greater in the first and fifth quintiles, while men are the majority in the second, third and fourth quintiles (Figure C.4). However, Figure C.5 shows that for all quintiles of the occupational rank, men earn more on average than women. Therefore, one should expect different effects of migration on labour outcomes by gender. Tables 4.4 and 4.5 present the results separately for men and women, respectively.

As expected, the first-stage analysis shows a positive and significant coefficient on origin state homicide measure for men (Table 4.4), and the F-statistic is greater than 10 for all specifications. The second-stage results present a significant effect of migration. The first column shows that migrants induced by shocks on homicide earn -83.1% less than non-migrants. Moreover, migrants are almost 91% more likely to change occupation. However, migrants are less likely to change to a higher rank occupation (-67.7%) and more likely to move to a lower rank occupation (73.2%) than non-migrants, columns (3) and (4), respectively. Additionally, the likelihood of changing to an education well-matched occupation is 41.8% smaller for migrants.

These results are consistent with the hypothesis that migrants were in a less favourable situation before moving, and thus when they are forced to move due to an increase in violence, their labour market outcomes are negative (e.g. Smits, 2001).

[15] The F-statistic is sufficiently high in all specifications.
[16] See for instance, Goldin and Katz (2002); Marianne (2011); Pan (2015).

Table 4.4: FEIV estimation - Men

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	FEIV	FEIV	FEIV	FEIV	FEIV
First stage					
Ln homicide rate origin t-1 (deviation)	0.012***	0.012***	0.012***	0.012***	0.012***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
First stage F statistics	23.96	24.00	23.45	23.45	24.03
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000
Second stage					
Migrant	-0.831**	0.909***	-0.677**	0.732***	-0.418*
	(0.324)	(0.351)	(0.288)	(0.270)	(0.237)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	yes	yes	yes	yes	yes
R-squared	0.139	0.285	0.324	0.077	0.156
Observations	475,237	475,237	475,237	475,237	475,237

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.

Note: All specifications use IV-2SLS with the deviation of the homicide rate in the state of origin in t-1 as an instrument for migrant. Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Table 4.5: FEIV estimation - Women

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	FEIV	FEIV	FEIV	FEIV	FEIV
First stage					
Ln homicide rate origin t-1 (deviation)	0.016***	0.016***	0.016***	0.016***	0.016***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
First stage F statistics	40.19	39.73	39.46	39.46	39.77
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000
Second stage					
Migrant	-0.198	0.311	0.243	-0.982***	0.373*
	(0.214)	(0.251)	(0.190)	(0.234)	(0.206)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	yes	yes	yes	yes	yes
R-squared	0.303	0.318	0.407	0.028	0.203
Observations	434,379	434,379	434,379	434,379	434,379

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.

Note: All specifications use IV-2SLS with the deviation of the homicide rate in the state of origin in t-1 as an instrument for migrant. [1]Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Table 4.5 presents the estimates for the subgroup of female workers. Similarly to men, the first-stage analysis shows a positive and significant coefficient on origin state homicide measure, and F-statistic is greater than 10 for all specifications. However, the second-stage results are different. Columns (1), (2) and (3) are not significant at the usual levels. Contrary to men, column (4) shows a negative and significant result for women. That means female migrants are 98.2% less likely to move to a lower rank occupation than non-migrants. Additionally, migrants are more likely to move to an education well-matched occupation (column 5).

Overall, it seems that the results from Table 4.3 are mostly due to the effect of homicide shocks on male migration, especially columns (1) and (2). Regarding columns (3), (4) and (5), the effects of migration on men and women are in different directions, which may be cancelling each other.

4.5.3 Empirical estimates by education level

This subsection presents the analysis of internal migration by education level. In Brazil, violence is higher in poorer regions, where the share of low-educated workers is greater than the share of high-educated, e.g. the Northeast region (IPEA, 2017, 2018). Table 4.6 shows the results for those with less than higher education and Table 4.7 contains the results for those with higher levels of education.

Table 4.6: FEIV estimation - Low-educated

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	FEIV	FEIV	FEIV	FEIV	FEIV
First stage					
Ln homicide rate origin t-1 (deviation)	0.011***	0.011***	0.011***	0.011***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
First stage F statistics	32.88	32.84	32.11	32.11	32.86
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000
Second stage					
Migrant	-0.269	1.009***	-0.631***	0.132	-0.601***
	(0.217)	(0.320)	(0.241)	(0.201)	(0.219)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	yes	yes	yes	yes	yes
R-squared	0.271	0.255	0.329	0.260	0.155
Observations	729,944	729,944	729,944	729,944	729,944

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.
Note: All specifications use IV-2SLS with the deviation of the homicide rate in the state of origin in t-1 as an instrument for migrant. [1]Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Starting with the low-educated, the first-stage results show a positive and significant effect of homicide rate measure at the origin, and as reported in Table 4.6, the F-statistic is greater than 10 for all specifications. Regarding the second-stage, column (1) shows that although the coefficient is negative, the impact of migration on hourly wages is not significant at the usual levels. Column (2) on its turn, presents a positive and significant coefficient, that is, migrants are more likely to change occupation than non-migrants. Moving to occupations rank, column (3) indicates that low-educated migrants are less likely to move to a higher rank occupation (-63.1%) than low-educated non-migrants. Conversely, the coefficient is positive for moving to a lower rank occupation, but it is not significant (column 4). Lastly, as column (5) demonstrates, the likelihood of moving to an educational well-matched occupation is smaller (-60.1%) for low-educated migrants compared to their counterparts.

Table 4.7: FEIV estimation - High-educated

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	FEIV	FEIV	FEIV	FEIV	FEIV
First stage					
Ln homicide rate origin t-1 (deviation)	0.017***	0.017***	0.017***	0.017***	0.017***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
First stage F statistics	16.15	16.28	16.43	16.43	16.30
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000
Second stage					
Migrant	0.389	0.166	-0.426	0.0504	-0.518*
	(0.405)	(0.385)	(0.319)	(0.248)	(0.300)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	yes	yes	yes	yes	yes
R-squared	0.342	0.362	0.468	0.266	0.150
Observations	179,672	179,672	179,672	179,672	179,672

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.

Note: All specifications use IV-2SLS with the deviation of the homicide rate in the state of origin in t-1 as an instrument for migrant. [1]Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Turning attention to high-educated, the first-stage results are also positive and significant, and the F-statistic is greater than 10 for all specifications (Table 4.7). In contrast to low-educated estimates, the second-stage results for high-educated workers are not significant at the usual levels. The exception is column (5), that shows that high-educated migrants are less likely to move to an educational well-matched occupation (-51.8%) than non-migrants. Although the coefficients are not significant for the other specifications, note that their signals are similar to the naïve estimates in Table 4.2. Particularly, for high-educated the effect of migration on hourly wage is positive.

Overall, the results suggest that the estimates from Table 4.3 are derived from the low-educated workers. This is consistent with the Brazilian context, where violence is concentrate em peripheral and poorer areas (IPEA, 2017, 2018).

4.5.4 Empirical estimates by income group in the origin state

So far, the subgroup analysis has used individuals characteristics to separate the sample. In this subsection, I separate the sample according to the origin state's GDP in 2006.[17] Table 4.8 presents the results for the origin states in the top tercile of the GDP, and Table 4.9 shows the results for the middle and bottom terciles.[18]

Table 4.8: FEIV estimation - Top tercile of the GDP

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	FEIV	FEIV	FEIV	FEIV	FEIV
First stage					
Ln homicide rate origin t-1 (deviation)	0.02***	0.02***	0.02***	0.02***	0.02***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
First stage F statistics	68.67	68.74	68.46	68.46	68.62
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000
Second stage					
Migrant	0.693***	0.326*	-0.00724	-0.268**	-0.102
	(0.171)	(0.178)	(0.141)	(0.131)	(0.135)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	yes	yes	yes	yes	yes
R-squared	0.251	0.353	0.420	0.251	0.220
Observations	675,823	675,823	675,823	675,823	675,823

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.
Note: All specifications use IV-2SLS with the deviation of the homicide rate in the state of origin in t-1 as an instrument for migrant. [1]Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

As Table 4.8 depicts, the first-stage results are positive and significant, and the F-statistic is greater than 10 for all specifications. Regarding the second-stage, column (1) shows that migrants earn more than non-migrants (69.3%). Column (2) indicates that migrants are 32.6% more likely to change occupation, corroborating previous expectations. Additionally, column (4) shows that migrants are -26.8% less likely to move to a lower rank occupation compared to non-migrants. Columns (3) and (5) are not significant at the usual levels.

Moving to Table 4.9, the first-stage results also show a positive and significant coefficient for the instrument, and F-statistic is also greater than 10 for all specifications. However, the second-stage results are different from the rich origin states. To begin with, column (1) indicates that migrants earn less than non-migrants (-93.27%). Columns (2) and (3) show that migrants are less likely to change occupation, and are less likely to move to a higher rank occupation, respectively. Moving to a lower rank occupation also presents a

[17]The GDP was obtained from the Brazilian Institute of Geography and Statistics (IBGE).
[18]I also divided the sample according to the origin state's GDP per capita in 2006. The results remain broadly unchanged.

negative coefficient, but it is not significant at the usual levels. Moreover, column (5) shows that migrants are less likely to move to an education well-matched occupation, -79.7%.

Table 4.9: FEIV estimation - Middle and bottom terciles of the GDP

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	FEIV	FEIV	FEIV	FEIV	FEIV
First stage					
Ln homicide rate origin t-1 (deviation)	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
First stage F statistics	12.15	12.42	12.31	12.31	12.32
Prob > F	0.0005	0.0004	0.0005	0.0005	0.0004
Second stage					
Migrant	-2.699***	-1.591**	-1.006**	-0.196	-0.797*
	(0.942)	(0.657)	(0.480)	(0.354)	(0.456)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	yes	yes	yes	yes	yes
R-squared	-0.917	0.082	0.269	0.249	0.098
Observations	233,793	233,793	233,793	233,793	233,793

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.
Note: All specifications use IV-2SLS with the deviation of the homicide rate in the state of origin in t-1 as an instrument for migrant. [1]Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Overall, the results suggest that migrants from richer states are positively selected, while migrants from poorer states are negatively selected. Moreover, negative selection may be an indication of forced migration in poor states.

4.5.5 Violence-induced migrants

The estimates of the IV approach are the local average treatment effect (LATE) (Imbens and Angrist, 1994). Thus, it is possible that the migrants captured in the main estimates, i.e. those who respond to shocks in the homicide rate, are different from the average migrants depicted in Table 4.1. As previously discussed, violence in Brazil affects mainly poorer individuals and is the main cause of death of young men (IPEA, 2018). Thus, to shed light on who are those violence-induced migrants, I restricted the sample to migrants and regressed the instrument on hourly wages and on an indicator variable for men, high education, undereducated, overeducated, industry, firm size and for the region (destination and origin). If people who migrate when violence increases are poorer and less educated, one should expect a negative relationship between violence and migrant's wages and education.

As depicted in Table 4.10, the deviation from the homicide rate in the state of origin is negative and significant in column (1). That is, the elasticity is around -0.03%, suggesting that violence-induced migrants earn less on average than other migrants. Column (2) shows that a 1% increase in the deviation from the homicide rate leads to an increase in the probability of the migrant being a man of 0.0156%. Regarding education (column 3), the coefficient for violence is negative and significant, which means that a 1% increase in the violence measure decreases the likelihood of being highly educated by 0.0233%. Also, violence-induced migrants are more likely to be undereducated than the regular migrant (column 4), but there is no difference in the likelihood of being overeducated (column 5). Columns (6) and (7) show that the LATE-compliers are less likely to work on the industry sector, and more likely to be in the services sector. Moreover, they are 0.0258% more likely to work in big firms, that is, firms with 500 or more employees than the average migrant.

In addition, Table 4.11 shows that migrants affected by violence are more likely to live at the North, Northeast and Midwest regions than the average migrant, and less likely to live at the Southeast and South regions, which are the richer regions in Brazil. Regarding their origin, the pattern is similar to the destination, that is, violence-induced migrants are more likely to come from states in the North, Northeast and Midwest, but less likely to come from Southeast and South.

Therefore, corroborating previous expectations, violence-induced migrants, on average, earn less, are less likely to be highly educated, more likely to come from and live in poorer regions and more likely to be male than other migrants.

Table 4.10: Violence-induced migrants - Socioeconomic characteristics

	Dependent variable:							
	Ln wageh	Male	High educ	Undereducated	Overeducated	Industry	Services	Big firms
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln homicide rate origin t-1 (deviation)	-0.0276***	0.0156***	-0.0233***	0.00598***	-0.000249	-0.0200***	0.0199***	0.0258***
	(0.00555)	(0.00415)	(0.00365)	(0.00205)	(0.00446)	(0.00448)	(0.00448)	(0.00431)
State dummies	yes	yes	yes	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes	yes	yes	yes
R-squared	0.087	0.007	0.019	0.005	0.003	0.019	0.019	0.017
Observations	83,050	83,050	83,050	83,050	83,050	83,050	83,050	83,050

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.

Notes: The sample consists of migrants only. The dependent variable in column (1) is continuous and after the worker migrated. All other dependent variables are dummy variables. Big firms (in column 8) means firms with 500 or more employees.

Table 4.11: Violence-induced migrants - Geographic regions

	Dependent variable:									
	Destination					Origin				
	North	Northeast	Midwest	Southeast	South	North	Northeast	Midwest	Southeast	South
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ln homicide rate origin t-1 (deviation)	0.0546***	0.198***	0.121***	-0.233***	-0.140***	0.0974***	0.333***	0.215***	-0.521***	-0.125***
	(0.00227)	(0.00436)	(0.00338)	(0.00539)	(0.00404)	(0.00255)	(0.00448)	(0.00350)	(0.00581)	(0.00423)
Year dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R-squared	0.009	0.047	0.024	0.042	0.029	0.027	0.123	0.077	0.205	0.023
Observations	83,050	83,050	83,050	83,050	83,050	83,050	83,050	83,050	83,050	83,050

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.

Notes: The sample consists of migrants only. All dependent variables are dummy variables.

4.6 Robustness

In this section, I perform robustness checks to determine how sensitive the results are to changes in the specifications.

4.6.1 Alternative measures of the instrument

The analyses so far have used the deviation of the logarithm of the homicide rate in the state of origin from the logarithm of the average homicide rate of the 3 most violent states at t-1 as an instrument for migration. Robustness checks show that results are robust to using the logarithm of the homicide rate in the state of origin, the deviation from the mean of all states and the homicide rate in the state of origin z-score. The results of the are not shown because they do not change compared to those presented in Table 4.3, the difference is in the perception of criminality by individuals, and thus in the interpretation.

4.6.2 Smaller sample

The dataset only contains information on individuals who are in the formal labour market. Thus, I repeat the estimation for a smaller sample, i.e. using only those workers who are observed 6 or more times. Table C.5 shows that the results are in line with those previously presented in Table 4.3. The coefficients of the second-stage are greater in magnitude, and the likelihood of moving to a higher rank becomes significant.

4.6.3 Alternative rank definition

The analysis of moving occupation status uses quintiles to check whether a worker moved up or down on the rank of occupations. Table C.6 shows the results when instead of quintiles, the occupations in each state are separated into sextiles or deciles. First, considering moving to a higher rank, the coefficient is still negative, but it becomes greater in magnitude and statistically significant as more intervals are considered. On the other hand, moving to a lower rank remains not significant at the usual levels, but as more intervals are taken into account the coefficient becomes greater.

