ESSAYS ON EDUCATION AND LABOR MARKETS IN LATIN AMERICA

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

David Jose Jaume

August 2018

ESSAYS ON EDUCATION AND LABOR MARKETS IN LATIN AMERICA

David Jose Jaume, Ph.D.

Cornell University 2018

This dissertation studies the labor market effect of different educational policies in Latin America. The first two chapters are focused on a market level analysis. The first chapter develops a framework to evaluate the labor market effects of different types of educational expansions in four labor market outcomes: (1) the occupational structure of employment; (2) the assignment of workers with different level of education to occupations; (3) the wage level for each educational group; and (4) the wage gaps between educational groups. I evaluate three policy experiments consistent on increasing secondary schooling, increasing higher education, and increasing both. In the second Chapter, I apply the framework to study the case of Brazil, a country that underwent a major educational expansion during the period 1995-2014. I provide some new stylized facts for Brazil on the inter-linkages between changes in education, occupations, and wages over the period of 1995-2014— changes in outcomes (1)-(4). I found that: (a) the occupational structure of employment improved, but that improvement was very small when compared to the extension of the educational expansion; (b) the conditional occupational attainment declined for each educational group— primary or less, secondary, and university; (c) and average wages increased but not for all educational groups since wages of primary educated workers increased while wages of more educated workers declined; (d) there were large reductions in inequality as measured by educational wage gaps. Then, I show that the model's predictions for the Brazilian educational expansion are qualitatively consistent with the patterns observed in the data. I further demonstrate that, after calibrating the model, the educational expansion in Brazil

was of utmost importance for generating the observed quantitative changes in the labor market. In the last chapter, I moved from a market level to an individual level analysis to evaluate the effect of a negative educational shock on workers' lifetime earnings. In particular, I examine how school disruptions caused by teacher strikes in Argentina affect students' long-run outcomes by exploiting cross-cohort variation in the prevalence of teacher strikes within and across provinces in Argentina in a difference-in-difference framework. I find robust evidence that teacher strikes worsen the future labor market outcomes of students when they are between 30 and 40 years old.

BIOGRAPHICAL SKETCH

David Jaume is a PhD (candidate) in Economics at Cornell University and a visiting researcher at the Center for Distributive, Labor and Social Studies (CEDLAS) at the Universidad Nacional de La Plata, Argentina. His work is focused on education, labor markets and income distribution in developing countries. In his recent research, David studies the labor market effects of different educational policies in Latin America. He has published in the areas developing economics and economics of education, including a book on the Growth, Employment, and Poverty nexus in Latin America (Oxford University Press), and has worked on several research projects with international organizations such as the World Bank, UNU-WIDER, and CEPAL. David previously worked as a research fellow at CEDLAS, Universidad Nacional de La Plata, where he received his master's degree in economics.

This document is dedicated to my wife and to my family, who showed me unconditionally support during this exciting but stressful process.

ACKNOWLEDGEMENTS

I am especially grateful to Gary Fields, my advisor and mentor, who taught me a lot about economics during the five years I spent at Cornell under his guidance. He also showed me what is required to be an excellent academic professor and I aspired to resemble him one day. I am honored that during this time he has not only been my advisor and coauthor, but also a very close friend.

I am also very grateful to Julieta Caunedo, Ravi Kanbur, and Victoria Prowse. Their support, guidance, and invaluable discussions on previous versions of my dissertation have been extremely helpful. I wish to stress Julieta's important supportive role during the stressful time of the Job Market, her kindness and advise made the entire process much more manageable.

I also thank to several scholars and classmates that contribute to improve the first two chapter of this dissertation. In particular, I thank Arnab Basu, Kaushik Basu, Francine Blau, Nancy Chau, Guillermo Cruces, Maria Fitzpatrick, Jorgen Harris, Lawrence Khan, Ravi Kanbur, Elissa Keller, Michael Paris, Victoria Prowse, Evan Riehl, Nicholas Sanders, Mallika Thomas, Erik Thorbecke, Mariana Viollaz as well as conference participants at the North East Universities Development Conference, the annual meeting Society for the Study of Economic Inequality, the Research Institute for Development, Growth and Economics, the North American Summer Meeting, the World Bank, the Bank of Mexico, EPGE-Rio, and Leicester University for valuable comments and suggestions.

For the third chapter, I am greatly indebted to Alexander Willén who has being an excellent coauthor of this chapter. I also thank Jonathan Guryan and two anonymous referees for their helpful comments. I would like to thank, Jason Cook, Douglas Miller, Evan Riehl, Lucas Ronconi as well as seminar participants at Cornell University, Universidad Nacional de La Plata, the Association for Education Finance and Policy, and the 8th Bolivian

Conference on Development Economics, for valuable comments. I would further like to thank Gustavo Torrens for access to historic data on province-specific GDP in Argentina, and to the Argentinian Ministry of Education for data on teacher wages. I gratefully acknowledge financial support to work on this chapter from the Department of Economics at Cornell University (Award in Labor Economics).

TABLE OF CONTENTS

Biographical Sketch . . . iii
Dedication . . . iv
Acknowledgements . . . v
Table of Contents . . . vii
List of Tables . . . ix
List of Figures . . . xi

1 The Labor Market Effects of an Educational Expansion: A Theoretical Model . . . 1
1.1 Introduction . . . 1
1.2 The Task-based Model . . . 7
1.2.1 Environment . . . 7
1.2.2 Initial Equilibrium with Given Educational Attainments . . . 10
1.2.3 Decomposition of the Changes in Wages . . . 20
1.3 The Labor Market Effects of an Educational Expansion . . . 23
1.3.1 Welfare analysis: changes in the CDF of the wage distribution, poverty, and inequality . . . 33
1.3.2 The importance of the comparative advantage across tasks . . . 39
1.4 Conclusions . . . 43

2 The Labor Market Effects of an Educational Expansion: An Application to Brazil . . . 45
2.1 Introduction . . . 45
2.2 Data . . . 53
2.3 Changes in the Brazilian labor market between 1995-2014 . . . 55
2.3.1 Education . . . 55
2.3.2 Wage distribution . . . 60
2.3.3 Occupational structure of employment . . . 61
2.4 Comparison of the qualitative predictions of the model and the changes in the data . . . 68
2.5 Calibration . . . 70
2.6 Quantifying the Effects of the Educational Expansion in Brazil . . . 75
2.6.1 Changes in the Occupational Structure and the Wage Distribution . . . 75
2.6.2 Where did all the increase in high education go? . . . 82
2.6.3 Output changes and the declining effects of an educational expansion 83
2.6.4 Decomposition of changes in wages . . . 86

2.6.5 Robustness Checks . . . 87
2.7 Discussion: of alternative explanations . . . 93
2.7.1 Decline in the Quality of Education . . . 94
2.7.2 Other factors . . . 98
2.8 Conclusions . . . 99

3 The Long-run Effects of Teacher Strikes. Evidence from Argentina 102
3.1 Introduction . . . 102
3.2 Background & Theoretical Predictions of Teacher Strikes . . . 106
3.2.1 The Argentinian Education System . . . 106
3.2.2 Teacher Strikes in Argentina . . . 107
3.2.3 Theoretical Predictions . . . 109
3.3 Prior Literature on Teacher Strikes . . . 112
3.4 Data . . . 114
3.4.1 Teacher Strikes . . . 114
3.4.2 Long-run Outcomes . . . 118
3.4.3 Local Labor Market Controls . . . 121
3.4.4 Short-run Outcomes . . . 123
3.5 Empirical Methodology . . . 124
3.6 Results . . . 127
3.6.1 Long-term Effects of Teacher Strikes . . . 127
3.6.2 Heterogeneous Treatment Effects . . . 139
3.6.3 Robustness & Sensitivity Analysis . . . 139
3.6.4 Short-run effects . . . 143
3.7 Discussion and Conclusion . . . 147

A Chapter 1 of appendix 150
A.1 Appendix . . . 150
A.1.1 Proof of Lemma 1 . . . 150
A.1.2 Proof of propositions 1-3 . . . 151

B Chapter 2 of appendix 161
B.1 Oaxaca Decomposition of Average Wages . . . 161
B.2 Additional tables . . . 167

C Chapter 3 of appendix 171

LIST OF TABLES

1.1 Summary of the labor market effects of different educational expansion . . . 25
1.2 A numerical example on the differential effect of Δh for different scenarios of comparative advantage across tasks. 41

2.1 Changes in the educational attainment of the workforce and the Brazilian wage distribution between 1995 and 2014 56
2.2 Employment shares by education and occupation of total employment . . . 62
2.3 Employment shares by occupations and mean occupational ranking for each educational group 66
2.4 Qualitative observed changes vs model predictions 69
2.5 Initial moments in the data (1995) and the model 74
2.6 The labor market effects of an educational expansion: data vs model 76
2.7 Declining effect on output per worker of successive educational expansions . 85
2.8 Decomposition of changes in wages with the calibrated model 87
2.9 Model results with different calibrations 92
2.10 Flat-point on earnings: Cross sectional evidence 97
2.11 Decomposition of the changes in wages on price and quantity effects 98

3.1 **Days of teacher strikes during primary school by birth cohort and birth province** 117
3.2 Effect of Strike Exposure on Individual Outcomes 129
3.3 Heterogeneous effects of strike exposure on wages, earnings and educational attainment 134
3.4 Effect of Strike Exposure on Socioeconomic Outcomes 137
3.5 Intergenerational Treatment Effects 138
3.6 Robustness and Sensitivity Checks 141
3.7 Short-Term Effects of Strike Exposure (12-17 Year Olds) 146

B.1 Oaxaca decomposition of changes in mean wages 166
B.2 Correlation between the ranking of occupations using different samples . . . 168
B.3 Changes in the occupational structure of employment in Brazil between 1995 and 2014 169
B.4 Labor market effects of the Brazilian educational expansion under different functional forms 170

C.1 **Cross-province mobility of 13 year olds** 172
C.2 Dependant variable means 173

C.3 Effect of Strike Exposure on Individual Outcomes; two-dimensional fixed effects 174
C.4 P-values from Wild Cluster Bootstrap Standard Errors Method 175
C.5 Heterogeneous Treatment Effects of Strike Exposure by School Grade 176
C.6 Effect of controlling for non-teacher strikes and GDP 177
C.7 Effect of local labor market conditions on teacher strikes 178
C.8 Effect of teacher wages on teacher strikes 179
C.9 Heterogeneous Treatment Effects on Short-Term Outcomes (12-17 Year Olds) 180
C.10 Effect of Strikes During Secondary School 181
C.11 Differential Effect of Exposure to Long Lasting Strikes within an School Year 182

LIST OF FIGURES

1.1 Equilibrium in the model: thresholds and relative wages 19
1.2 Equilibrium in the model: employment and real wages 22
1.3 Effects of an educational expansion 29
1.4 Cumulative density function (CDF) before and after each policy experiment 35
1.5 Growth incidence curve of different type of educational expansions 36
1.6 Four different scenarios of comparative advantage schedules 42

2.1 Changes in enrollment rates in Brazil 58
2.2 Changes in the occupational composition of the workforce 1995-2004 64
2.3 Occupational downgrading for each educational group between 1995-2014 . . 67
2.4 Changes in relative inequality: Lorenz Curves 80
2.5 Data vs model predictions: wages year by year 81
2.6 Evolution of output per worker in the data and in the model 84
2.7 Calibration of productivity parameters year by year 89
2.8 Comparison of 1995 calibration vs year by year calibration 91
2.9 Data vs model predictions with alternative calibration : wages year by year 93

3.1 Variation in Teacher Strikes 1977-2014 108
3.2 Data structure for a subsample of birth cohorts 118
3.3 Correlation between teacher strikes and student outcomes 121

B.1 Changes in log wages by occupation. Period 1995-2014 167

C.1 Correlation between teacher strikes and teacher wages 171

CHAPTER 1

THE LABOR MARKET EFFECTS OF AN EDUCATIONAL EXPANSION: A THEORETICAL MODEL

1.1 Introduction

Increases in human capital through formal education are widely cited as a key driver of economic development (Hanushek and Woessmann, 2008, 2012). Most countries have experienced or are experiencing a rapid increase in the educational attainment of their workforce. Between 1990 and 2010 the worldwide average enrollment rates in secondary and university education grew by around 25 percentage points each (Barro and Lee, 2013). These educational expansions are uneven in the sense that they raise the educational level of some workers while leaving that of others unchanged. Those workers receiving the additional education are expected to increase their human capital, earn higher wages, and have access to better jobs. But an educational expansion likely affects the rest of the workforce as well, positively or negatively. By generating general equilibrium effects, an influx of more educated workers can transform the labor market by reshaping the occupational composition of employment (what jobs are available in the economy and who performs them) together with changing the entire wage distribution (the wages on these jobs). The direction and magnitude of these effects are by no means obvious. For example, workers who remain with low levels of education may be hurt if they are displaced to worse jobs, where wages remain stagnant or decline, or they may become better off if their wages increase as result of a higher demand for their jobs (generated by general equilibrium effects) coupled with a lower labor supply. Given the welfare implications of these different scenarios, it is of utmost importance for policymakers to better understand how the labor market responds to different educational

expansions.

This chapter provides a theoretical framework to study the labor market effects of an educational expansion using a task-based model. The model assigns workers with three levels of education—low, medium, and high—to a continuum of occupations that vary in complexity and are combined to produce a final good. Workers' types differ in their productivity when performing each occupation so that there is an optimal assignment of workers' types to occupations. The key assumption of the model is that the comparative advantage of more educated workers relative to less educated workers increases with the complexity of the occupation. This assumption ensures positive assortative matching, where more educated workers are optimally assigned to more complex occupations.

In equilibrium, the model generates four labor market outcomes: (1) an occupational structure of employment (the share of workers employed in each occupation); 2) an assignment of workers' types to occupations; (3) a wage level for each educational group; and (4) wage gaps between educational groups. I analyze the effects on these equilibrium outcomes under three alternative policy experiments: (*i*) a "*type–h*" educational expansion defined as an increase in the share of high educated workers through a decline in the share of low educated workers; (*ii*) a "*type–m* educational expansion defined as an increase in the share of medium educated workers through a reduction in the share of low educated workers; and (*iii*) a "*type–h&m*" educational expansion characterized by an increase in the shares of both high and medium educated workers.

I find that the predicted effects depend on whether the increase took place in high (*type–h*) or medium education (*type–m*). When there is an increase in both (*type–h&m*), one of these effects dominates depending on the extent of the changes in relative supplies—high to medium and medium to low.

There are several differences among each type of educational expansion. With respect to occupational outcomes (1) and (2), an increase in higher education lowers the conditional occupational attainment for each educational group, while an increase in secondary education improves the occupational attainment of high educated workers and expands the range of occupations where medium educated workers are employed. With respect to outcome (3), an increase in higher education raises wages of low educated, lowers wages of high educated, and generates ambiguous changes in wages of medium educated. On the other hand, educational expansions focused on medium education decline wages of medium educated and generate ambiguous changes in wages for the remaining groups. Finally, I find an increase in higher education always declines wage inequality as measured by wage gaps of workers with different educational level, outcome (4), while an educational expansion focused on medium education reduces the wage gap between medium and low educated workers but increases the gap between high and medium educated.

I also find some similarities between the different types of educational expansions. When any educational expansion takes place, there is always a decline of the occupational attainment for low educated workers, the occupational composition of employment presents small changes when compared to the extent of the educational expansion, and there is a decline in wages for at least one educational group.

The task-based approach used here is not new to the economic literature. This type of model was first proposed by Autor *et al.* (2003), and it has been broadly used since to analyze the effects of technological changes.[1] To the best of my knowledge, this is the first paper to use this framework to investigate in detail the labor market effects of different types

[1]Some variations of this approach are present in Acemoglu and Autor (2011); Autor and Dorn (2013); Goos *et al.* (2014); Beaudry *et al.* (2016); Burstein *et al.* (2016); Deming (2017); Acemoglu and Restrepo (2018). See Autor (2013) for a comprehensive discussion on the task approach.

of educational expansions.[2]

The version of the model I used here is closely related to Acemoglu and Autor (2011). I differ from Acemoglu and Autor (2011) by shifting the focus of the analysis from technological changes to educational expansions. To that end, I incorporate three dimensions. First, I modify the model to be more suitable to study educational expansions by adding the restriction of a measure of workers in the economy equal to one. Thus, if the number of workers in one group increase there must be a corresponding decline in one of the other educational groups. This modification is crucial to correctly account for changes in skill supplies resulting from an educational expansion. Second, I investigate changes in a broader set of outcomes by studying changes in the overall occupational composition of employment and in real wages.[3] Third, I provide an empirical strategy to calibrate the model and quantitatively assess the importance of changes in education on the patterns observed in the data. This empirical strategy is presented in Chapter 2, where I use data from Brazil to calibrate the model.

The model presented here provides unique predictions that differ from other theoretical models previously used to analyze the labor market effects of an educational expansion. In my framework, wages and occupations (overall and for each educational group) are interlinked, and they are simultaneously determined in equilibrium. On the contrary, most of the existing literature on the labor market effects of an increase in education is concentrated either on its effect on the wage distribution or on the effects on occupations, usually considering one of these dimensions as fixed.[4] A descriptive analysis of this literature can be found

[2] Teulings (2005) and Costinot and Vogel (2010) also use a similar model to study the effects of changes in the skill distribution of the labor force on the wage distribution and the occupational composition of employment, but do not refer to increases in education. These models include a continuum of workers skills which is not explicitly related to education. This makes the effects of changes in the distribution of skills difficult to interpret in terms of an educational expansion.

[3] Acemoglu and Autor (2011) analyze the distribution of workers across occupations and relative wages.

[4] Some models allows some wages and some occupations to change, while others remain fixed Fields (2018).

in Fields (1995), who presents the labor market consequences of an educational expansion under different scenarios depending on the functioning of the labor market.

Fields (1995) distinguishes between two set of models. First, there are models with stratified labor markets by workers' educational level, where wages fluctuate when supply changes, but the occupations performed by a given worker type are fixed.[5] In this context, an educational expansion will shift the labor supplies in each market, but it does not change the occupations performed by each type of worker.[6] On the contrary, in the model presented here, an educational expansion affects the occupational attainment of each type of worker since it is endogenously determined in equilibrium.

Second, there are other models that consider labor market to be segmented, where wages in the high earning sector are fixed and only a fixed number of positions or slots are available.[7] The first contribution to this literature is Fields (1974), who considers that high educated workers are hired preferentially for the best jobs available in the economy. Fields (1974) finds that an educational expansion increases the number of preferential hiring taking place, leaving fewer high-wage jobs to unskilled workers who are increasingly employed in the low-wage sector defined as fall-back jobs. Given that low educated workers are bumped into to low-wage jobs, Fields (1995) refers to these as "bumping models" of the labor market. Fields (2018) expands the previous model to the study of social and private returns to education under different scenarios of fallback jobs.[8] In these models, wages of low educated workers

[5]Becker (1964) presents the first version of this model. Fields (1995) refers to it as he "standard textbook model" for its broad use to studying stratified labor markets.

[6]Other recent examples of more elaborated versions of the textbook model are papers that use the Katz and Murphy (1992) relative supply and demand framework to study changes in wage inequality (Goldin and Katz, 2009; Gasparini *et al.*, 2011a; Messina and Silva, 2017). These model can be interpreted as implicitly assuming stratified labor markets since workers with a given educational level produce only one input.

[7]To consider fixed wages in one sector of the economy is a common assumption in the segmented labor markets literature, such as the novel contributions of Lewis (1954) and Harris and Todaro (1970).

[8]There are three cases: a) everyone has access to a fallback job with a fixed wage, b) wages fall on fallback jobs as labor supply increases, and c) there is a limited number of fallback jobs that will also hire

usually do not increase with an educational expansion as these workers are increasingly employed on fallback jobs for which wages are either fix or decline with increases in supply, and they are more likely to be unemployed because they are bumped by high educated workers. Similar results are also present in Lazear *et al.* (2016) by using a model where hiring into posted job slots is based on comparative advantage, so that the likelihood of being hired into a given position not only depends on the worker's skill but also on the skills of other applicants. These predictions contrast with the ones of the framework developed in this chapter. To mention some key differences, the number of jobs in high-paid occupations is not fixed in my model and wages of low educated always increases when there is an educational expansion centered on high educated workers.

Another strand of the literature studies the effects of increases in education on wage premiums and inequality. It is a well-known result that an increase in the relative supply of skills causes a decline in wage premiums if the technology of the production function is constant (Katz and Murphy, 1992; Goldin and Katz, 2009; Gasparini *et al.*, 2011a; Acemoglu and Autor, 2011). This paper adds to the analysis of wage gaps by including the effects on real wages for each educational group, for different percentiles of the wage distribution, and the occupational structure of employment.

In my framework, an educational expansion in high education generates lower occupational attainment for each educational group. This result relates to the literature on overeducation, where there is also an allocation problem of deciding which workers will perform which jobs, but there are mismatches arising from incomplete markets such as imperfect information or job-search frictions (Sattinger, 1993; Leuven and Oosterbeek, 2011; Sattinger, 2012). In this literature, there are Pareto improvements to be made by switching workers across occupations, and policy interventions that ameliorate these frictions are desirable. In

more educated workers preferentially.

my model, more educated workers also end up in occupations where their skills have a lower productivity differential when compared to less educated workers, but they are optimally assigned to those occupations by labor market forces (supply and demand) and there is no wage or productivity gain to be made by switching workers across occupations. My finding indicates that more educated workers may be increasingly employed in occupations where their acquired human capital through formal education is less relevant in terms of productivity differentials to perform those occupations. This lower occupational attainment results from labor market forces (demand and supply) acting in an environment where education expands and technology is fixed, as opposed to being originated by labor market frictions.

The rest of this chapter is organized as follows: Section 2 presents the model; Section 3 contains the predictions of the model for different policy experiments; and Section 4 concludes.

1.2 The Task-based Model

1.2.1 Environment

I follow Acemoglu and Autor (2011) set up of the *Ricardian* "task-based" model, henceforth the *AA model*. I modify the model slightly by adding a restriction on the measure of workers which has to be equal to one. This small modification turns out to be key to study the effects of different types of educational expansions in the next section. However, this modification has no impact on the initial equilibrium conditions, which are identical to theirs. I follow Acemoglu and Autor (2011) derivations below, adding a formal expression for the wage levels which are one of the four outcomes I study in this chapter.

In this economy, there is a unique final good produced by combining a continuum of tasks represented in the unit interval $[0, 1]$.[9] The final good is produced by combining these different tasks.[10] The production function is Cobb-Douglas, mapping the tasks produced to the final good. Let Y denote the production of the unique final good and let $y(i)$ be the production level of the task i, then:

$$Y = exp\left[\int_0^1 ln\, y(i)\, di\right]. \tag{1.1}$$

The economy is closed in the sense that all tasks have to be produced within the economy (no trade in tasks). The tasks can be produced by using three type of workers —low, medium, and high educated—which are perfect substitutes, in terms of efficiency units, in the production of each task i. Workers' types differ in their productivity to perform each task, where productivity is defined as the amount of output of task i produced by a worker in a given time period. This constitutes the only difference among workers such that all workers with a given educational level are equally productive in the production of tasks. Consider A_J to be the factor-augmenting technology of input J and $\alpha_J(i)$ to be the task-specific productivity of input J in task i, with $J = \{L, M, H\}$. The production of each task i is defined as:

$$y(i) = A_L\alpha_L(i)L(i) + A_M\alpha_M(i)M(i) + A_H\alpha_H(i)H(i),$$

where $L(i)$, $M(i)$ and $H(i)$ are the amount of workers employed in producing task i with low, medium, and high education, respectively. There are two key assumptions in the model. First, tasks can be ordered by a unidimensional level of complexity represented by the index

[9] A task relates here to the main activity involved in a occupation. For example, for drivers their task production in a given period is the "number of miles driven", for teachers it is the "increases of students' human capital", and for researchers it is the "number of publishable papers produced".

[10] The same equilibrium conditions will hold if each task is interpreted as a differentiable good purchased by consumers whose utility function is identical to the production function of the final good considered here. This has already been noted by Costinot and Vogel (2010).

$i \in [0,1]$, where 0 represents the least complex task and 1 stands for the most complex task. Second, the comparative advantage of more educated workers relative to less educated workers increases with the complexity of the task and they are continuously differentiable on i. That is,

$$\frac{\partial \frac{\alpha_M(i)}{\alpha_L(i)}}{\partial i} > 0, \ \frac{\partial \frac{\alpha_H(i)}{\alpha_M(i)}}{\partial i} > 0. \tag{1.2}$$

These are key assumption of the model because they establish the structure of comparative advantage across tasks, ensuring positive assortative matching. In equilibrium, more educated workers are optimally assigned to more complex tasks. I further assume that more educated workers have an absolute advantage in the production of all tasks. That is, $A_L\alpha_L(i) < A_M\alpha_M(i) < A_H\alpha_H(i) \ \forall i \ \in (0,1)$. This assumption does not affect the assignment of workers across tasks and ensures that a more educated labor force always produces more output.[11]

Let l, m and h be the share of low, medium, and high educated workers in the economy. I further assume that there is no unemployment so that labor market clearing requires that:

$$\int_0^1 L(i)di = l; \ \int_0^1 M(i)di = m; \ \int_0^1 H(i)di = h. \tag{1.3}$$

Finally, I assume that there is a measure of workers equal to 1 in the economy,

$$l + m + h = 1; \tag{1.4}$$

Note that because of the assumption of a measure 1 of workers, $L(i)$, $M(i)$ and $H(i)$ are interpreted as the share of the population with a specific level of education employed in task i. The labor supply for each of the three types can be defined by two parameters, e.g. h

[11]The assumption on the absolute advantage is important in Section 1.3 to provide unambiguous predictions of directional changes in real wages for each educational group. It has no impact on the results of this section.

and m. When either h or m goes up there must be a corresponding decrease in the share of workers with another level of education. This is the most significant difference between the environment of my model and that of Acemoglu and Autor (2011).[12] This restriction is a fundamental feature of an educational expansion where some percentage of the workforce will move from one educational level to another, increasing the supply of a more educated group by reducing the supply of a less educated group.

1.2.2 Initial Equilibrium with Given Educational Attainments

Results from the AA model

The wage and employment determination mechanisms in this model are similar to the traditional demand and supply textbook model with stratified labor markets. There are three labor markets, one for each educational group. The labor supply in each market is fixed, exogenously determined, and equal to l,m, and h. Because there is no unemployment, the employment level for each type equals supply. There is a downward sloping labor demand for each type of worker that is going to be determined below. All workers of the same educational level receive the same wage even when they perform different tasks, and wages can go up or down to clear the market. The final good market and all labor markets are competitive in the sense that firms take as given the price of the final good and the wage of each educational group. The market-level labor demand from profit maximizing firms is the value of the marginal product of labor in producing the final good. The important difference with the traditional textbook model is that the demands for each worker type are interrelated, and the market clearing equilibrium in the three markets needs to be solve

[12] In Acemoglu and Autor (2011) the number of low, medium, and high educated are independent from one another so that one can increase while the rest remain fixed.

simultaneously.

The competitive equilibrium in this economy consists of an assignment of workers' types to tasks $L(i), M(i), H(i)$ and real wages for each type of worker W_L, W_M, W_H, such that producers maximize profits and labor markets clear, given the supply of skills and the productivity across tasks of each type of worker.

The price of the final good is considered to be the numeraire, that is:

$$exp\left[\int_0^1 ln\, p(i)\, di\right] = 1. \tag{1.5}$$

The structure of this economy implies that the same equilibrium conditions are derived by considering that there is only one firm producing the final good and each of the tasks, or that some firms specialize in the production of task i and, in turn, sell the task production for a price $p(i)$ to the a firm or many firms producing the final good.[13] If there is one firm producing both all tasks and the final good, the same process will take place within the firm. To make exposition clear, I express the firm problem as two separable problems: the final good producers problem and the task producers problem.