4.6.4 Alternative instrument

In this subsection, I use an alternative instrument for migration. Similarly to the violence measure, I now use the deviation of soy exports at the state of origin from the average soy exports of the 3 biggest exporting states in the previous period.[19] Note that this analysis considers only migrants coming from the 9 origin states that soy is the main exporting product.[20] Table C.7 present the results using soy exports deviation as an instrument for migration. The first-stage analysis indicates whether people are more likely to move after a negative shock on soy exports. The results show a positive and significant coefficient for the new instrument, indicating that people are 0.007% more likely to move from their original state in response to a 1% decrease in the soy exports value. The F-statistic is greater than 10 for all specifications. The second-stage results show a positive and significant coefficient for migrants in column (1). That is, migrants earn more than non-migrants on average (80.8%). The coefficients are not significant at the usual levels in columns (2) to (5), however, not that the signal is positive only for changing to a lower rank occupation (column 4).

In order to compare the results of the violence measure instrument with the soy exports instrument, I re-estimate the effects of a positive shock on violence, restricting the sample to the same 9 origin states where soy is the main exporting product. Table C.8 present the results. The first-stage analysis indicates that an increase of 1% in the perception of violence at the origin, increases the likelihood of migration by 0.043%. Note that F-statistic is greater than 10 for all specifications. The second-stage results show that migrants are positively selected and earn more on average than non-migrants (column 1). Additionally, migrants are -46.8% less likely to move to a higher rank occupation than non-migrants. The coefficients for columns (2), (4) and (5) are not significant at the usual levels. Moreover, note that coefficients in each column have the same signal in Tables C.7 and C.8. Thus, it seems that for the same sample, negative shocks on soy exports have similar effects as positive shocks on violence.

In addition, I re-estimate the analysis of positive violence shocks, excluding the soy exporting states at the origin. Table C.9 present the results. As expected, the first-stage results suggest that an increase in the perception of violence in the origin state increases the likelihood of moving to a different state. The second-stage results show that migrants earn on average less than non-migrants (column 1), and are more likely to change occupations (column 2). In addition, column (4) suggest that migrants are less likely to move to lower rank occupations. Columns (3) and (5) are not significant at the usual levels.

Thus, the results without soy states are in line with those previously presented in Table 4.3. Moreover, these results corroborate previous findings in Subsection 4.5.4, that migrants are positive or negative selected depending on their origin state.

[19] Soy exports value are free-on-board (US$) and the data was obtained from the Brazilian Ministry of Industry, Foreign Trade and Services for the period 2006-2013.

[20] The states where soy is the main exporting product are Roraima (RR), Tocantins (TO), Piauí (PI), Distrito Federal (DF), Goiás (GO), Mato Grosso (MT), Mato Grosso do Sul (MS), Paraná (PR) and Rio Grande do Sul (RS). They account for 19% of the sample.

4.7 Conclusion

This paper investigates the effect of internal migration on job match transitions focusing on young people who are entering the labour market. Using a rich panel dataset from Brazil, I employ an instrumental variable approach to deal with migrants' self-selection. In order to get exogenous variation, I then use violent crime shocks at the origin state.

The results show that without controlling for self-selection, internal migration has a modest, but positive effect on earnings. Moreover, migrants are more likely to move to an education well-matched occupation, but also more likely to move to lower rank occupations. Conversely, when the endogeneity is controlled for, the results show a positive and significant first-stage relationship, which indicates that people are more likely to move from their original state after an increase in its homicide rate compared to the 3 most violent states. The second-stage estimates exhibit a negative impact of internal migration on earnings. Moving on the rank of occupations and getting a well-matched occupation is no longer significant. In addition, I separate the sample into different groups in order to understand the negative selection of the instrument. I find heterogenous results, with a stronger effect on men and low-educated workers. Moreover, I find that migrants who came from poorer states are negatively selected, while migrants who came from richer states are positively selected, suggesting that forced migration occurs only in poorer states. Furthermore, I find that LATE-complying migrants are different from other migrants and that on average, they earn less, are less likely to be men, highly educated, and more likely to come from poorer regions.

Contrary to most studies in the literature, this paper focus on a developing country with high criminality and poor labour market conditions. In this context, public policies are indeed relevant. The results suggest that the increase in the state's perception of violence pushes migrants away. Particularly, affecting low-educated migrants and those coming from a poorer state. Moreover, policy makers need to account for the effects of forced migration on inequality on both origin and destination. Forced migrants may take generations to catch up with their receiving population, as they are usually low-skilled and competing with natives may be challenging for them. Thus, policies aiming to reduce violence and protect the population may decrease the displacement of the more vulnerable workers. Additionally, policies should aim on the integration of this kind of migrant in society. Otherwise, forced migration may impact other amenities at the destination, such as access to public goods.

Finally, it is important to highlight that this paper considers short-run effects only. In the long run, workers may migrate and perform a lower salary job at the beginning, as long as they have better expectations for the future (see Smits, 2001). However, further research is needed to understand these long-run patterns.

Chapter 5

Conclusions

In this thesis, I study important topics in labour economics in Brazil. This research is divided into three chapters, and each one investigates issues related to one of the following topics: educational mismatch, minimum wage policies and internal migration.

The first chapter examines the incidence of educational mismatch and analyses its effects on wages in Brazil. By using a large employer-employee panel dataset, I find evidence that a quarter of the Brazilian formal labour market is overeducated and a quarter is undereducated. Also, after controlling for workers' heterogeneity, I find that undereducated workers earn 4.31% more, on average than their colleagues who hold a well-matched job. Conversely, overeducated workers earn -3.97% less than their counterparts who hold a well-matched job. However, the penalty for overeducation is less than half of the premium for going to university (incomplete higher education premium is 9.63%), indicating that for the individual is worth acquiring higher education. Moreover, I find that eliminating the educational mismatch by changing workers' education would decrease aggregate wages by -0.32%. However, eliminating the educational mismatch by reassigning workers to well-matched jobs would increase aggregate wages by 0.25%. These results suggest that given the symmetry of the mismatch observed in the dataset, over and undereducation effects are cancelling each other.

The second chapter examines the effects of changes the minimum wage policy on the investments in education, through changes in the wage distribution and skill premia. By using data for people who are already in the labour market, first, I analyse how changes in two different measures of the minimum wage policy, an absolute and a relative minimum wage measure, affect the skill premium. Then, I estimate the effect of these two measures on investments in education. I find that an increase in the absolute minimum wage is associated with a decrease of the skill premia, particularly for higher education, diminishing the probability of acquiring more schooling by -0.56%. The discouragement is bigger for states with higher absolute minimum wage. Moreover, the effect is stronger for overeducated, female and for workers on the bottom of the wage distribution. On the contrary, an increase in the relative minimum wage is associated with an increase in the skill premia for higher education, but there is no effect on lower and intermediate education. Moreover, the relative minimum wage has no significant effect on educational investments on average. However, there is evidence of positive effect for educational mismatched workers, and for workers on the 25th, 50th and 75th percentiles of the wage distribution. These results suggest that minimum wage policies that deteriorate the skill premia, particularly for secondary education and above, will discourage investments in education, confirming basic economic intuition.

The third chapter investigates the labour market effect of internal migration in Brazil by instrumenting migration with violent crime shocks at the origin state. By using data for young people entering the labour market, I find in the first-stage analysis that people are more likely to leave their original state after a positive shock on violence perception. The results in the second-stage analysis show that on average, violence-induced migrants are negatively selected, and earn less than non-migrants. In particular, the effects of migration are stronger for men and low-educated workers. However, the type of selectivity of migrants depends on their state of origin. I find that migrants who came from richer states are positively selected compared to non-migrants. On the contrary, migrants who came from poorer states are negatively selected. Thus, these results may suggest forced migration only for workers who migrated from poorer states.

Together these chapters contribute to the existing literature in several ways. First, this research contributes to the scarce literature on educational mismatch in developing countries (e.g Mehta et al., 2011; Reis, 2017) by analysing the effects of the growth in the Brazilian educational system and the increase of its workforce education attainment on the labour market. Second, it relates to the literature on redistributive effects of the minimum wage on the wage distribution, and on education acquisition (e.g Neumark and Wascher, 2003; Dolton et al., 2012; Butcher et al., 2012; Bárány, 2016). This research disentangles the distributional effects of minimum wage policies by providing empirical evidence that two measures of the real minimum wage have a different impact on the skill premia, and thus on education acquisition. In particular, this study contributes to the literature by investigating these effects in a developing country where in contrast to developed countries, minimum wages and education acquisition are not necessarily substitutes. Third, there is little research on the effect of crime on migration decision in developing countries with high homicide rate (e.g. Calderón-Mejía and Ibáñez, 2016). This research provides credible estimates of the effect of internal migration on job match transitions by instrumenting migration with violent shocks in the origin state. Fourth, it also relates to the literature that investigates the determinants of decreasing inequality and changes in wage premia for different educational levels in Brazil (e.g Barros et al., 2010; Barbosa et al., 2015; Engbom and Moser, 2017).

This research also emphasizes several policy debates. First, the occurrence of educational mismatch in a developing country like Brazil may indicate inefficiencies in the labour market, and contribute to misallocation of skills, raising questions about the relationship between education and the labour market. Thus, policies should focus not only on increasing the educational attainment of the workforce, but also should focus on providing good quality education in order to increase workers' productivity. In addition, they also should focus on increasing job opportunities, in which workers can fully explore their productivity. Second, minimum wage policies are relevant to reduce inequality. Hence, minimum wage policies require careful evaluation on the precise effect on the skill premia, as this research suggested, policies that erode the skill premia will discourage education acquisition for different parts of the wage distribution. Third, violence may have high socioeconomic costs, and the consequences of internal migration driven by violent shocks on job transitions have important implications for policy-makers. In particular, in developing countries with high criminality. Also, it is important to highlight that Brazil is a diverse country with heterogeneous states, and thus decentralised policies may be more efficient to reduce inequality among them.

Lastly, it is important to highlight that the results of this research hold only for the Brazilian formal labour market. Nevertheless, this thesis provides interesting avenues for future research to better understand issues related to the labour market. One potential direction is to investigate the link between educational and allocative mismatches by analysing the extent to which workers change firms, possibly indicating a poor match to begin with, and how these changes are rewarded by the labour market. This would allow us to test whether the educational mismatch is persistent and whether investing in education on-the-job is recompensed by the labour market.

Appendix A

Appendix for Chapter 2

A.1 Other tables

Table A.1: Descriptive statistics

Variable	Type	Description	Mean/Frequency
Wage/hour*	Continuous	R$	16.56
			(0.082)
Hour	Continuous	Hours worked	39.45
			(0.012)
Age**	Continuous	Years	41.96
			(0.015)
Tenure	Continuous	Months	124.78
			(0.184)
Gender	Categorical	Male	56.48
		Female	43.52
Class worker	Categorical	Blue collar	26.54
		White collar	73.46
Firm size	Categorical	Up to 19 employees	11.59
		20 to 99 employees	12.08
		100 to 499 employees	19.39
		500 or more	56.96
Educational mismatch	Categorical	Undereducated	25.67
		Well-matched	49.66
		Overeducated	24.67
Educational level	Categorical	Illiterate	0.15
		Incomplete Basic Education I	2.09
		Complete Basic Education I	3.93
		Incomplete Basic Education II	6.61
		Complete Basic Education II	11.20
		Incomplete Secondary School	4.74
		Complete Secondary School	40.02
		Incomplete Higher Education	3.27
		Complete Higher Education	27.47
		Master	0.43
		Doctorate	0.10
Sector***	Categorical	Industry	39.08
		Service	60.92

Source: RAIS, 2006-2013.
Notes: Linearized SE in parenthesis. *The wage was deflated by the 2013 prices from IPCA (Brazilian consumer price index); **Between 25 and 65 years; ***Agriculture was excluded.

Table A.2: Prevalence of over/undereducation (%)

Educational difference	**Year of observation**								
	2006	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**Total**
-7 levels	0.05	0.03	0.02	0.02	0.02	0.01	0.02	0.01	**0.02**
-6 levels	0.11	0.09	0.07	0.06	0.05	0.05	0.05	0.04	**0.06**
-5 levels	1.02	0.88	0.83	0.75	0.70	0.65	0.60	0.55	**0.75**
-4 levels	2.39	2.20	2.03	1.82	1.75	1.40	1.32	1.18	**1.76**
-3 levels	4.97	4.45	4.19	3.94	3.77	3.67	3.40	3.24	**3.95**
-2 levels	13.92	12.74	12.05	11.49	10.96	10.33	9.84	9.01	**11.29**
-1 level	9.38	8.86	8.50	7.83	7.51	7.19	6.91	6.45	**7.83**
0 level	46.88	48.55	49.24	50.11	50.52	50.31	50.68	50.96	**49.66**
1 level	5.55	5.54	5.56	5.46	5.39	5.29	5.28	5.39	**5.43**
2 levels	15.19	16.08	16.84	17.75	18.49	20.04	20.79	21.90	**18.39**
3 levels	0.30	0.30	0.30	0.31	0.37	0.38	0.39	0.40	**0.34**
4 levels	0.25	0.29	0.38	0.44	0.49	0.67	0.72	0.86	**0.51**
5 levels	0.00	0.000	0.001	0.003	0.001	0.003	0.003	0.012	**0.003**
6 levels	0.00	0.000	0.000	0.000	0.001	0.001	0.001	0.002	**0.001**
Total	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**

Source: RAIS, 2006–2013; CBO, 2010.

Table A.3: Wage equations

Variables	(1) POLS	(2) FE	(3) FE	(4) FE	(5) FE
Overeducation	-0.217***	-0.0571***	-0.0534***	-0.0641***	-0.0324***
	(0.00294)	(0.00490)	(0.00597)	(0.00573)	(0.00550)
Undereducation	0.156***	0.0430***	0.0471***	0.0403***	0.0446***
	(0.00370)	(0.00663)	(0.00819)	(0.00721)	(0.00687)
Female	-0.357***	-	-	-	-
	(0.00237)	-	-	-	-
Overeducation*Female	-	-	-0.00852	-	-
	-	-	(0.00842)	-	-
Undereducation*Female	-	-	-0.00835	-	-
	-	-	(0.0105)	-	-
Overeducation*Young	-	-	-	0.0123**	-
	-	-	-	(0.00487)	-
Undereducation*Young	-	-	-	0.00589	-
	-	-	-	(0.00522)	-
Overeducation*LHE	-	-	-	-	-0.0577***
	-	-	-	-	(0.0104)
Undereducation*LHE	-	-	-	-	0.00401
	-	-	-	-	(0.0204)
Tenure	0.00181***	0.000382***	0.000381***	0.000383***	0.000383***
	(3.53e-05)	(4.84e-05)	(4.84e-05)	(4.84e-05)	(4.83e-05)
Tenure^2	2.83e-07***	5.69e-08	5.75e-08	5.16e-08	5.91e-08
	(1.03e-07)	(1.69e-07)	(1.69e-07)	(1.69e-07)	(1.69e-07)
Less than Fundamental Education II	-0.625***	-0.0203*	-0.0215**	-0.0212**	-0.00987
	(0.00461)	(0.0105)	(0.0107)	(0.0105)	(0.0108)
Complete Fundamental Education II	-0.406***	-0.00122	-0.00113	-0.00175	0.00964
	(0.00351)	(0.00800)	(0.00803)	(0.00801)	(0.00818)
Incomplete Secondary School	-0.254***	-0.00893	-0.00959	-0.00943	-0.0107
	(0.00455)	(0.00994)	(0.00997)	(0.00994)	(0.00992)
Complete Secondary School	omitted	omitted	omitted	omitted	omitted
Incomplete Higher Education	0.444***	0.0706***	0.0705***	0.0698***	0.0980***
	(0.00641)	(0.0114)	(0.0114)	(0.0114)	(0.0142)
Complete Higher Education	0.897***	0.193***	0.192***	0.193***	0.214***
	(0.00313)	(0.00823)	(0.00833)	(0.00825)	(0.00970)
Master/Doctorate	1.530***	0.330***	0.330***	0.332***	0.376***
	(0.0157)	(0.0290)	(0.0291)	(0.0291)	(0.0308)
Other controls	Yes	Yes	Yes	Yes	
State dummies	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes
Constant	1.447***	-0.847***	-0.848***	-0.855***	-0.863***
	(0.0227)	(0.0406)	(0.0406)	(0.0408)	(0.0406)
Observations	507,612	507,612	507,612	507,612	507,612
Adjusted R-squared	0.549	0.928	0.928	0.928	0.928

Standard errors in parentheses clustered by individual; *** p<0.01, ** p<0.05, * p<0.1.

Notes: Other controls include tenure, age, firm size, sector and class of worker.

Table A.4: Wage equations - role of tenure

Variables	(1) POLS	(2) FE
Overeducation	-0.309***	-0.044***
	(0.00660)	(0.008)
Undereducation	0.0983***	0.034***
	(0.00694)	(0.010)
Tenure	0.00169***	0.000319***
	(5.06e-05)	(6.33e-05)
Tenure	6.10e-07***	2.22e-07
	(1.45e-07)	(2.08e-07)
Overeducation*Tenure	0.000623***	0.000138
	(8.48e-05)	(8.48e-05)
Undereducation*Tenure	-0.000187**	0.000114
	(8.13e-05)	(9.19e-05)
Overeducation*Tenure^2	-2.17e-06***	-4.55e-07*
	(2.53e-07)	(2.58e-07)
Undereducation*Tenure^2	4.55e-07*	-1.72e-07
	(2.38e-07)	(2.80e-07)
Other controls	Yes	Yes
State dummies	Yes	Yes
Year dummies	Yes	Yes
Constant	1.461***	-0.869***
	(0.0235)	(0.041)
Observations	507,612	507,612
Adjusted R-squared	0.551	0.9276

Standard errors in parentheses clustered by individual; *** p<0.01, ** p<0.05, * p<0.1.

Notes: Other controls include age, education, firm size, sector and class of worker.

Table A.5: Wage equations using levels of educational mismatch

Variables	(1) POLS	(2) FE	Variables	(1) POLS	(2) FE
Undereducation			Overeducation		
-7 levels	0.636***	0.118	+1 level	-0.187***	-0.00323
	(0.0902)	(0.0835)		(0.00771)	(0.0119)
-6 levels	0.378***	0.185***	+2 levels	-0.328***	-0.0560***
	(0.0352)	(0.0393)		(0.00496)	(0.00750)
-5 levels	0.208***	0.109***	+3 levels	-0.548***	-0.0743***
	(0.0128)	(0.0219)		(0.0185)	(0.0238)
-4 levels	0.272***	0.0957***	+4 levels	-0.930***	-0.144***
	(0.00972)	(0.0163)		(0.0186)	(0.0227)
-3 levels	0.210***	0.0822***	+5 levels	-1.227***	-0.969**
	(0.00740)	(0.0135)		(0.280)	(0.420)
-2 levels	0.103***	0.0578***	+6 levels	-1.732***	-0.127**
	(0.00518)	(0.00944)		(0.296)	(0.0521)
-1 level	0.101***	0.0553***	Other controls	Yes	Yes
			Constant	1.469***	-0.856***
				(0.0233)	(0.0408)
			Observations	507,612	507,612
			Adjusted R-squared	0.555	0.928

Standard errors in parentheses clustered by individual; *** p<0.01, ** p<0.05, * p<0.1.
Notes: Other controls include tenure, age, firm size, sector, class of worker, state and year.

Table A.6: Wage equations - role of tenure

Variables	(1) POLS	(2) FE
Overeducation	-0.181***	-0.0476***
	(0.00511)	(0.00771)
Undereducation	0.265***	0.0428***
	(0.00488)	(0.00874)
Overeducation*Female	-0.367***	0.00563
	(0.00512)	(0.00869)
Undereducation*Female	-0.254***	0.00168
	(0.00413)	(0.0105)
Overeducation*Young	-0.00392	0.0119**
	(0.00531)	(0.00531)
Undereducation*Young	0.0439***	0.00657
	(0.00433)	(0.00520)
Overeducation*LHE	0.255***	-0.0614***
	(0.00625)	(0.0104)
Undereducation*LHE	-0.0665***	-0.00751
	(0.0144)	(0.0209)
Other Controls	Yes	Yes
State dummies	Yes	Yes
Year dummies	Yes	Yes
Constant	1.392***	-0.866***
	(0.0237)	(0.0408)
Observations	507,612	507,612
Adjusted R-squared	0.528	0.928

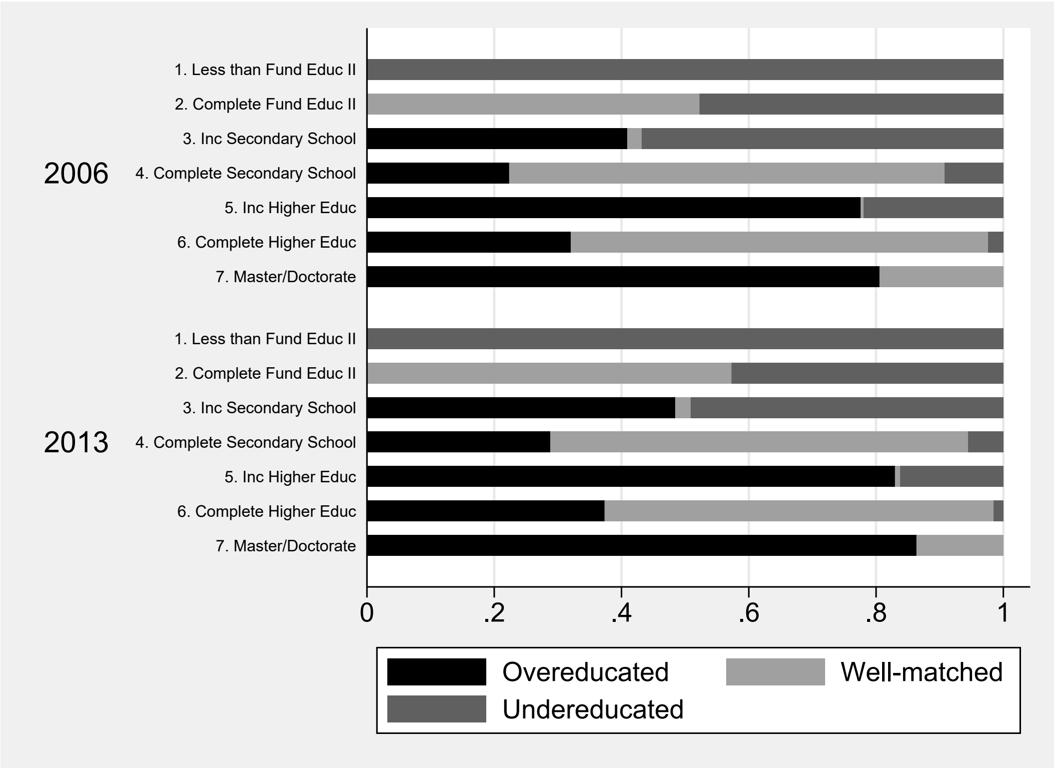
Standard errors in parentheses clustered by individual; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Notes: Other controls include age, tenure, education, firm size, sector and class of worker. The OLS regression also control for female.

$Overeducation^{*}Female$ ($Undereducation^{*}Female$) is a dummy that takes value one if the individual is overeducated (undereducated) and female; $Overeducation^{*}Young$ ($Undereducation^{*}Young$) is a dummy that takes value one if the individual is overeducated (undereducated) and up to 40 years old; $Overeducation^{*}LHE$ ($Undereducation^{*}LHE$) is a dummy that takes value one if the individual is overeducated (undereducated) and has less than higher education.

A.2 Other figures

Figure A.1: Educational mismatch over education



Source: RAIS, 2006–2013.

Appendix B

Appendix for Chapter 3

B.1 Other tables

Table B.1: Descriptive statistics

Variable	**Type**	**Description**	**Mean/Frequency**
Changed education	Categorical	Yes	6.79
		No	93.21
Minimum wage	Continuos	Relative minimum wage (index)	0.53
			(0.2047)
		Real minimum wage (index)	0.76
			(0.2580)
Mismatch level	Categorical	Undereducation	25.73
		Well-matched	44.49
		Overeducation	29.79
Tenure	Continuos	Months	71.03
			(85.87)
Age**	Continuos	Year	39.41
			(8.67)
Gender	Categorical	Female	39.16
		Male	60.84
Attained education	Categorical	Illiterate	0.20
		Incomplete Fundamental Education I	2.40
		Complete Fundamental Education I	4.16
		Incomplete Fundamental Education II	7.85
		Complete Fundamental Education II	12.90
		Incomplete Secondary School	6.10
		Complete Secondary School	39.16
		Incomplete Higher Education	3.97
		Complete Higher Education	22.68
		Master	0.49
		Doctorate	0.08
Firm size	Categorical	Up to 19 employees	20.21
		20 to 99 employees	20.87
		100 to 499 employees	21.90
		500 or more employees	30.67
Class of worker	Categorical	Blue collar	29.14
		White collar	70.86
Sector***	Categorical	Industry	42.03
		Service	57.97

Source: RAIS, 2006-2013.
Notes: Linearized SE in parenthesis. ** Between 25 and 65 years; *** Agriculture was excluded.

Table B.2: Effects of the absolute minimum wage on hourly wages

Variables	**(1) OLS**	**(2) FE**
Absolute minimum wage	0.00724	0.0489***
	(0.00582)	(0.00381)
Illiterate	-1.031***	-0.136***
	(0.0329)	(0.0241)
Incomplete Fundamental Education I	-0.824***	-0.0462***
	(0.0171)	(0.0117)
Complete Fundamental Education I	-0.758***	-0.0979***
	(0.0133)	(0.00893)
Incomplete Fundamental Education II	-0.569***	-0.0963***
	(0.0114)	(0.00780)
Complete Fundamental Education II	-0.423***	-0.0537***
	(0.00924)	(0.00621)
Incomplete Secondary School	-0.284***	-0.0265***
	(0.0132)	(0.00713)
Incomplete Higher Education	0.541***	0.0981***
	(0.0204)	(0.0114)
Complete Higher Education	0.982***	0.266***
	(0.0101)	(0.00744)
Master	1.760***	0.525***
	(0.0652)	(0.0561)
Doctorate	1.807***	0.336***
	(0.102)	(0.0780)
Absolute MW × Illiterate	0.313***	0.0479**
	(0.0387)	(0.0233)
Absolute MW × Incomplete Fundamental Education I	0.273***	-0.0198*
	(0.0216)	(0.0112)
Absolute MW × Complete Fundamental Education I	0.150***	0.00653
	(0.0151)	(0.00799)
Absolute MW × Incomplete Fundamental Education II	0.00625	0.00524
	(0.0127)	(0.00689)
Absolute MW × Complete Fundamental Education II	-0.0178*	-0.0167***
	(0.0106)	(0.00634)
Absolute MW × Incomplete Secondary School	0.0856***	0.0172**
	(0.0152)	(0.00779)
Absolute MW × Incomplete Higher Education	-0.00773	-0.0387***
	(0.0231)	(0.0124)
Absolute MW × Complete Higher Education	-0.0207*	-0.0759***
	(0.0124)	(0.00719)
Absolute MW × Master	-0.266***	-0.112
	(0.0808)	(0.0690)
Absolute MW × Doctorate	-0.169	0.0516
	(0.150)	(0.0937)
Other controls[1]	Yes	Yes
Observations	898,321	898,321
Adjusted R-squared	0.622	0.900

Dependent variable: hourly wage. Omitted education variable: complete secondary school. *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.

Note: [1]Other controls: tenure, age, attained education, educational mismatch, firm size, sector, class of worker, state, year. In the OLS regression, I also control for gender.

Table B.3: Effects of the relative minimum wage on hourly wages

Variables	**(1) OLS**	**(2) FE**
Relative minimum wage	-0.207***	-0.0397***
	(0.0195)	(0.0149)
Illiterate	-0.945***	0.0481
	(0.0719)	(0.0462)
Incomplete Fundamental Education I	-0.456***	-0.0436**
	(0.0330)	(0.0200)
Complete Fundamental Education I	-0.690***	-0.0315**
	(0.0234)	(0.0156)
Incomplete Fundamental Education II	-0.608***	-0.0812***
	(0.0196)	(0.0113)
Complete Fundamental Education II	-0.610***	-0.0725***
	(0.0144)	(0.00976)
Incomplete Secondary School	-0.417***	-0.0218*
	(0.0191)	(0.0130)
Incomplete Higher Education	0.344***	-0.0692***
	(0.0274)	(0.0227)
Complete Higher Education	1.015***	0.137***
	(0.0104)	(0.0112)
Master	1.415***	0.322***
	(0.0633)	(0.0515)
Doctorate	1.564***	0.504***
	(0.163)	(0.133)
Relative MW × Illiterate	0.248**	-0.258***
	(0.118)	(0.0771)
Relative MW × Incomplete Fundamental Education I	-0.298***	-0.0311
	(0.0512)	(0.0316)
Relative MW × Complete Fundamental Education I	0.0766**	-0.103***
	(0.0367)	(0.0232)
Relative MW × Incomplete Fundamental Education II	0.0687**	-0.0183
	(0.0308)	(0.0167)
Relative MW × Complete Fundamental Education II	0.300***	0.0111
	(0.0237)	(0.0157)
Relative MW × Incomplete Secondary School	0.359***	0.0169
	(0.0319)	(0.0213)
Relative MW × Incomplete Higher Education	0.356***	0.252***
	(0.0477)	(0.0374)
Relative MW × Complete Higher Education	-0.101***	0.139***
	(0.0185)	(0.0186)
Relative MW × Master	0.289**	0.230**
	(0.118)	(0.0977)
Relative MW × Doctorate	0.244	-0.250
	(0.306)	(0.223)
Other controls[1]	Yes	Yes
Observations	898,321	898,321
Adjusted R-squared	0.623	0.900

Dependent variable: hourly wage. Omitted education variable: complete secondary school. *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.

Note: [1]Other controls: tenure, age, attained education, educational mismatch, firm size, sector, class of worker, state, year. In the OLS regression, I also control for gender.