The producers of the final good have to choose how many tasks to buy to produce the final good in order to maximize profits, taken prices of both the final good and all tasks as given. The final good producer problem is:

$$\underset{y(i)}{\text{Max}}\ \ Y - \left[\int_0^1 y(i)p(i)\, di\right]. \tag{1.6}$$

The task producers observe the price of each task i and the wages of workers with different

[13] Because the production function is homogeneous of degree one on tasks, it makes no difference to assume that one firm is producing the final good or that there is N identical firms in the market. To see this, consider the case of N firms producing total output $Y^N = \sum^N exp\left[\int_0^1 ln\, y(i)/N\, di\right] = \sum^N exp^{lnY} exp^{ln-N} = Y$. The result follows from identical firms demanding the same amount of task $y(i)$ so that the total production of a task is equally distributed among them.

educational level, and decide how many workers of each type to hire for the production of task i. Because of the assumption that labor markets are competitive, all workers with a given educational level have to earn the same wage, independent of the task that they are performing. The task producer problem for task i is:

$$\underset{L(i),M(i),H(i)}{\text{Max}} \quad p(i)y(i) - w_L L(i) - w_M M(i) - w_H H(i). \tag{1.7}$$

The equilibrium conditions in this model are identical to those in Acemoglu and Autor (2011). The equilibrium is easy to characterize given the positive assortative matching that arises from the supermodularity in the production function and the assumption that the comparative advantage of more educated workers with respect to less educated workers is increasing in i. In particular, there exist two thresholds that determine which tasks are carried out by low, medium and high educated workers. Restating Lemma 1 from Acemoglu and Autor (2011):

Lemma 1. *In any equilibrium there exist* $\{I_L, I_H\} \in (0,1)$ *with* $I_L < I_H$ *such that for any* $i < I_L$, $L(i) > 0$ *and* $M(i) = H(i) = 0$*; for any* $i \in (I_L, I_H)$, $M(i) > 0$ *and* $L(i) = H(i) = 0$*; and for any* $i > I_H$,$H(i) > 0$ *and* $L(i) = M(i) = 0$.

Proof: See Appendix A.1.

These thresholds naturally arise from the assumption that comparative advantage is increasing in i, generating positive assortative matching. Intuitively, it is optimal for employers to use the most productive workers (high educated) for the more complex tasks over the interval $(I_H, 1)$ of tasks complexity, where they have a larger comparative advantage. It is

also in their profit-maximizing interest to employ low educated workers in the tasks of low complexity which belong to the interval $(0, I_L)$. And finally, medium educated workers will be employed in the remaining tasks with a medium level of complexity located in the interval (I_L, I_H). I show below that these thresholds I_L, I_H arise from solving the maximization problem of the firms, both the final good producers and the tasks producers, and the fact that competitive labor markets clear.

Workers of the same type may perform different tasks but earn the same wage because markets are competitive. Letting w_L, w_M, and w_H be the wage of low, mediu, and high educated respectively, from the first order condition of the tasks producers problem it is possible to state:

$$w_L = p(i)A_L\alpha_L(i) \qquad \forall i < I_L.$$

$$w_H = p(i)A_M\alpha_M(i) \qquad \forall i \in (I_L, I_H).$$

$$w_H = p(i)A_H\alpha_H(i) \qquad \forall i > I_H.$$

It implies that for two tasks i and i' that are produced using the same type of labor it must be the case that tasks' price differences exactly offset productivity differences. Therefore, there is one price for all tasks produced the same type of worker. Let P_L, P_M, and P_H be the price for tasks produced by low, medium, and high educated workers, respectively, then:

$$p(i)\alpha_L(i) = p(i')\alpha_L(i') \equiv P_L \qquad \forall i, i' \in (0, I_L). \tag{1.8}$$

$$p(i)\alpha_M(i) = p(i')\alpha_M(i') \equiv P_M \qquad \forall i, i' \in (I_L, I_H). \tag{1.9}$$

$$p(i)\alpha_H(i) = p(i')\alpha_H(i') \equiv P_H \qquad \forall i, i' \in (I_H, 1). \tag{1.10}$$

Because the technology of the production function is Cobb-Douglas, from the maximization problem of the final good producer, the expenditure across all tasks should be equalized

$(p(i)y(i) = p(i')y(i')$ for any i,i'). It implies that for any two task produced by workers of educational level $J = \{L, M, H\}$ the following condition must hold:

$$p(i)\alpha_J(i)J(i) = p(i')\alpha_J(i')J(i').$$

Additionally, given equation (1.8), $J(i) = J(i')$ for all i,i' performed by worker type $J = \{L, M, H\}$. Let $E(i)$ be the employment in each task i. In order to study the changes in the employment structure of the economy in the next section, it is useful to define *bottom-level*, *middle-level* and *top-level* tasks according to whether they were originally performed for low, medium or high educated workers, respectively, before the educational expansion took place. The employment levels for each of these categories are represented by $E_B(i)$, $E_M(i)$, and $E_T(i)$. Using the market clearing condition (1.3),

$$E_i = \begin{cases} E_B(i) = L(i) = \frac{(1-m-h)}{I_L}, & if\ \ 0 < i < I_L \\ E_M(i) = M(i) = \frac{m}{(I_H - I_L)}, & if\ \ i \in (I_L, I_H) \\ E_T(i) = H(i) = \frac{h}{(1-I_H)}, & if\ \ I_H < i < 1. \end{cases} \tag{1.11}$$

Equation (1.11) implies that there is a constant level employment among tasks performed by the same type of worker. Comparing now tasks performed by low educated workers and those of medium educated workers, and using the fact that total expenditure of employers in each task has to be the same to maximize profits (because of the production is Cobb-Douglas), it is possible to write

$$p(i)A_L\alpha_L(i)L(i) = p(i')A_M\alpha_M(i')M(i'), \qquad \text{for any } i < I_L, I_L < i' > I_H.$$

Similarly, by comparing task of high educated with those of medium educated we get

$$p(i)A_M\alpha_M(i)M(i) = p(i')A_M\alpha_M(i')M(i'), \qquad \text{for any } I_L < i > I_H, i' > I_H.$$

Using equations (1.8), (1.10), (1.9), (1.11), we can express the relative prices of tasks produced by different workers as:

$$\frac{P_M}{P_L} = \left(\frac{I_H - I_L}{I_L}\right)\left(\frac{A_L l}{A_H h}\right). \tag{1.12}$$

$$\frac{P_H}{P_M} = \left(\frac{1 - I_H}{I_H - I_L}\right)\left(\frac{A_M m}{A_H h}\right). \tag{1.13}$$

Consider also that the cost of producing task I_L with low or medium educated workers must be the same or employers could increase profits by using the less expensive worker in this task. Correspondingly, the cost of producing task I_H with medium and high educated workers must be equalized. We can express these conditions as follows:

$$p(I_L)A_L\alpha_L(I_L)L(I_L) = p(I_L)A_M\alpha_M(I_L)M(I_L).$$

$$p(I_H)A_M\alpha_L(I_H)M(I_H) = p(I_H)A_H\alpha_H(I_H)H(I_H).$$

By using (1.11),

$$\frac{(I_H - I_L)}{I_L}\frac{\alpha_L(I_L)}{\alpha_M(I_L)} = \frac{A_M}{A_L}\frac{m}{(1 - m - h)}. \tag{1.14}$$

$$\frac{(1 - I_H)}{I_H - I_L}\frac{\alpha_M(I_H)}{\alpha_H(I_H)} = \frac{A_H}{A_M}\frac{h}{m}. \tag{1.15}$$

These equations provide a unique mapping between the thresholds I_H and I_L, and the relative supply and relative productivities of high and low educated workers across different tasks. These thresholds are uniquely determined. To see this, consider the case of equation (1.14). If I_L is equal to zero, the left-hand-side (LHS) goes to infinite; if I_L is equal to I_H the LHS is equal to zero. The LHS is decreasing in I_L given the assumption that more educated workers are more productive in more complex tasks; the right-hand-side is positive and does not depends on I_L. Taken together, there is a unique value of $I_L \in (0, I_H)$ that solves (1.14), given a value of I_H. Using similar arguments, there is a unique value of $I_H \in (I_L, 1)$

that solves equation (1.15). Finally, there is a unique pair $I_L, I_H \in (0,1)$ that solves both equations simultaneously.

Note that (1.14) and (1.15) together with (1.11) imply that $E_B(i) > E_M(i) > E_T(i)$. This means that the employment share in tasks performed by low educated workers (those of lower complexity) is higher than the employment share in tasks performed by medium educated workers, which in turn are higher than those of high educated workers. The intuition is that because high educated workers are more productive in all tasks, employers find it profitable to use them in a broader set of tasks, given that each task is equally important to produce the final good because the elasticity of substitution across tasks in the production function is equal to 1.

Wages in the equilibrium are equal to the value of the marginal products of different type of workers because of the assumption that markets are competitive:

$$W_L = P_L A_L. \tag{1.16}$$

$$W_M = P_M A_M. \tag{1.17}$$

$$W_H = P_H A_H. \tag{1.18}$$

By equations 1.12 and 1.13, relative wages can be expressed as follows:

$$\frac{W_M}{W_L} = \frac{P_M A_M}{P_L A_L} = \frac{(I_H - I_L)}{I_L}\frac{(1-h-m)}{m}. \tag{1.19}$$

$$\frac{W_H}{W_M} = \frac{P_H A_H}{P_M A_M} = \frac{(1-I_H)}{I_H}\frac{m}{h}. \tag{1.20}$$

Wage levels in the AA model

All the results presented in the previous section can be found in Acemoglu and Autor

(2011). In this section, I derived a functional form for the wage levels to finish characterizing the equilibrium.

First, it is necessary to compute price levels. Due to the choice of the numeraire, $\left[\int_0^1 ln\, p(i)\, di\right] = 0$ from equation (1.6) and equations (1.8), (1.9) and (1.10),

$$\int_0^{I_L} (ln\, P_L\ - ln\, \alpha_L(i))\ di + \int_{I_L}^{I_H} (ln\, P_M\ - ln\, \alpha_M(i))\ di + \int_{I_H}^{1} (ln\, P_H\ - ln\, \alpha_H(i))\ di = 0. \tag{1.21}$$

This equation together with (1.19) and (1.20) solve for the price levels P_L, P_M and P_H. Having estimated prices, real wages are computed using (1.16), (1.17) and (1.18). Let AVP be the average productivity in the production of tasks in the economy, $C_{HM}(i)$ be the comparative advantage of high with respect to medium educated workers at task i, and $C_{ML}(i)$ be the comparative advantage of medium relative to low educated workers at task i. Wage levels can be expressed in terms of AVP, $C_{HM}(I_H)$, $C_{ML}(I_L)$ and the thresholds levels I_H and I_L, that is:

$$ln\, W_H = AVP + I_H C_{HM}(I_H) + I_L C_{ML}(I_L) \tag{1.22}$$

$$ln\, W_M = AVP - (1 - I_H) C_{HM}(I_H) + I_L C_{ML}(I_L) \tag{1.23}$$

$$ln\, W_L = AVP - (1 - I_H) C_{HM}(I_H) - (1 - I_L) C_{ML}(I_L) \tag{1.24}$$

where,

$$AVP = \int_0^{I_L} ln\, A_L \alpha_L(i)\, di + \int_{I_L}^{I_H} ln\, A_M \alpha_M(i)\, di + \int_{I_H}^{1} ln\, A_H \alpha_H(i)\, di.$$

and

$$C_{HM}(j) = ln\, A_H \alpha_H(j) - ln\, A_M \alpha_M(j)$$

$$C_{ML}(j) = ln\, A_M \alpha_M(j) - ln\, A_L \alpha_M(j).$$

The wages for each type of worker have two components, one is common for all workers and the other contains two terms that are specific to each type of worker. The common factor is the average productivity of tasks' production in the economy. The type-specific component adds (subtracts) to the average labor productivity depending on how much more (less) productive workers of a type are exactly at the task thresholds, weighted by the number of tasks each type performs.

From taking the difference of equations (1.22)-(1.24), the wage gaps in the model are equal to the productivity differential at the task thresholds. That is:

$$ln\, W_H - ln\, W_M = C_{HM}(I_H) \tag{1.25}$$

$$ln\, W_M - ln\, W_L = C_{ML}(I_L) \tag{1.26}$$

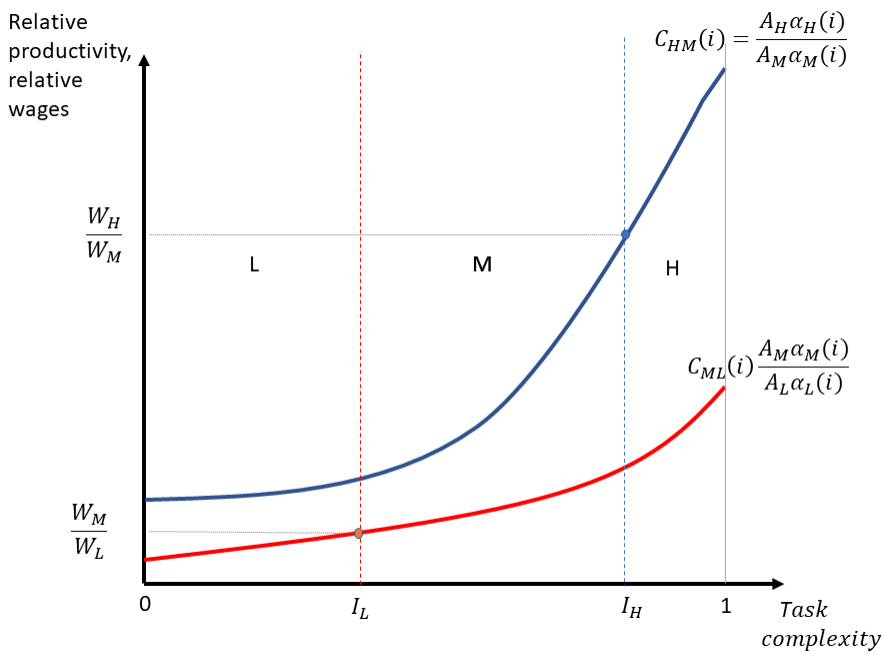
$$ln\, W_H - ln\, W_L = C_{HM}(I_H) + C_{ML}(I_L). \tag{1.27}$$

The four labor market outcomes of interest for this paper are defined by equations (1.11), (1.14), (1.15), and (1.22)-(1.27). They determine, respectively, the employment distribution across tasks in the entire economy; the distribution of workers' type to tasks; the wage levels for each type; and the wage gaps between types.

An intuitive depiction of the equilibrium is shown in Figure 1.1. The figure contains the level of task complexity (i) on the horizontal axis and the relative productivities and wages on the vertical axis. The thresholds I_L and I_H define the assignment of workers' types to tasks, such that low educated workers perform tasks below I_L, medium educated workers those between I_L and I_H, and high educated workers carry out the most complex tasks above I_H. The function that defines the comparative advantage across tasks, $C_{HM}(i)$ and $C_{ML}(i)$, are increasing in the complexity of the task defined by i, and relative wages are defined by the comparative advantage (relative productivity) among different types at the thresholds

I_L and I_H. Employment and wages of equilibrium are shown in Figure 1.2. Panel A presents the employment share in each task and the average task for workers of a given educational group. Panel B displays the wage levels for different percentiles of the wage distribution divided by 100. Considering that wages are the same for all workers of a given educational type and that $W_L < W_M < W_H$, the lowest l percentiles are low educated, percentiles l to $m + l$ are medium educated, and percentiles $l + m$ to 1 are high educated.

Figure 1.1: Equilibrium in the model: thresholds and relative wages



Notes: The figure shows an infinite number of tasks in the horizontal axis indexed by their level of complexity from 0 (the less complex tasks) to 1 (the most complex tasks) and the relative productivity between workers of different educational level across tasks on the vertical axis. The function $C_{HM}(i)$ represents the comparative advantage of high educated with respect to medium educated workers at task i and $C_{ML}(i)$ states for the comparative advantage of medium relative to low educated workers at task i. Both are increasing in i by assumption. The threshold levels I_L and I_H determine the tasks performed by low (L), medium (M) and high educated (H). In equilibrium, $C_{HM}(I_H)$ ($C_{ML}(I_L)$) equals relative wages of high and medium educated workers (medium and low educated workers), otherwise employers can profit from switching workers across occupations.

1.2.3 Decomposition of the Changes in Wages

To better understand the effect on workers' wages of any change in the parameters of the model (e.g. skills supplies), it is useful to decompose changes in wages into three different parts. Consider the optimal production of each task $y(i)$ from the equilibrium condition in equation (1.11). If the optimal production of a task (i) is replaced in the production function Y of equation (1.1), we get

$$Y = exp(AVP)(\frac{1-m-h}{I_L})^{I_L}(\frac{m}{I_H - I_L})^{I_H - I_L}(\frac{h}{1-I_H})^{1-I_H}.$$

The production function takes a Cobb-Douglas form with three inputs—low, medium, and high educated— where coefficients sum to one. This functional form ensures that employers optimally spend a fixed share of the total production on each input, and that share is equal to the input-specific coefficient in the production function.[14] Wages can then be expressed as:

$$W_L = \frac{I_L Y}{(1-m-h)}; \quad W_M = \frac{(I_H - I_L)Y}{m}; \quad W_H = \frac{(1-I_H)Y}{h};$$

while log wages consist of three components,

$$\begin{aligned} ln\, W_L &= \underbrace{ln\, I_L}_{\text{Tasks}} + \underbrace{ln\, Y}_{\text{Output}} - \underbrace{ln\,(1-m-h)}_{Supply} \\ ln\, W_M &= \underbrace{ln\,(I_H - I_L)}_{\text{Tasks}} + \underbrace{ln\, Y}_{\text{Output}} - \underbrace{ln\, m}_{Supply} \\ ln\, W_H &= \underbrace{ln\,(1-I_H)}_{\text{Tasks}} + \underbrace{ln\, Y}_{\text{Output}} - \underbrace{ln\, h}_{Supply}\ . \end{aligned}$$

Note that average wages $\bar{W} = (1-m-h)W_L + mW_M + hW_H = Y$. Any increase in the production of the final good raises the average wage for the entire economy.

[14] This property of the task model has also been noted by Acemoglu and Restrepo (2018).

This implies that the change in wages of any educational group arising from a change in one of the parameters of the model, e.g. the skill supplies, can be decomposed into three effects:

$$\Delta ln\, W_L = \underbrace{ln\, I'_L - ln\, I_L}_{\text{Displacement effect}} + \underbrace{ln\, Y' - ln\, Y}_{\text{Productivity effect}} - \underbrace{(ln\,(1 - m' - h') - ln\,(1 - m - h))}_{\text{Supply effect}} \tag{1.28}$$

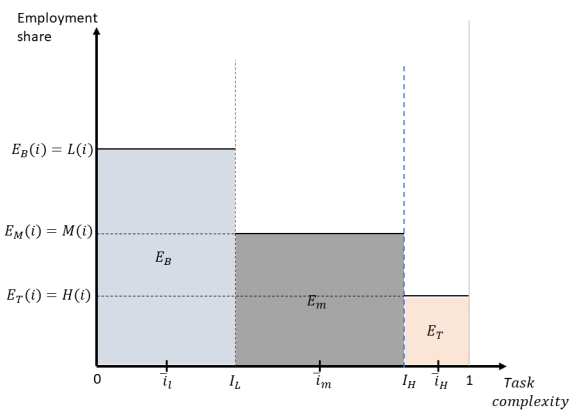
$$\Delta ln\, W_M = \underbrace{ln\,(I'_H - I'_L) - ln\,(I_H - I_L)}_{\text{Displacement effect}} + \underbrace{ln\, Y' - ln\, Y}_{\text{Productivity effect}} - \underbrace{(ln\, m' - ln\, m)}_{\text{Supply effect}} \tag{1.29}$$

$$\Delta ln\, W_H = \underbrace{ln\, I'_H - ln\, I_H}_{\text{Displacement effect}} + \underbrace{ln\, Y' - ln\, Y}_{\text{Productivity effect}} - \underbrace{(ln\, h' - ln\, h)}_{\text{Supply effect}} \tag{1.30}$$

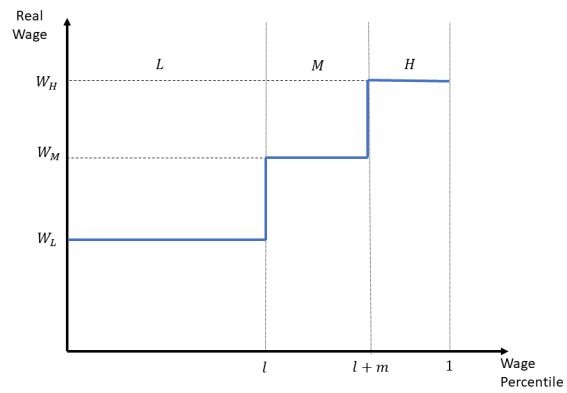
The *Displacement effect* refers to the changes in the share of tasks produced by a workers of a given educational level. The larger the increase in the share of tasks performed by workers of a given educational level, the higher the increase in wages. The *productivity effects* consist on the relationship between wages and changes in the production of the final good in an economy, and it is the same for all educational levels. Higher output is associated to an increase in wages. Finally, the *supply effect* accounts for the direct effect of a change in the percentage of workers of a given educational level. As usual, an increase in supply reduces wages. Next, I show that an educational expansion generates all three of these effects and that they are important to understanding the impact on wages.

Figure 1.2: Equilibrium in the model: employment and real wages

(a) Panel A: Employment share across tasks



(b) Panel B: Wage levels



Notes: Panel A shows tasks in the horizontal axis indexed by their level of complexity from 0 (the less complex tasks) to 1 (the most complex tasks) and the employment share in each task on the vertical axis. Panel B displays the wage levels for different percentiles of the wage distribution divided by 100. Lowest l percentiles are low educated, percentiles l to $m + l$ are medium educated, and percentiles $l + m$ to 1 are highly educate.

1.3 The Labor Market Effects of an Educational Expansion

In order to study the effects of an educational expansion on the four labor market outcomes of interest, it is useful to start with some definitions. First, an educational expansion is defined as:

Educational expansion: There exist an educational expansion between time t and t' if $m'_t \geq m_t$ and $h'_t \geq h_t$, with at least one of those relationship holding with strict inequality.[15]

Second, in order to study the changes in the employment structure of the economy, it is useful to define the total employment share in *bottom-level*, *middle-level* and *top-level* tasks according to whether they were originally performed for low, medium or high educated workers, respectively, before the educational expansion took place. These employment shares are represented by E_B, E_M, and E_T and are defined as:

$$E_B = \int_0^{I_L} E_B(i)\, di; \;\; E_M = \int_{I_L}^{I_H} E_M(i)\, di; \;\; E_T = \int_{I_H}^{1} E_T(i)\, di. \tag{1.31}$$

Note that in the initial equilibrium before any educational expansion, $E_B = l$, $E_M = m$ and $E_T = h$.

Third, to simplify the analysis on the changes in the occupational attainment for a given educational group, I define the average complexity of the tasks performed by a low, medium and high educated workers as $\bar{i}_L$, $\bar{i}_M$ and $\bar{i}_H$ respectively. Given that employment shares are constant for the tasks performed by the same educational group, the average tasks only depends on the task thresholds I_L and I_H and can be written as:

$$\bar{i}_L = \frac{I_L}{2}; \;\; \bar{i}_M = I_L + \frac{I_H - I_L}{2}; \;\; \bar{i}_L = I_H + \frac{1 - I_H}{2}. \tag{1.32}$$

[15] In the rest of the paper the subscript t is dropped to ease notation.

I study three policy experiments of educational expansions: (i) *type–h* consisting in an increase in h by decreasing l; (ii) *type–m* consisting in an increase in m by decreasing l; and (iii) *type–h&m* consisting in an increase in both m and h. *Type–h&m* is a mixture of the first two types and can be further divided into two sub-experiments (iii.a) or (iii.b) depending on whether it is the effects of *type–h* or *type–m* that dominates. Which one dominates depends on the relative importance of the increase in high and medium education, as well as the curvature of the comparative advantage schedules.[16] The summary of the labor market effects of these different educational expansions is shown in Table 1.1.

The effects of an educational expansion on the equilibrium conditions of the model will be solved sequentially below: when an educational expansion takes place, equations (1.14) and (1.15) determine the new allocation of workers to tasks. Then, equation (1.11) solve for the overall employment composition of the economy. These equations only depend on the relative supply of skills, which changes differently with each type of educational expansion, and the comparative advantages across tasks which are assumed to be fixed. Finally, equations (1.22)-(1.27) determine the new wage levels and relative wages. In other words, I first determine the effect of each policy experiment on the task thresholds, and the effect on wages follows from those changes.

[16]Below I show that for *type–m* to dominate it is necessary that the percentage increase in m is higher than that of h.

Table 1.1: Summary of the labor market effects of different educational expansion

	(i) Type–h ($\Delta h > 0$)	(ii) Type–m ($\Delta m > 0$)	(iii) Type–h&m ($\Delta m > 0$,$\Delta h > 0$)	
			(iii.a) $\downarrow I_H$	(iii.b) $\uparrow I_H$
Panel A: Changes in thresholds				
I_L	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
I_H	$\downarrow$	$\uparrow$	$\downarrow^*$	$\uparrow^*$
Changes in task composition of employment				
Bottom-level (E_B)	$\Delta E_B > -\Delta h$	$\Delta E_B > -\Delta m$	$\Delta E_B > -(\Delta m + \Delta h)$	$\Delta E_B > -(\Delta m + \Delta h)$
Medium-level (E_M)	$\downarrow$	$\Delta E_M < \Delta m$	$\Delta E_M < \Delta m$	$\Delta E_M < \Delta m$
Top-level (E_T)	$\Delta E_T < \Delta h$	$\uparrow$	$\Delta E_T < \Delta h$	$\Delta E_T > \Delta h$
Panel C: Changes in mean task index				
Low educated ($\bar{i}_L$)	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
Medium educated ($\bar{i}_M$)	$\downarrow$	$\uparrow\downarrow$	$\downarrow$	$\uparrow\downarrow$
High educated ($\bar{i}_H$)	$\downarrow$	$\uparrow$	$\downarrow$	$\uparrow$
Panel D: Changes in wages				
Low educated (W_L)	$\uparrow$	$\Delta W_L > \Delta W_M$	$\uparrow$	$\Delta W_L > \Delta W_M$
Medium educated (W_M)	$\Delta W_H < \Delta W_M < \Delta W_L$	$\downarrow$	$\Delta W_H < \Delta W_M < \Delta W_L$	$\downarrow$
High educated (W_H)	$\downarrow$	$\Delta W_H > \Delta W_M$	$\downarrow$	$\Delta W_H > \Delta W_M$
Panel E: Changes in wage gaps				
W_H/W_M	$\downarrow$	$\uparrow$	$\downarrow$	$\uparrow$
W_M/W_L	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
W_H/W_L	$\downarrow$	$\uparrow\downarrow$	$\downarrow$	$\uparrow\downarrow$

Notes: The table shows the theoretical predictions of the model under four different policy experiments. Policy experiment (i) consists of an increase in high education. Policy experiment (ii) refers to increase in medium education. Policy experiment (iii) refers to an increase in both medium and high, and it is further separated between case (iii.a) when the threshold I_H declines, and case (iii.b) when the threshold I_H increases. The share of low educated workers is reduced by the same amount in all policy experiments. $\uparrow$ denotes an increase for any value in the parameters of the model; $\downarrow$ denotes a decrease; $\uparrow\downarrow$ denotes that it can increase or decrease. Other cells are filled with lower or upper bound to changes in the outcome variable. * Given by assumption.

I postulate one proposition for each policy experiment (i)-(iii).

Proposition 1 *Under a type–h educational expansion, policy experiment* (i)*, the share of high educated workers increases from* h *to* h' *with a corresponding decline in the share of low educated such that* $\Delta h = h' - h = -\Delta l > 0$*, holding the share of medium educated constant, generating:*

(i.1) *Changes in occupational attainment by educational group:* $I'_L < I_L$*,* $I'_H < I_H$ *such that* $\{\Delta\bar{i}_L, \Delta\bar{i}_M, \Delta\bar{i}_H\} < 0$*;*

(i.2) *Changes in the occupational structure:* $\Delta E_B > \Delta l$*,* $\Delta E_M < 0$*, and* $\Delta E_T < \Delta h$*;*

(i.3) *Changes in real wages:* $\Delta W_L > 0$*,* $\Delta W_H < 0$*, and* $\Delta W_M \gtrless 0$ *with* $\Delta W_H < \Delta W_M < \Delta W_L$*;*

(i.4) *Changes in relative wages:* $\{\Delta\frac{W_M}{W_L}, \Delta\frac{W_H}{W_M}, \Delta\frac{W_H}{W_L}\} < 0$.

Proof: See Appendix A.3.

Result (i.1) indicates that, under policy experiment (i), the average worker in each educational group ends up more concentrated in occupations of lower complexity than before (lower occupational attainment). Result (i.2) establishes that the employment share of bottom-level occupations declines less than the reduction in the share of low educated workers because some medium educated workers start to carry out some of these tasks. Similarly, the employment share in top-level occupations increases by less than the change in the share of workers with high education because some of these workers start to perform medium-level tasks. Results (i.3) and (i.4) arise from the general equilibrium effects of changes in supply and demand for workers with different skill levels. Result (i.3) indicates that wages of low

educated workers increase, wages of high educated workers decline, and changes in wages of medium education workers may increase or decline, but the change is bounded below and above by the changes in wages of high educated and low educated respectively. Finally, Result (i.4) indicates that all wage gaps between more educated workers with respect to less educated workers decline.

These results are intuitively appealing. When the share of high educated workers increases, they become more abundant and their wages fall in the tasks in which they were originally employed, while low educated workers become less abundant and their wages increase in the tasks they were initially performing. Therefore, it becomes profitable for firms to start hiring high educated workers on tasks previously performed by medium educated workers, and to use medium educated workers in tasks previously performed by low educated workers. Then, all workers' types end up more concentrated in occupations of lower complexity than before, generating lower conditional occupational attainment for each educational group. The changes in the task composition of employment for the entire economy are small when compared to the educational expansion because more educated workers start to perform occupations of lower complexity.

To interpret the direction of the changes in wages, it is useful to use the decomposition from equations (1.28)-(1.30). Wages for high educated workers decline because the negative effect of the increase in supply dominates the displacement effect and the productivity effect. The increase in wages of low educated workers is due to the supply effect combined by the productivity effect dominating the negative effect of the displacement effect. In other words, the decline in supply coupled with a higher demand for the tasks they perform (generated by general equilibrium effects) dominates over the fact that they are displaced towards tasks of lower complexity.[17] Changes in wages of medium educated workers can be positive or

[17] Although the productivity to perform each task does not change with an educational expansion for any

negative but they are always bounded below and above by those of low and high educated workers, respectively. In the decomposition, any negative effect on wages of medium educated workers from the displacement effect would be partially or totally offset by the increase in productivity in the economy. Wages gaps of more educated workers with respect to lower educated workers decline. The workers that benefit the most from an increase in the share of high educated are those who get educated (their wages raise to those of high educated workers), but also the remaining low educated for who wages increase. The workers that are hurt the most are those that already were high educated.

Panel (A) in Figure 1.3 displays some of the effects of policy experiment (i). When the share of higher educated workers increase, their wages decline and it becomes profitable for employers to hire them for tasks of lower complexity depicted by a decline in I_H, moving medium educated workers downward in the occupational ladder which in turn also displaced low educated workers towards tasks with lower complexity which translate into a decline in I_L. Wage gaps decline given the fall in the comparative advantages of high to medium and that of medium to low educated at each threshold, depicted by a downward movement along the comparative advantages curves.

I discuss next the effect of an increase in secondary education, policy experiment (ii).

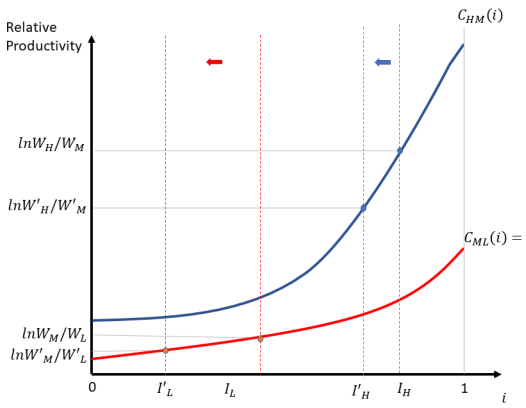
Proposition 2 *Under a type–m educational expansion, policy experiment* (ii)*, the share of medium educated workers increases from* m *to* m' *with a corresponding decline in the share of low educated such that* $\Delta m = m' - m = -\Delta l > 0$*, holding the share of high educated constant, generating:*

(ii.1) *Changes in occupational attainment by educational group:* $I'_L < I_L$, $I'_H > I_H$,

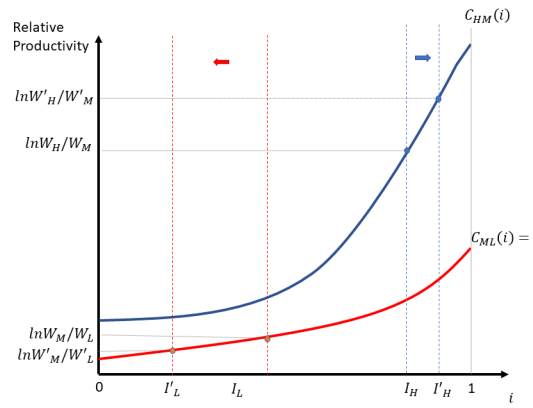
group, there is an implicit complementarity in the production of the final good by combining occupations in a Cobb-Douglas production function with an elasticity of substitution equal to one.

Figure 1.3: Effects of an educational expansion

(a) Panel A: Policy experiment (i) (Δh)



(b) Panel B: Policy experiment (ii) (Δm)



Notes: The figure shows the theoretical effects of an educational expansion on the assignment of workers' type to task and relative wages. Panel A shows to the effect of policy experiment (i) consisting of an increase in the share of high educated workers and a corresponding decrease in low educated workers. Panel B displays the effects of policy experiment (ii) increasing the share of medium educated workers by a corresponding decline in low educated workers.

$(I'_H - I'_L) > (I_H - I_L)$, *and* $(I'_L + I'_L)/2 \lesseqgtr (I_L + I_H)/2$ *such that* $\Delta\bar{i}_L < 0$, $\Delta\bar{i}_M \gtreqless 0$, *and* $\Delta\bar{i}_H > 0$;

(ii.2) *Changes in the occupational structure:* $\Delta E_T > 0$, $\Delta E_B > \Delta l$, *and* $\Delta E_M < \Delta m$;

(ii.3) *Changes in real wages:* $\Delta W_M < 0$, *and* $\{\Delta W_L, \Delta W_H\} \lesseqgtr 0 > \Delta W_M$;

(ii.4) *Changes in relative wages:* $\Delta W_H/W_M > 0$, $\Delta W_M/W_L < 0$, *and* $\Delta W_H/W_L \lesseqgtr 0$.

Proof: See Appendix A.3.

Result (ii.1) establishes that low educated workers are displaced to occupations of lower complexity, high educated workers are employed in more complex tasks, and medium edu-

cated workers take over a larger number of tasks in the middle of the task distribution such that the average task that they perform may increase or decrease. Result (ii.2) follows from result (ii.1). It shows that the employment share in top-level tasks increases because a share of medium educated workers is incorporated into these tasks. Similarly, employment share in bottom-level tasks declines less than the reduction in the share of low educated workers because some medium educated workers start carrying out some of those tasks. Because a share of medium educated workers start to perform new tasks, changes in employment in medium-level tasks are lower than the increase in the share of medium educated workers. Results (ii.3) and (ii.4) arise from the general equilibrium effects of changes in supply and demand for workers with different skill levels. Wages of medium educated workers decline, while wages of the other groups may increase or decrease but the changes are always higher than the decline in wages of medium educated. The changes in real wages decline the wage gap between medium and low educated, and increase the wage gap of high to medium.

These results can be intuitively interpreted as follows. When the share of medium educated workers increases, they become more abundant and their wages fall. It becomes profitable for employers to start using them in tasks previously performed by low and by high educated workers. Therefore, there is an expansion of the number of tasks carried out by medium educated. As the tasks performed by medium educated expands, there is a displacement of low educated workers towards tasks of lower complexity and of high educated workers to tasks of higher complexity. Note that in this case the employment share in top-level tasks expands because there is an influx of medium educated workers into these tasks. In terms of wages, the supply effect dominates for medium educated and their wages decrease. For low and high educated workers, wages can increase or decrease. For high educated workers, they are now concentrated into more complex tasks for which they are more productive, but the price for these tasks diminishes because they are more abundant now.

Therefore, the value of their marginal product (because we assume markets are competitive, wages equal the value of their marginal product) may increase or decrease. On the contrary, low educated workers are now concentrated in tasks of lower complexity, where they are less productive (although having the comparative advantage on those tasks, productivity is assumed to increase with the complexity of the task). But the prices for these tasks may increase because they are relatively more scarce than before. Therefore, the value of their marginal product (wages) may increase or decrease depending on which effect is stronger. Relative wages of medium with respect to other workers decline given their large drop in wages, while that of high educated with respect to low educated may increase or decrease. The workers that benefit the most from an increase in the share of medium educated are those who get educated (their wages raise to those of medium educated workers) and the workers that are hurt the most are those that already have medium education (their wages will have the largest fall).

In terms of the decomposition approach from equations (1.28)-(1.30), wages of medium educated workers decline due to the supply effect dominates. For low educated, the displacement effect is larger than before because medium educated workers are more substitute of low educated workers than high educated, the productivity effect is lower than under policy experiment (i), and the supply effect dominance is no longer ensured as it was under policy experiment (i). For high educated workers, the effect is ambiguous because the supply effect is zero by definition, they are performing a lower share of occupations which diminishes their wages, while the increase in productivity pushes their wages up.

Panel (B) in Figure 1.3 depicts the effects of policy experiment (ii). As medium education expands, wages for medium educated workers decline and they start to perform a broader set of tasks in the economy, pushing low educated workers towards tasks of lower complexity

(depicted by a decline in I_L) and driving high educated workers to tasks of higher complexity (portrayed by an increase in I_H). The wage gaps of medium to low declines and that of high to medium educated increases due to the downward movement along the curve C_{ML} and upward movement along the curve C_{HM}, respectively.

Turning now to policy experiment (*iii*), the effects of a simultaneous increase in medium and high educated are summarized in the following preposition.

Proposition 3 *Under a type–h&m* educational expansion, policy experiment (*iii*), the share of medium and high educated workers increases from m to m' and from h to h', with a corresponding decline in the share of low educated, such that $\Delta m = m' - m > 0$, $\Delta h = h' - h > 0$, and $\Delta m + \Delta h = -\Delta l$, generating:

(iii.1) $\Delta I_H \lesseqgtr 0$. If $\Delta I_H > 0$ it must be the case that $\frac{m'}{h'} > \frac{m}{h}$.

(iii.2) If $\Delta I_H < 0$, preposition 1 holds.

(iii.3) If $\Delta I_H > 0$, preposition 2 holds.

Proof: See Appendix A.3.

The proposition states that if the share of medium and high educated workers increases at the same time, there is a mixture of the two cases discussed above. Result (iii.1) shows that the threshold separating tasks of medium and high educated may increase or decrease, depending on the extent of the changes in relative supply and on the technology reigning the production function in the economy. It also establishes that for I_H to increase it is necessary (but not sufficient) that the relative supply of medium with respect to high educated workers increases. Results (iii.2) and (iii.3) determine that the changes in the equilibrium of the model

are dominated by the increase in h or the increase in m depending on the directional change in I_H. In other words, which one dominates depends on whether in the new equilibrium high educated workers displace medium educated workers towards tasks of lower complexity ($\Delta I_H < 0$, effect of the increase in high education dominates), or if medium educated workers displace high educated workers in some of the tasks they used to perform ($\Delta I_H > 0$, effect of the increase in medium education dominates).

By looking at propositions 1,2 and 3 and across columns of Table 1.1, there are some common patterns of adjustment across all the policy experiments. First, there is a decline of the occupational attainment for low educated workers under any educational expansion (I_L always decreases). Second, the changes in the occupational composition of employment are small when compared to changes in the educational composition of employment due to the new assignment of workers' types to tasks when there is any educational expansion in the economy. Finally, wages always decline for at least one educational group.

Corollary 1: When an educational expansion takes place, there is always a decline of the occupational attainment for low educated workers, small changes in the occupational composition of employment when compared to the educational expansion, and a decline in wages for at least one educational group.

1.3.1 Welfare analysis: changes in the CDF of the wage distribution, poverty, and inequality

This section studies the effects of the three different policy experiments on the wage distribution by looking into cumulative density functions (CDF), growth incidence curves (that

reflect wage changes at each percentile of the anonymous wage distribution), poverty, and inequality.[18] In the model, the CDF is estimated using the share of workers with different educational level and their respective wages, before and after each educational expansion. Some of the directional changes in wages are ambiguous and depend on the parameters of the model, according to prepositions 1-3. When this is the case, there are two opposite forces pushing wages up and down. Wages in those cases may not change much due to the compensating forces. In this section, I make the simplifying assumption that wages remain constant when the predictions of the model are ambiguous. Figure 1.4 shows the CDF before and after the educational expansion for each policy experiment. The changes in the CDF are better characterized by the growth incidence curves of wages (GIC), which is depicted in Figure 1.5 for each policy experiment. The GIC shows the dollar change in wages in the vertical axis and the percentile of the wage distribution (divided by 100) in the horizontal axis. The label "J'" indicates the value of any variable J after each policy experiment.

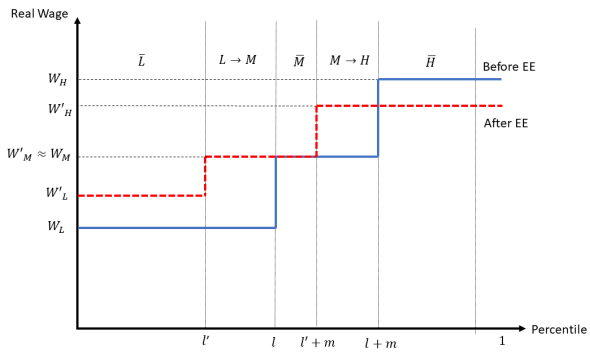
Panel (A) in Figures 1.4 and 1.5 displays the effect on the CDF and the GIC of policy experiment (i). An educational expansion in high education increases wages across the entire wage distribution, except for the wages at the top, those that originally were high educated. The new high educated workers move up in the wage distribution on top of those with medium education, increasing the educational level of most percentiles in the middle of the wage distribution. There is an increase in wages for the lower percentiles that remain with low education ($\bar{L}$), as predicted by the model. Although the share of medium educated workers does not change, they move to the left of the wage distribution, increasing the wages of percentiles that move from being low to medium educated ($L \rightarrow M$). Some percentiles remain medium educated and its wages could increase or decrease ($\bar{M}$), and are depicted

[18]Note that I am not considering a full welfare analysis which should take into account that an educational expansion is costly and that it has to be financed by a tax increase or other changes in government expenditures.

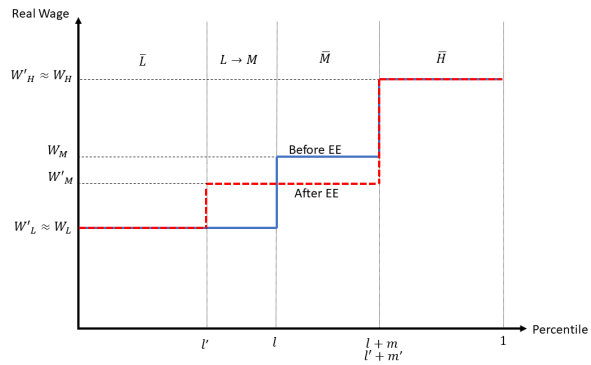
by a dashed line. Wages for percentiles the move from medium to high educated increase ($M \rightarrow H$) because more educated workers have higher earnings (albeit falling due to the educational expansion). Finally, the wage falls for the percentiles that already were high educated located at the top of the distribution ($\bar{H}$).

Figure 1.4: Cumulative density function (CDF) before and after each policy experiment

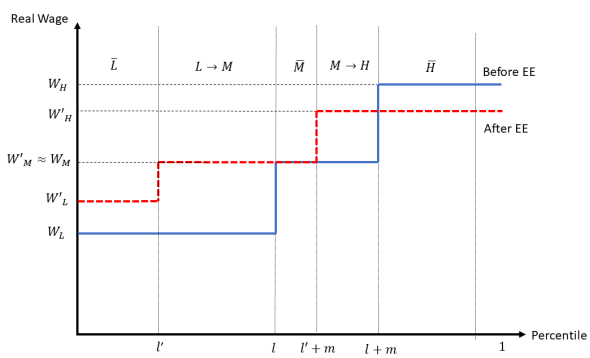
(a) Policy experiment (i) $(\Delta h > 0)$



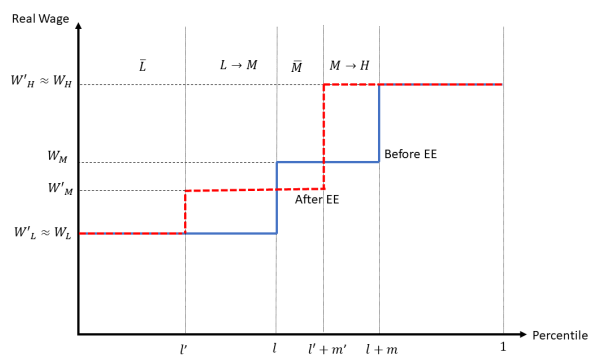
(b) Policy experiment (ii) $(\Delta m > 0)$



(c) Pol. experiment $(iii.a)$: Δh dominates



(d) Pol. experiment $(iii.b)$: Δm domintes

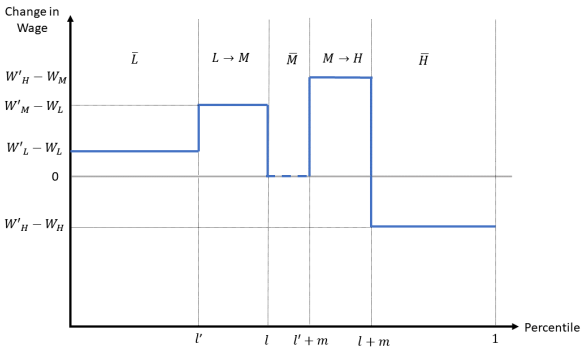


Notes: The figure shows the Cumulative Density Function (inverted) of the distribution of wages before and after the educational expansion under the different policy experiments.

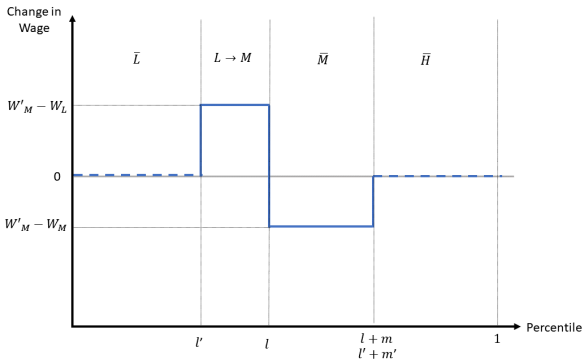
Panel (B) shows the effects of policy experiment (ii). With an expansion in medium education, the wages of percentiles that were originally low educated and are now medium educated ($L \rightarrow M$) increase to the new level of medium educated wages ($W'_M - W_L$), while

Figure 1.5: Growth incidence curve of different type of educational expansions

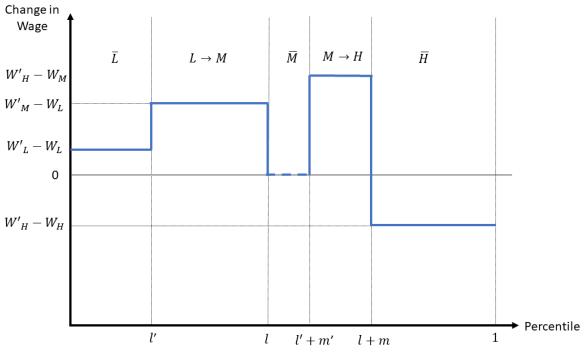
(a) Policy experiment (i) $(\Delta h > 0)$



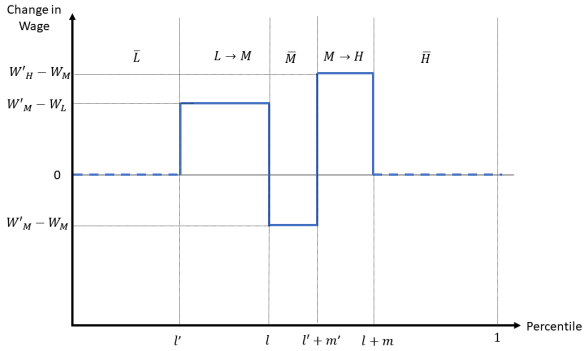
(b) Policy experiment (ii) $(\Delta m > 0)$



(c) Pol. experiment $(iii.a)$: Δh dominates



(d) Pol. experiment $(iii.b)$: Δm domintes



Notes: The figure shows the wage change (in absolute value) after different types of educational expansions at each percentile of the wage distribution. It is the difference in dollars between the curves *Before* and *After* from Figure 1.4.

that of percentiles that were already medium educated ($\bar{M}$) decline as predicted by the model. The changes in wages for those percentiles that remain low and high educated ($\bar{L}$ and $\bar{H}$, respectively) are ambiguous and I characterized them with a dashed line in the GIC. It is clear from the figure which anonymous percentiles benefit and which are hurt with an educational expansion in medium education. Not surprisingly, the percentiles where the educational level increases benefit by a rise in their wages, while percentiles that already had

wages of medium educated are hurt when wages fall as supply increases.

Panel (C) presents the effect of policy experiment (*iii*) when the increase in supply of high educated dominates (case *iii.A* in Table 1.1). The effects in the wage distribution are similar to those in Panel (A), with the only difference of a larger share of percentiles that were originally low educated are now medium educated ($L \to M$).

Finally, Panel (D) shows the effect of policy experiment (*iii*) when the increase in medium education prevails (case *iii.b* in Table 1.1). In addition to the effects in Panel (B), there is an increase in the wage of percentiles that become high educated ($M \to H$). Note that the larger the increase in the share of high educated the smaller the number of percentiles that remain medium educated ($\bar{M}$). If the increase in high educated is large enough, these percentiles for which wage diminishes disappear ($\Delta h > m$). This is the only policy experiment in the context of the model where a wage distribution after an educational expansion can first order dominate the original wage distribution.[19]

Looking across panels of Figures 1.4 and 1.5, I find that an educational expansion concentrated in medium education mostly changes the middle of the wage distribution, while an educational expansion concentrated in high education reshapes the entire wage distribution. I conclude that the labor market effects of policies that increase basic education as opposed to policies that increases higher education are very different, and that increases in higher education are more favorable to lower percentiles of the income distribution.

I turn now to the welfare analysis of the different policy experiments presented here using an abbreviated welfare function. Because an educational expansion always declines the wages of at least one educational group (see corollary 1), in most educational expansions

[19]It is also necessary that the parameters of the model are such that the changes in wages of low and high educated workers, which are ambiguous in the model, are not negative.

there is no welfare improvement characterized by a first order dominance.[20] But a first order dominance is a very demanding welfare criterion to evaluate the policy experiments analyzed here. A more practical welfare analysis can be performed by constructing an abbreviated welfare function, with welfare depending negatively on poverty and inequality, and positively on the total production of the economy.[21] Next, I analyze the predictions of the model on these three relevant welfare inputs.

Changes in poverty follow from the changes in the CDF of the wage distribution depicted in Figure 1.4. The poverty rate is defined as the percentage of workers below a wage poverty line. If the wage poverty line is anywhere between W_L and W_M, the poverty rate declines under policy experiments (i) and $(iii.a)$, but it may increase in policy experiments (ii) and $(iii.b)$ if the new wage for medium educated workers is below the wage poverty line. Note also that other poverty indexes that aggregate poor individuals according to how far they are from the poverty line will further decline under policy experiment (i) and $(iii.a)$ given the increase in wages of those that remain in poverty (workers with low education, considering that the wage poverty line is lower than their new wage level). This is not ensured in experiments (ii) and $(iii.b)$ because wages of low educated workers may fall and the gap between wages and the poverty line may increase.

In terms of inequality, the most natural measures of inequality in this context are the wage gaps. Inequality declines under policy experiments (i) and $(iii.a)$, given that all wage gaps diminish. For policy experiments (ii) and $(iii.b)$, changes in inequality are ambiguous due to a decline in one wage gap (W_M/W_L) and an increase in another (W_H/W_M).

[20] The only exception being case $(iii.b)$ if no percentiles remain medium educated and technology is such that wages of low and high educated do not decline.

[21] In the model, total production is equal to average wages. See section 1.2.3.

Finally, total production increases with any educational expansion.[22] This results from the assumption that more educated workers also have absolute advantage in the production of tasks, so that a more educated labor force that is fully employed under the equilibrium conditions of the model always produces more output.

Corollary 2: Without considering the cost of education, an educational expansion that increases the share of high educated workers (or when it dominates) always improves welfare measured by an abbreviated social welfare function which depends on poverty (negatively), inequality (negatively), and total output (positively). This may not be the case of an expansion in the share of medium educated workers (or when it dominates) since some indexes of inequality and poverty may increase while output always increases.

1.3.2 The importance of the comparative advantage across tasks

This section discusses the importance of the slope of the comparative advantage across tasks to determine the extend of the effects of different educational expansions on the four labor market outcomes of interest: the structure of employment across tasks, the task's assignment of workers of different type, wage levels, and wage gaps. To that end, I evaluate here the effects of an increase in the share of workers with high education under different parameterization of the comparatives advantages curves. The assumptions in the model related to the comparative advantages across tasks is that they must be a function that is continuous, differentiable, and increasing in i, according to equation (1.2). However, the slopes of these relationships are crucial to determine the magnitude of the effects of an educational expansion on each labor market outcomes. In particular, if the comparative

[22] Given that there is a measure of workers equal to 1, total production is equal to production per worker and to average wages.

advantages are steep, the changes in the thresholds are small and the changes in wages are large. On the contrary, if the comparative advantages are flat, an educational expansion results in large changes in the thresholds and small changes in wages.

To see this, consider four different combinations of comparative advantage schedules. The comparative advantage of high educated with respect to medium educated workers, $C_{HM}(i)$, could largely or barely increase with i. Let each of these cases be *steep* $C_{HM}(i)$ and *flat* $C_{HM}(i)$, respectively. Similarly, let *steep* $C_{ML}(i)$ and *flat* $C_{ML}(i)$ be the case of comparative advantage of medium with respect to low largely and barely increasing with i, respectively. There are four possible combinations of steep and flat comparative advantages across tasks. Table 1.2 present a numerical example of the labor market effects of an increase in h for each of the four cases.[23] Figure 1.6 depicts each ones of the comparative advantage curves considered in this example.

By looking across columns, it is clear that in the case of flat comparative advantages, an educational expansion produces large changes in occupations and small changes in wages. On the contrary, in the case of steep comparative advantages, an educational expansion mainly affects workers' wages, with a small impact on their assignment to different occupations.

The economic intuition behind these results is the following. If the curves are flat, employers find it profitable to use the more abundant high educated workers in many new tasks that were previously performed by medium educated workers, with only a small cost in terms of lower relative productivity on those tasks. That is, the demand for high educated workers is more elastic due to the easiness of switching to occupations of lower complexity for

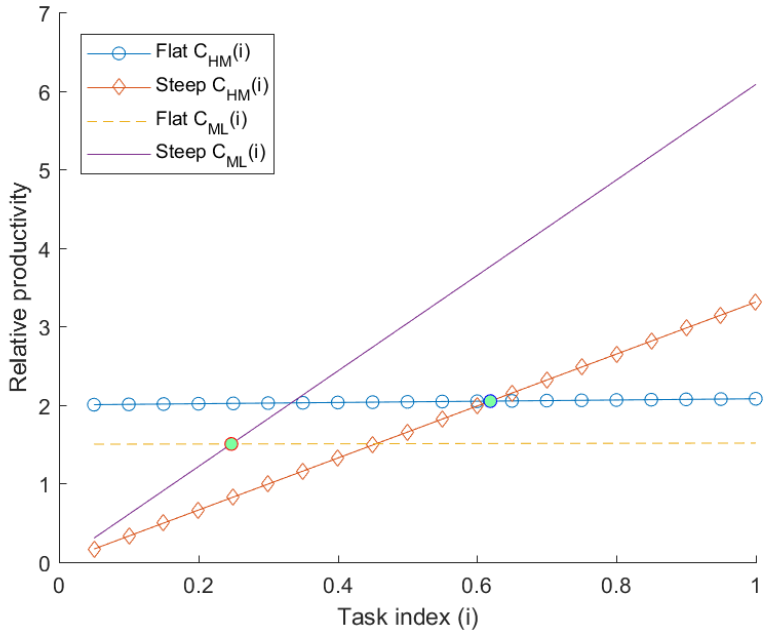
[23]The table and figure show the effect of an increase in h of 10 percent points, from 0.2 to 0.3. The rest of the parameters are fixed and they are defined as follows: $m = 0.4$, $A_H = 4$, $A_L = 2$, $A_M = 2.5$. In the flat scenario,$\alpha_L = 1 + 0.0001 * i$; $\alpha_M = 1 + 0.01 * i$, $\alpha_H = 1 + 0.05 * i$. In the steep scenario, $\alpha_L = i$; $\alpha_M = i^2$, $\alpha_H = i^3$, while A_L, A_M, and A_H are estimated so that the thresholds and the initial wage is the same than in the flat scenario.