Table B.4: Descriptives by gender

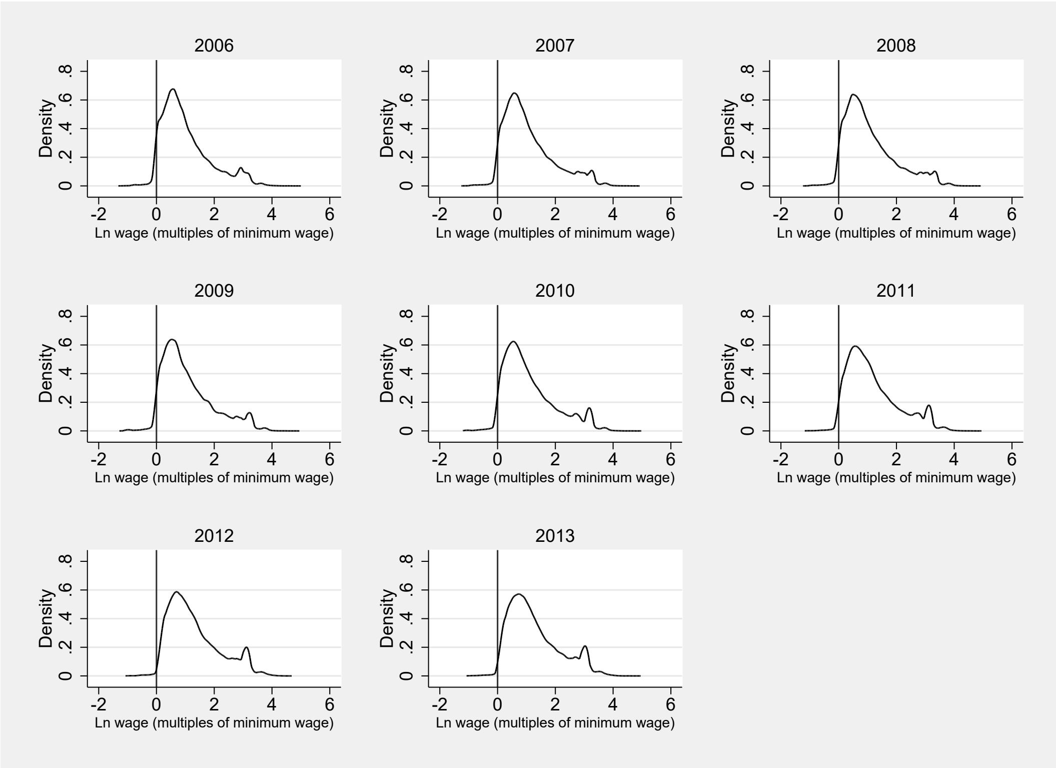
Variable	**Male**	**Female**	**t-statistic of mean comparison test**	**p-value**
Wage/hour (BRL)	10.96	9.01	51.61	0.0000
	(18.12)	(17.47)		
Tenure (months)	67.42	76.63	-50.71	0.0000
	(84.95)	(86.99)		
Age**	39.63	39.06	31.04	0.0000
	(8.76)	(8.51)		
Share that changed education (%)	7.34	5.96	22.20	0.0000
Share that changed firm (%)	17.94	15.95	21.31	0.0000
Share that changed occupation (%)	14.02	14.06	-0.4924	0.6224
Education (%)				
Illiterate	0.26	0.12	14.32	0.0000
Incomplete Fudamental Education I	3.01	1.45	48.17	0.0000
Complete Fundamental Education I	5.10	2.71	56.59	0.0000
Incomplete Fudamental Education II	9.59	5.16	78.01	0.0000
Complete Fundamental Education II	15.38	9.04	89.65	0.0000
Incomplete Secondary School	7.11	4.52	51.18	0.0000
Complete Secondary School	38.19	40.67	-23.92	0.0000
Incomplete Higher Education	3.39	4.86	-35.59	0.0000
Complete Higher Education	17.61	30.55	-150.00	0.0000
Master	0.28	0.83	-36.85	0.0000
Doctorate	0.08	0.09	0.00	0.0378
Educational mismatch (%)				
Undereducated	28.53	21.40	77.24	0.0000
Well-matched	41.13	49.68	-81.43	0.0000
Overeducated	30.34	29.92	14.58	0.0000
Firm size (%)				
Up to 19 employees	20.18	20.25	-0.89	0.3730
20-99 employees	21.53	19.85	19.49	0.0000
100-499 employees	23.10	20.04	34.89	0.0000
500-more employees	29.58	32.37	-28.59	0.0000
Sector (%)				
Industry	49.93	29.74	197.10	0.0000
Services	50.07	70.26	-200.00	0.0000

Source: RAIS, 2006-2013.

Notes: Linearized SE in parenthesis. ** Between 25 and 65 years.

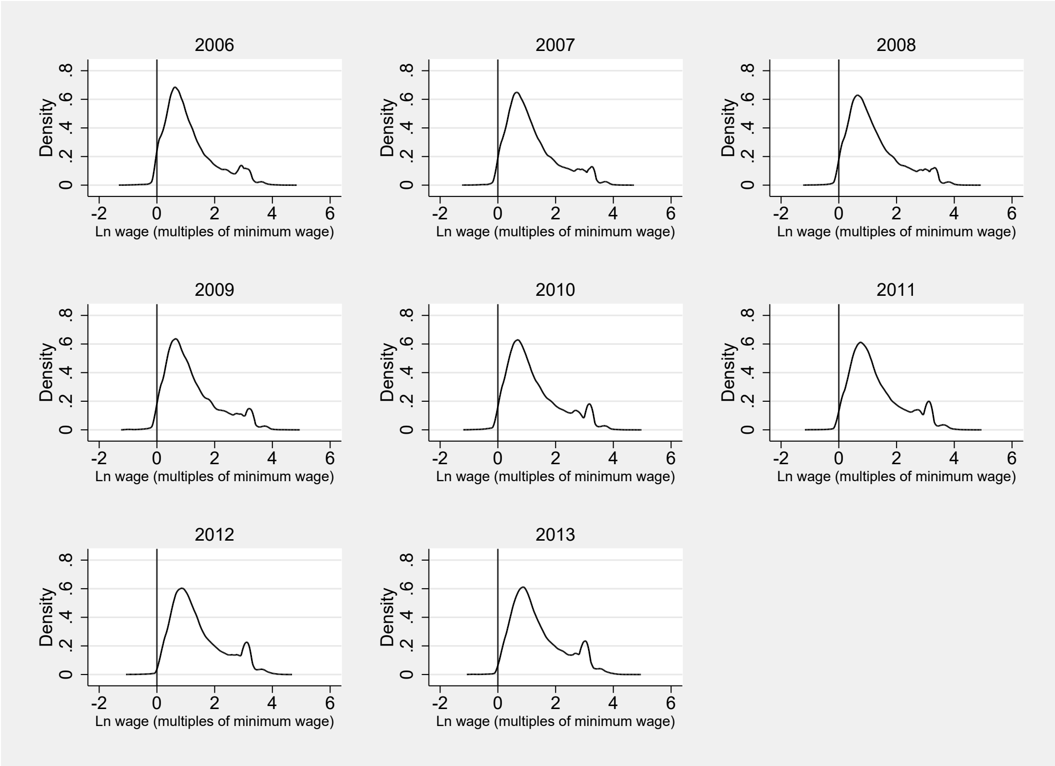
B.2 Other figures

Figure B.1: Earnings distribution by year – All workers



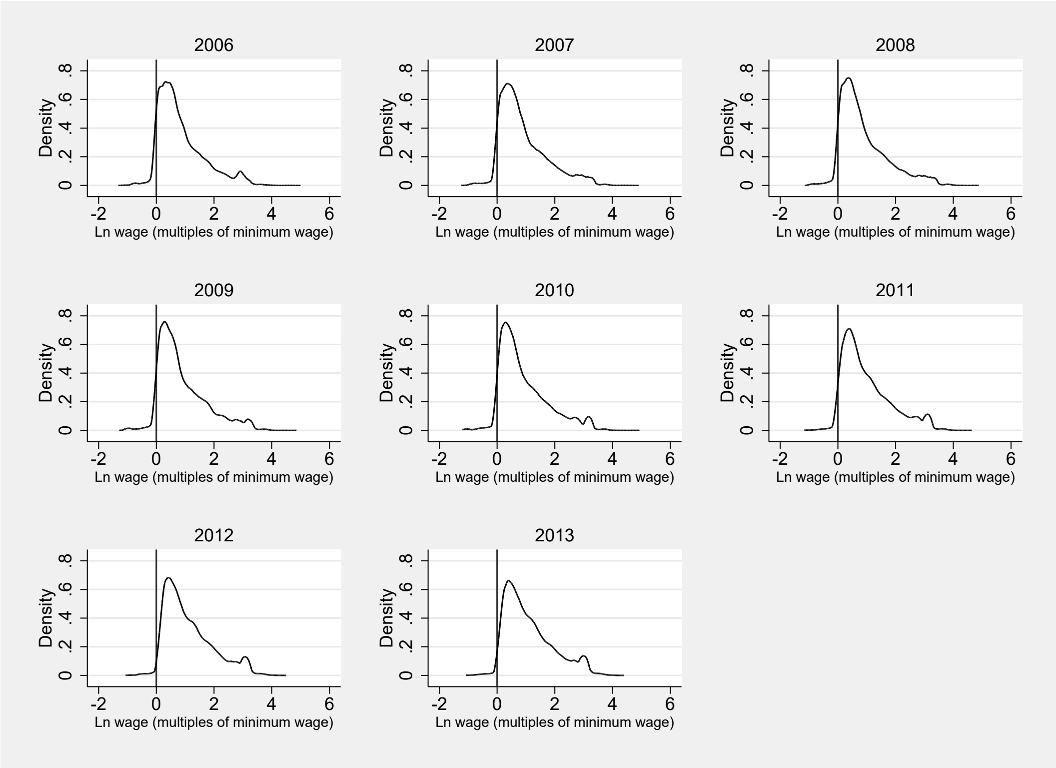
Source: RAIS, 2006-2013.

Figure B.2: Earnings distribution – Male



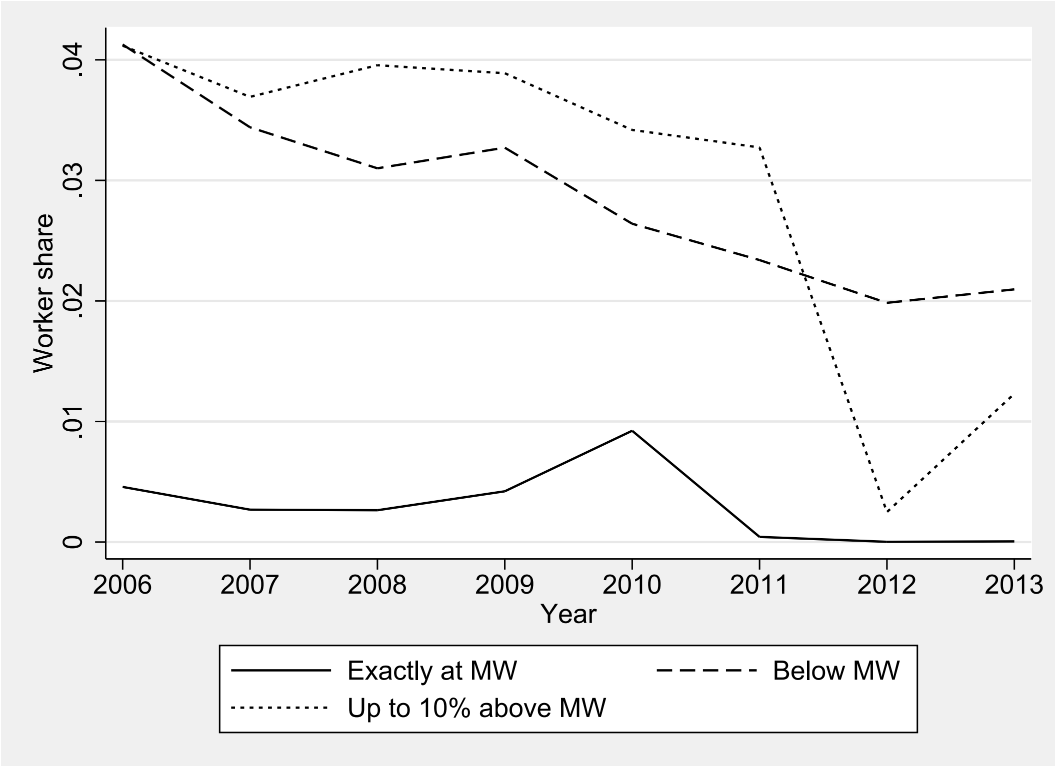
Source: RAIS, 2006-2013.

Figure B.3: Earnings distribution – Female



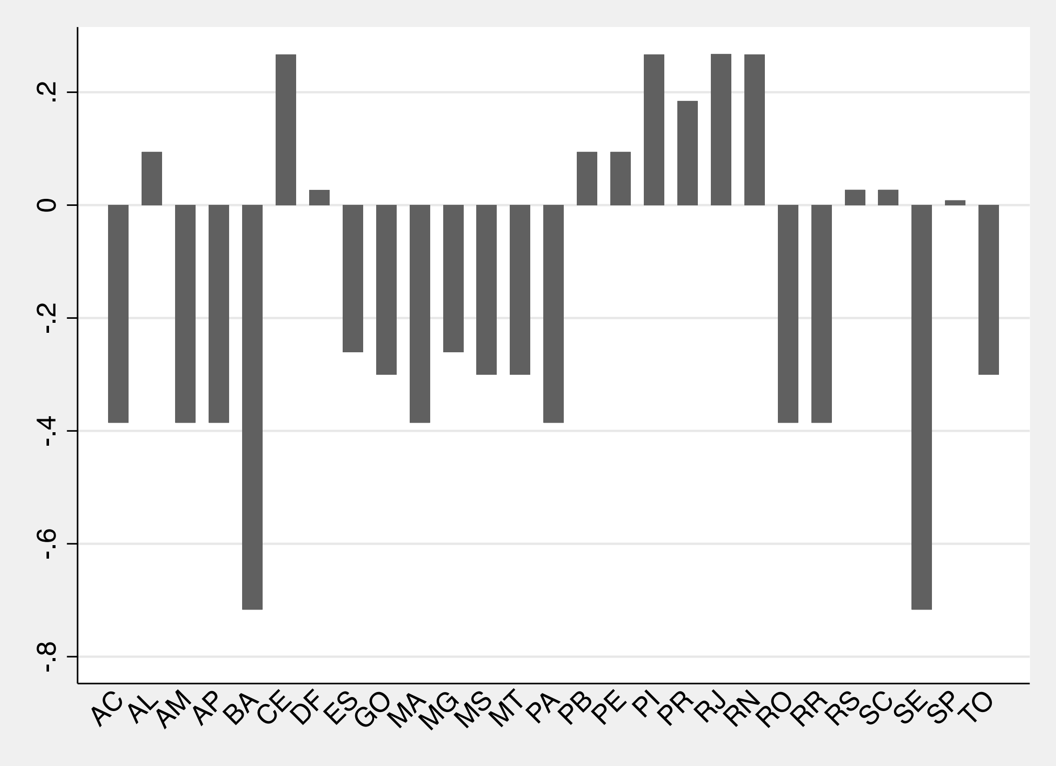
Source: RAIS, 2006-2013.

Figure B.4: Bindingness of the minimum wage, 2006-2013



Source: RAIS, 2006-2013.

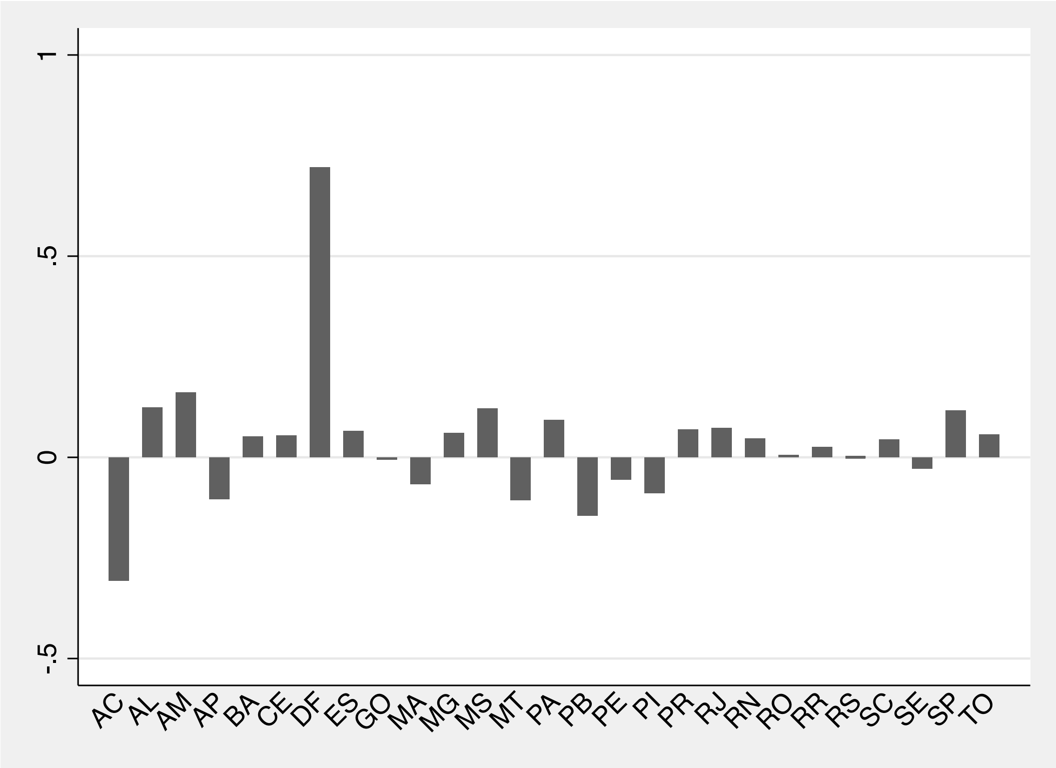
Figure B.5: Effects of changes in the absolute minimum wage on education acquisition by state (2006-2013), in percentages



Source: RAIS, 2006-2013.

Note: The graph report the estimated effect of changes in the absolute minimum wage index on the probability of acquiring education on-the-job over the period 2006-2013 for each Brazilian state.

Figure B.6: Effects of changes in the relative minimum wage on education acquisition by state (2006-2013), in percentages



Source: RAIS, 2006-2013.

Note: The graph report the estimated effect of changes in the relative minimum wage index on the probability of acquiring education on-the-job over the period 2006-2013 for each Brazilian state.

Appendix C

Appendix for Chapter 4

C.1 Other tables

Table C.1: Top ranked occupations in the states of Maranhão and São Paulo

	Maranhão		São Paulo	
Rank	CBO code	Title	CBO code	Title
1	2424	Public defenders	2423	Police chiefs
2	2251	Clinical physicians	2544	State and municipal tax inspectors
3	2140	Environmental engineers	2541	Federal tax inspectors
4	2142	Civil engineers	2153	Aircraft pilots
5	2543	Labour market tax auditors	2424	Public defenders
6	2143	Electrical and electronic engineers	2251	Clinical physicians
7	7101	Supervisors of mining extraction	2021	Automation control engineers
8	2144	Mechanical engineers	2252	Surgeons
9	2147	Mining engineers	2412	Public Attorneys
10	2410	Lawyers	2144	Mechanical engineers

Note: Rankings are calculated based on the average wage for the occupational family. The 1st in the rank means the highest-paying occupation.

Table C.2: Worst ranked occupations in the states of Maranhão and São Paulo

	Maranhão		São Paulo	
Rank	CBO code	Title	CBO code	Title
1	7733	Woodworking operators	7771	Shipwrights
2	7681	Manual weaving workers	6320	Forest workers
3	6224	Producers of flowers and ornamental plants	5167	Astrologers
4	5114	Tour guides	6228	Aromatic plants producers
5	7612	Wiring operators	6126	Stimulating plants producers
6	3112	Technicians in the chemical industry	5114	Tour guides
7	7731	Woodworking machine operators	8412	Salt workers
8	7522	Flat glass processing workers	5198	Sex workers
9	7831	Rail transport operators	7640	Footwear manufacturer workers
10	2627	Interpret musicians	6322	Forest extractors to produce gums and resin

Note: Rankings are calculated based on the average wage for the occupational family. The 1st in the rank means the lowest-paying occupation.

Table C.3: OLS estimation

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	OLS	OLS	OLS	OLS	OLS
Migrant	0.0551***	0.0417***	0.0231***	0.0435***	0.0150***
	(0.00272)	(0.00172)	(0.00133)	(0.00116)	(0.00126)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	no	no	no	no	no
R-squared	0.558	0.259	0.208	0.144	0.295
Observations	909,616	909,616	909,616	909,616	909,616

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parenthesis clustered by individual.

Notes: Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Table C.4: IV estimation

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	IV	IV	IV	IV	IV
First stage					
Ln homicide rate origin t-1 (deviation)	0.049***	0.049***	0.049***	0.049***	0.049***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
First stage F statistics	214.60	215.61	215.61	215.61	215.61
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000
Second stage					
Migrant	-0.368***	0.267***	-0.200***	0.326***	0.0207
	(0.0548)	(0.0482)	(0.0486)	(0.0408)	(0.0360)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	no	no	no	no	no
R-squared	0.241	0.209	0.104	0.024	0.185
Observations	909,616	909,616	909,616	909,616	909,616

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parenthesis clustered by individual.

Notes: All specifications use IV-2SLS with the deviation of the homicide rate in the state of origin in t-1 as an instrument for migrant. Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Table C.5: FEIV estimation - Smaller sample

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	FEIV	FEIV	FEIV	FEIV	FEIV
First stage					
Ln homicide rate origin t-1 (deviation)	0.013***	0.013***	0.013***	0.013***	0.013***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
First stage F statistics	37.94	38.08	37.38	37.38	37.79
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000
Second stage					
Migrant	-0.740***	0.742***	-0.397*	-0.234	-0.125
	(0.255)	(0.266)	(0.203)	(0.179)	(0.190)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	yes	yes	yes	yes	yes
R-squared	0.221	0.288	0.396	0.252	0.211
Observations	627,902	627,902	627,902	627,902	627,902

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.

Notes: All specifications use IV-2SLS with the deviation of the homicide rate in the state of origin in t-1 as an instrument for migrant. Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Table C.6: FEIV estimation - Alternative rank definition

	Sextiles		Deciles	
	Higher rank	Lower rank	Higher rank	Lower rank
	(1)	(2)	(3)	(4)
Variables	FEIV	FEIV	FEIV	FEIV
First stage				
Ln homicide rate origin t-1 (deviation)	0.015***	0.015***	0.015***	0.015***
	(0.002)	(0.002)	(0.002)	(0.002)
First stage F statistics	68.04	68.04	68.04	68.04
Prob > F	0.0000	0.0000	0.0000	0.0000
Second stage				
Migrant	-0.390**	-0.0374	-0.421***	0.162
	(0.152)	(0.133)	(0.154)	(0.140)
Other controls	yes	yes	yes	yes
State dummies	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes
Worker fixed effects	yes	yes	yes	yes
R-squared	0.413	0.284	0.396	0.252
Observations	909,616	909,616	909,616	909,616

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parenthesis clustered by individual.

Notes: All specifications use IV-2SLS with the deviation of the homicide rate in the state of origin in t-1 as an instrument for migrant. Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Table C.7: FEIV estimation - IV: Soy exports deviation (negative shock)

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	FEIV	FEIV	FEIV	FEIV	FEIV
First stage					
Ln soy exports rate origin t-1 (deviation)	0.007***	0.007***	0.007***	0.007***	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
First stage F statistics	16.23	16.15	16.20	16.20	16.21
Prob > F	0.0001	0.0001	0.0001	0.0001	0.0001
Second stage					
Migrant	0.808**	-0.0558	-0.306	0.0119	-0.129
	(0.350)	(0.362)	(0.298)	(0.248)	(0.265)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	yes	yes	yes	yes	yes
R-squared	0.257	0.347	0.407	0.264	0.207
Observations	172,652	172,652	172,652	172,652	172,652

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.

Notes: All specifications use IV-2SLS with the deviation of the soy exports rate in the state of origin in t-1 as an instrument for migrant. Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Table C.8: FEIV estimation - Soy exporting states only

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	FEIV	FEIV	FEIV	FEIV	FEIV
First stage					
Ln homicide rate origin t-1 (deviation)	0.043***	0.043***	0.043***	0.043***	0.043***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
First stage F statistics	38.00	37.92	37.97	37.97	38.05
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000
Second stage					
Migrant	0.877***	-0.371	-0.468**	0.111	-0.270
	(0.282)	(0.265)	(0.222)	(0.181)	(0.211)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	yes	yes	yes	yes	yes
R-squared	0.242	0.336	0.392	0.262	0.197
Observations	172,652	172,652	172,652	172,652	172,652

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parenthesis clustered by individual.

Notes: All specifications use IV-2SLS with the deviation of the homicide rate in the state of origin in t-1 as an instrument for migrant. Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Table C.9: FEIV estimation - Excluding soy exporting states

	Ln hourly wage	Occupation	Higher rank	Lower rank	Well-matched
	(1)	(2)	(3)	(4)	(5)
Variables	FEIV	FEIV	FEIV	FEIV	FEIV
First stage					
Ln homicide rate origin t-1 (deviation)	0.007***	0.007***	0.007***	0.007***	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
First stage F statistics	15.12	15.20	14.67	14.67	15.04
Prob > F	0.0001	0.0001	0.0001	0.0001	0.0001
Second stage					
Migrant	-2.169***	1.647***	0.0667	-0.765**	0.168
	(0.676)	(0.582)	(0.317)	(0.352)	(0.310)
Other controls	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes
Occupation fixed effects	yes	yes	yes	yes	yes
Worker fixed effects	yes	yes	yes	yes	yes
R-squared	-0.566	0.105	0.413	0.119	0.217
Observations	727,709	727,709	727,709	727,709	727,709

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parenthesis clustered by individual.

Notes: All specifications use IV-2SLS with the deviation of the homicide rate in the state of origin in t-1 as an instrument for migrant. Other controls include tenure, age, education, firm size, sector. Column (1) also contains a dummy variable for over and undereducation. Columns (2), (3), (4) and (5) include a dummy variable for the lag of over and undereducation.

Table C.10: Frequency of migrants

	State in t																										
State in t-1	**RO**	**AC**	**AM**	**RR**	**PA**	**AP**	**TO**	**MA**	**PI**	**CE**	**RN**	**PB**	**PE**	**AL**	**SE**	**BA**	**MG**	**ES**	**RJ**	**SP**	**PR**	**SC**	**RS**	**MS**	**MT**	**GO**	**DF**
RO	0	17	30	3	21	1	0	5	2	2	1	4	4	0	2	5	23	3	13	43	34	13	7	19	73	18	18
AC	24	0	11	1	4	0	0	1	0	1	1	1	0	0	0	1	4	2	7	13	3	1	6	3	8	7	7
AM	26	2	0	29	70	6	3	7	1	19	4	7	16	2	4	12	30	3	53	155	15	9	10	5	5	17	12
RR	0	0	22	0	8	1	1	2	1	3	0	0	2	0	0	0	2	0	4	9	0	1	0	1	0	4	6
PA	26	2	95	14	0	52	48	174	18	36	3	4	40	4	7	39	140	12	97	283	29	19	11	9	49	117	51
AP	2	0	5	1	36	0	1	3	0	4	0	1	1	0	0	2	9	0	7	17	2	2	1	2	3	2	6
TO	5	0	4	2	47	1	0	42	2	3	0	1	9	3	1	11	23	4	16	44	5	3	3	3	16	125	38
MA	14	2	11	6	144	5	44	0	75	51	8	1	48	5	1	24	61	9	49	154	24	15	4	8	31	101	42
PI	2	0	3	0	19	1	6	99	0	43	4	12	23	0	0	12	12	5	14	117	5	3	5	4	11	13	46
CE	6	1	17	5	20	1	3	37	32	0	62	41	140	13	11	102	42	5	103	362	15	49	14	8	11	15	46
RN	2	0	10	0	8	1	4	19	22	92	0	52	99	11	12	40	32	8	62	201	9	8	13	1	17	13	25
PB	0	0	1	0	6	0	3	7	6	29	45	0	109	12	6	10	11	4	76	146	10	7	3	0	1	7	13
PE	7	0	12	1	29	1	2	36	16	112	72	118	0	97	22	189	54	18	150	492	34	27	14	4	21	34	46
AL	4	0	1	0	8	0	2	5	0	19	4	17	125	0	28	44	70	7	30	288	22	9	5	20	27	27	7
SE	0	0	4	0	4	0	0	1	2	12	10	5	29	32	0	111	14	2	25	76	7	6	8	4	4	7	8
BA	6	2	9	0	24	2	7	22	8	80	29	17	195	23	110	0	222	112	203	921	56	44	41	10	11	75	83
MG	34	12	50	2	133	9	50	69	23	58	22	29	77	32	10	236	0	292	649	2004	175	121	111	62	70	250	219
ES	3	0	2	1	7	0	1	13	4	1	6	5	21	2	3	68	208	0	211	202	29	16	11	6	4	5	19
RJ	14	4	27	6	57	2	13	47	16	75	28	61	119	15	16	157	441	173	0	1393	104	76	88	25	27	79	101
SP	47	10	123	6	162	6	32	169	91	265	113	127	402	94	63	620	1472	186	1474	0	1017	429	339	264	167	376	528
PR	12	5	9	0	17	0	6	9	3	10	3	9	17	5	1	37	97	19	115	947	0	560	165	77	71	40	34
SC	15	0	6	0	5	2	5	15	5	26	3	6	19	1	1	35	61	11	76	409	556	0	344	23	25	37	22
RS	8	4	6	0	8	1	8	6	2	17	5	7	16	5	3	21	62	9	83	382	144	447	0	12	19	37	27
MS	13	4	0	0	7	0	5	5	2	4	2	0	7	3	2	3	48	6	27	293	88	31	22	0	57	31	20
MT	52	3	6	1	37	3	13	24	15	8	1	2	11	10	2	10	44	3	36	222	96	30	22	44	0	103	46
GO	18	5	4	1	59	2	85	50	7	18	7	6	24	4	2	50	188	7	70	420	39	40	20	32	107	0	436
DF	22	7	9	3	32	7	26	42	25	30	3	5	53	1	6	117	179	16	130	385	31	38	18	14	28	419	0

Table C.11: Correlation between crime rate IV and labour market outcomes

	Crime rate	Wage/h	Occupation	Higher rank	Lower rank	Well-matched
Crime rate	1					
Wage/h	-0.1235	1				
Occupation	-0.0326	-0.0342	1			
Higher rank	-0.0208	0.0201	0.5919	1		
Lower rank	-0.0153	-0.0488	0.4187	-0.127	1	
Well-matched	-0.0235	0.0272	0.5298	0.4282	0.1293	1

Note: Crime rate IV is the deviation of the homicide rate in the state of origin in t-1.