Table 1.2: A numerical example on the differential effect of Δh for different scenarios of comparative advantage across tasks.

	Case 1	**Case 2**	**Case 3**	**Case 4**
Comparative adv. schedule				
C_H	*Flat*	*Steep*	*Steep*	*Flat*
C_L	*Flat*	*Steep*	*Flat*	*Steep*
Panel A: Thresholds				
Initial I_L	0.304	0.304	0.304	0.304
Initial I_H	0.686	0.686	0.686	0.686
Final I_L	0.212	0.255	0.224	0.224
Final I_H	0.566	0.610	0.598	0.586
Panel B: Changes in occup. composition of employment				
Bottom-level (E_B)	0.40	-4.48	-1.44	-1.16
Medium-level (E_M)	-2.10	0.33	-1.99	-1.59
Top-level (E_T)	1.71	4.15	3.43	2.75
Panel C: Changes in mean occup. ranking				
Low educated ($\bar{i}_L$)	-0.046	-0.025	-0.040	-0.028
Medium educated ($\bar{i}_M$)	-0.107	-0.063	-0.084	-0.079
High Educated ($\bar{i}_H$)	-0.060	-0.038	-0.044	-0.050
Panel D: % Changes in wages				
Low educated	0.2	16.3	5.1	16.2
Medium educated	0.2	-2.8	5.0	-5.4
High Educated	-0.3	-13.6	-8.4	-5.7
Panel E: Changes in wage gaps				
W_H/W_M	-0.57	-10.90	-12.82	-7.25
W_M/W_L	-0.34	-16.91	-0.35	-3.54
W_H/W_L	-0.90	-25.96	-13.13	-10.53

Notes: The table shows the theoretical predictions of the model under four different scenarios of comparative advantage across tasks. All scenarios start with the same thresholds I_L and I_H, and simulate the effect of an increase in 10 percentage points in the share of high educated (and a corresponding reduction in the share of low educated). Column one shows the scenario when both comparative advantage schedules are flat, that is, they barely increase with the complexity of the task denoted by i. Column two displays the scenario of steep comparative advantage schedules. Columns 3 and 4 show a combination of steep and flat comparative advantage schedules. Parameters are calibrated so that initial wages and thresholds are the same in each column.

Figure 1.6: Four different scenarios of comparative advantage schedules



Notes: The figure shows $C_{HM}(i)$ and $C_{ML}(i)$ under a flat or a steep scenario of increase in comparative advantage across tasks (i). Parameters are chosen such that initial I_H and I_L are the same scenario (depicted by a point at the intersection of the flat and steep comparative advantage).

which they are almost as productive as in the occupations they were originally performing. When the comparative advantages are steep, wages have to diminish sharply before employers find it profitable to start employing high educated workers in a small number of tasks of lower complexity. In this case, the labor demand is more inelastic and wages react more to changes in supply.

Note that in all cases an educational expansion of high educated workers increases the employment in Top-level occupations. That is, a *type–h* educational expansion always improves the occupational structure of employment, at least in the four numerical examples analyzed here. These improvement will be larger with more steep comparative advantage curves since employers will have less incentive to use high educated workers in tasks of lower complexity.

1.4 Conclusions

To increase the educational level of the workforce is usually one of the priorities of every developing country and of the Sustainable Development Goals of the United Nations. In some countries, public policies can and have been put in place to rapidly increase the educational level of the workforce. In this context, it is important to develop an understanding of the general equilibrium effects of these policies. Increases in education may not only impact the labor market outcomes of those who are being educated, but also of those that remain uneducated or the ones that already have a given educational level, positively or negatively.

This paper developed simple framework to evaluate the labor market effects of different types of educational expansions in four labor market outcomes: the occupational structure of employment; the assignment of workers' with different level of education to occupations; the wage level of workers with different education; and the wage gaps between educational groups. In particular, I focused on the effects of three different policy experiments of increasing secondary schooling, increasing higher education, or both at the same time. I find that the labor market impacts of increases in secondary schooling largely differ from increases in higher education, and I also find that one of these effects will dominate with a simultaneous increase in both.

The main difference between the educational policies with respect to occupational outcomes is that an increase in higher education declines the conditional occupational attainment for each educational group, while an increase in secondary education improves the occupational attainment of high educated workers and expands the range of occupations where medium educated workers are employed. With respect to wages, an increase in higher education raises wages of low educated, lowers wages of high educated, and produces am-

biguous changes in wages of medium educated. On the other hand, an educational expansion focused on medium education declines wages of medium educated and generates ambiguous changes in the rest. Finally, I find the an increase in higher education always declines wage inequality as measured by wage gaps of workers with different educational level, while an increase in medium education reduces the wages gap between medium and low educated workers but increases the gap between high and medium educated.

I also find some similarities between the different types of educational expansions. When an educational expansion takes place, there is always a decline on the occupational attainment of low educated workers, small changes in the occupational composition of employment when compared to the educational expansion, and a decline in wages for at least one group.

Finally, I performed a welfare analysis of the different educational policies. I conclude that, without considering the cost of education, an educational expansion that increases the share of high educated workers (or when it dominates) always improves welfare measured by an abbreviated social welfare function which depends on poverty (negatively), inequality (negatively), and total output (positively). This may not be the case of an expansion in the share of medium educated workers (or when it dominates) given that inequality and poverty do not unambiguously decline.

CHAPTER 2

THE LABOR MARKET EFFECTS OF AN EDUCATIONAL EXPANSION: AN APPLICATION TO BRAZIL

2.1 Introduction

The aim of this chapter is to apply the theoretical framework built in the previous chapter to study the case of Brazil. During the period 1995-2014, Brazil underwent several major educational reforms that largely expanded the educational level of its workforce. To evaluate the labor market impact of this educational expansion, I proceed as follows. First, I provide some stylized facts for Brazil on the inter-linkages between changes in education, occupations, and wages over the period of 1995-2014. In particular, I document changes in four labor market outcomes in which this research is focused on: 1) the occupational structure of employment (the share of workers employed in each occupation); 2) the assignment of workers with different level of education to occupations; 3) wages of workers with different education; and 4) wage gaps between educational groups. I also document changes in wage-poverty and other measures of wage-inequality resulting from these labor marker outcomes. Then, I use the model developed in Chapter 1 to evaluate if the model's predictions for the Brazilian educational expansion are qualitatively consistent with the patterns observed in the data. Finally, I calibrate the model and assess if the educational expansion is quantitatively relevant to explain the labor market changes observed in Brazil during the analyzed period.

An exceptionally large educational expansion took place in Brazil between 1995 and 2014 accompanied by some remarkable changes in the labor market market. The Brazilian educational expansion is reported to be one of the fastest expansions in history (Bruns *et al.*, 2011). The share of workers with secondary education doubled from 20.5 to 40.0 percent, the

share with university level grew from 11.3 to 23.6 percent, and the share of workers with only primary education or less halved from 68.1 to 36.4 percent. With respect to the four labor market outcomes of interest, I find that the occupational structure of employment improved, but that improvement was very small when compared to the educational expansion. For example, there was an increase in employment of only 1.9 percentage points in the one third of occupations with the highest wages despite the expansion of 13 percentage points in the share of high educated workers. I also find that there was a marked decline in the occupational attainment within each of the educational groups—primary or less, secondary, and university—defined as an increase in employment in occupations of lower wage-ranking.[1] Despite the common decline in occupational attainment across educational groups, average wages large increased (27.8 percent) but not for all groups: wages of primary educated workers increased (37.3 percent), wages of secondary educated declined (10.5 percent), and wages of university educated fell (21.6 percent). These changes in wages imply large reduction in wages gaps between educational groups as well as marked reductions in wage-poverty and other measures of wage inequality.[2]

Taken together, these facts are puzzling for standard theories evaluating the labor market effects of an educational expansion but they are consistent with the model developed in Chapter 1. A large increase of employment in high-rank occupations would have been expected as result of the educational expansion if labor markets were fragmented as in the broadly use relative demand and supply framework. In this theory, workers with different educational levels participate in separate labor markets with specific occupations such that

[1] I rank ISCO-88 at 3-digit occupations between 0 and 1 according to their median wages over the entire period, with 0 representing the occupation with the lowest median wage and 1 representing the one with the highest median wage. The average ranking diminished by 15.2 percent, 28.2 percent and 8.5 percent for primary, secondary and university educated workers, respectively, between 1995 and 2014.

[2] The magnitude of the reductions in inequality and poverty depend on the index being considered. For example, the Gini index declined by 0.085 points and the wage poverty rate (using a 2.4 USD dollar-a-day at 2011 Purchasing Power Parity wage poverty line) diminished by 16.4 percentage points.

educational expansion shifts the labor supplies in each market without changing the occupations performed by each type of worker (Becker, 1964; Katz and Murphy, 1992; Goldin and Katz, 2009; Gasparini *et al.*, 2011a; Messina and Silva, 2017). Another strand of the theory ties wages to occupations, predicting that average wages should have fallen for each educational group given the widespread decline in occupational attainment (Fields, 1995; Lazear *et al.*, 2016).[3] Contrary to these predictions, the occupational structure of employment remained relatively fixed when compared to the educational expansion while changes in wages were heterogeneous. The failure of these standard models to trace back the changes in the labor market to the educational expansion may lead to the conclusion that other factors, such as the increase in the minimum wage (Engbom and Moser, 2017) or the commodity boom (Adão, 2015), played a more important role in changing labor markets outcomes than the educational expansion. However, I show in this chapter that all the observed labor market changes are perfectly consistent with the predictions of the model developed in chapter 1, which contains elements of these two theories, suggesting that the educational expansion was the most relevant factor behind the changes in the labor markets of Brazil over the period 1995-2014.

The model qualitatively predicts all the observed labor market changes in the occupational structure of employment and the wage distribution that took place in Brazil. Consider the effects of the *type-H* educational expansion, consisting of an increase in higher education by reducing workers with primary or less, which produces the same effects than a *type-H&M* when H dominates but is easier to understand (the general predictions of a *type-H* expansion are the same than for *type-H&M* when H dominates, as shown in Table 1.1). When the share of workers with high education increases, they become more abundant and their wages fall in the occupations in which they were originally employed, while the supply of

[3]See Chapter 1 for more details on these models.

low educated workers declines and their wages increase in the occupations they were initially employed. Therefore, it becomes profitable for firms to start hiring high educated workers in the best occupations that were previously performed by medium educated workers, and to use medium educated workers in the best occupations previously performed by low educated workers, generating lower occupational attainment for all educational groups. The occupational composition of employment for the whole economy only changes slightly given that the high (medium) educated workers move into occupations they did not use to perform, such that the employment in initially highly-paid (low-paid) occupations increases (declines) by less than the increase in the share of high educated workers. Wages of high educated workers decline due to the increase in supply. Wages of low educated workers rise because of the decline in supply coupled with a higher demand for the occupations they perform, given that there is an increase in the production of other occupations.[4] The model also predicts that changes in wages of medium educated workers are always between those of low and high educated workers and these changes could be positive or negative (the negative effect of the lower occupational attainment could be partially or totally offset by an increase in the value of the new occupations that they perform). Inequality declines, as measured by the wage gaps. Poverty diminishes because of two effects. First, some of the previously poor workers escape poverty by becoming more educated, which allows them to earn higher wages, even when the wages of high educated workers are lower then before. Second, wages of the remaining low educated also increase due to the general equilibrium effects originated by the educational expansion. All these patterns in occupations, wages, poverty, and inequality exactly match the stylized facts from Brazil.

In the final part of the paper, I examine how much of the observed changes in the labor

[4] Although labor productivity at each occupation does not change with an educational expansion for any educational group, there is an implicit complementarity in the production of the final good by combining occupations in a Cobb-Douglas production function with an elasticity of substitution equal to one.

market can be explained by the educational expansion through the lens of the model. To that end, I calibrate the model using the data from 1995 as the baseline year. Then, I estimate the effects of the Brazilian educational expansion on the four labor market outcomes of interest, as well as in wage-poverty and other measures of wage-inequality, by isolating the effects of an increase in the education level of the workforce on occupations and the wage distribution when all the other factors, such as technology and educational quality, are held constant. By comparing the equilibrium outcomes in the model with the ones observed in the data, I found that the model's predictions of the effect of the Brazilian educational expansion are remarkably accurate. I conclude that the increase in education was of utmost importance to explain the changes in the Brazilian labor market in the last two decades.

In particular, changes in the average ranking of occupations in the model follow closely that of the data: the model predicts a decline in the occupational attainment of all educational groups, with a larger decline for medium educated workers (the average occupational ranking declines by 0.126 in the model compared to 0.128 in the data). With respect to the overall composition of employment, the general prediction of the model is that the occupational structure improves, but its changes are small when compared to the large educational expansion, although the model predicts larger changes than the ones in the data. For example, the model predicts an increase of 7.6 percentage points in the share of high-paid occupations, while in the data it only increases by 1.9 percentage points.[5] For real wages, average wages rise 29.3 percent in the model and 27.8 percent the data; wages of low educated workers increase 46.9 percent in the model compared to 37.3 percent in the data; wages of medium educated workers decline 10.9 percent in the model and 10.5 percent in the data; and wages of high educated workers decrease 22.2 percent in the model and 21.6

[5]One of the reasons for this difference is that the model do not account for low and medium educated workers that initially were performing high-paid occupations and were replaced by high educated workers after the educational expansion. I discuss this at length in Section 2.6.2.

percent in the data. A precise prediction of the changes in the wage gaps follows directly from the accuracy of the model to predict changes in wages for each group. For example, the model predicts a decline in 47.0 percent in the wage gap between high and low educated workers compared to a decline of 42.9 percent in the data. I also show that these predictions are robust to using different years and functional forms to calibrate the parameters of the model.

Finally, by using the calibrated model to run counterfactuals, I show that the effect of increases in education on average wages declines rapidly with successive educational expansions when technology is fixed. I find that a further educational expansion over a much more educated workforce in 2014 will have less than a third of the effect it had in 1995 when the workforce was less educated. The reason for this is that more educated workers are increasingly employed in tasks of lower complexity, where their relative productivity diminishes, reducing the additional impact of the increase in education on the total output of the economy and, therefore, on average wages. From this exercise, I conclude that further educational expansions in Brazil will have a much lower impact than previous expansions, and this is especially true for increases in medium education.

The main contribution of this chapter is to provide evidence that the increase in education was the main factor behind the changes in the occupational structure of employment and the wage distribution in Brazil between 1995 and 2014. There is a heated debate in the literature that studies the contributing factors to the decline in inequality and wage gaps of workers with different educational levels in the specific case of Brazil. These factors include: educational upgrading which declined returns to education (Barros *et al.*, 2010; Gasparini *et al.*, 2011a; López-Calva *et al.*, 2016; Alvarez *et al.*, 2017); increases in the minimum wage that were spread throughout the wage distribution (Engbom and Moser, 2017); and the fall

in returns to experience which compressed the wage distribution (Ferreira *et al.*, 2016).[6] In my results, I provide additional support to the first factor by showing that the educational expansion observed in Brazil not only explains most of the changes in key aspects of the wage distribution, but it is also consistent with the changes in the occupational structure of employment in the analyzed years, which has receive much less attention in the literature.

Other papers have empirically studied the effects of an educational expansion on wages of all educational groups by exploiting regional variation within the United States (Rauch, 1993; Acemoglu and Angrist, 2000; Moretti, 2004a). These papers find that an increase in the supply of college graduates raises wages of all educational groups, even that of college graduates, attributing these effects to productivity spillovers. In this paper, I find that an increase in the share of high educated workers also raises the wages of low educated workers, but wages of high educated decline. The main reason for this discrepancy could be that the empirical strategy of these papers only captures small increases in the share of high educated workers, as opposed to the major increase that took place in Brazil, where the supply effect is more likely to dominate. Another reason could be that productivity spillovers require directed technological changes for the educational group that now is more abundant (Acemoglu, 2002), which may take place over a longer period of time in developing countries than in developed ones.[7]

This paper also contributes to the literature that studies the changes in the occupational structure of employment. My results contribute to better understand the effect of educational

[6] Other studies are focused on different factors. Ulysseas (2014) and Meghir *et al.* (2015) study the role of formal-informal gaps. Alvarez (2017) analyzes agricultural wage gaps. Dix-Carneiro and Kovak (2015) look at the effect of the trade liberalization during the nineties. Adão (2015) relates the decline in inequality to shocks in commodity prices.

[7] Khanna (2015) uses similar techniques exploiting regional and cohort variation to evaluate the general equilibrium effects of an educational reform in India which largely expanded the access to primary education in half of the country. He finds that the reform largely increased the share of high educated workers, increasing the wages of low educated workers and declining the wages of high educated workers. This is consistent with the supply effects dominating in developing countries at least in the short-run or medium-run.

expansions on the occupational composition of employment, a factor that has not received much attention in this literature. In developed countries the changes in occupational structure are characterized by a hollowing-out of middle-wage occupations, with a corresponding increase in low and high wage occupations, leading to job polarization (Katz *et al.*, 2006; Goos *et al.*, 2014). This distinctive pattern on occupational changes has been related to a rapid change in technology in a context where education is relatively fixed (Acemoglu and Autor, 2011; Autor and Dorn, 2013; Deming, 2017; Beaudry *et al.*, 2016; Burstein *et al.*, 2016).[8] In the case of Brazil, I find that job polarization has not taken place. Far from that, the occupational structure has been particularly rigid and the educational expansion was mostly associated with lower occupational attainment, meaning that more educated workers are increasingly employed in occupations previously performed by lower educated workers. I show that the patterns observed in Brazil are expected in countries where education is increasing rapidly, and it may offset the effects of technological changes that are usually associated with job polarization.[9]

The rest of this chapter is organized as follows: Section 2.2 describes the data; Section 2.3 contains the stylized facts for Brazil; Section 2.4 calibrates the model with Brazilian data; Section 2.5 presents the results from the calibrated model; Section 2.6 discuss alternative explanations; and Section 2.7 concludes.

[8] In developed countries, a large expansion in education took place in the past century, and technological progress is the main force behind recent changes in the assignment of workers' types to occupations and their corresponding wages. In particular, there is an special interest in the effects of labor-saving technologies that have polarized the labor market in developed countries (Autor, 2014).

[9] Maloney and Molina (2016) finds that Brazil is not an exception and the labor markets in most developing countries show no evidence of job polarization.

2.2 Data

This paper uses data from the Pesquisa Nacional por Amostra de Domicílios (PNAD), a nationwide household survey for the years 1995 to 2014. Workers are classified into low, medium and high educated. *Low educated* workers are those with completed primary education or less (less than 9 years of schooling). *Medium educated* workers are those with some or complete secondary education (between 9 and 11 years of schooling). *High educated* workers are those with some or complete university or another tertiary education (more than 11 years of schooling).

Hourly wages are expressed in 2005 purchasing power parity (PPP) dollars, and are estimated by dividing monthly labor earnings by hours worked in the corresponding month. The analysis is performed on employed workers between 18 and 55 years old, and living in regions that have been surveyed throughout the entire period.[10] Observations are re-weighted to hold constant the demographic composition within each educational group at the levels of 1995.[11]

Codes of different occupations are harmonized using the International Standard Classification of Occupations of 1988 (ISCO-88). In the household surveys occupations are classified according to the *Classificação Brasileira de Ocupações* (CBO), which changed between PNAD 1996-1999 and PNAD 2001-2014. I follow Salardi (2014) for recoding the occupations from the CBO for each period into ISCO-88 at 3 digits.

[10] I exclude regions from rural north as they were incorporated in the sample in 2004.

[11] I follow DiNardo *et al.* (1996) by constructing 20 cells within each educational group: two genders, five age categories and two sub-levels of education. I fix the share of each cell to the value of 1995 (initial year). This procedure avoids confounding changes in average wages of a particular educational group with a change in its gender composition, its age structure, or more disaggregated educational attainment within each educational group. Results are robust to not implementing this re-weighting procedure and are available upon request.

To rank occupations in the data, I follow the standard procedure in the literature of using the wage percentile of each occupation (Katz *et al.*, 2006; Autor and Dorn, 2013; Beaudry *et al.*, 2016; Deming, 2017). The wage percentile rank is usually estimated using the mean wage of the occupation in the baseline year. I opt for a different rank procedure that is better suited for my analysis: the median wage percentile rank of occupations taking into account the entire period 1995-2014. I prefer to use the median instead of the mean to avoid the influence of outliers. Moreover, I use all years between 1995 to 2014 for three reasons. First, it allows me to have more observations in each occupation. Second, the rank is less susceptible to changes in the characteristics of the occupations that may impact wages of particular occupations.[12] Third, the model presented in Section 1.2 predicts that changes in occupations' wages throughout the entire period are informative on its ranking.[13]

Occupations are usually considered in the literature as the best description a researcher has on the type of job a worker performs.[14] Some occupations are assumed to be more complex than others. For example, the tasks performed by an electrical engineer are arguably of higher complexity than that of a housekeeper. Although the tasks' content of different occupations may differ in more than one dimension (for example, routine manual, routine cognitive, and non-routine cognitive tasks), the relative wage of an occupation is assumed to be an summary indicator of how complex that occupation is when compared to other

[12]For example, clerks may had a decline in their relative position when compared to other occupations because of the change in the task content of their job. Clerks were in percentile 51 in 1995 and declined to percentile 17 in 2014, while using my classification with date of the entire period they are in percentile 27. Except for some extreme cases like this one, the percentile rank of an occupation is similar using different classifications such as the median wage in 1995, 2004, 2014, or the median wage during the entire period 1995-2014. The correlation of the ranking produced by these different alternatives and the one used in this paper is always above .94, as shown in Appendix Table B.2.

[13]For example, the model predicts that the first occupations to be overtook by high educated workers from medium educated workers with an educational expansion are those of higher complexity previously performed by medium educated workers. Wages in those occupations increased to the level of high educated workers.

[14]See Autor (2013) for a detailed discussion on the information content of occupations.

occupations in the economy.[15] Therefore, the use of wages to rank occupations provides useful information regarding their relative position in terms of how complex their tasks content is.[16]

2.3 Changes in the Brazilian labor market between 1995-2014

2.3.1 Education

During the 1995-2014 period, there was a major increase in the educational attainment of the Brazilian workforce. Panel A in Table 2.1 shows the share of workers with low (primary or less), medium (some secondary), and high education (some university) in 1995 and 2014. In only 19 years, the share of workers with medium and high education doubled, while the share of low educated workers halved. The share of low educated workers decreased from 68.1 percent in 1995 to 36.4 percent in 2014, while the share of medium educated increased from 20.5 to 40.0 percent and the share of high educated workers rose from 11.3 to 23.6 percent.[17]

To contextualize the extent and pace of the educational expansion that took place in

[15]This approach has been extensively used in the literature (Acemoglu and Autor, 2011; Autor and Dorn, 2013; Beaudry *et al.*, 2016). However, there is an increasing literature that further discriminates among different tasks within an occupation by matching occupations with other datasets with detail tasks' information (for example, the Dictionary of Occupational Titles for the United States). This information is not available for the case of Brazil.

[16]It could be the case that wages in the public sector or in unionized jobs introduce noise to this ranking in Brazil since it may not reflect the relative position of an occupations in terms of the complexity of its tasks. To show that this is not driving the results of this paper, Appendix Table B.2 shows that the ranking is practically the same (correlation of .98) when dropping workers that belong to a union or that work in the public sector.

[17]The extent of the educational expansion is practically the same when using the share of total hours worked instead of the share of employed workers.

Table 2.1: Changes in the educational attainment of the workforce and the Brazilian wage distribution between 1995 and 2014

	Brazil		
	1995	2014	Change
Panel A:Educational attainment (share of the workforce)			
Low educated (less than secondary)	0.68	0.36	-0.32
Medium educated (some secondary)	0.21	0.40	0.19
High educated (some university)	0.11	0.24	0.13
Panel B: Average Wages			
Total workers	2.36	3.02	27.8%
Low educated	1.47	2.02	37.3%
Medium educated	2.76	2.47	-10.5%
High educated	7.00	5.49	-21.6%
Panel C: Wage gaps			
W_M/W_L	1.88	1.22	-34.8%
W_H/W_M	2.54	2.22	-12.4%
W_H/W_L	4.76	2.72	-42.9%
Panel D: Inequality			
Gini	0.445	0.360	-0.085
Theil			
Total	0.327	0.219	-0.109
Between educ. groups	0.071	0.051	-0.020
Atkinson (Inequality aversion = 2)			
Total	0.572	0.440	-0.132
Between educ. groups	0.140	0.062	-0.078
Panel E: Wage poverty (wage line: 2.4)			
FGT(0)	0.626	0.462	-0.164
FGT(1)	0.305	0.144	-0.162
FGT(2)	0.183	0.065	-0.118

Notes: The table shows the levels and changes in the educational attainment of the workforce and in different outcomes of the wage distribution estimated from PNAD 1995 and 2014. The column changes displays the absolute changes in each variable between 2014 and 1995 unless is preceded by % indicating percentage change.

Brazil, it is useful to compare it with changes in enrollment rates in the rest of the world over a similar period of time. Panel (A) of Figure 2.1 shows the changes in enrollment rates in secondary and university education across countries, according to the data from Barro and Lee (2013). Brazil is the country with the largest increase in secondary school enrollment, going from 16.0 percent in 1990 to 86.2 percent in 2010, closely followed by other Latin

American countries and Portugal.[18] In university education, the enrollment rate in Brazil increased 20 percentage points, around the average of the countries considered in the data. The educational expansion in Brazil between 1990 and 2010 was also large from a historical point of view for the country. Panel (B) in Figure 2.1 plots the evolution of enrollment rates since 1975. Except for a moderate increase in primary education, enrollment rates were practically constant in Brazil between 1975 and 1990. From 1990 to 2010 enrollment in secondary schooling soared, while university enrollment increased steadily after 1995. The substantial increase in enrollment had an impact on the educational level of the labor force as younger and more educated cohorts replaced older less educated cohorts. According to Bruns *et al.* (2011), the rise in the educational attainment of Brazil's labor force since 1995 has been one of the fastest on record in history.

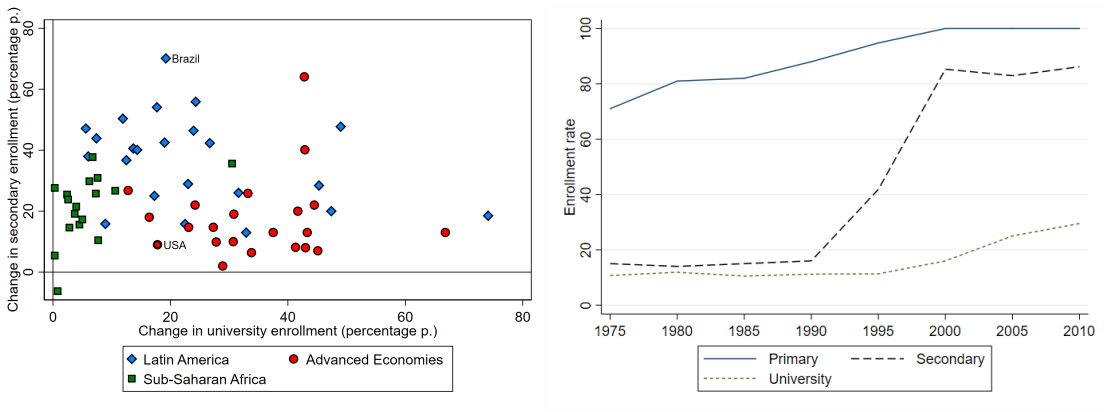
Brazil implemented several important educational reforms to increase the educational level of the population. These reforms included, among others: Financial reforms such as an increase in public expenditure in education from 2.0 percent of GDP in 1995 to over 5.0 percent in 2010 and a redistribution of resources towards poor municipalities in primary education in 1996 (a reform called Fundef) and secondary education in 2007 (Fundeb); increases in accountability by implementing nation wise standardized tests to track the learning progress over time in primary and secondary schooling; and reducing the direct and indirect cost of schooling by creating more schools and universities (less travel time) together with conditional cash transfers (Bolsa familia) and other scholarships (for example, ProUni) that particularly benefited poor households.[19] Bruns *et al.* (2011) estimate that, as a consequence of the complementarity between these reforms, a six-year-old Brazilian child starting school in 2010 from the bottom quintile of the income distribution will, on average,

[18]The enrollment rate is defined as the ratio of students at a given level of schooling in the designated age group to the total population of that age group.

[19]Bruns *et al.* (2011) and OECD (2011) provide a detailed description of these reforms.

Figure 2.1: Changes in enrollment rates in Brazil

(a) Across countries comparison (1990-2010) **(b)** Historical comparison



Notes: Panel (a) shows the changes in enrollment rates across countries between 1990 and 2010 in secondary and university/tertiary education. Panel (b) shows the historical evolution of enrollment rates for primary, secondary and university education in Brazil. Enrollment ratios are defined as the ratio of students at a given level of schooling in the designated age group to the total population of that age group.
Source: Barro and Lee (2013).

complete more than twice as many years of schooling than her parents have.[20]

Data from PNAD shows the large increase in secondary schooling between 1995 and 2014: the share of the population between 20 and 24 years old that finished secondary schooling increased from 24.0 to 63.3 percent during this period. The demand for higher education skyrocketed. According to the data from Instituto Nacional de Estudos e Pesquisas

[20]It is possible that the large increase in educational expenditure after 1995 affected labor markets indirectly by introducing changes in tax policy. For example, tax revenues as percentage of GDP increased from 11 percent in 1995 to 14.6 percent in 2001 (World Development Indicators, 2018) in part because of an increase in tax on incomes, profits and capital gains that represented 13.5 percent of total revenues in 1995 and 25.3 percent in 2001. These taxes may have distorted capital and labor supply decisions (Lledo, 2005). However, most of the increase in taxes was introduced by the national government, while the increase in educational expenditure took place at a municipality and regional level because of the implementation of Fundef (Bruns *et al.*, 2011), making it difficult to disentangle how the educational expansion was actually financed. Due to this limitation, instead of building a general equilibrium framework including the public sector, I consider the educational expansion as exogenous of the labor market. I acknowledge that this is a limitation of this study, which is restricted to estimating the direct labor market effects of the educational expansion instead of a full welfare analysis considering the cost and other indirect effects originated by it.

Educacionais Anísio Teixeira (INEP), the number of students in tertiary education increased from 1.7 million in 1995 to 6.6 million in 2015. But the selection process became tougher: 19 percent of the candidates were accepted in 1995, as opposed to only 14 percent in 2015. Between 1995 and 2015, the number of candidates increased from 2.6 to 14.3 million. The supply and diversity of higher level institutions also increased. The number of institutions offering university/tertiary education went from 894 in 1995 to 2,364 in 2014, driven by the deregulation of higher education which foster the creation of a large number of for-profit private institutions (Ferreyra *et al.*, 2017).

Despite the large educational expansion that took place in secondary and university level, the available evidence points out that the quality of education in Brazil remain relatively unchanged. In secondary education, the students' scores in the Program for International Student Assessment (PISA) increased slightly between 2000 and 2015 in each of the evaluated subjects (science, mathematics, and reading), and the scores in 2009 first-order dominate those of the year 2000, with the highest increase taking place at the bottom quintile (Bruns *et al.*, 2011).[21] The scores are low by OECD standards, but they have not deteriorated in a time of a large educational expansion, which has been recognized as a remarkable achievement (Bruns *et al.*, 2011). Regarding tertiary education, selection to top universities remains highly competitive, as was discussed previously. For example, there are 16 applicants for each accepted candidate in UNICAMP and USP (two of the largest universities in Brazil). According to INEP, approximatively 90 percent of the students in university/tertiary education attend an institution that applied some kind of selection in 2014.[22]

[21]PISA is regarded as one of the best measures of student outcomes. It is constructed to assure the comparison of results across countries and within a country in different periods.

[22]Changes in education quality can also be inferred by comparing data on wages of workers with the same educational level but that were educated at different periods. In Section 2.7.1, I follow Bowlus and Robinson (2012) decomposition to disentangle what part of the changes in wages for each educational group can be explained by a change in education quality (the quantity of human capital inherent to each education level) and which part corresponds to a change in market prices. I find that practically all the changes in wages come from a change in the market price of human capital for each educational level, consistent with

2.3.2 Wage distribution

Along with the educational expansion, the wage distribution in Brazil changed in several important ways between 1995 and 2014. Panels B-E in Table 2.1 presents changes in different dimensions of the wage distribution. The average wage for the total workforce largely increased by a 27.8 percent during this period, but wages did not increase for all educational groups. While the average wage of low educated workers increased by 37.3 percent, the average wage of medium and high educated workers fell by 10.5 and 21.6 percent respectively.[23]

Wage inequality declined sharply in Brazil, contrary to what happened in the United States and other developed countries since the 1980s (Cingano, 2014). The wage gaps between workers with different educational levels diminished, and all the indexes of relative income inequality declined.[24] In 1995, the average wage of high educated workers was 2.5 times that of medium educated, and 4.8 times the average wage of low educated workers. These gaps declined to 2.2 and 2.7 in 2014, respectively. The Gini index went down 0.085 points, from 0.445 in 1995 to 0.360 in 2014. To better understand the importance of changes in wages for workers with different educational levels in the reduction of inequality, I compute other indexes that can be decomposed in between-group and within-group inequality. The Theil index diminished 0.109, with 20 percent of that fall explained by changes in inequality between the three educational groups—low, medium, and high educated—considered in this paper. The importance of the between-group inequality is larger in the case of the Atkinson index with an inequality aversion parameter equal to 2, which puts more weight on changes at the bottom of the wage distribution. The index declined 0.132, with 59.2 percent

the assumption that the human capital content, which relates to educational quality, remained relatively unchanged.

[23]The increase in the overall average wage was driven by the wage increment of low educated workers, and the raise in the share of workers receiving wages of medium and high levels of education (the wage bill increased for these two groups despite the fall in average wages).

[24]See Barros *et al.* (2010) for a detail characterization of the decline in inequality in Brazil.

of that decline due to falling inequality between workers with different educational level. Wage-poverty, defined as the share of workers with wages below USD 2.4 at 2011 Purchasing Power Parity, diminished by 16.4 percentage points from 62.6 percent to 46.2 percent.[25]

2.3.3 Occupational structure of employment

The large educational expansion in Brazil was not matched by a corresponding occupational upgrading between 1995 and 2014. On the contrary, the occupational structure of employment was practically fixed. Table 2.2 displays the employment shares and its changes for different occupation categories and educational groups for 1995 and 2014. Three occupational groups are considered based on the 82 occupations from the ISCO-88-3 digit classification according to the median wage over the period 1995-2016: *Bottom-third* refers to the 27 occupations with lower wage, *medium-third* consists of the next 27 occupations, and *top-third* represents the remaining 28 occupations with the highest wages in the economy. Panel A and B show the employment shares of the total workforce in each occupation-education cell for the years 1995 and 2014 respectively, while Panel C contains the changes in each cell. For example, 48.8 percent of all workers were low educated and employed in bottom-third occupations in 1995 but only 28.4 percent of the workforce fell into that category in 2014, which represents a decline of 20.5 percentage points. Note that the total row share contains the share of employment in each occupational group and the total column share equals the share of workers in each educational group.

By looking at the changes in the total row share in Panel C, it is clear that the occu-

[25] I consider a wage poverty line of USD 2.4 given that it is the income needed for a family of four to not be consider poor by the USD 3.1 dollar-a-day moderate poverty line used by the world Bank if one of the member of the household works at least 40 hours a week.

Table 2.2: Employment shares by education and occupation of total employment

	Low educated	Medium educated	High educated	Total row share
Panel A: 1995				
Bottom-third	48.8	6.4	1.0	**56.2**
Middle-third	14.7	9.7	3.7	**28.0**
Top-third	4.5	4.9	6.4	**15.8**
Total column share	**68.0**	**21.0**	**11.0**	**100.0**
Panel B: 2014				
Bottom-third	28.4	22.1	4.2	**54.7**
Middle-third	6.6	14.0	6.9	**27.5**
Top-third	1.0	3.9	12.8	**17.7**
Total column share	**36.0**	**40.0**	**24.0**	**100.0**
Panel C: Change 1995-2014 (p.p.)				
Bottom-third	-20.5	15.7	3.3	**-1.5**
Middle-third	-8.1	4.3	3.3	**-0.5**
Top-third	-3.4	-1.1	6.5	**1.9**
Total column share	**-32.00**	**19.00**	**13.00**	**0.00**

Notes: The 82 occupations from the ISCO-88-3 digit classification are divided into 3 groups of equal number of occupations according to their median wage over the period 1995-2016. Bottom-third refers to the 27 occupations with lower average wage, medium-third are the next 27, and the remaining 28 are classified as top-third occupations. Workers are classified into low (less than secondary), medium (some secondary), and high educated (some university). Panel A and B show the employment share in each education-occupation cell of the total workforce for 1995 and 2014 respectively. Panel C contains the changes in the employment share in each cell between 2014 and 1995.

pational structure of employment improved between 1995 and 2016 in the sense there was an increase in employment in high-wage occupations. But it also becomes evident that improvement was very small when compared to the large educational expansion. The share of top-third occupations only increased by 1.9 percentage points while the share of low-third and middle-third occupations slightly diminished by 1.5 and 0.5 percent respectively. These small changes contrast with the large educational expansion that took place in Brazil, as shown in the large changes in the total column share. It is striking that the share of top-third occupations only increased 1.9 percentage points while the share of high educated workers increased by 13 percentage points.[26]

[26]The relatively small change in the occupational composition of employment is also evident when looking into the 1 digit ISCO-88 classification. The Appendix Table B.3 displays the changes in the occupational composition of employment overall and for each educational level between 1995 and 2014 and occupational

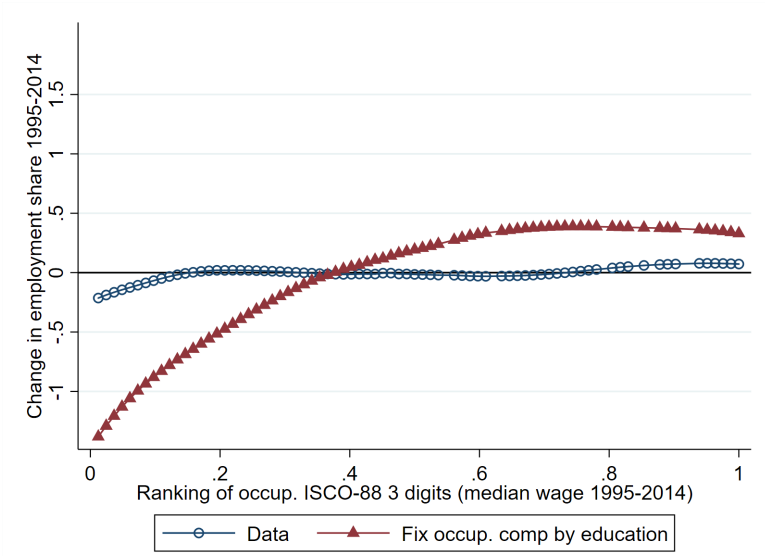
The small improvement in the occupational composition are consistent with more flexible ways of classifying occupations. To be as flexible as possible, I follow Acemoglu and Autor (2011) and Autor and Dorn (2013), among others, in estimating smoothed regressions of the changes in employment share over a ranking of occupations based on the median wage between 1995 and 2014.[27] Figure 2.2 shows the locally weighted regression of changes in employment share along the occupational ranking. The blue-circle line in the figure shows the observed change in employment shares across occupations for the total workforce. With this more flexible approach, I find that there is an increase in high-ranking occupations and a corresponding decline in low-ranking occupations, but the changes in employment are small when compared to the increase in education. To see this, consider the red-triangle line in Figure 2.2. This line displays the result of a thought experiment which simulates the changes in occupations between 1995 and 2014 if workers with a given educational level in 2014 were distributed across occupations as in 1995. This exercise creates a counterfactual on how the occupational structure of employment would have looked like in 2014 if the increases in education had not changed the assignment of workers' type to occupations. The employment share in low ranking occupations would have largely declined, and the employment share would have increased for occupations in the middle and at the top of the ranking. The difference between the blue-circle and the red-triangle lines is due to the lower occupational attainment within each educational level, to which I now turn.[28]

categories are ordered from higher pay to lower pay according to their median wage during the period 1995-2014. The changes for the overall workforce are characterized by a small increase in the share of employment in top ranking occupations and a small decline in low ranking occupations. For example, the share of professionals and managers only increased by 2.5 percentage points, from 13 to 15.5 percent, between 1995 and 2014.

[27] The ranking is robust to using other measures to construct the occupational ranking, such as the average wage of the initial or final period in each occupation. For more details see Section 2.2.

[28] It is also noteworthy that the slightly positive slope in Brazil differs from to the U-shaped pattern that has been estimated for the United States and other developed countries (Acemoglu and Autor, 2011; Goos *et al.*, 2014), where occupations in the middle declined and those at both ends of the occupational distribution increased in the last three decades.

Figure 2.2: Changes in the occupational composition of the workforce 1995-2004



Notes: Occupations at 3 digit level from ISCO-88 are ranked according to their median wage from lower to higher in the horizontal axis. The blue line (circle) in the figure is a locally weighted smoothing regression of the changes in employment shares (percent points) across the occupational ranking based on the median wage of each occupation over the period 1995-2014. The red line (triangle) displays a counterfactual change in the occupational composition of employment between 1995 and 2014 if the occupational composition within each educational group is fixed to the levels of 1995.

Along with the large educational expansion and the small improvement in the occupational structure of employment I find that there was a clear deterioration of the conditional occupational attainment for each educational group, especially for medium educated workers. Table 2.3 displays the share of employment across different occupational categories for each educational group. As expected, in 1995 most of the low educated workers (71.8 percent) were employed in bottom-third occupations, while most of medium and high educated were in middle-third and top-third occupations respectively (46.0 and 57.9 percent). When looking at changes between 1995 and 2014, all educational groups experienced a sizable increase in employment in the bottom-third occupations. For low educated workers, the share of employment in the bottom-third occupations increased by 7.0 percentage points. For medium, the share in bottom-third occupations practically doubled increasing 24.9 per-

centage points. In the case of high educated workers, the share in bottom-third occupations more than doubled from 8.7 percent in 1995 to 17.6 percent in 2014. This lower occupational attainment is also depicted in Figure 2.3 which shows locally weighted regression of changes in employment for each educational group. Each educational groups lost employment shares in higher rank occupations and increased their employment share in lower rank occupations so that the average occupational rank declined, and this pattern was much stronger among medium educated workers.[29]

In summary, there are four stylized facts that characterize the changes in the Brazilian labor market between 1995 and 2014 along with the large educational expansion. First, average wages increased but not for all groups since wages of low educated workers increased and wages of medium and high educated workers fell. Second, wage gaps among workers with different educational levels largely declined. Third, the occupational structure of employment improved but that improvement was surprisingly small when compared to the large educational expansion that took place during this period. Finally, there was a large deterioration in the conditional occupational attainment for each educational group. I argue in the rest of the paper that the educational expansion was the main factor generating these distinct patterns in wages and occupations.

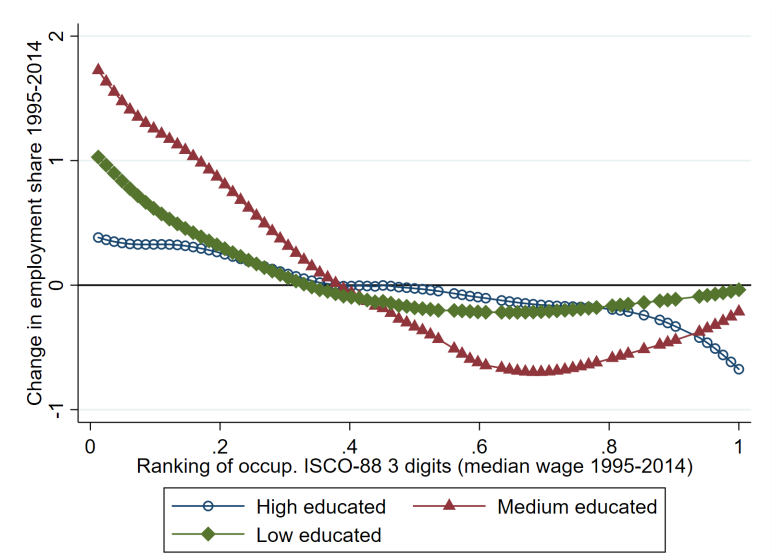
[29]In the Appendix B.1 I explore in detail the relationship between changes in average wages of each educational group and their occupational composition of employment. Using the Oaxaca-Blinder decomposition, I disentangle what part of the changes in averages wages is due to an occupational composition effect (changes in the occupations workers perform) and what part is due to a pay structure effect (changes in wages within occupations). I find that for all workers the composition effect is small and positive, consistent with small improvements in the occupational composition of the economy; while the increase in wages is driven by higher wages at each occupation, especially those that started with a lower wage. I also find that all educational levels experienced a negative occupational composition effect, consistent with lower occupational attainment for each educational level. For low educated workers wages within each occupation largely increased, compensating for the negative composition effect, and average wages increased. For medium educated workers, wages increased only in a small number of occupations but not enough to compensate for the negative composition effect. In the case of high educated workers, wages declined within occupations, reinforcing the negative composition effect, and average wages largely fell. The analysis indicates that changes in both between and within occupations played a relevant role to explain the observed pattern of wages in Brazil.

Table 2.3: Employment shares by occupations and mean occupational ranking for each educational group

	Low educated	**Medium educated**	**High educated**	**Total workers**
Panel A: 1995				
Bottom-third	71.8	30.5	8.7	56.2
Middle-third	21.6	46.0	33.5	28.0
Top-third	6.6	23.5	57.9	15.8
Total	100.0	100.0	100.0	100.0
Mean occupational ranking	0.240	0.449	0.670	0.335
Panel B: 2014				
Bottom-third	78.8	55.3	17.6	54.7
Middle-third	18.3	35.0	29.0	27.5
Top-third	2.9	9.6	53.5	17.7
Total	100.0	100.0	100.0	100.0
Mean occupational ranking	0.203	0.320	0.612	0.268
Panel C: Change 1995-2014 (p.p.)				
Bottom-third	7.0	24.9	8.9	-1.5
Middle-third	-3.3	-11.0	-4.5	-0.5
Top-third	-3.7	-13.9	-4.4	1.9
Total	0.0	0.0	0.0	0.0
Mean occupational ranking	-0.037	-0.129	-0.058	0.007

Notes: The 82 occupations from the ISCO-88-3 digit classification are divided into 3 groups of equal number of occupations according to their average wage over the period 1995-2016. Bottom-third refers to the 27 occupations with lower average wage, medium-third are the next 27, and the remaining 28 are classified as top-third occupations. Workers are classified into low (less than secondary), medium (some secondary), and high educated (some university). Panel A and B show the employment share for each educational group in different occupations and the mean ranking for 1995 and 2014 respectively. Panel C contains the changes in employment shares and mean ranking for each educational group between 2014 and 1995. The mean occupational ranking refers to the mean for each educational group of their occupation percentile according to the median wage over the period 1995-2014 divided by 100.

Figure 2.3: Occupational downgrading for each educational group between 1995-2014



Notes: The figure plots a locally weighted smoothing regression of the changes in employment shares (percent points) across the occupational ranking based on the median wage of each occupation over the period 1995-2014. Workers are classified into low (less than secondary), medium (some secondary), and high educated (some university). Occupations are ranked as in Figure 2.2.

2.4 Comparison of the qualitative predictions of the model and the changes in the data

This section compares observed changes in labor market outcomes in Brazil with the predictions of the model developed in Chapter 1 when *only* the Brazilian educational expansion takes place, holding constant the rest of the parameters of the model. That is, I explore the effects of changes in m and h in the magnitudes observed in Brazil, holding constant the technology of the production function entirely characterized by the set $\{A_L, A_M, A_H, \alpha_L, \alpha_M, \alpha_H\}$. I first discuss the qualitative results, followed by a quantitative assessment of the model's predictions arising from the calibrated model in the next sections.

Table 2.4 displays the directional changes in the data for Brazil between 1995 and 2014, and the qualitative results of the model when there is an increase in the shares of medium and high educated workers.[30] The qualitative predictions of the model exactly match the changes observed in the data: the occupational composition of the workforce presents small changes when comparted to the educational expansion, so that the share of employment in tasks originally performed by high (low) educated workers increases (decreases) by less than the increase in the share of high educated workers (decline in the share of low educated workers); the average complexity of the task performed by each type of worker declines (lower educational attainment for each educational group); wages of low educated increase, wages of high educated workers decline, and changes in wages of medium educated workers are between those of low and high educated; and wage gaps of more educated workers with respect to less educated workers fall.

[30]Section 1.2 described two possible scenarios when both m and h increases, depending on whether I_H increases or decreases. The necessary condition for I_H to increase is that the ratio of the share of high education workers to the share of medium educated workers declines, which is not the case in Brazil. Therefore, the only relevant case here is a decline in I_H.

Table 2.4: Qualitative observed changes vs model predictions

	Brazil 1995-2014	**Model (iii.a)** $\Delta m > 0$ **and** $\Delta h > 0$
Panel A: Changes in thresholds		
I_L		$\downarrow$
I_H		$\downarrow$
Panel B: Changes in occup. composition of employment		
Bottom-third (O_B)	$(\downarrow)\Delta O_B > -(\Delta m + \Delta h)$	$\Delta E_B > -(\Delta m + \Delta h)$
Middle-third (O_M)	$(\downarrow)\ \Delta O_M < \Delta m$	$\Delta E_M < \Delta m$
Top-third (O_H)	$(\uparrow)\Delta O_H < \Delta h$	$\Delta E_T < \Delta h$
Panel C: Changes in mean occup. ranking		
Low educated	$\downarrow$	$\downarrow$
Medium educated	$\downarrow$	$\downarrow$
High educated	$\downarrow$	$\downarrow$
Panel D: Changes in wages		
Low educated	$\uparrow$	$\uparrow$
Medium educated	$(\downarrow)\ \Delta W_H < \Delta W_M < \Delta W_L$	$\Delta W_H < \Delta W_M < \Delta W_L$
High educated	$\downarrow$	$\downarrow$
Panel E: Changes in wage gaps		
W_H/W_M	$\downarrow$	$\downarrow$
W_M/W_L	$\downarrow$	$\downarrow$
W_H/W_L	$\downarrow$	$\downarrow$

Notes: The table shows the qualitative changes in the labor market in Brazil during the years 1995-2014 and the changes predicted by the model under policy experiment (iii.a). The table contents should be interpreted as follows: $\uparrow$ denotes an increase for any value in the parameters of the model; $\downarrow$ denotes a decrease; $\uparrow\downarrow$ denotes that it can increase or decrease; $\{E_H, E_M, E_B\}$ refer to the categories in Table 1.1 that classify occupations in the model. Other cells are filled with lower or upper bound to changes in the outcome variable.

In the context of the model, only a supply shock as the one that took place in Brazil can generate the patterns on occupations and wages observed in the data. Changes in the labor market can arise from supply shocks or from technological shocks in the model. I already showed that a supply shock as the one that took place in Brazil matches all the patterns observed in the data. It remains to be shown that technological changes in the model cannot generate these patterns. First, if the productivity parameters of all workers increase by the same ratio (factor-neutral technical change where A_L, A_M, and A_H experience the same

proportional increase), there is no effect on the assignment of workers to tasks, and wages of all types increase in the same proportion, leaving wage gaps unaffected. Second, a skill-biased technological change, portrayed in the model by an increase in A_H, generates a pattern of lower occupational attainment for each group consistent with the data, but it also predicts an increase in wages of high educated workers, which is at odds with the data.[31] The only technological change that can reduce the wage gaps is an increase in the productivity of low educated workers, that is, depicted by an increase in A_L in the model, but it would also imply occupational upgrading instead of lower occupational attainment: less educated workers will start to perform tasks of higher complexity. A *type-H&M* educational expansion where H dominates is the only plausible change in the model consistent with all the patterns observed in the data.

2.5 Calibration

The procedure to calibrate the parameters of the model with data from Brazil is fairly simple. The set of parameters that need to be estimated are the skill supplies and the productivity across tasks for each worker type. I use the baseline year, the data from 1995, to calibrate all these parameters.

The parameters corresponding to the educational level of the workforce come directly from the data. The values of l, m, and h are the share of the employed workers with low (less than secondary), medium (some secondary), and high education (some university) in the survey of 1995.

[31] This type of technological change has been identified as the main factor influencing the labor markets during the 1990s in the United States and other developed countries (Acemoglu, 1998; Bekman *et al.*, 1998).

What remains to be estimated are the factor-specific augmenting technologies common to all tasks A_L, A_M, and A_H, and the specific productivities across tasks $\alpha_L(i)$, $\alpha_M(i)$, and $\alpha_H(i)$. To that end, I use the average wage for low, medium and high educated workers in 1995. My identifying assumption is to assume a functional form for the functions $\alpha_L(i)$, $\alpha_M(i)$, and $\alpha_H(i)$ that simplifies the equilibrium conditions in the model.[32]

With wage gaps and relative supplies in 1995 it is possible to estimate the thresholds levels I_L and I_H from the equilibrium conditions of the model. Restating equations (1.12), (1.13), it is possible to express:

$$\frac{(1 - I_H^*)}{I_H^* - I_L^*} = \frac{W_H}{W_M}\frac{h}{m}, \tag{2.1}$$

$$\frac{(I_H^* - I_L^*)}{I_L^*} = \frac{W_M}{W_L}\frac{m}{(1 - m - h)}. \tag{2.2}$$

where I_L^* and I_H^* are the task thresholds in 1995. In terms of Figure 1.1, these equations solve for the thresholds levels and the wage gaps corresponding to each threshold. Note that $C_{HM}(i)$ is a function that goes through the point where I_H meets with the relative wage W_H/W_M, and it defines the relative productivity across all tasks of high educated workers with respect to medium educated workers for all $i \in (1, 0)$. Similarly, $C_{ML}(i)$ defines the relative productivity of medium to low educated workers across all tasks and it pass through the point where I_L intersects the relative wage W_M/W_L. I assume a functional form for these functions. In particular, let

$$\alpha_L(i) = i;\ \alpha_M(i) = i^2;\ \alpha_H(i) = i^3.$$

[32]In Appendix Table B.4 I show that the general results of the model are robust to assuming other functional forms.

Then, the functions that determine the relative productivity across tasks are:

$$C_{HM}(i) = ln\,\frac{A_H\alpha_H(i)}{A_M\alpha_M(i)} = ln\,\frac{A_H}{A_M}i,$$

$$C_{ML}(i) = ln\,\frac{A_M\alpha_M(i)}{A_L\alpha_L(i)} = ln\,\frac{A_M}{A_L}i.$$

The above expressions combined with condition $W_H/W_M = C_{HM}(I_H)$ (equation (1.25)) and $W_M/W_L = C_{ML}(I_L)$ (equation (1.26)) solve for $\frac{A_H}{A_M}$ and $\frac{A_M}{A_L}$

$$\frac{W_H}{W_M} = \frac{A_H}{A_M}I_H^*, \Rightarrow A_H = \frac{W_H}{W_M}\frac{A_M}{I_H^*}$$

$$\frac{W_M}{W_L} = \frac{A_M}{A_L}I_L^*, \Rightarrow A_M = \frac{W_M}{W_L}\frac{A_L}{I_L^*}.$$

Finally, I use a simulated method of moments to target A_L so that equation (1.24) solves for the observed average wage for low educated workers in 1995. As result of this calibration, $A_L = 2.81$; $A_M = 12.39$ and $A_H = 47.36$, which implies that $C_{HM}(i) = ln\,3.82i$ and $C_{ML}(i) = ln\,4.42i$.

The calibrated model using these parameters matches the initial moments in the data relatively well. Table 2.5 contains the results of the calibration for the targeted and non-targeted moments. Panel A and B show that the supply of skills and average wages in 1995 are used to calibrate the model, so that its values perfectly match that of the model. The the overall distribution of occupations and the average task performed by each educational level are not targeted in the model and are shown in Panel C and Panel D, respectively. According to the model the average task of low educated workers is 0.21 in 1995, while in the data it is 0.24, while the average tasks in 1995 for medium and high educated are slightly overestimated in the model.[33] The model matches that most of the employment is concentrated on occupations of lower ranking (52.9% of total employment is located into

[33] I consider that the ranking of an occupation is informative of its complexity, such that a higher ranked

bottom-third occupations), while there is a much lower share of employment in high-rank occupations where high educated workers are employed (11.2% of total employment is into top-third occupations). The reason for the low share of employment into the top occupations in the model results from the optimal decisions of profit maximizers employers to use the most productive workers in a larger range of tasks, diminishing the employment share on those tasks.