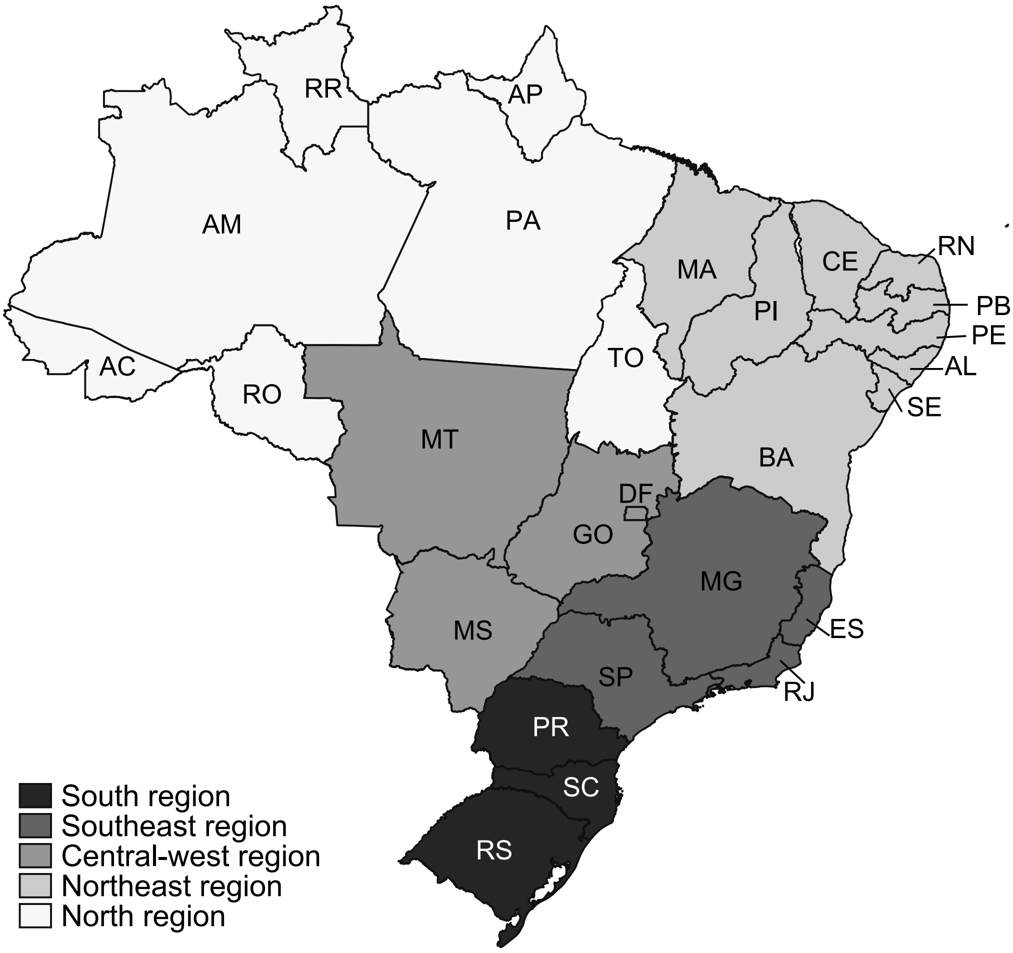
Table C.12: Moran's I for the crime rate instrument variable

	2006	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
Moran's I	0.12562	0.09161	0.08407	0.08441	0.13235	0.19287	0.22396
E(I)	-0.03846	-0.03846	-0.03846	-0.03846	-0.03846	-0.03846	-0.03846
SE(I)	0.10114	0.09995	0.10067	0.10054	0.10033	0.0995	0.09889
Z(I)	1.62234	1.30133	1.21714	1.2221	1.70253	2.32504	2.65373
p-value	0.10473	0.19314	0.22355	0.22167	0.08866	0.02007	0.00796

Note: H_0: Spatial Randomization

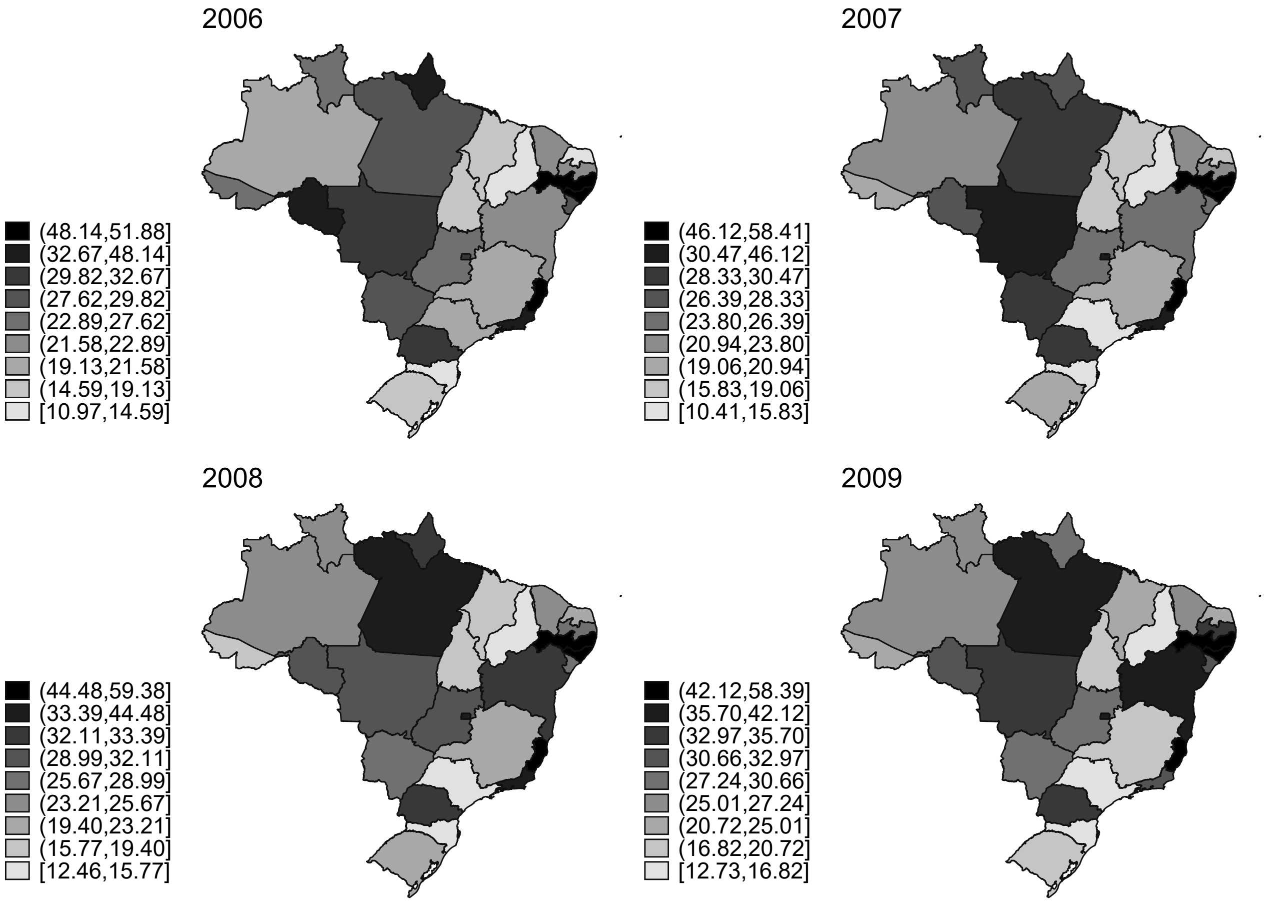
C.2 Other figures

Figure C.1: Map of Brazilian states and regions



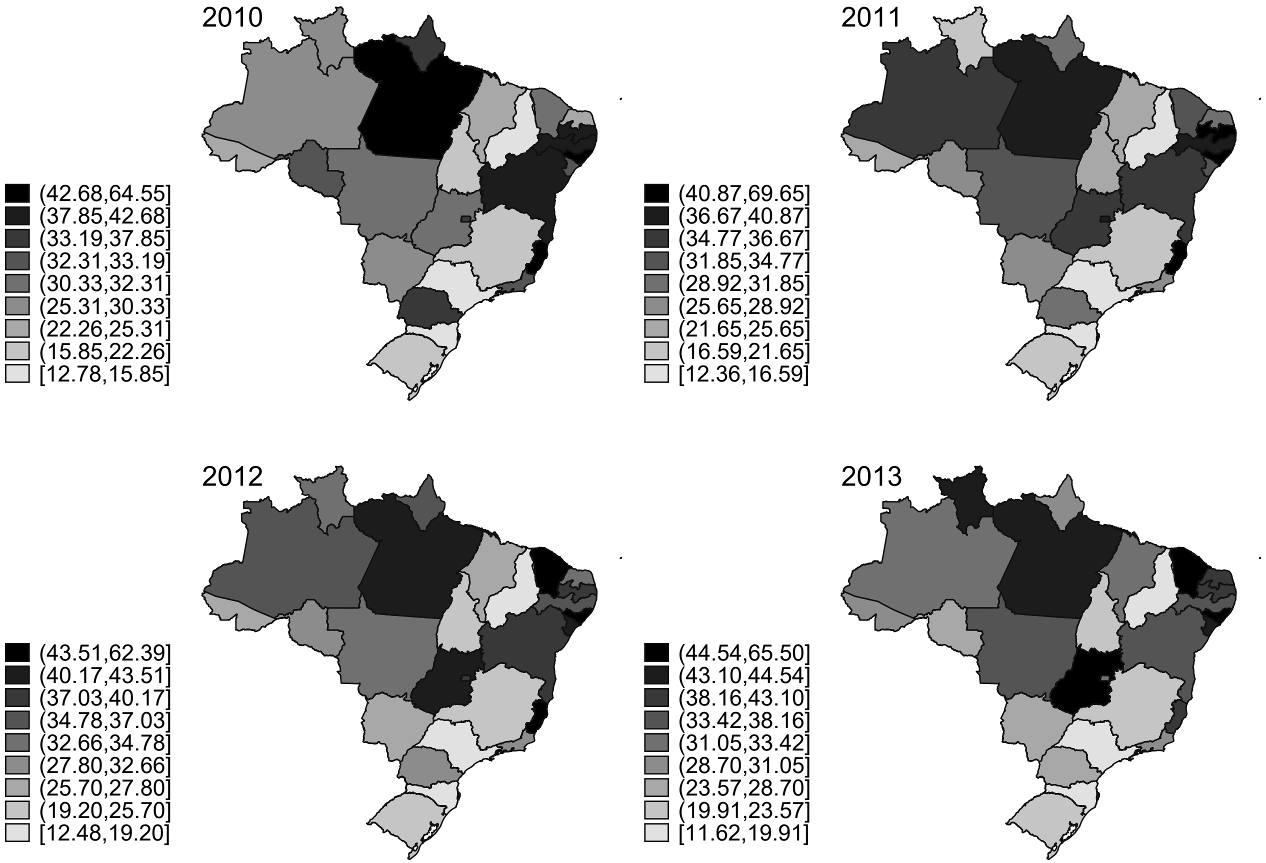
States: Acre (AC), Amazonas (AM), Rondônia (RO), Roraima (RR), Amapá (AP), Pará (PA), Tocantins (TO), Maranhão (MA), Piauí (PI), Ceará (CE), Rio Grande do Norte (RN), Paraíba (PB), Pernambuco (PE), Alagoas (AL), Sergipe (SE), Bahia (BA), Mato Grosso (MT), Goiás (GO), Distrito Federal (DF), Mato Grosso do Sul (MS), Minas Gerais (MG), Espírito Santo (ES), Rio de Janeiro (RJ), São Paulo (SP), Paraná (PR), Santa Catarina (SC), Rio Grande do Sul (RS).

Figure C.2: Map of the homicide rate in the Brazilian states, 2006-2009



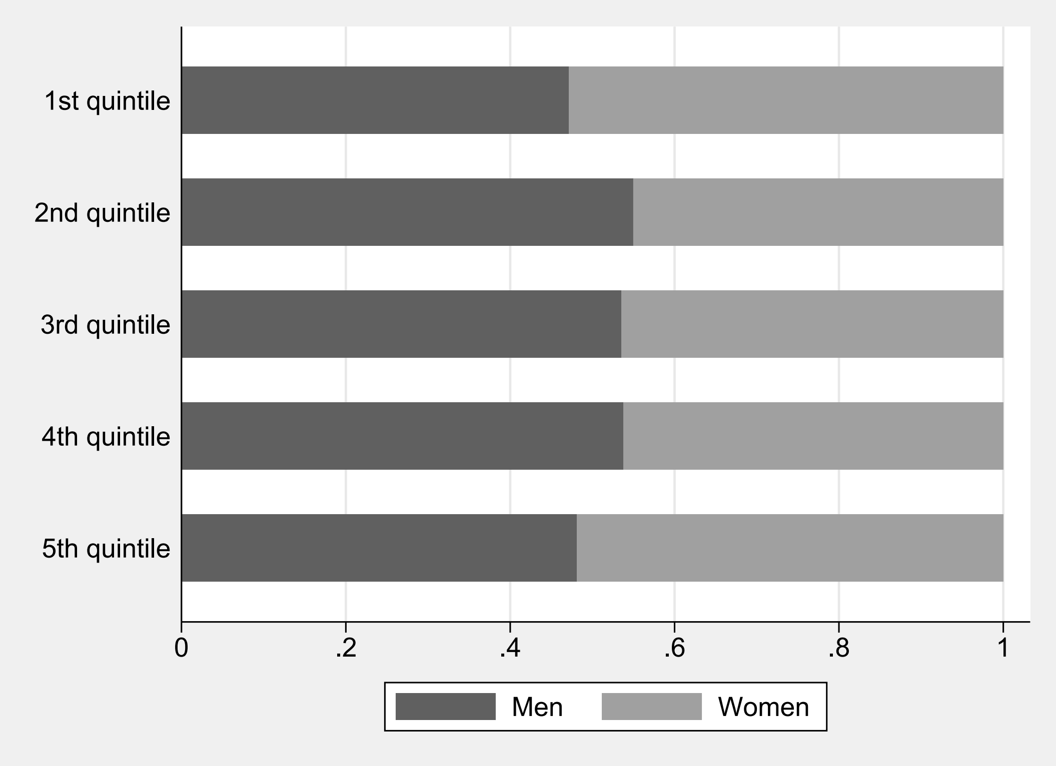
Source: DATASUS, 2006-2009.

Figure C.3: Map of the homicide rate in the Brazilian states, 2010-2013



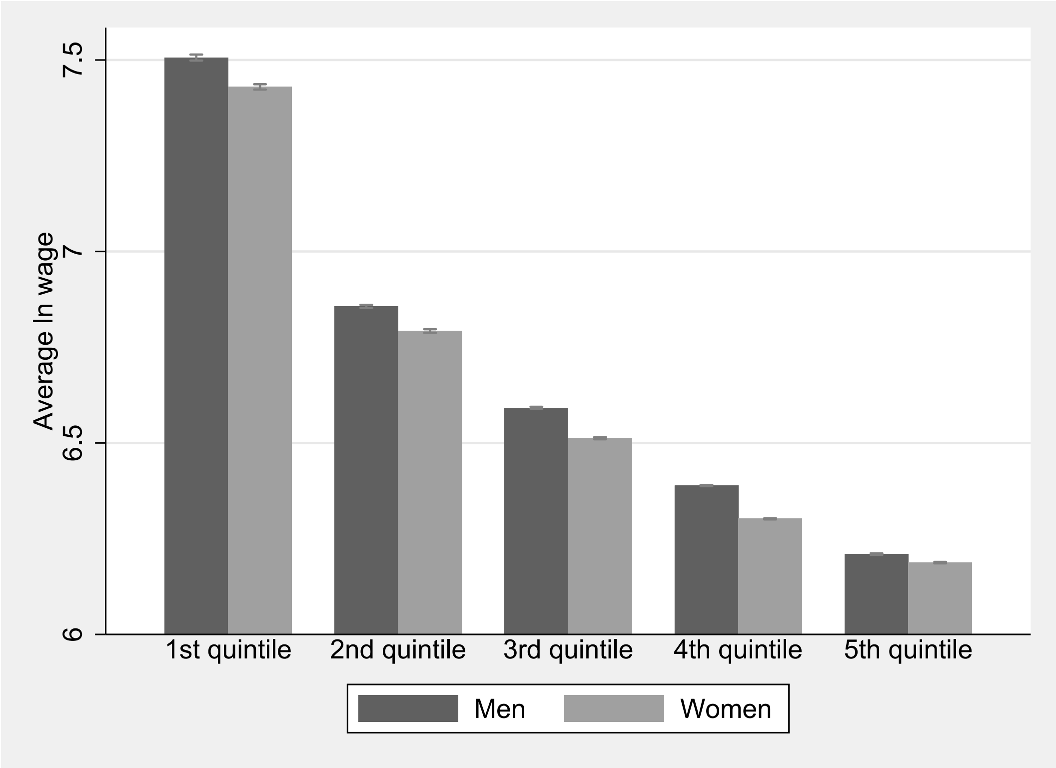
Source: DATASUS, 2010-2013.