Panel E displays the wage-inequality predicted by the model. Wage gaps are not presented since they are directly estimated from the average wages that are targeted, and perfectly matched by the model. The inequality of the wage distribution in the model is, as expected, lower than that of the data, given that in the model there is no inequality within each educational group. The Gini index in the model is 30 percent lower than that of the data, but the level of inequality in the model is higher than the between component of the Theil and the Atkinson(2) indexes. The model does a better job on fitting different indexes of wage poverty with a poverty line of $2.4. I argue in the next section that although the model cannot capture the total level of inequality at a given point in time because it only produces three different wages for the total workforce (one for each educational group), it is useful to understand changes in inequality that arise from variations in these wages and in the share of workers who earns them.

occupation is more complex than a lower ranked occupation. In a strict sense, there is a correspondence between the ranking of the occupation and the index of task complexity in the model, but one index can be a monotonic transformation of the other, preserving the order but not necessarily the distance among each other.

Table 2.5: Initial moments in the data (1995) and the model

	Brazil 1995	**Model**
Panel A: Initial thresholds		
Initial I_L		0.424
Initial I_H		0.664
Panel B: Wages (targeted)		
Low educated	1.47	1.47
Medium educated	2.76	2.76
High educated	7.00	7.00
Panel C: Occup. composition of employ. (not-targeted)		
Bottom-third	0.562	0.529
Middle-third	0.279	0.359
Top-third	0.159	0.112
Panel D: Av. occ. ranking (not-targeted)		
Low educated	0.243	0.212
Medium educated	0.454	0.544
High Educated	0.669	0.832
Panel E: Inequality (not-targeted)		
Gini	0.445	0.271
Theil		
Total	0.327	0.179
Between educ. groups	0.071	
Atkinson (Inequality aversion = 2)		
Total	0.572	0.205
Between educ. groups	0.140	
Panel F: Wage poverty- wage line $2.4 (not-targeted)		
FGT(0)	0.626	0.709
FGT(1)	0.305	0.275
FGT(2)	0.183	0.106

Notes: Estimations for Brazil comes from PNAD 1995. The model is calibrated for the year 1995 according to Section 2.5. 82 occupations at ISCO-88 3 digit level are ranked from 0 to 1 according to its median wage for the period 1995-2014. The average ranking is computed for each educational group. Bottom-third refers to the 27 occupations with lower average wage, medium-third are next 27, and the remaining 28 are classified as top-third occupations. Workers are classified into low (less than secondary), medium (some secondary), and high educated (some university).

2.6 Quantifying the Effects of the Educational Expansion in Brazil

2.6.1 Changes in the Occupational Structure and the Wage Distribution

This section compares observed changes in labor market outcomes in Brazil between 1995 and 2014 with the predictions of the model when the Brazilian educational expansion takes place, holding constant the rest of the parameters of the model. That is, I explore the effects of changes in m and h in the magnitudes observed in Brazil, holding constant the technology of the production function entirely characterized by the set $\{A_L, A_M, A_H, \alpha_L, \alpha_M, \alpha_H\}$.

I already established that a supply shock as the one that took place in Brazil matches all the qualitative patterns observed in the data. I turn now to the quantitative assessment of the effects of an educational expansion through the lens of the model. To this end, I estimate the changes in several outcomes of interest in the calibrated model generated by the educational expansion that took place in Brazil. The educational expansion consists of: a decrease in l from 0.68 to 0.36, an increase in m from 0.21 to 0.40, and an increase in h from 0.11 to 0.24. Table 2.6 presents the actual and simulated changes in the occupational structure of employment and the wage distribution. The table compares the changes observed in the data between 1995 and 2014 with those predicted by the model as results of the educational expansion. The objective of this exercise is to estimate how much of the changes in the data can be predicted by the model when *only* the educational level of the workforce increases as it did in Brazil while all other parameters (technology) are held constant. The last column of the table shows the difference between the model fit and the observed changes in the data.

Table 2.6: The labor market effects of an educational expansion: data vs model

	Data (1995-2014) (1)	**Model** (2)	**Difference** (2)-(1)
Panel A: Changes in thresholds			
I_L		-0.167	
I_H		-0.084	
Panel B: Changes in occup. composition of employment			
Bottom-third	-0.015	-0.072	-0.057
Middle-third	-0.005	-0.004	0.001
Top-third	0.019	0.076	0.057
Panel C: Changes in mean occup. ranking			
Low educated	-0.037	-0.083	-0.046
Medium educated	-0.129	-0.126	0.003
High Educated	-0.058	-0.042	0.016
Panel D: % Changes in wages			
Total workforce	27.8	29.3	1.5
Low educated	37.3	46.9	9.6
Medium educated	-10.5	-10.9	-0.4
High Educated	-21.6	-22.2	-0.6
Panel E: % Changes in wage gaps			
W_H/W_M	-12.4	-12.7	-0.3
W_M/W_L	-34.8	-39.3	-4.5
W_H/W_L	-42.9	-47.1	-4.2
Panel F: Changes in welfare and inequality			
First order dominance	Yes	No	
Gini	-0.085	-0.082	0.003
Theil	-0.109	-0.099	0.010
Between component	-0.020	-0.099	-0.011
Atkinson(2)	-0.132	-0.088	0.044
Between component	-0.078	-0.088	0.010
Panel G: Changes in wage Poverty (wage line 2.4)			
FGT(0)	-0.164	-0.331	-0.167
FGT(1)	-0.162	-0.237	-0.075
FGT(2)	-0.118	-0.103	0.015

Notes: Estimations for Brazil comes from PNAD 1995-2014. The estimations for the model comes from simulating the Brazilian educational expansion between 1995-2014 on the calibrated model, holding constant the rest of the parameters. The Brazilian educational expansion consists of: an increase in the share of medium educated from 20.5 to 40.0 percent, an increase in the share of high educated workers from 11.3 to 23.6 percent, and a decline in the share of low educated workers from 68.1 to 36.4 percent. The third column estimates the difference between the changes in the data and in the model.

I found that the model's predictions of the effect of an educational expansion are remarkably accurate. The model is not only able to predict the qualitative changes but it also makes a close quantitative prediction of the observed changes in most of the outcomes of interest.

Panel B displays the results regarding the overall composition of employment. The general prediction of the model is that the educational expansion generated an improvement in the occupational composition of the workforce, but that improvement was small when compared to the extent of the educational expansion, which match the patterns observed in the data. The model predicts larger changes than the ones observed in the data, but the same qualitative results (a decline in the bottom-third and an increase in the top-third occupations). When the educational expansion takes place, high and medium educated workers moved down to tasks of lower complexity, implying that not all the high and medium educated workers find a job in the same tasks workers with that educational level used to carry out, and the overall occupational composition of employment improves slightly, but its changes are small when compared to the educational expansion. The see this clearly, consider that the model predicts a decline of only 7.2 percentage points in bottom-third occupations, despite the fact that the share of low educated workers declined by 32 percentage points. At the same time, the model predicts an increase of 7.6 percentage points in the share of top-third occupations despite the increase in 13 percentage points in the share of high educated workers.[34]

Panel C shows that the changes in the average ranking of occupations in the model follow closely that of the data: there was a decline in the conditional occupational attainment for each educational group, and it was larger for medium educated workers. For medium educated workers, the average occupational ranking declines by 0.126 in the model compared

[34] I discuss the predictions of the model for high educated workers at length in Section 2.6.2.

to 0.128 in the data. For high educated workers, the average ranking declines by 0.042 in the model and by 0.057 in the data. The model predicts a larger decrease in the average task of low educated workers (0.083) than the one observed in the data (0.037). Part of the reason is that the model cannot explain the decline in the share of employment in agriculture that is due to some structural change taking place during these years, which decreased the share of low educated workers in agricultural-related occupations that are at the bottom of the ranking.[35]

Whit respect to real wages, Panel D shows that the model's predictions are very close to actual changes overall and for each educational group. Average wages rise 29.3 percent in the model and 27.8 percent the data; wages of low educated workers increase 46.9 percent in the model compared to 37.3 percent in the data; wages of medium educated workers decline 10.9 percent in the model and 10.5 percent in the data; and wages of high educated workers decrease 22.2 percent in the model and 21.6 percent in the data. The model predicts that wages of medium educated workers decrease less than wages of high educated workers, as observed in the data, despite the fact that medium educated workers experienced the largest decline in occupational attainment (both in the data and the model). A precise prediction of the changes in the wage gaps follows directly from the power of the model to predict changes in wages, as depicted in Panel E. The wage gap of medium to low falls 39.4 percent in the model and 34.8 percent in the data. The wage gap of high to medium declines 12.7 and 12.4 in the model and the data respectively. Finally, the wage gap of high to low declines 47.0 percent in the model and 42.9 percent in the data.

[35]Structural change may also be driven, in part, by an educational expansion. Consider a model with two sectors, agriculture and non-agriculture. In the agriculture sector, labor productivity declines with the number of workers in that sector. The non-agriculture sector produces a final good combining an infinite number of tasks, as in the model presented here. Low educated workers can be employed in the agriculture sector or in tasks of low complexity in the non-agriculture sector. An educational expansion such as the one that took place in Brazil may increase the wages of low educated workers in tasks of lower complexity in the non-agriculture sector, pulling workers out of agriculture.

Changes in inequality and poverty in the model are also close to those of the data. With respect to inequality, Figure 2.4 displays the Lorenz curves for Brazil in the data (Panel A) and in the model (Panel B). The Lorenz curve in 2014 is always above the one in 1995, which implies that any index of relative inequality decreased during this period. The same pattern is observed in the model. In terms of specific indexes of inequality portrayed in Panel F of Table 2.6, the Gini coefficient declined 0.085 points in the data and 0.082 in the model. The reduction in the Theil index is also similar, but the changes in the between education groups component of inequality (directly related to the changes in the model) is overestimated. This is not the case in the Atkinson index with an inequality aversion parameter equal to 2, which puts more weight to changes at the bottom of the wages distribution. In this case, the between groups inequality declined 0.078 in the data while in the model it decreased 0.088.[36] The fact that the fit of the model is better for the changes in the Atkinson (2) than in the Theil index is consistent with the model not being able to capture some of the movement at the top of the income distribution related to the dispersion of wages among high educated workers. Taking the average fit across the indexes of inequality, the model predicts 85 percent of the observed changes.[37]

The generalized increase in wages at the bottom of the wage distribution reduced poverty in the data and in the model, as depicted by Panel G. Wages for the percentiles that remained low educated increased and wages rose even more for those percentiles that switched from low to medium and high education. This represents the bottom 70 percent of the wage distribution in the model (the percentage that originally was low educated). However, changes in the poverty rate are overestimated in the model given its sensitivity to different poverty lines. The model better predicts changes in the poverty gap—FGT(1)—and the

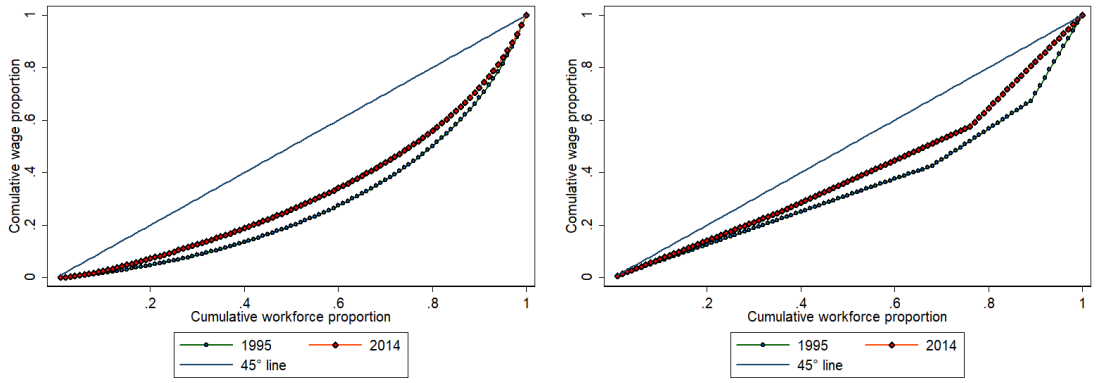
[36] In the model there is no within group inequality, therefore total inequality is equal to between-group inequality.

[37] If only the between components and the Gini index are taking into account, the fit of the model diminishes slightly to 68.4.

Figure 2.4: Changes in relative inequality: Lorenz Curves

(a) Brazil: 1995 and 2015

(b) Model: Before and after the educational expansion

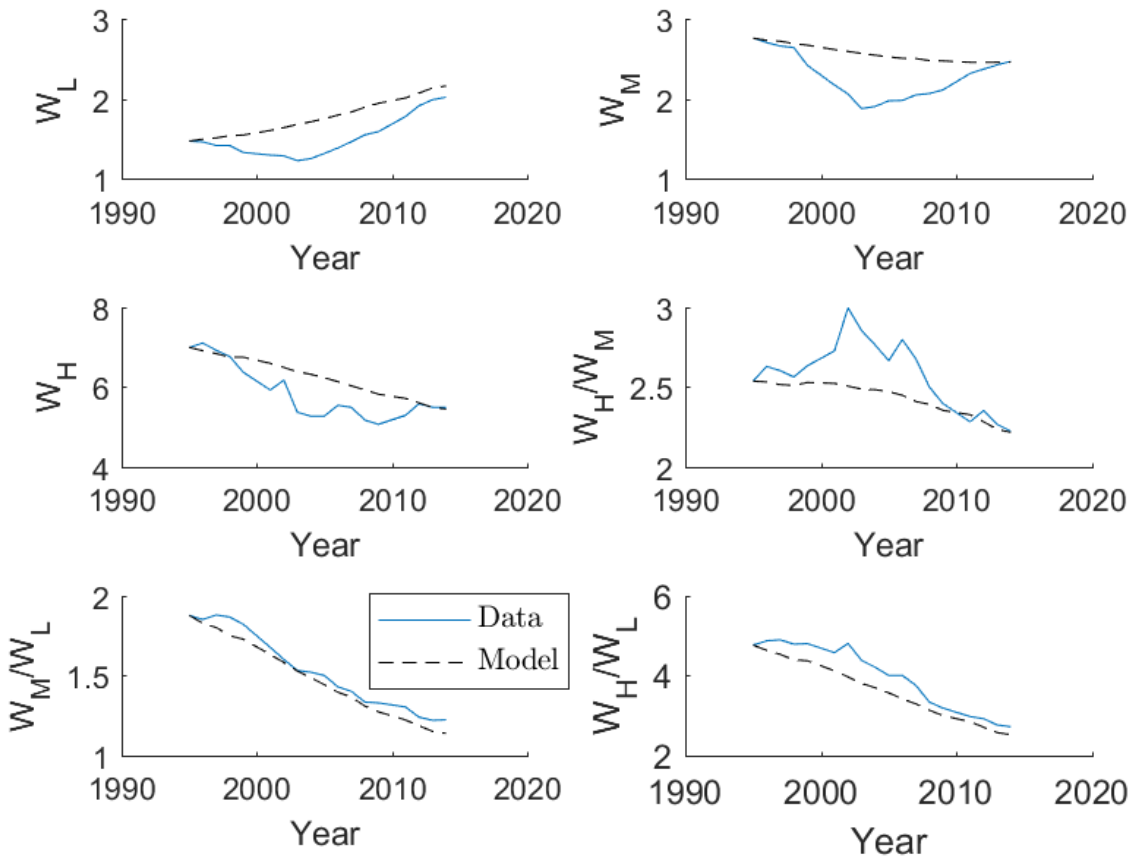


Notes: Panel (a) shows Lorenz curves for Brazil in 1995 and 2014. Panel (b) shows the Lorenz curves in the calibrated model before and after the educational expansion.

depth of poverty—FGT(2). On average, the model fits 68.4 percent of the changes in wage poverty in Brazil measure by FGT(0)-FGT(2).

Although the model accurately predicts most of the patterns in the occupational structure of employment and the wage distribution observed in Brazil between 1995 and 2014, it has some limitations to predict the year by year changes in a satisfactory way. Figure 2.5 shows the evolution of key variables and the predictions from the model. Brazil had a financial crisis in 1999 that lasted until 2003. The average wages during this period dropped overall and for the three educational groups. The model does not contain any structure to analyze business cycles, so it is expected to perform poorly during this period. The crisis mainly affected the levels of real wages, but it may be argued that the wage ratios were less affected if the crisis equally hurt workers with any educational level. This seems to be the case because even during this period of crisis the model closely predicts the evolution of wage gaps.

Figure 2.5: Data vs model predictions: wages year by year



Notes: The figure shows the evolution of wages and wages gaps labor market outcomes in the data and in the model year by year. The model simulates expansions in education as the one that took place year by year in Brazil, holding all the other parameters in the model constant at the levels of 1995.

In summary, the model developed in the previous chapter and calibrated in this chapter allows isolating the effects of an increase in the education level of the workforce on occupations and the wage distribution when all the other factors, such as technology and educational quality, are held constant. This exercise resulted in estimations that follow very closely the observed changes in the data, providing tentative evidence that the educational expansion in Brazil was the main factor behind the observed changes in the labor market between 1995 and 2014.

2.6.2 Where did all the increase in high education go?

It is striking that the share of high educated workers in Brazil increased by 13.0 percentage points during the period 1995-2014 while the occupational structure of employment remained relatively fixed, presenting only a small improvement. This section discusses in which occupations end up the high educated workers after the educational expansion and the strengths and limitations of the model to adequately characterize the observed patterns.

When the educational expansion took place in Brazil, there were three labor market effects in place to absorb the large increase in the share of high educated workers, according to Table 2.2. First, there was a small but still sizable increase in employment in the top-third occupations, which accounts for 1.9 percentage point of the increase in the share of high educated workers. Second, there was a decline in the occupational attainment of some high educated workers who started to be employed in occupations previously performed by workers with lower levels of education. This effect accounts for 6.6 percentage points.[38] Finally, some high educated workers replaced workers with lower educational level in jobs at the top-third occupations that already existed before the educational expansion. This effect accounts for the remaining 4.5 percentage points.[39]

The framework developed in the first two chapters of this dissertation is able to account for new jobs being created in the top-third occupations and for the lower occupational attainment of high educated workers. But the framework is silent about high educated workers replacing middle and low educated workers that were initially employed in top-third occupations before the educational expansion. The reason is that the model predicts thresholds

[38]This number comes from Table 2.2, which shows the increase in the total employment share of high educated workers in bottom-third and middle-third occupations.

[39]Table 2.2 shows that there was a decline if 3.4 and 1.1. percentage points of employment in top-third occupations for low and medium educated respectively.

that perfectly separates the occupations performed by workers with different educational level, while in the data that is not necessarily true and workers with different education overlap across occupations (see Table 2.2). The model abstracts from the fact that labor markets have frictions and that workers with a given educational level are heterogeneous. These factors are necessary to rationalize the overlap of workers with different education into the same occupation, and future research should be encouraged to incorporate these factors into this type of models. Because of this limitation, the model predicts a larger job creation in the top-third occupations than what is observed in the data. Nonetheless, the version of the model presented here accurately predicts two of the three effects that change the employment of high educated workers, and it rationalizes one important mechanism generated by the educational expansion: high educated workers take over the best jobs that workers with lower education used to perform, which is validated by the data.

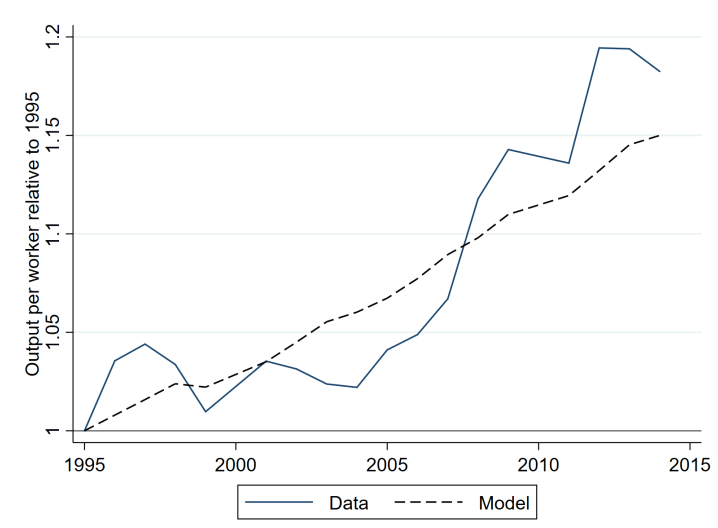
2.6.3 Output changes and the declining effects of an educational expansion

An expansion in education has distributional effects by changing the wages and shares of workers with different educational levels. It also affects the amount of output produced because more educated workers are more productive in the tasks that they start to perform after the educational expansion. In this section, I evaluate whether the model is informative with respect to the changes in output per worker observed in Brazil. Then, I evaluate different counterfactuals to study the differences between an increase in education in 1995 and in 2014.

The trend in output per worker is correctly captured by the model. Figure 2.6 shows the

evolution of output per worker in Brazil and in the calibrated model.[40] Output per worker in Brazil increased 18 percent between 1995 and 2014 and 15 percent in the model, accounting for 82 percent of the observed change.[41] This result suggests that the large investment in education in Brazil resulted in a considerably higher output.[42]

Figure 2.6: Evolution of output per worker in the data and in the model



Notes: The figure shows the evolution of output per worker in the data for Brazil. Output per worker is defined as GDP over employ workers. obtained from the World Development Indicators (2017). Output per worker in the model is defined as the total amount of the final good that is produced according to equation (1.1) in the calibrated model and taking into account the yearly changes in skill supplies from the data.

Despite the large effect of education on output during the period 1995-2014, I find that further increases in education are predicted to have a much lower impact. I estimate the effects of an increase in 1 percentage point in the share of medium or high educated workers in 1995 and 2014. The results are presented in Table 2.7. An increase of 1 percentage point

[40] Output per workers is measured as the coefficient of gross domestic product and the number of employed workers. The statistics are taken from World Development Indicators (WDI).

[41] Changes in average productivity in the model can be estimated from (1.1).

[42] My estimation may underestimate the real impact in economy-wide productivity if a more abundant supply of high educated workers might lead to skill-based technological change (Acemoglu, 2002), or increases in human capital might generate positive externalities in terms of productivity that are not taken into account here, such as reduction of crime and better democratic institutions in the long term (Moretti, 2004a).

in the share of medium educated in 1995 raises output by 0.5 percent. However, the same increase in 2014 increases output by only 0.1 percent, an impact five times lower than in 1995. In the case of an increase of 1 percentage point in the share of high educated workers, output increases by 2.3 percent in 1995 and only 1.1 percent in 2014. These are still sizable effects, but they account for less than half of the impacts in 1995.[43] In short, the return to a dollar invested in education declines as the labor force becomes more educated.

Table 2.7: Declining effect on output per worker of successive educational expansions

	1995	**2014**	**Difference**
Effect of 1 p.p increase in			
Share of medium educated (m)	0.54%	0.09%	**-0.45**
Share of high educated (h)	2.30%	1.06%	**-1.25**
Share of both (m&h)	2.83%	1.13%	**-1.69**

Notes: The table presents the effects on output per worker of an increase of 1 percentage point in m or h in the model under the different equilibrium conditions in 1995 (with a less educated labor force) and 2014 (with a more educated labor force). The results have to be interpreted as the predictions of the model of increasing the supply of skills by 1 percentage point while decreasing the share of low educated workers in the same magnitude.

The previous result implies that further increases in secondary schooling might not continue to foster increases in output and average wages at the same rate as before. This result must be interpreted with caution for several reasons. First, in order to increase h without reducing m, as implied in the last two columns of Table 2.7, the share of workers with at least secondary education has to increase as fast as h (otherwise there would be a decline in m). Second, it does not necessarily mean that the government should stop investing in

[43]The changes in real wages for each group and in relative wages are similar in 1995 than in 2014. But the share of people affected by these changes is very different. For example, an increase in m in 2014 will decrease wages of 2/3 of the workforce while it only decreased wages on 1/3 in 1995. The educational composition of the workforce in the baseline year matters when estimating the effect of an educational expansion on output, even when the effects in the wage gaps are very similar in different years (there is a movement along the curves of comparative advantage in Figure 1.3).

secondary schooling. Instead, it shows that investing *only* in secondary schooling will have a small impact on output, given the restrictions on the demand side by holding the technology constant. That the positive effect on output declines rapidly with successive educational expansions when the demand side remains fixed implies that policymakers have to focus on both supply and demand. Finally, the effects on wages or output estimated here do not have to be interpreted as social returns to education. It is well-documented that education generates several positive externalities that might be difficult to quantify, including (but not limited to): decrease in crime, positive effects in health, and better family planning.[44]

2.6.4 Decomposition of changes in wages

In Chapter 1 of this dissertation I showed that the changes in log wages can be decomposed into three effect: the displacement effect, the productivity effects, and the supply effect. I further mention that an educational expansion as the one that took place in Brazil generates all three. Next, I present the quantitative results of the decomposition for Brazil, using the calibrated model.

The result of this decomposition are shown in Table 2.8. The table displays the change in log wage that is predicted by the model and its decomposition in log points into the three different effects identified in (1.29). I find that the supply effects dominates, but the displacement and productivity effects are also quantitatively important. For low educated workers, most of the increase in wages that is generated by the supply effect is counterbalanced by the displacement of those workers towards a lower share of tasks (those of less complexity in the economy). Only one third of the increase in log wages is actually coming from the

[44]See Moretti (2004b) for a detailed discussion on human capital externalities.

difference between the supply effect and the displacement effect, while the remaining two thirds are generated by the increase in the productivity in the economy. In the case of medium and high educated workers, the supply effect of the educational expansion was very large and negative. However, most of it, although not all, was counterbalanced by positive displacement and productivity effects.

Table 2.8: Decomposition of changes in wages with the calibrated model

	Low	**Medium**	**High**
Predicted change in log wages	**0.38**	**-0.12**	**-0.25**
Decomposition:			
Displacement effect	-0.51	0.30	0.22
Productivity effect	0.26	0.26	0.26
Supply effect	0.63	-0.67	-0.73

Notes: The table shows decomposition of changes in wages according to equation (1.29) after simulating the Brazilian educational expansion between 1995-2014 on the calibrated model, holding constant the rests of the parameters.

2.6.5 Robustness Checks

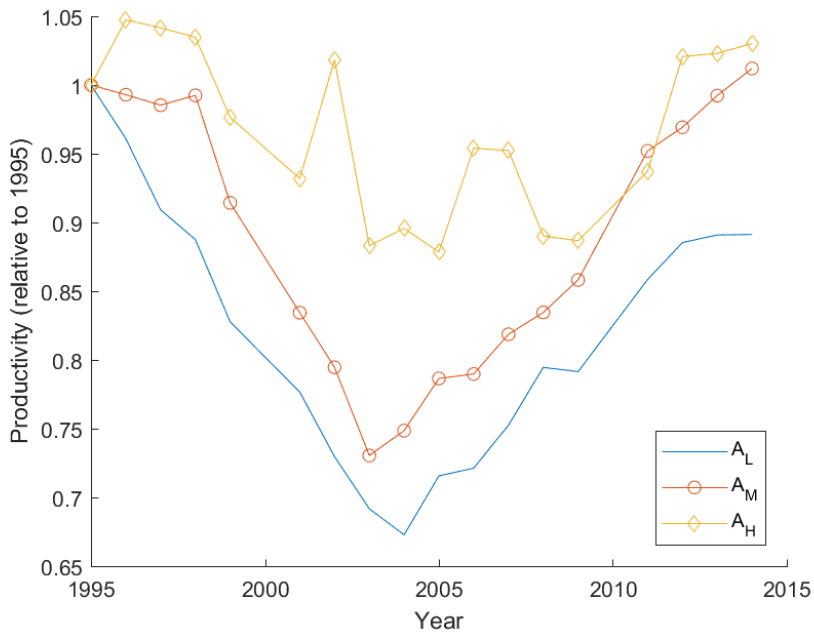
The comparative advantages across tasks are crucial for estimating the labor market effects of an educational expansion, as discussed in Section 1.3.2. Therefore, the results presented above are sensitive to the calibration procedure. In this section, I evaluate different alternatives to calibrate the parameters of interest. I show that the results are robust to using these alternatives. First, I show the results are robust to using other years to calibrate the model instead of 1995 data. Second, I discuss whether the linear assumption for $\alpha_L(i)$, $\alpha_M(i)$ and $\alpha_H(i)$ is reasonable according to the patterns observed in the data.

In principle, any year could be used to calibrate the parameters A_L, A_M and A_H as

in Section 2.5. Figure 2.7 shows the value of the productivity parameters with respect to 1995, when parameters are estimated for each year separately, using data on wages and skill supplies from each specific year. All productivity parameters decreased by around 25 percent between 1995 and 2004, and increased thereafter reaching similar levels in 2014 with respect to 1995. These patterns are consistent with the precise predictions of the model for wage levels in 2014—assuming that productivity parameters were the same in 1995 and 2014—and the lower fit for the years between 1995 and 2014. The model is more suitable to explain changes over an extensive period of time because it lacks the flexibility to accommodate for short-term effects related to the business cycle, e.g. financial crises. This is an important limitation of the model given that business cycles in Latin America (including Brazil) tend to be more pronounced than in other regions of the world (Mejia-Reyes, 1999). However, it is worth noticing that because all the productivity parameters tend to move in the same direction and extent, the ratios A_M/A_L and A_H/A_M do not change much during the analyzed period. These ratios determine the assignment of workers to tasks in the model, as well as the relative wages. This is the reason why the model does a better job at predicting the trends in relative wages than in real wages throughout the analyzed period.

To evaluate the sensibility of the results to different values of the parameters A_L, A_M and A_H, I compare the results of four different alternatives: 1) the base calibration using 1995 data; 2) the mean value of A_L, A_M and A_H from the series 1995-2014; 3) the minimum value of the series 1995-2014; and 4) the maximum value of the series 1995-2014. The results are shown in Table 2.9. From looking across the columns of the table, I find that the initial and final tasks thresholds I_L and I_H estimated in the model do not change much under these different calibrations. Note that the wages of 1995 are no longer targeted under estimations 2-4. More importantly, the predictions of proportional changes in real wages, relative wages, occupations within each educational group and overall distribution of occupations are robust

Figure 2.7: Calibration of productivity parameters year by year



Notes: The figure shows the values of the productivity parameters if the calibration exercise is performed separately using year by year data on the share of workers with different educational level and their average wages.

to the different calibration exercises. What matters for estimating these effects are the ratios A_M/A_L and A_H/A_M, which are similar across all alternatives.

The second assumption that I discuss here is whether the linearity for $\alpha_L(i)$, $\alpha_M(i)$ and $\alpha_H(i)$ is consistent with the data when the assumption of no change in productivity holds. Yearly data on wage gaps and skill supplies allows estimating year-by-year changes on thresholds levels according to Section 2.5. Let c_J be equal to $exp(C_J)$ for $J = HM, ML$. Then, it is possible to estimate $c_{HM}(I_H)$ and $c_{ML}(I_L)$ for the different thresholds and to check whether these points lie on a linear curve. We can then define:

$$c_{HM}(I_{Ht}) = \frac{A_H}{A_M}\frac{\alpha_H(I_{Ht})}{\alpha_M(I_{Ht})}, \quad c_{ML}(I_{Lt}) = \frac{A_M}{A_L}\frac{\alpha_M(I_{Lt})}{\alpha_L(I_{Lt})}, \; for\; t = \{1995, ..., 2014\}.$$

The identifying assumption in Section 2.5 was that $\frac{\alpha_H(I_{Ht})}{\alpha_M(I_{Ht})} = \frac{\alpha_M(I_{Lt})}{\alpha_L(I_{Lt})} = i$. If this assump-

tion is consistent with the data, the year-by-year points estimated by the model should lie close to the curve with no constant and with the slopes estimated in Section 2.5. Figure 2.8 shows the results of this exercise. In the case of the curve $c_{ML}(i)$, the points lie very close to the estimated curve. However, in the case of $c_{HM}(i)$ it seems that the slope is larger than the one predicted by the model, and it goes mostly below the points estimated year-by-year. There are two reasons why the curve $c_{HM}(i)$ may not fit very well the points in the data. First, the assumption of the function $\alpha_H(I_H)/\alpha_M(I_H) = i$ may be accurate, but A_H might have changed, which seems to be the case according to the previous exercise for the years in between 1995 and 2014 but not for the end points (in fact, the first and the last point of the series lie in the curve $c_{HM}(i)$ calibrated in 1995). Second, the function $\alpha_H(I_H)/\alpha_M(I_H)$ is misspecified and should be approximated by a linear function with a constant $(c + bi)$. I checked how sensible the results are to the linear function being misspecified. To that end, I considered each of the yearly estimated points as an observation in an OLS regression. The result is a linear function that minimizes the sum of the quadratic distance of the points to the linear function. This exercise provides an alternative estimation for $c_{HM}(i)$ and $c_{ML}(i)$ based on the data of the entire period. Figure 2.9 shows the predictions of the model in some key variables when this calibration is used. The results do not significantly differ from those in Figure 2.5.[45]

The main takeaway from this exercise is that the results are robust to using different functional forms for the productivity schedules across tasks that are also consistent with the data. A reason for this is that in Brazil the threshold I_H did not diminish much since both m and h increased during this period.[46] Therefore, moderate changes in the functional form

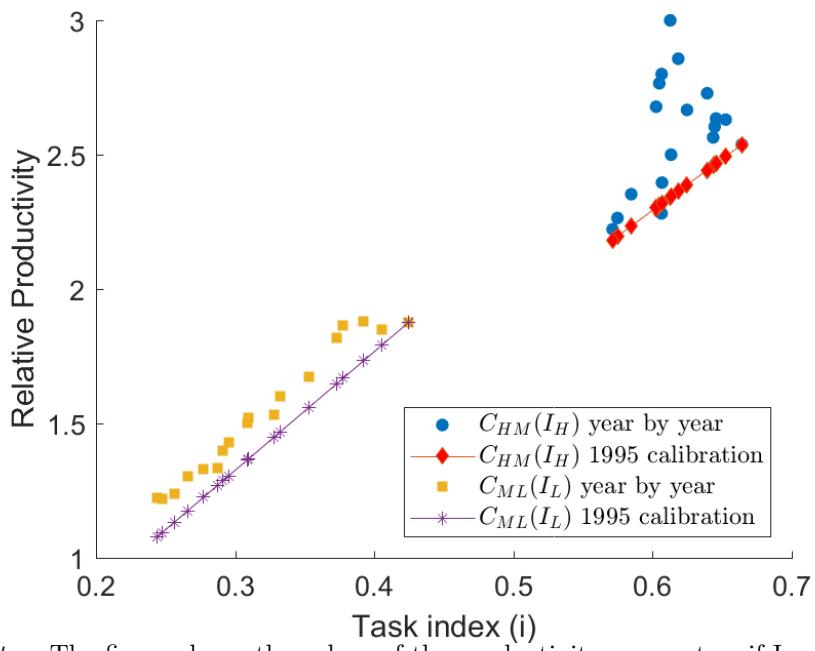
[45]The model also matches, overall, the changes observed in Brazil when using different functional forms to define the comparative advantages across occupations instead to the linear assumption in Section 2.5. The Appendix Table B.4 shows the main results from the model when other fucntional forms are used. The specific prediction varies following Table 1.3.2, that is, for more steep slopes in C_{HM} and C_{MH} the model predicts larger changes in wages and lower changes in the occupational structure of employment.

[46]I_H went from 0.66 to 0.58, compared to a decline twice as big in I_L from 0.42 to 0.26.

of the curve $C_{HM}(i)$ have a small impact on the results.

Although the specific value of the model's predictions changes when using alternative years to calibrate the model as well as different functional forms to fit the comparative advantages across tasks, the same general conclusion holds: the model matches very closely the changes in the occupational structure of employment and in the wage distribution observed in Brazil.

Figure 2.8: Comparison of 1995 calibration vs year by year calibration



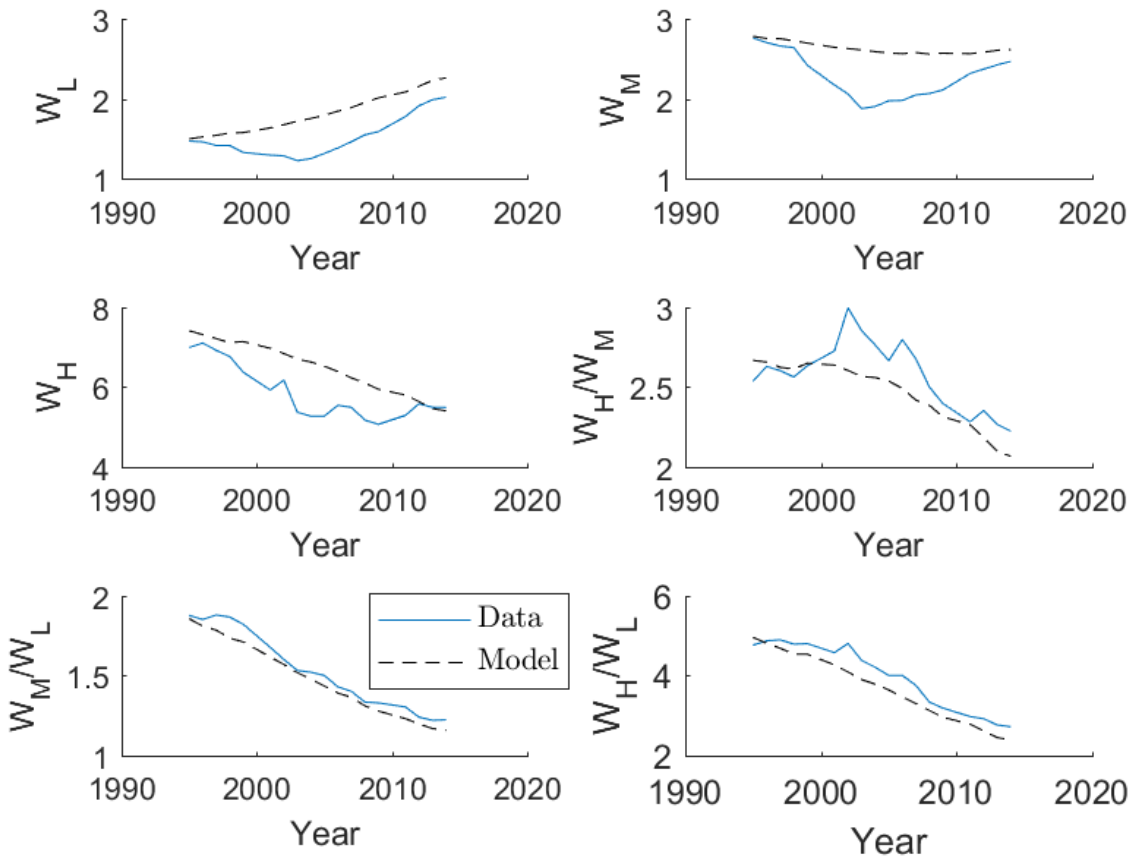
Notes: The figure shows the values of the productivity parameters if I calibrate the model using the average wages and educational levels of each year.

Table 2.9: Model results with different calibrations

	Brazil	Model calibration			
		1995	**mean** 1995-2014	**min** 1995-2014	**max** 1995-2014
Panel A: Thresholds					
Initial I_L		0.424	0.420	0.419	0.421
Initial I_H		0.664	0.661	0.652	0.669
Final I_L		0.257	0.254	0.254	0.255
Final I_H		0.580	0.578	0.567	0.586
Panel B: Average occup. ranking 1995					
Low educated	0.243	0.212	0.210	0.209	0.210
Medium educated	0.454	0.544	0.541	0.535	0.545
High Educated	0.669	0.832	0.831	0.826	0.834
Panel C: Changes in occup. composition of employment					
Bottom-third	-0.015	-0.072	-0.074	-0.072	-0.076
Middle-third	-0.005	-0.004	-0.003	-0.009	0.001
Top-third	0.019	0.076	0.077	0.081	0.075
Panel D: Changes in mean occup. ranking					
Low educated	-0.037	-0.083	-0.083	-0.082	-0.083
Medium educated	-0.129	-0.126	-0.125	-0.125	-0.125
High Educated	-0.058	-0.042	-0.042	-0.043	-0.041
Panel E: Average wage 1995					
Low educated	1.47	1.47	1.31	1.11	1.53
Medium educated	2.76	2.76	2.51	2.05	2.99
High Educated	7.00	7.00	6.37	5.54	7.25
Panel F: % Changes in wages					
Low educated	37.3	46.9	47.3	47.8	47.0
Medium educated	-10.5	-10.9	-10.8	-10.4	-11.0
High Educated	-21.6	-22.2	-22.1	-22.1	-22.0
Panel G: Changes in wage gaps					
W_H/W_M	-12.4	-12.7	-12.7	-13.1	-12.4
W_M/W_L	-34.8	-39.3	-39.5	-39.4	-39.5
W_H/W_L	-42.9	-47.1	-47.1	-47.3	-47.0

Notes: Estimations for Brazil comes from PNAD 1995-2014. The productivity parameters of the model are calibrated in different ways, according to the values in Figure 2.7. The model simulates an educational expansion between 1995-2014 following Table 2.1. Workers are classified into low (less than secondary), medium (some secondary), and high educated (some university).

Figure 2.9: Data vs model predictions with alternative calibration : wages year by year



Notes: The figure shows the evolution of wages and wages gaps labor market outcomes in the data and in the model year by year as in Figure 2.5. In this case, the calibration of the comparative advantage parameters is obtained by the OLS regressions over the points depicted in Figure 2.7 instead of using the data on 1995.

2.7 Discussion: of alternative explanations

In this section, I discuss other plausible explanations that have been mentioned in the literature that could generate some of the labor market changes in the four outcomes studied in this chapter.

2.7.1 Decline in the Quality of Education

When the access to education increases rapidly in a country, it is plausible that the quantity of human capital generated through schooling declines. Two factors may contribute to this decline. First, when the access to education expands the quality of each lesson may go down due to different factors, such as classroom congestions or lower teacher quality.[47] Second, the level of unobserved ability of the marginal student might diminished as education expands.[48]. If these factors are at play, the quantity of human capital after the educational expansion may be lower for a given educational attainment, and the decline in educational quality could potentially explain the decrease in wages and the lower occupational attainment for workers with secondary and university education.

In this section, I test this hypothesis by evaluating whether the observed changes in wages of medium and high educated workers are driven by market forces (supply and demand) or by a change in the human capital that each of these educational levels represent. I follow Bowlus and Robinson (2012) and Heckman *et al.* (1998) who compute changes in the prices of human capital by looking into yearly changes in median wages of cohorts for which the quantity of human capital should have remained constant. By comparing the changes in prices with the observe changes in wages, I conclude that most of the changes in wages were generated by changes in the price of human capital. This results support the hypothesis that the decline in wages of medium and high educated workers were mostly driven by the market forces described in the previous sections instead of a decline in educational quality.

[47]For example, Chaudhury *et al.* (2006) report that school enrollment rates have largely increased in developing counties since 1960, especially in primary schooling, but many children either leave school at a very young age or learn little while in school. This is unlikely to be the case in Brazil, since the educational expansion was accompanied by a large increase in school expenditures and teacher's educational requirements.

[48]For this argument to be valid it is necessary to assume that the distributions of abilities remains unchanged across cohorts. It may not be the case if health outcomes of young children are improving, or if households conditions are changing such that parents are more educated, poverty is less prevalent, etc.

In the theory and the model presented above, an educational level is equivalent to a quantity of human capital that the workers possess. In another strand of the literature related to human capital models, wages are defined as the product of efficiency units of human capital that each worker possesses and its price. Let w_{it} be the wage of worker i in time i, λ_t be the market level price for a unit of human capital, and E_{it} be the quantity of human capital. It is possible to define

$$w_{t,i} = \lambda_t E_{t,i}.$$

Changes in the log of median wages between t and t' for an educational level s from cohort c can be defined as:

$$ln\, Median\, w_{t',s,c} - ln\, Median\, w_{t,s} = ln\, \lambda_{t',s,c} - ln\, \lambda_{t,s,c} + ln\, Median\, E_{t',s,c} - ln\, Median\, E_{t,s,c}.$$

This equation shows that change in the log of median wages can be decomposed into a market price effect given by the first term and a human capital or quantity effect reflected in the second term. To estimate the incidence of each of these two terms in changes in wages, Heckman *et al.* (1998) looks into changes in wages for cohorts whose level of human capital does not change over a short period of time. In other words, they look into the age groups for which the quantity of human capital is assumed to remain constant so that the second term in the previous equation is equal to zero. These age groups are said to be in a "flat point" with respect to their life-cycle earnings. Let define the cohorts of educational level s for which human capital remain constant from t to t'. The changes in prices for an educational group s can be defined as an average of the changes in prices for each cohort in the flat point like follows

$$ln\, \lambda_{s,t'} - ln\, \lambda_{s,t} = \frac{1}{n} \sum_c ln\, \lambda_{t',s,c} - ln\, \lambda_{t,s,c} = \frac{1}{n} \sum_c ln\, Median\, w_{t',s,c} - ln\, Median\, w_{t,s,c}, \quad (2.3)$$

where n is the number of cohorts consider to be in the flat spot.

To find the ages groups, I follow Bowlus and Robinson (2012) defining the *flat point* according to potential years of experience for each educational group by looking into the cross-sectional life earning cycle of wages. I find that the life earnings cycle is flat in most years when looking into low educated between 43-52 years old, between 48-57 years for medium educated workers, and 50-61 for high educated workers.[49]. Table 2.10 replicates Table 1 in Bowlus and Robinson (2012). A *flat point* is found if earnings do not change with age. I found that to be the case for most of the years, except for the first four years for low educated workers. As indicated by Bowlus and Robinson (2012), this could generate a downward bias in the price series of low educated workers, since the decline in prices for those years might be partly due to decline in the quantity of human capital. In this case, the price series I estimate here can be interpreted as a lower bound. Once the price series is obtained, the quantity effect is estimated as the residual of wage effects minus the price effect.

Table 2.11 summarizes the results of this decomposition exercise for the period 1995-2014. I found that the price effects explain most of the changes in wages between 1995 and 2014. Most of the increase in wages for low educated workers and the decrease for high educated workers (97% and 84% respectively) is explained by the price effect as opposed to a quantity effect. For medium educated workers, price effects are larger than the changes in wages, suggesting that the quantity of education embedded in medium educated workers is actually higher in 2014 than in 1995. This evidence is also consistent with results from international standardized tests in secondary education which find an increase in test scores during the

[49] I perform this analysis only for men, working more than 35 hours a week. The reason to restrict the sample to men is that at this age range participation in the labor force decline rapidly for women, and price effect may be biased due to selection effects into the labor force

Table 2.10: Flat-point on earnings: Cross sectional evidence

	Low educated 43-52 years old	Mediun Educated 48-57 years old	High educated 51-60 years old
1995	-0.00965**	0.00746	0.0300*
1996	-0.0115***	-0.0133	0.00505
1997	-0.0125***	-0.00230	-0.0178
1998	-0.00683*	-0.00235	0.000176
1999	0.00116	-0.00728	-0.0307*
2002	-0.00200	0.00943	0.00564
2003	0.00261	0.00323	-0.00511
2004	0.00508	0.00702	-0.00456
2005	0.00389	0.0118	-0.00290
2006	0.00303	0.00248	-0.00344
2007	0.00242	0.0113*	0.0104
2008	0.00211	0.00570	0.00191
2009	0.00389	0.00354	0.00122
2011	-0.00125	-0.0000646	0.00331
2012	0.000809	0.00264	0.0131
2013	0.00269	0.00763	-0.00329
2014	-0.00121	0.00523	0.0166*

Note: Each cell in the table shows the returns to an extra year of experience in a mincer equation of low wages. The sample is restricted only to employed males. Workers are classified into low (less than secondary), medium (some secondary), and high educated (some university). * Significant at 10 percent. ** Significant at 5 percent. *** Significant at 1 percent.

2000s.[50].

The evidence presented in this section is consistent with the theoretical model presented above where the changes are originated by market level effects as a response to changes in the supply of skills instead of changes in educational quality. It has been a remarkable achievement for Brazil to have one of the largest educational expansions in history without decreasing the quality of is education. However, this quality is still relatively low to inter-

[50] Bruns *et al.* (2011) contain a detailed analysis of the evolution of test scores in Brazil when compared to other countries as a measure of the evolution of quality of education. The authors conclude that, although the scores are still far below OCDE levels, Brazil has made a fast and sustained progress during the 2000s.

national standards and one of the main challenges that the educational system faces today is to increase educational quality.

Table 2.11: Decomposition of the changes in wages on price and quantity effects

	Low	Medium	High
Median log wages 1995	0.504	1.197	2.209
Median log wages 2014	0.776	1.038	1.820
Change in wages 1995-2014	0.272	-0.160	-0.389
Prices effect	0.264	-0.396	-0.326
Quantity effect	0.008	0.237	-0.063

Notes: Median wages are estimated for males working more than 30 hours per week. Price effect is estimated from adding the year-by-year price effects from equation (2.3). Quantity effect is the residual of the total change in median wages minus the price effect. Workers are classified into low (less than secondary), medium (some secondary), and high educated (some university).

2.7.2 Other factors

Of the four labor market outcomes analyzed in this paper, the one that has received more attention in the literature, by far, is the reduction in wage inequality. Among other reasons, it has been documented that wage inequality declined because of increases in the minimum wage that spread throughout the wage distribution (Engbom and Moser, 2017), a fall in returns to experience which compressed the wage distribution (Ferreira *et al.*, 2016), a decline in the agricultural-non agricultural wage gap (Alvarez, 2017), and to shocks in commodity prices that benefited low skill workers in Brazil (Adão, 2015).[51]

Although each of this factors might have contributed to the decline in the wage gaps

[51]There are others non-competitive labor market factors that have been consider to study the recent increase in inequality in the developed world, such as the rise of superstar firms (Dorn *et al.*, 2017) and the interaction between employer power and labor saving technical changes (Chau and Kanbur, 2018). These factors seem to be less important for the case of Brazil where inequality largely fell.

among workers with different educational levels, my results suggest that the educational expansion was the single most important factor. However, my results can be related to the other channels that have been identified in the literature. For example, it could be possible that the increase in the minimum wage was effective in raising wages and did not decrease formal employment because the labor market for low educated workers became much tighter as a result of the supply and demand mechanisms generated by the educational expansion. Moreover, if returns to experience are larger in more complex occupations (occupations of high rank according to my model), returns to experience may have declined for each educational group because of the lower occupational attainment arising from the educational expansion. Additional research that examines the relationship between these various sources of changes in the wage distribution should be encouraged.

2.8 Conclusions

In this chapter, I first documented the major educational expansion of the Brazilian workforce between 1995 and 2014 and the remarkable changes in labor market outcomes that took place in the same period. I found that the occupational structure of employment presents a small improvement, but remained surprisingly fixed when compared to the educational expansion; workers of all educational groups—primary or less, secondary, and university—experienced a decline in occupational attainment; and wages largely increase overall but not for all groups since wages of primary educated workers increased while wages of more educated workers declined, bringing forth large reductions in wage gaps and other inequality measures as well as a stark reduction in wage-poverty.

Then, I showed that the model from Chapter 1 reproduces, qualitatively, all the patterns

in occupations and wages observed in the data with an educational expansion as the one that took place in Brazil (*type-h&m*), suggesting that the educational expansion may have been an important contributing factor to the observed labor market changes. Finally, I calibrated the model to quantitatively assess how much of the observed changes in the data can be explained by the educational expansion. I find that the changes predicted by the model are remarkably accurate. I also show that these predictions are robust to different calibration strategies and that are not driven by a decline in educational quality.

I conclude that the increase in educational attainment of the workforce was of utmost importance to the notable changes in the Brazilian labor market in the last two decades. My results reassert educational expansions as a key driver of economic development. In the case of Brazil, it largely improved the income distribution by reducing wage poverty and wage inequality, and it increased mean labor earnings. But I also find that the return to a dollar invested in education declines as the labor force becomes more educated. For example, I estimated that the effect of an increase in one percentage point in medium education in 2014 generates around one-fifth of the effect in average wages with respect to its effects in 1995. This indicates the necessity of educational expansions to be accompanied by policies directed to increase the job opportunities in occupations where more educated workers can exploit the productivity differentials that they have obtained through education. Otherwise, more educated workers will end up in occupations where schooling adds little value to workers' productivity, undermining the positive effect of an educational expansion on economic development.

Given the extent and speed of educational expansions around the world, it is relevant to improve our understanding of its effect on the entire labor market, not just for those workers receiving additional education. The theoretical model introduced in Chapter 1

together with the empirical application presented in this chapter provide a new framework to study the labor markets effects of different educational expansions. This framework contains novel theoretical predictions, with a special focus on the relationship between changes in the occupational structure of employment and the wage distribution. All these predictions can be tested empirically using data on wages and occupations already available in most countries. This Chapter showed an application of the framework to Brazil and it can also be used to study the labor market effects of educational expansions in many other countries as well.

CHAPTER 3

THE LONG-RUN EFFECTS OF TEACHER STRIKES. EVIDENCE FROM ARGENTINA

(This chapter was written in collaboration with Alexander Willén, Cornell University)

3.1 Introduction

Teacher industrial action is a prevalent feature of public education systems across the globe; during the past few years teacher strikes have been observed in countries such as Argentina, Canada, Chile, China, France, Germany, India, Israel, Lebanon, Mexico, Russia and the United States (e.g. Charleston, Seattle, East St. Louis, Pasco, Prospect Heights and Chicago). A shared belief among policymakers across several of these countries is that teacher strikes disrupt learning and negatively affect student educational attainment. In some countries this sentiment has led to the enactment of legislation that severely restricts teachers' right to strike.[1] However, despite the prevalence and debates surrounding teacher strikes, there is very little empirical research that credibly examines how they impact student long-term outcomes.

In this paper, we construct a new data set on teacher strikes in Argentina and use this to

present the first evidence in the literature on the effects of teacher strikes on students' long-run outcomes. Between 1983 and 2014 Argentina experienced a total of 1,500 teacher such strikes, with substantial variation across time and provinces with substantial variation

[1]For example, even though 33 states in the US have passed duty-to-bargain laws that require districts to negotiate with a union, only 13 states allow teachers to go on strike in the event of a bargaining impasse (Colasanti, 2008).

across time and provinces, making this an interesting case for the study of teacher strikes. We analyze the relationship between exposure to strikes during primary school and relevant education, labor market and other socioeconomic outcomes when the affected cohorts are between 30 and 40 years old. We also investigate if the effects that we estimate carry over to these individuals' children.

To identify the effect of teacher strikes, we rely on a dose-response difference-in-difference method that examines how education and labor market outcomes changed among 30 to 40 year olds who were exposed to more days of teacher strikes during primary school compared to 30 to 40 year olds who were exposed to fewer days of teacher strikes during primary school.[2] The sources of variation we exploit come from within-province differences in strike exposure across birth cohorts and within-cohort differences in strike exposure across provinces. On average, provinces lost 372 instructional days due to strikes during this period (6.7 percent of total instructional days), ranging from 188 days in La Pampa to 531 days in Rio Negro.[3]

The main assumptions underlying our estimation strategy are that there are no shocks (or other policies) contemporaneous with teacher strikes that differentially affect the various cohorts and that the timing of teacher strikes is uncorrelated with prior trends in outcomes across birth cohorts within each province. We show extensive evidence that our data are consistent with these assumptions.

We find robust evidence that school disruptions caused by teacher strikes worsen future labor market outcomes: being exposed to the average incidence of teacher strikes during primary school (88 days) reduces labor market wages for males by 2.82 percent and monthly labor market earnings for females by 4.22 percent. The prevalence of teacher strikes in

[2]We focus on this age range because existing literature suggests that labor market outcomes at this age are informative about lifetime outcomes (e.g. Haider and Solon (2006))

[3]There are 180 instructional days per year in Argentina.

Argentina means that the effect on the economy as a whole is substantial: A back-of-the-envelope calculation suggests an aggregate annual earnings loss of $3265 million. This is equivalent to the cost of raising the average employment income of all primary school teachers in Argentina by 87.1 percent.

In addition to adverse wage and earnings effects, our results reveal negative effects on several other education and labor market outcomes. Specifically, we find that the average incidence of teacher strikes leads to a 14.7 percent and 10.4 percent increase in unemployment for males and females respectively. In the case of females, we also find a 7.48 percent increase in the probability of not working or studying relative to the respective mean. For males, we find evidence that teacher strikes causes individuals to sort into lower-skilled occupations. Examining short- and long-run educational outcomes show that the labor market effects are driven, at least in part, by a reduction in educational attainment. Using data on students who have just finished primary school we demonstrate that many of these adverse effects are visible immediately after individuals have finished primary school. Finally, we document significant intergenerational treatment effects: children of individuals exposed to teacher strikes during primary school suffer negative education effects as well.

Our paper contributes to the existing literature in several important ways. First, no other paper has examined the effects of school disruptions on student long-run outcomes. Given the large literature demonstrating that short-run program effects on student outcomes can be very different from any long-run effects (e.g. Chetty *et al.* (2011); Deming *et al.* (2016); Lovenheim and Willén (2016)), this is of great value to policy makers. Second, the frequency and prevalence of teacher strikes that we exploit is much greater than that which has been used in earlier studies. This allows us to obtain more precise estimates, and examine a richer set of outcomes. Third, this paper makes use of a novel data set which we have created

based on information from annual business reports on the Argentine economy. This data is a great tool for other researchers interested in questions centering on teacher strikes and industrial action.