Figure C.4: Share of workers by gender and occupational rank



Source: RAIS, 2006-2013.

Figure C.5: Average wage by gender and occupational rank



Source: RAIS, 2006-2013.

Bibliography

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. (2017). When should you adjust standard errors for clustering? Technical report, National Bureau of Economic Research.

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). Minimum wages and employment in France and the United States. Technical report, National Bureau of Economic Research.

Abramitzky, R. (2009). The effect of redistribution on migration: Evidence from the israeli kibbutz. *Journal of Public Economics*, 93(3-4):498–511.

Abramitzky, R., Boustan, L. P., and Eriksson, K. (2012). Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review*, 102(5):1832–56.

Acemoglu, D. and Pischke, J.-S. (1999). The structure of wages and investment in general training. *Journal of Political Economy*, 107(3):539–572.

Acemoglu, D. and Pischke, J.-S. (2003). Minimum wages and on-the-job training. In *Worker Well-Being and Public Policy*, pages 159–202. Emerald Group Publishing Limited.

Aguayo-Tellez, E., Muendler, M.-A., and Poole, J. P. (2010). Globalization and formal sector migration in Brazil. *World development*, 38(6):840–856.

Akgündüz, Y. E., van den Berg, M., and Hassink, W. (2018). The impact of the Syrian refugee crisis on firm entry and performance in Turkey. *The World Bank Economic Review*, 32(1):19–40.

Alba-Ramirez, A. (1993). Mismatch in the spanish labor market: overeducation? *Journal of Human Resources*, pages 259–278.

Alba-Ramírez, A. and Blázquez, M. (2004). Types of job match, overeducation and labour mobility in Spain. In Büchel, F., De Grip, A., and Mertens, A., editors, *Overeducation in Europe: Current issues in theory and policy*, pages 65–90. Edward Elgar: Cheltenham.

Albrecht, J. and Vroman, S. (2002). A matching model with endogenous skill requirements. *International Economic Review*, 43(1):283–305.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Antman, F. M. (2011). The intergenerational effects of paternal migration on schooling and work: What can we learn from children's time allocations? *Journal of Development Economics*, 96(2):200–208.

Arulampalam, W., Booth, A. L., and Bryan, M. L. (2004). Training and the new minimum wage. *The Economic Journal*, 114(494).

Ashenfelter, O., Harmon, C., and Oosterbeek, H. (1999). A review of estimates of the schooling/earnings relationship, with tests for publication bias. *Labour Economics*, 6(4):453–470.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2005). Rising wage inequality: the role of composition and prices. *NBER working paper series*, 11628.

Baker, S. R. (2015). Effects of immigrant legalization on crime. *American Economic Review*, 105(5):210–13.

Bárány, Z. L. (2016). The minimum wage and inequality: the effects of education and technology. *Journal of Labor Economics*, 34(1):237–274.

Barbosa, A. d. F., Cacciamali, M. C., Rodgers, G., and Tatei, F. (2015). Minimum wage in Brazil: a useful policy tool to reduce wage inequality? *Global Wage Debates: Politics or Economics?*, pages 11–27.

Barros, R., De Carvalho, M., Franco, S., and Mendonça, R. (2010). Markets, the State, and the Dynamics of Inequality in Brazil. *Declining inequality in Latin America: A decade of progress*, pages 134–74.

Barros, R. P. d., Mendonça, R., dos Santos, D. D., and Quintaes, G. (2001). Determinantes do desempenho educacional no Brasil. *Pesquisa e Planejamento Econômico*, 31(1).

Bartolucci, C., Villosio, C., and Wagner, M. (2018). Who migrates and why? evidence from italian administrative data. *Journal of Labor Economics*, 36(2):551–588.

Bauer, T. K. (2002). Educational mismatch and wages: a panel analysis. *Economics of Education Review*, 21(3):221–229.

Becker, G. S. (1964). Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*.

Becker, G. S. and Chiswick, B. R. (1966). Education and the distribution of earnings. *The American Economic Review*, 56(1/2):358–369.

Becker, S. O. and Ferrara, A. (2019). Consequences of forced migration: A survey of recent findings. *Labour Economics*.

Blanchard, O. J., Diamond, P., Hall, R. E., and Murphy, K. (1990). The cyclical behavior of the gross flows of US workers. *Brookings Papers on Economic Activity*, 1990(2):85–155.

Blattman, C. and Annan, J. (2016). Can employment reduce lawlessness and rebellion? A field experiment with high-risk men in a fragile state. *American Political Science Review*, 110(1):1–17.

Borjas, G. J. (1985). Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of Labor Economics*, 3(4):463–489.

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *The American Economic Review*, pages 531–553.

Borjas, G. J. (2014). *Immigration economics*. Harvard University Press.

Borjas, G. J., Bronars, S. G., and Trejo, S. J. (1992). Self-selection and internal migration in the United States. *Journal of Urban Economics*, 32(2):159–185.

Borjas, G. J., Kauppinen, I., and Poutvaara, P. (2018). Self-selection of emigrants: Theory and evidence on stochastic dominance in observable and unobservable characteristics. *The Economic Journal*, 129(617):143–171.

Brown, C., Gilroy, C., and Kohen, A. (1982). The effect of the minimum wage on employment and unemployment. *Journal of Economic Literature*, 20(2):487–528.

Bruns, B., Evans, D., and Luque, J. (2011). *Achieving world-class education in Brazil: The next agenda*. World Bank Publications.

Büchel, F. and Van Ham, M. (2003). Overeducation, regional labor markets, and spatial flexibility. *Journal of Urban Economics*, 53(3):482–493.

Burdett, K. and Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, pages 257–273.

Butcher, T., Dickens, R., and Manning, A. (2012). Minimum wages and wage inequality: some theory and an application to the UK.

Cahuc, P. and Michel, P. (1996). Minimum wage unemployment and growth. *European Economic Review*, 40(7):1463–1482.

Calderón-Mejía, V. and Ibáñez, A. (2016). Labour market effects of migration-related supply shocks: evidence from internal refugees in Colombia. *Journal of Economic Geography*, 16(3):695–713.

Camargo, J. M. (1984). Minimum wage in Brazil theory, policy and empirical evidence. Technical Report 67, Working Paper, Pontíficia Universidade Católica do Rio de Janeiro.

Card, D. (1992). Do minimum wages reduce employment? A case study of California, 1987–89. *Industrial Labour Relations Review*, 46(1):38–54.

Card, D. and Krueger, A. B. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *The American Economic Review*, pages 772–793.

Card, D. and Krueger, A. B. (2000). Minimum wages and employment: a case study of the fast-food industry in new jersey and pennsylvania: reply. *American Economic Review*, 90(5):1397–1420.

Carroll, D. and Tani, M. (2013). Over-education of recent higher education graduates: new Australian panel evidence. *Economics of Education Review*, 32:207–218.

Cerqueira, D. (2012). Mortes violentas não esclarecidas e impunidade no Rio de Janeiro. *Economia Aplicada*, 16(2):201–235.

CGEE (2016). Mestres e doutores 2015: Estudos da demografia da base técnico-científica brasileira. *Brasília, DF: Centro de Gestão e Estudos Estratégicos.*

Chairassamee, N. (2018). Crimes and moving decision in the United States: A conditional logit approach. *Applied Economics Journal*, 25(1):1–14.

Chiquiar, D. and Hanson, G. H. (2005). International migration, self-selection, and the distribution of wages: Evidence from mexico and the united states. *Journal of political Economy*, 113(2):239–281.

Chiswick, B. R. (1978). The effect of Americanization on the earnings of foreign-born men. *Journal of Political Economy*, 86(5):897–921.

Corseuil, C. H. L. and Foguel, M. N. (2016). Expansão econômica e aumento da formalização das relações de trabalho: uma abordagem através das contratações. In Barbosa Filho, F. d. H., Ulyssea, G., and Veloso, F., editors, *Causas e consequências da informalidade no Brasil*, pages 131–152. Elsevier, Rio de Janeiro.

Corseuil, C. H. L., Foguel, M. N., Gonzaga, G., and Ribeiro, E. P. (2014). A rotatividade dos jovens no mercado de trabalho formal brasileiro. Technical report, Instituto de Pesquisa Econômica Aplicada (IPEA).

Cullen, J. B. and Levitt, S. D. (1999). Crime, urban flight, and the consequences for cities. *Review of Economics and Statistics*, pages 159–169.

Davis, S. J. and Haltiwanger, J. (1992). Gross job creation, gross job destruction, and employment reallocation. *The Quarterly Journal of Economics*, 107(3):819–863.

Dedecca, C. S. (2015). A redução da desigualdade e seus desafios. Technical report, Texto para Discussão, Instituto de Pesquisa Econômica Aplicada (IPEA).

Détang-Dessendre, C., Drapier, C., and Jayet, H. (2004). The impact of migration on wages: empirical evidence from French youth. *Journal of Regional Science*, 44(4):661–691.

Di Pietro, G. and Urwin, P. (2006). Education and skills mismatch in the Italian graduate labour market. *Applied Economics*, 38(1):79–93.

Diaz, M. D. M. and Machado, L. (2008). Overeducation e undereducation no Brasil: incidência e retornos. *Estudos Econômicos (São Paulo)*, 38(3):431–460.

DiNardo, J., Fortin, N. M., and Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica*, 64(5):1001–1044.

Dix-Carneiro, R., Soares, R. R., and Ulyssea, G. (2018). Economic shocks and crime: Evidence from the Brazilian trade liberalization. *American Economic Journal: Applied Economics*, 10(4):158–95.

Dolton, P., Bondibene, C. R., and Stops, M. (2015). Identifying the employment effect of invoking and changing the minimum wage: A spatial analysis of the UK. *Labour Economics*, 37:54–76.

Dolton, P., Bondibene, C. R., and Wadsworth, J. (2012). Employment, Inequality and the UK National Minimum Wage over the Medium-Term. *Oxford Bulletin of Economics and Statistics*, 74(1):78–106.

Dolton, P. and Vignoles, A. (2000). The incidence and effects of overeducation in the UK graduate labour market. *Economics of Education Review*, 19(2):179–198.

Dolton, P. J. and Silles, M. A. (2008). The effects of over-education on earnings in the graduate labour market. *Economics of Education Review*, 27(2):125–139.

Duncan, G. J. and Hoffman, S. D. (1981). The Incidence and Wage Effects of Overeducation. *Economics of Education Review*, 1(1):75–86.

Dustmann, C., Fasani, F., Frattini, T., Minale, L., and Schönberg, U. (2017). On the economics and politics of refugee migration. *Economic Policy*, 32(91):497–550.

Ehrenberg, R. G. and Marcus, A. J. (1982). Minimum wages and teenagers' enrollment-employment outcomes: A multinomial logit model. *Journal of Human Resources*, pages 39–58.

Engbom, N. and Moser, C. (2017). Earnings Inequality and the Minimum Wage: Evidence from Brazil. Technical report, CESifo Working Paper.

Engel, S. and Ibáñez, A. M. (2007). Displacement due to violence in Colombia:A household-level analysis. *Economic Development and Cultural Change*, 55(2):335–365.

Enrico, M. (2011). Local labor markets. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor economics*, volume 4, pages 1237–1313. Elsevier.

Esteves, L. A. (2009). Incompatibilidade escolaridade-ocupação e salários: evidências de uma empresa industrial brasileira. *Revista Brasileira de Economia*, 63(2):77–90.

Fairris, D. and Pedace, R. (2004). The impact of minimum wages on job training: an empirical exploration with establishment data. *Southern Economic Journal*, pages 566–583.

Fiess, N. M. and Verner, D. (2003). *Migration and human capital in Brazil during the 1990s*. The World Bank.

Firpo, S. P. and Pieri, R. G. d. (2018). The labor market in Brazil, 2001-2015. *IZA World of Labor*, (441).

Freeman, R. (1976). *The Overeducated American*. New York: Academic Press.

Frenette, M. (2004). The overqualified Canadian graduate: the role of the academic program in the incidence, persistence, and economic returns to overqualification. *Economics of Education Review*, 23(1):29–45.

Gabriel, P. E. and Schmitz, S. (1995). Favorable self-selection and the internal migration of young white males in the United States. *Journal of Human Resources*, pages 460–471.

Gasparini, L., Galiani, S., Cruces, G., and Acosta, P. (2011). Educational upgrading and returns to skills in Latin America. *Documentos de Trabajo del CEDLAS*.

Gautier, P. A. (2002). Unemployment and search externalities in a model with heterogeneous jobs and workers. *Economica*, 69(273):21–40.

Giuliano, L. (2013). Minimum wage effects on employment, substitution, and the teenage labor supply: Evidence from personnel data. *Journal of Labor Economics*, 31(1):155–194.

Goldin, C. and Katz, L. F. (2002). The power of the pill: Oral contraceptives and women's career and marriage decisions. *Journal of Political Economy*, 110(4):730–770.

Gonzaga, G. and Cayres Pinto, R. (2014). Rotatividade do trabalho e incentivos da legislação trabalhista. In Veloso, F. and Bonelli, R., editors, *Panorama do mercado de trabalho no Brasil*. Editora FGV, Rio de Janeiro: IBRE.

Gonzaga, G., Maloney, W. F., and Mizala, A. (2003). Labor turnover and labor legislation in Brazil. *Economia*, 4(1):165–222.

Graham, D. H. (1970). Divergent and convergent regional economic growth and internal migration in Brazil: 1940-1960. *Economic Development and Cultural Change*, 18(3):362–382.

Green, F. and McIntosh, S. (2007). Is there a genuine under-utilization of skills amongst the over-qualified? *Applied Economics*, 39(4):427–439.

Groeneveld, S. and Hartog, J. (2004). Overeducation, wages and promotions within the firm. *Labour Economics*, 11(6):701–714.

Groot, W. and Van Den Brink, H. M. (2000). Overeducation in the labor market: a meta-analysis. *Economics of Education Review*, 19(2):149–158.

Grossberg, A. J. and Sicilian, P. (1999). Minimum wages, on-the-job training, and wage growth. *Southern Economic Journal*, pages 539–556.

Haltiwanger, J., Hyatt, H., and McEntarfer, E. (2018). Who moves up the job ladder? *Journal of Labor Economics*, 36(S1):S301–S336.

Hara, H. (2017). Minimum wage effects on firm-provided and worker-initiated training. *Labour Economics*.

Hartog, J. (2000). Over-education and earnings: where are we, where should we go? *Economics of Education Review*, 19(2):131–147.

Hashimoto, M. (1982). Minimum wage effects on training on the job. *The American Economic Review*, 72(5):1070–1087.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the Econometric Society*, pages 153–161.

Hersch, J. (1991). Education match and job match. *The Review of Economics and Statistics*, pages 140–144.

Herzog Jr, H. W., Schlottmann, A. M., and Boehm, T. P. (1993). Migration as spatial job–search: a survey of empirical findings. *Regional Studies*, 27(4):327–340.

Hsieh, C. T., Hurst, E., Jones, C. I., and Klenow, P. J. (2013). The allocation of talent and US economic growth. *NBER Working Papers*, (18693).

Hsieh, C. T. and Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Huttunen, K., Møen, J., and Salvanes, K. G. (2018). Job loss and regional mobility. *Journal of Labor Economics*, 36(2):479–509.

Iammarino, S. and Marinelli, E. (2015). Education–job (mis) match and interregional migration: Italian university graduates' transition to work. *Regional Studies*, 49(5):866–882.

Ibáñez, A. M. and Vélez, C. E. (2008). Civil conflict and forced migration: The micro determinants and welfare losses of displacement in Colombia. *World Development*, 36(4):659–676.

IBGE (2010). Instituto Brasileiro de Geografia e Estatística – Censo Demográfico 2010. *Rio de Janeiro: IBGE.*

IBGE (2013). Instituto Brasileiro de Geografia e Estatística – Síntese de Indicadores Sociais. *Estudos e Pesquisas – Informação Demográfica e Socioeconômica*, (32).

IBGE (2014). Instituto Brasileiro de Geografia e Estatística – Síntese de Indicadores Sociais. *Rio de Janeiro: IBGE*, (34).

IBGE (2016). Instituto Brasileiro de Geografia e Estatística – Pesquisa Nacional por Amostra de Domicílio – PNAD Contínua. *Brasília (DF), Trimestral.*

Imbens, G. W. and Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2):467–475.

INEP (2013a). Instituto Nacional de Estudos e Pesquisas Educacionais – Censo do Ensino Superior. *Ministério da Educação.*

INEP (2013b). Instituto Nacional de Estudos e Pesquisas Educacionais – Censo Escolar da Educação Básica. *Ministério da Educação.*

IPEA (2017). Instituto de Pesquisa Econômica Aplicada, Fórum Brasileiro de Segurança Pública – Atlas da Violência. *Rio de Janeiro.*

IPEA (2018). Instituto de Pesquisa Econômica Aplicada, Fórum Brasileiro de Segurança Pública – Atlas da Violência. *Rio de Janeiro.*

Ipeadata (2017). Dados Macroeconômicos e Regionais. http://www.ipeadata.gov.br/Default.aspx. Accessed on 11/11/2017.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy*, 87(5, Part 1):972–990.

Kaestner, R. and Malamud, O. (2014). Self-selection and international migration: New evidence from mexico. *Review of Economics and Statistics*, 96(1):78–91.

Korpi, T. and Tåhlin, M. (2009). Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000. *Labour Economics*, 16(2):183–193.

Lechthaler, W. and Snower, D. J. (2008). Minimum wages and training. *Labour Economics*, 15(6):1223–1237.

Lee, D. S. (1999). Wage inequality in the United States during the 1980s: Rising dispersion or falling minimum wage? *The Quarterly Journal of Economics*, 114(3):977–1023.

Lemos, S. (2004). Minimum wage policy and employment effects: Evidence from Brazil. *Economia*, 5(1):219–266.

Leuven, E. and Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. *Handbook of the Economics of Education*, 4:283–326.

Lindley, J. and McIntosh, S. (2009). A panel data analysis of the incidence and impact of over-education. *Sheffield Economic Research Paper Series SERP*, (Nr. 2008009).

Machin, S. and Manning, A. (1994). The effects of minimum wages on wage dispersion and employment: Evidence from the UK Wages Councils. *ILR Review*, 47(2):319–329.

Maloney, W. and Mendez, J. (2004). Measuring the impact of minimum wages. Evidence from Latin America. In *Law and employment: lessons from Latin America and the Caribbean*, pages 109–130. University of Chicago Press.

Manacorda, M., Sánchez-Páramo, C., and Schady, N. (2010). Changes in returns to education in Latin America: The role of demand and supply of skills. *ILR Review*, 63(2):307–326.

Marianne, B. (2011). New perspectives on gender. In *Handbook of Labor Economics*, volume 4, pages 1543–1590. Elsevier.

Martine, G. (1990). Brazil. In Nam, C. B., Serow, W. J., and Sly, D., editors, *International Handbook on Internal Migration*, pages 31–46. New York: Greenwood Publishing Group.

Mattila, J. P. (1978). Youth labor markets, enrollments, and minimum wages. In *Proceedings of the 31st Annual Meeting, Industrial Relations Research Association Series*, pages 134–40.

Mavromaras, K., McGuinness, S., O'Leary, N., Sloane, P., and Wei, Z. (2013). Job mismatches and labour market outcomes: panel evidence on university graduates. *Economic Record*, 89(286):382–395.

McGuinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys*, 20(3):387–418.

McGuinness, S. and Bennett, J. (2007). Overeducation in the graduate labour market: A quantile regression approach. *Economics of Education Review*, 26(5):521–531.

McGuinness, S. and Pouliakas, K. (2017). Deconstructing theories of overeducation in europe: A wage decomposition approach. In *Skill Mismatch in Labor Markets*, pages 81–127. Emerald Publishing Limited.

McGuinness, S. and Sloane, P. J. (2011). Labour market mismatch among UK graduates: An analysis using REFLEX data. *Economics of Education Review*, 30(1):130–145.

Mehta, A., Felipe, J., Quising, P., and Camingue, S. (2011). Overeducation in developing economies: How can we test for it, and what does it mean? *Economics of Education Review*, 30(6):1334–1347.

Mendes, P. S., Gonçalves, E., and Freguglia, R. (2017). Determinantes da mobilidade interfirmas e inter-regional de trabalhadores no Brasil formal. *Economia Aplicada*, 21(2):223.

Mincer, J. and Leighton, L. S. (1981). Effect of minimum wages on human capital formation. In Rottenberg, S., editor, *The economics of legal minimum wages*, pages 155–173. American Enterprise Institute for Public Policy Research.

Mincer, J. A. (1974). Individual Acquisition of Earning Power. In *Schooling, Experience, and Earnings*, pages 5–23. New York: National Bureau of Economic Research, Inc.

Monteiro, J. and Rocha, R. (2017). Drug battles and school achievement: evidence from Rio de Janeiro's favelas. *Review of Economics and Statistics*, 99(2):213–228.

Moraga, J. F.-H. (2011). New evidence on emigrant selection. *The Review of Economics and Statistics*, 93(1):72–96.

Morales, J. S. (2018). The impact of internal displacement on destination communities: Evidence from the Colombian conflict. *Journal of Development Economics*, 131:132–150.

Morrison, A. R. (1993). Violence or economics: what drives internal migration in Guatemala? *Economic Development and Cultural Change*, 41(4):817–831.

MTE (2010). Ministério do Trabalho e Emprego – Classificação Brasileira de Ocupações: CBO – 2010 – 3^a ed. Brasília: MTE, SPPE, 2010. v. 1 828 p.

Murray, J., de Castro Cerqueira, D. R., and Kahn, T. (2013). Crime and violence in Brazil: Systematic review of time trends, prevalence rates and risk factors. *Aggression and Violent Behavior*, 18(5):471–483.

Nakosteen, R. A., Westerlund, O., and Zimmer, M. (2008). Migration and self-selection: Measured earnings and latent characteristics. *Journal of Regional Science*, 48(4):769–788.

Neri, M., Gonzaga, G., and Camargo, J. (2000). Efeitos informais do salário mínimo e pobreza. [The informality sector and its effects on minimum wages and poverty]. *Revista de Economia Política*, 18(1):120–40.

Neumark, D. and Wascher, W. (1995a). Minimum wage effects on employment and school enrollment. *Journal of Business & Economic Statistics*, 13(2):199–206.

Neumark, D. and Wascher, W. (1995b). Minimum-wage effects on school and work transitions of teenagers. *The American Economic Review*, 85(2):244–249.

Neumark, D. and Wascher, W. (2001). Minimum wages and training revisited. *Journal of Labor Economics*, 19(3):563–595.

Neumark, D. and Wascher, W. (2003). Minimum wages and skill acquisition: Another look at schooling effects. *Economics of Education Review*, 22(1):1–10.

Neumark, D., Wascher, W. L., et al. (2007). Minimum wages and employment. *Foundations and Trends in Microeconomics*, 3(1–2):1–182.

OECD (2013). *Education at a Glance 2013: Highlights*. OECD Publishing.

Oosterbeek, H. and Van Ophem, H. (2000). Schooling choices: Preferences, discount rates, and rates of return. *Empirical Economics*, 25(1):15–34.

Ordine, P. and Rose, G. (2009). Overeducation and instructional quality: A theoretical model and some facts. *Journal of Human Capital*, 3(1):73–105.

Pan, J. (2015). Gender segregation in occupations: The role of tipping and social interactions. *Journal of Labor Economics*, 33(2):365–408.

Pauli, R. C. d., Nakabashi, L., and Sampaio, A. V. (2012). Mudança estrutural e mercado de trabalho no Brasil. *Revista de Economia Política*, 32(3):459–478.

Queiroz, B. L. and Golgher, A. B. (2008). Human capital differentials across municipalities and states in Brazil. *Population Review*, 47(2).

Quinn, M. A. and Rubb, S. (2006). Mexico's labor market: The importance of education-occupation matching on wages and productivity in developing countries. *Economics of Education Review*, 25(2):147–156.

Ramalho, H. M. d. B. and Queiroz, V. d. S. (2011). Migração interestadual de retorno e autosseleção: evidências para o Brasil. *Pesquisa e Planejamento Econômico*, 41(3).

Ramos, R. and Sanromá, E. (2013). Overeducation and Local Labour Markets in Spain. *Tijdschrift voor Economische en Sociale Geografie*, 104(3):278–291.

Ravn, M. O. and Sørensen, J. R. (1999). Schooling, training, growth and minimum wages. *The Scandinavian Journal of Economics*, 101(3):441–457.

Reis, M. C. (2017). Educational mismatch and labor earnings in Brazil. *International Journal of Manpower*, 38(2).

Restuccia, D. and Rogerson, R. (2013). Misallocation and productivity. *Review of Economic Dynamics*, 16(1):1–10.

Ribeiro, E. P. (2010). Fluxo de empregos, fluxo de trabalhadores e fluxo de postos de trabalho no Brasil. *Revista de Economia Política*, 30(3):401–419.

Robst, J. (1995). College quality and overeducation. *Economics of Education Review*, 14(3):221–228.

Rocha, L. P. d., Pero, V. L., and Corseuil, C. H. (2018). Turnover, learning by doing, and the dynamics of productivity in Brazil.

Rosen, S. (1972). Learning and experience in the labor market. *Journal of Human Resources*, pages 326–342.

Rubb, S. (2003). Post-college schooling, overeducation, and hourly earnings in the United States. *Education Economics*, 11(1):53–72.

Ruiz, I. and Vargas-Silva, C. (2013). The economics of forced migration. *The Journal of Development Studies*, 49(6):772–784.

Ruiz, I. and Vargas-Silva, C. (2018). Differences in labour market outcomes between natives, refugees and other migrants in the UK. *Journal of Economic Geography*, 18(4):855–885.

Rumberger, R. W. (1987). The impact of surplus schooling on productivity and earnings. *Journal of Human Resources*, pages 24–50.

SAE (2018). Secretaria Especial de Assuntos Estratégicos da Presidência da República – Custos Econônomicos da criminalidade no Brasil. *Relatório de Conjuntura*, (4).

Sahota, G. S. (1968). An economic analysis of internal migration in Brazil. *Journal of Political Economy*, 76(2):218–245.

Sánchez-Sánchez, N. and McGuinness, S. (2015). Decomposing the impacts of overeducation and overskilling on earnings and job satisfaction: an analysis using REFLEX data. *Education Economics*, 23(4):419–432.

Santos, A. M. (2002). Overeducation no mercado de trabalho Brasileiro. *Revista Brasileira de Economia de Empresas (Brazilian Journal of Business Economics)*, 2(2):1–22.

Santos Júnior, E. d. R. d., Menezes-Filho, N., and Ferreira, P. C. (2005). Migração, seleção e diferenças regionais de renda no Brasil. *Pesquisa e Planejamento Econômico*, 35(3).

Sattinger, M. (1993). Assignment models of the distribution of earnings. *Journal of Economic Literature*, 31(2):831–880.

Schmertmann, C. P. (1992). Estimation of historical migration rates from a single census: Interregional migration in Brazil 1900–1980. *Population Studies*, 46(1):103–120.

Schumann, M. (2017). The effects of minimum wages on firm-financed apprenticeship training. *Labour Economics*.

Schwartzman, S. (2004). Equity, quality and relevance in higher education in Brazil. *Anais da Academia Brasileira de Ciências*, 76(1):173–188.

Shrestha, M. (2017). *Push and pull: A study of international migration from Nepal.* The World Bank.

Sicherman, N. (1991). "Overeducation" in the Labor Market. *Journal of Labor Economics*, pages 101–122.

Sicherman, N. and Galor, O. (1990). A theory of career mobility. *Journal of Political Economy*, pages 169–192.

Simão, A. R. A. (2009). Sistema de vigilancia e fiscalização do trabalho no Brasil: efeitos sobre a expansão do emprego formal no período 1999-2007. Technical report.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of Political Economy*, 70(5, Part 2):80–93.

Sloane, P. J. (2003). Much ado about nothing? what does the overeducation literature really tell us. In Büchel, F., De Grip, A., and Mertens, A., editors, *Overeducation in Europe; Current issues in theory and policy*, pages 11–45. Edward Elgar Cheltenham.

Sloane, P. J., Battu, H., and Seaman, P. T. (1999). Overeducation, undereducation and the british labour market. *Applied Economics*, 31(11):1437–1453.

Smith, J. P. and Welch, F. (1978). The Overeducated American? A Review Article. *ERIC Working Paper*, (Nr. ED175340).

Smits, J. (2001). Career migration, self-selection and the earnings of married men and women in the Netherlands, 1981-93. *Urban Studies*, 38(3):541–562.

Sousa, F. L. d. (2014). Does crime affect migration flows? *Papers in Regional Science*, 93:S99–S111.

Souza, P. R. and Baltar, P. E. (1979). Salário mínimo e taxa de salários no Brasil. *Pesquisa e Planejamento Econômico*, 9(3).

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, pages 355–374.

Stewart, M. B. (2004). The employment effects of the national minimum wage. *The Economic Journal*, 114(494).

Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear IV regression. In Stock, J. H., Yogo, M., and Andrews, D. W., editors, *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, pages 80–108. Cambridge University Press, New York.

Teulings, C. N. (2003). The contribution of minimum wages to increasing wage inequality. *The Economic Journal*, 113(490):801–833.

Thurow, L. C. (1975). Generating Inequality. *New York: Basic Books.*

Tsai, Y. (2010). Returns to overeducation: A longitudinal analysis of the US labor market. *Economics of Education Review*, 29(4):606–617.

Turner, M. D. and Demiralp, B. (2001). Do higher minimum wages harm minority and inner-city teens? *The Review of Black Political Economy*, 28(4):95–116.

UNDP (2009). Human Development Report 2009: Overcoming Barriers: Human Mobility and Development. *United Nations Development Programme.*

UNODC (2013). *United Nations Office on Drugs and Crime – Global study on homicide 2013: trends, contexts, data.* Number 14.IV.1. United Nations publication.

Venhorst, V. A. and Cörvers, F. (2018). Entry into working life: Internal migration and the job match quality of higher-educated graduates. *Journal of Regional Science*, 58(1):116–140.

Verdugo, R. R. and Verdugo, N. T. (1989). The impact of surplus schooling on earnings: Some additional findings. *Journal of Human Resources*, pages 629–643.

Verhaest, D. and Omey, E. (2006). Discriminating between alternative measures of over-education. *Applied Economics*, 38(18):2113–2120.

Waldorf, B. and Yun, S. D. (2016). Labor migration and overeducation among young college graduates. *Review of Regional Research*, 36(2):99–119.

Yankow, J. J. (2003). Migration, job change, and wage growth: a new perspective on the pecuniary return to geographic mobility. *Journal of Regional Science*, 43(3):483–516.

Yap, L. (1976). Internal migration and economic development in Brazil. *The Quarterly Journal of Economics*, 90(1):119–137.