It is important to highlight that the pervasive level of teacher strikes during our analysis period is not a deviation from the norm in Argentina, and current student cohorts are exposed to similar levels of strikes. This cements the relevance of our paper and highlights the urgency of implementing reforms that can reduce the prevalence of teacher strikes in the country. One policy could be to introduce labor contracts that extend over several years, and only allow teachers to strike if a bargaining impasse is reached when renewing these multi-year contracts. This would eliminate sporadic teacher strikes while still allowing teachers to use industrial action as a tool to ensure fair contracts.

The rest of this paper proceeds as follows: Section 2 provides an overview of the education system in Argentina and offers theoretical predictions of how teacher strikes may affect student outcomes; Section 3 discusses pre-existing research; Section 4 introduces the data; Section 5 presents our empirical strategy; Section 6 discusses our results; and Section 7 concludes.

3.2 Background & Theoretical Predictions of Teacher Strikes

3.2.1 The Argentinian Education System

Education in Argentina is the responsibility of the provinces and is divided into four levels: kindergarten, primary education, secondary education and tertiary education.[4] Primary education begins the calendar year in which the number of days the child is 6 years old is maximized, and comprises the first seven years of schooling. During our analysis period, only primary education was mandatory in Argentina (Alzúa *et al.*, 2015).[5] Since then, compulsory schooling has grown to include secondary education as well, increasing the length of mandatory education from 7 to 12 years. Public education is financed through a revenue-sharing system between the provinces and the federal government, and is free at all levels.

The fraction of students that attended private school at the primary level during our analysis period was approximately 0.2, and this fraction was held relatively constant across the years that we examine. Since 2003, however, private enrollment at the primary level has increased substantially. Existing research suggests that this increase is driven by high- and middle-income families migrating from public to private schools, leading to an increase in socioeconomic school segregation (Gasparini *et al.*, 2011b; Jaume, 2013).[6]

[4]Primary education was decentralized in 1978 and secondary education was decentralized in 1992. However, the national government remains highly involved in terms of setting curriculum, regulations and financing.

[5]The youngest cohort in our main analysis sample finished primary school in the year prior to the implementation of the "Federal Education Law" (1998; approved in 1993) which extended mandatory education to encompass secondary schooling as well.

[6]A commonly held belief is that individuals perceive private education as superior due to the fact that teacher strikes are much less pronounced at private institutions, but existing literature finds no effect of teacher strikes on the likelihood of being enrolled at a public institution (Narodowski and Moschetti, 2015). We examine this in detail in Section 6.4.

3.2.2 Teacher Strikes in Argentina

The presence of unions, collective bargaining and labor strikes in Argentina can be traced back to the early years of the 20th century, except for the years during which the country was subject to military dictatorships (Confederacion de Educadores Argentinos, 2009).[7] Following the most recent reinstatement of democracy in 1983, industrial action has quickly regained its status as a pervasive feature of the Argentine labor market. Since then, public sector teachers have been the most active protesters in the country, making up 35 percent of all strikes in Argentina (Etchemendy, 2013). In comparison, private school teachers account for less than 4 percent of total strikes in the country. The occupation with the second largest incidence of strikes in modern times is public administration, accounting for 25 percent of all strikes (Chiappe, 2011; Etchemendy, 2013).

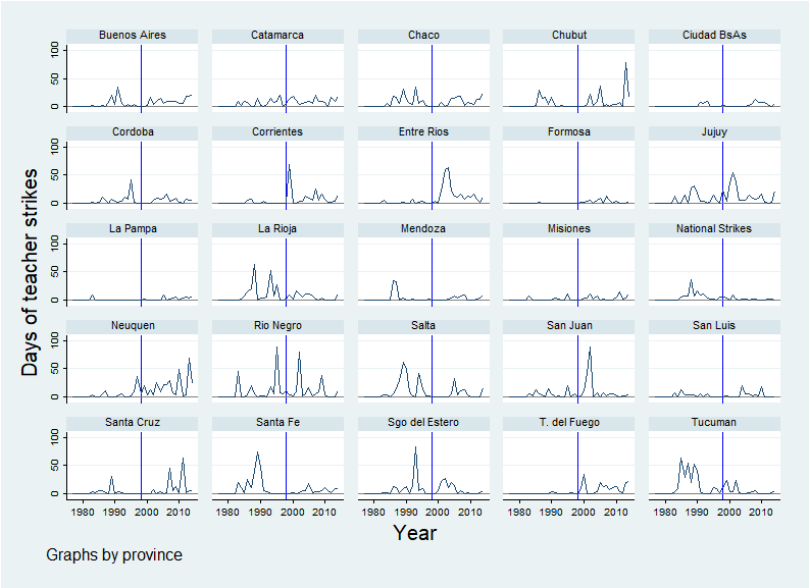
Teacher unions are typically organized at the provincial level, and variation in teacher strikes across time and provinces is substantial. On average, provinces have lost 372 instructional days due to teacher strikes between 1983 and 2014 (6.7 percent of total instructional days), ranging from 188 days in La Pampa to 531 days in Rio Negro, with a standard deviation of 109 days.[8] The pervasive level of strikes during our analysis period is not a deviation from the norm in Argentina, and current students are exposed to similar levels of strikes. To illustrate this point, Panel A of Figure 3.1 shows the variation in the number of days of school disruptions caused by teacher strikes by province from 1977 to 2014, and Panel B of Figure 3.1 displays the number of strikes by province during the same period (a strike can last for a couple of hours or for several weeks).

Although no study has examined the effect of teacher strikes on student outcomes in

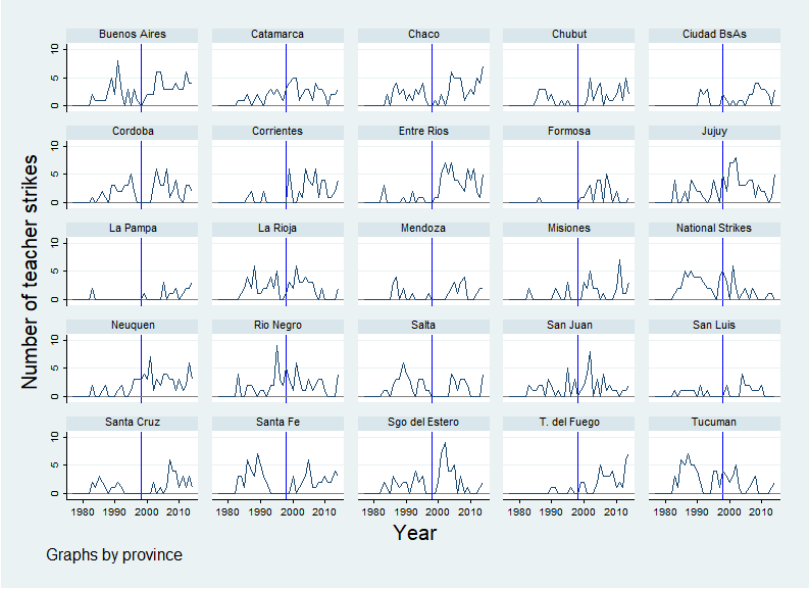
[7] During the dictatorships, labor strikes were prohibited and collective bargaining limited.
[8] There are 180 instructional days per year in Argentina.

Figure 3.1: Variation in Teacher Strikes 1977-2014

(a) Days of teacher strikes



(b) Number of teacher strikes



Notes: Authors' tabulations from annual reports on the Argentine economy published by Consejo Tecnico de Inversiones (1977-2014). Panel A shows the evolution of days teacher strikes for each province and at a national level. Panel B displays the number of teacher strikes. The vertical line indicates the two sub-samples used for the estimation of long-run (left) and short-run (right) outcomes.

Argentina, several studies have tried to disentangle the factors underlying the prevalence of teacher strikes in the country. The results are mixed:Murillo and Ronconi (2004) finds that teacher strikes are more common in provinces where union density is high and political relations with the local government is tense, while Narodowski and Moschetti (2015) concludes that strikes display an erratic behavior without any discernible trends or explanations. What these studies have in common is that they both emphasize the lack of a relationship between local labor market conditions and teacher strikes. This is important for our analysis since our main identifying assumption is that there are no shocks contemporaneous with teacher strikes that differentially affect the different cohorts (see Section 6.3).

3.2.3 Theoretical Predictions

This paper exploits variation in teacher strikes within and across provinces in Argentina to identify the reduced form effect of teacher strikes on students. As such, it is not a full analysis of the benefits and costs associated with teacher strikes, but rather a partial equilibrium analysis that uses teacher strikes to measure the effect of school disruptions on students' long-term outcomes.[9] Nevertheless, the success of education policies ultimately depends on how they impact the long-run outcomes of students, and identifying the net effect of teacher strikes on student outcomes is therefore of great independent importance. In this section, we list and discuss a large subset of relevant factors that may be influenced by teacher strikes, specifying which ones our empirical strategy is able to pick up.[10]

[9]Specifically, we are not measuring the benefits or costs in a dynamic general equilibrium setting. For example, the ability of teachers to strike may give them leverage in negotiations, leading to changes in working conditions that attract a different quality of teachers or affecting investments in schooling.

[10]Many of the predictions of the effects of teacher strikes are related to the underlying reasons for teachers to strike. It is therefore difficult to determine the generalizability of our results to other countries and settings, as teachers in, for example, the US may strike for other reasons than those that lead teachers in Argentina to strike. In a companion paper, we build a political economy model that aspires to identify

The main way in which strikes can affect student outcomes is by reducing the time students spend in school. Theoretical as well as empirical research provide clear predictions that reduced instructional time lowers academic achievement (Cahan and Davis, 1987; Cahan and Cohen, 1989; Neal and Johnson, 1996; Lee and Barro, 2001; Gormley and Gayer, 2005; Cascio and Lewis, 2006; Luyten, 2006; Pischke, 2007; Marcotte, 2007; Sims, 2008; Marcotte and Hemelt, 2008; Leuven *et al.*, 2010; Fitzpatrick *et al.*, 2011; Hansen, 2011; Rivkin and Schiman, 2015; Goodman, 2014).

In addition to reducing effective instructional time, teacher strikes can (1) affect teacher effort, (2) alter resource levels and allocation, (3) affect academic expectations and graduation requirements, (4) alter the value of a diploma, (5) change the value differential between a public and a private degree, and (6) change the composition of teachers. The direction and magnitude of the effects flowing through these channels will depend on the nature and outcome of the strike. For example, if the unions go on a short strike to raise wages and are successful, the strike will likely lead to an increase in teacher effort and productivity. This could also lead to an improvement in the composition of the teacher workforce in the long-run. However, if the strike is in effect for several months before the two sides reach an agreement, academic expectations and graduation requirements may be adjusted downwards with the potential implications of a reduction in the value of a diploma and an increase in the value differential between a public and a private degree. Further, the increase in teacher pay may be financed through a reallocation of resources from other inputs that enter the education production function, and this can lead to a reduction in educational quality.

The discussion above makes clear that the effect of teacher strikes on education production can be both positive and negative, and the resulting predictions of the effects of teacher

the most common drivers of teacher strikes in Argentina. The outcome of that paper should be used to determine the generalizability of the results in the current paper to other countries and settings.

strikes on student outcomes are therefore ambiguous. With respect to the current study, it is important to note that we use teacher strikes to measure the effect of school disruptions on student long-term outcomes through a partial equilibrium analysis. To the extent that the above factors impact current students, they will contribute to the effects that we estimate. However, our estimation strategy does not permit us to pick up any general equilibrium effects that strikes may have on the future education system of Argentina.

Two factors augment the theoretical ambiguity associated with the effect of teacher strikes on student outcomes. First, there may be treatment heterogeneity across students. The most likely source of heterogeneity relates to the socioeconomic characteristics of the students' families: wealthy parents will be able to move their children to private institutions if they believe the strikes to hurt their children. If this behavior is sufficiently pervasive it may lead to a segregated school system with additional adverse effects on the students from poor families that are left behind.[11] Another source of treatment heterogeneity relates to when during primary school children are exposed to strikes. Ample research suggests that younger children are more susceptible to policy interventions in general, and children who lose several weeks of instructional time in first grade may therefore suffer more than children who lose the same amount of days in the final grade of primary school (Meisels and Shonkoff, 2000; Cunha and Heckman, 2007; Chetty *et al.*, 2016). We explore both of these potential heterogeneity effects in Section 6.4.

Second, teacher strikes may have important effects on non-educational outcomes. The reason is that teacher strikes reduce effective instructional time. Unless parents can make alternative educational arrangements (which will depend on whether it was an expected or unexpected strike, and on the resources that the parents possess), this will lead to an

[11]This effect may be further augmented if teachers from poorer districts are more likely to join teacher unions and participate in strikes.

increase in leisure time and to an increase in the risk of engaging in bad behavior and criminal activity (Anderson, 2014; Henry *et al.*, 1999). This can directly impact the future education and labor market outcomes of children. Though we are unable to look directly at the relationship between teacher strikes and engagement in bad behavior and criminal activity, to the extent that this occurs and affects the long-run labor market and education outcomes of students, it will be a part of the effects captured by our estimation strategy.

3.3 Prior Literature on Teacher Strikes

The majority of the existing research on teacher strikes is cross sectional with identification strategies that are vulnerable to omitted variable bias (Caldwell and Jeffreys, 1983; Zirkel, 1992; Thornicroft, 1994; Zwerling, 2008; Card *et al.*, 2010; Johnson, 2011). Specifically, students, teachers and schools subject to strikes may be systematically different from those that are not on dimensions that we cannot observe. If these differences have independent effects on the outcomes that are being examined, this will bias the results. Further, these studies have focused on contemporaneous effects (test scores) of teacher strikes that are of very short duration. These two factors significantly limit our understanding of the consequences associated with school disruptions caused by teacher strikes.

Abstracting away from potential identification issues, the results from the above studies are mixed. While some studies find no association between strikes and student outcomes (e.g. Thornicroft (1994); Zirkel (1992); Zwerling (2008), others find marginally statistically significant and negative effects (e.g. Caldwell and Jeffreys (1983); Johnson (2011)). Taken together, these studies suggest that school disruptions caused by teacher strikes have a

minimal impact on student outcomes.[12]

To the best of our knowledge, only two studies that look at the effect of teacher strikes on student outcomes have relied on research designs that are not cross sectional: Michele and Dinand (2010) and Baker (2013). Michele and Dinand (2010) exploit an institutional reform in Belgium in 1990 that led to substantial and frequent strikes in the French-speaking community but not in the Flemish-speaking community of the country. By comparing the difference in education outcomes between individuals in school to those not in school in the French-speaking community to that same difference in the Flemish-speaking community, the authors find some evidence in favor of strikes causing a reduction in education attainment and an increase in class repetition. Though interesting, this study is not able to examine if the identified education effects carry over to the labor market, if there are non-educational effects of teacher strikes or if there are intergenerational treatment effects. Further, the point estimates in Michele and Dinand (2010) provide the intent-to-treat effect of exposure to all strikes in 1990 among students in all grade school years. This makes it difficult to extrapolate the marginal effect of teacher strikes on students in specific school grade years.

Baker (2013) evaluates the effect of teacher strikes on student achievement in Ontario by comparing the change in test score between grade 3 and 6 for cohorts exposed to a strike to the corresponding change for cohorts that were not subject to a strike. The results suggest that strikes that lasted for more than 10 days and took place in grade 5 or 6 have statistically and economically significant negative effects on test score growth, while strikes that occurred in grades 2 or 3 do not. However, data limitations prevent the author from examining long-run education and labor market effects – one of the main contributions of the current analysis.

[12]It should be noted that these studies – just as the current paper – are unable to look at any general equilibrium effects associated with teacher strikes, and therefore do not provide a complete analysis of the benefits and costs associated with teacher strikes.

There is no existing research that has explored the long-run educational attainment and labor market effects of teacher strikes. Further, no study has been able to examine if there are intergenerational treatment effects associated with teacher strike exposure. These gaps in the literature prevent us from fully understanding the dynamics of teacher industrial action, and whether the net effect of such policies is beneficial or harmful to students. This cements the importance of our empirical investigation on the topic.

3.4 Data

3.4.1 Teacher Strikes

Data on teacher strikes are obtained from annual reports on the Argentine economy published by Consejo Técnico de Inversiones (CTI). These reports provide province- and sector-specific information on strikes per month, and we use information from 1977 to 1998 to construct our data set. We assume that children begin school the calendar year they turn 6, and graduate from primary school at the age of 12. This means that we have information on exposure to teacher strikes while in primary school for children born between 1971 and 1985.[13]

In our main analysis, we restrict attention to teacher strikes in primary school. This decision is based on the fact that there are multiple levels of selection that would complicate the analysis on teacher strikes in secondary school. In particular, four factors stand out. First, during our analysis period only primary education was mandatory (less than 60 percent

[13]The assumption that children attend primary school between the ages of 6 and 12 leads to some measurement error in treatment assignment because children start primary school the calendar year in which the number of days they are 6 years old is maximized. This assumption will thus cause a slight attenuation bias in our results. Using household survey data on the educational attainment of 6 year olds between 2003-2015, we estimate that 70 percent of individuals in our sample are assigned to the right cohort.

attended secondary school). Second, the strike data used here concerns public primary school teachers. While secondary school teachers may participate in these strikes, the institutional features of the Argentinian education system make it unlikely that they will. Third, private enrollment was 53 percent higher in secondary schooling (28.7 percent) for the cohorts in our sample and strikes are less prevalent in the private sector, reducing the percentage of students whose school day is disrupted by the strikes. Fourth, if school disruptions in primary school affects educational attainment, it may impact who enters secondary education causing selection bias (something we show in Section 6.3). The first three of these four factors would introduce noise and cause attenuation bias while the fourth factor would bias the results upward. All of these factors would make it difficult to identity the causal effects of teacher strikes on student outcomes during secondary school based on our empirical strategy. To demonstrate this point, we show results from specifications that use strike exposure during both primary and secondary school as the treatment variable in Appendix Table C.10. As will be discussed in Section 6, these results are much noisier and less negative than our preferred results that examine the effect of teacher strikes only in primary school.[14]

Table 3.1 depicts our identifying variation. Looking across the table, there is substantial variation both within provinces over time and across provinces in any given year. Table 3.1 also shows that the average number of days of teacher strikes that these cohorts were exposed to during primary school is 40 (3.2 percent of primary school).[15] If one takes national teacher strikes into account this number increases to 88 (6.98 percent).[16] As discussed in Section 2, strikes were prohibited during the military junta of 1977-1983. This explains why the oldest

[14]Only using variation in teacher strike exposure in primary school means that cohorts who were in secondary school during high strike years are part of the "control" groups. If there is some correlation between strikes in primary and secondary schools, then this would bias our results toward finding no effects. However, given the factors listed in the text and the fact that our main results for primary school remain significant after controlling for strikes during secondary school we believe that such correlation is unlikely.

[15]Primary school in Argentina is comprised of 1260 instructional days, 180 days per year.

[16]We do not consider national teacher strikes when constructing our treatment measure as they are completely subsumed by the cohort fixed effects that we use. See Section 5.

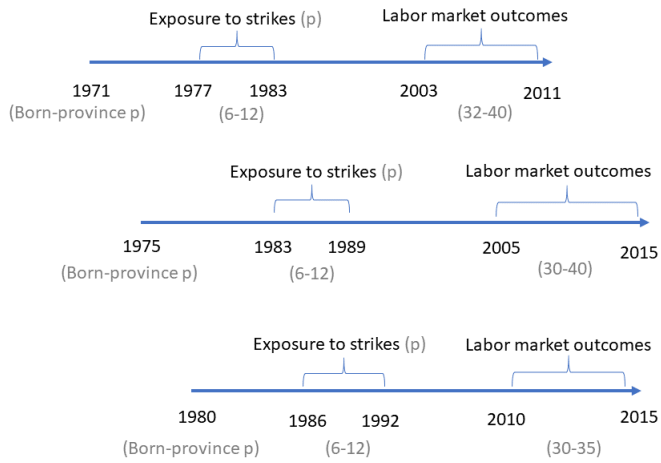
cohorts in our sample are exposed to relatively fewer days of teacher strikes.

Table 3.1: Days of teacher strikes during primary school by birth cohort and birth province

	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	**Mean**
Buenos Aires	2	2	3	5	6	14	35	36	71	76	74	77	69	52	50	**38**
Catamarca	9	11	21	29	29	30	45	38	36	29	35	42	51	56	55	**34**
Chaco	0	5	5	23	40	45	76	88	88	91	108	97	103	74	62	**60**
Chubut	0	0	2	31	45	62	65	82	82	80	53	39	23	20	3	**39**
Ciudad Bs.As.	0	0	0	0	0	0	0	0	7	13	22	22	22	22	22	**9**
Cordoba	1	1	2	13	19	19	27	30	32	34	34	35	76	70	66	**31**
Corrientes	0	0	0	5	12	12	12	12	16	16	11	4	4	4	4	**7**
Entre Rios	6	6	6	6	6	6	8	2	2	10	10	11	15	13	13	**8**
Formosa	0	0	0	2	2	2	2	2	2	2	0	0	0	0	0	**1**
Jujuy	12	12	12	27	27	54	85	91	95	98	83	87	75	49	31	**56**
La Pampa	9	9	9	9	9	9	9	0	0	0	0	0	0	0	0	**4**
La Rioja	0	1	9	24	44	107	107	110	112	110	147	134	99	98	95	**80**
Mendoza	0	0	0	35	68	68	72	72	72	74	39	6	6	2	3	**34**
Misiones	7	7	7	7	7	7	7	0	3	5	5	5	15	15	15	**7**
Neuquen	4	4	4	9	19	19	19	15	17	22	17	7	9	17	53	**16**
Rio Negro	45	45	45	49	68	73	73	30	31	31	45	31	114	121	125	**62**
Salta	4	8	8	13	27	56	118	163	168	170	165	193	178	117	69	**97**
San Juan	5	7	19	23	27	27	41	41	40	30	25	21	40	26	27	**27**
San Luis	7	7	19	22	25	28	31	24	29	17	19	16	13	10	10	**18**
Santa Cruz	4	6	12	17	19	19	49	46	47	42	37	35	35	5	4	**25**
Santa Fe	19	29	31	56	67	106	180	207	203	205	180	169	130	56	10	**110**
Sgo del Estero	2	3	3	16	27	29	38	48	47	62	132	126	132	123	111	**60**
T. Del Fuego	0	0	0	0	0	0	0	4	6	6	6	6	6	8	4	**3**
Tucuman	4	13	76	105	159	179	232	269	264	201	172	118	109	65	26	**133**
Mean	**6**	**7**	**12**	**22**	**31**	**40**	**55**	**59**	**61**	**59**	**59**	**53**	**55**	**43**	**36**	**40**

Notes: Authors' tabulations from annual reports on the Argentine economy published by Consejo Tecnico de Inversiones (1977-2014). The table shows the total days of exposure to teacher strikes at ages 6-12 for each birth year- birth province cell. Cohorts 71-85 correspond to the 30-40 year old respondents in the 2003-2015 EPH for which outcomes variables are available.

Figure 3.2: Data structure for a subsample of birth cohorts



Notes: Example of three cohorts that are part of our main analysis.

3.4.2 Long-run Outcomes

Our main outcome data come from the 2003-2015 waves of the Encuesta Permanente de Hogares (EPH), a household survey representative of the urban population of Argentina (91 percent of the population). We restrict out analysis to individuals between the ages of 30 and 40 because these individuals are typically on a part of their life-cycle where current earnings are reflective of lifetime earnings (e.g. Haider and Solon (2006); Böhlmark and Lindquist (2006). Figure 3.2 shows a visual depiction of the data structure for a sample of birth cohorts.[17]

Critical to our identification strategy is our ability to link respondents to their province

[17]The birth cohorts range from 1971 to 1985. These are the only cohorts that are between 30 and 40 years old when the outcomes of interest are measured (2003-2015) for which we can perfectly calculate exposure to teacher strikes during primary school. This means that we do not have a balanced panel of age observations across the EPH waves. In Section 6.3 we show that limiting our analysis to EPH waves 2011-2015 for which we have a balanced panel has no impact on our results.

of birth, because teacher strikes may lead to selective sorting across provinces, especially if exposure to strikes affects school quality. Teacher strikes could also impact post-primary school mobility patterns if strike-induced education effects affect one's access to national labor markets. Relying on birth province rather than current province of residence eliminates these endogenous migration issues. It is still the case that a fraction of respondents will be assigned the wrong treatment dose as families can move across provinces such that birth province is different from the province in which the child attended primary education. However, Appendix Table C.1 shows that the province of residence is the same as the birth province for 93 percent of 13 year olds in Argentina.[18]

To construct our analysis sample, we collapse the data on the birth province – birth year – EPH year level. Aggregation to this level is sensible because treatment varies on the birth province – birth year level. Appendix Table C.2 provides summary statistics of the outcome variables we use in our analysis. For educational attainment, we generate dummy variables for completion of secondary education and for having obtained at least a bachelor's degree. These indicators are constructed from a years of education variable that we also use to examine the educational attainment effect of strike exposure. With respect to labor market outcomes, we look at the proportion of people that are unemployed, out of the labor force and dedicated to home production (neither studying nor working). To construct a measure of occupational skill we follow Lovenheim and Willén (2016) and calculate the fraction of workers in each 3-digit occupation code that has more than a high school degree. We use this to rank occupations by skill level to examine if strike exposure leads individuals to sort into lower-skilled occupations.[19] We also use the EPH measures of hours worked and earnings. With respect to earnings, we consider both the log of hourly wage and log of total labor

[18] In Section 6.3 we further show that our results are robust to excluding the five provinces with the highest migration rates.

[19] We also construct two alternative measures of occupational skill based on average years of education and average wage in the occupation. The results are robust to these alternative measures.

earnings. Since teacher strikes may affect labor force participation and unemployment, we also study the effect on the level of total labor earnings, which includes individuals with zero earnings.

Preliminary evidence on the relationship between teacher strikes and student outcomes is displayed in Figure 3.3, which plots the predicted years of schooling (Panel A) and labor earnings (Panel B) as a function of the number of days of teacher strikes during primary school.[20] There is clear suggestive evidence of a strong linear negative correlation between teacher strikes and later-in-life outcomes: For each 180 days of teacher strikes (one year of primary school) labor earnings are reduced by 6.7 percent, and years of education decline by 3.1 percent, relative to the sample means.[21] Though instructive, it is important to note that causal inference cannot be made from these graphs.

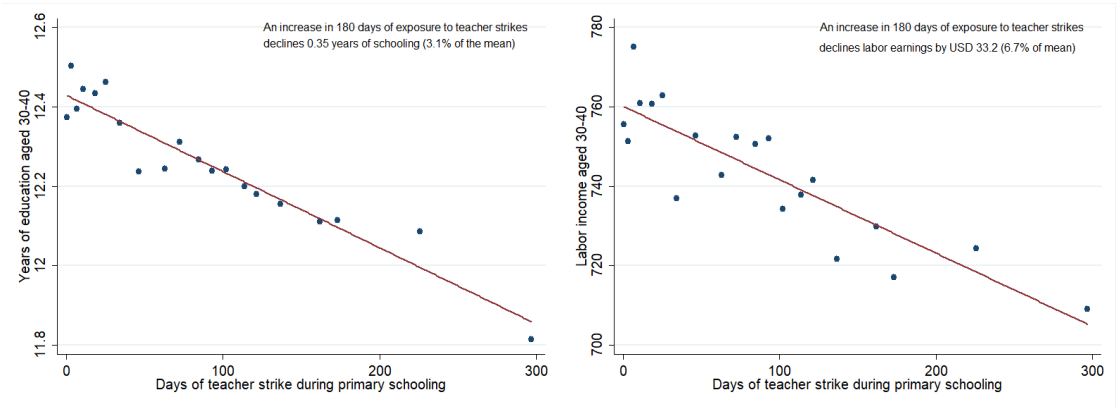
We also examine the effect of strikes on several sociodemographic outcomes: the likelihood of being the household head; the likelihood of being married; the number of children in the household; the age of the oldest child; the education level of the partner; and the per capita income of the household. In addition, we analyze intergenerational effects by examining the effect of teacher strikes on two education outcomes of children to individuals who were exposed to strikes in primary school. We first construct a dummy variable that equals 1 if the child is not delayed at school (age of the child minus years of education plus 6 is greater than zero). We then construct a variable of the educational gap of the child, defined by years of schooling plus 6 minus age. We collapse these variables at the household level.

[20] The figures are obtained through a model that includes birth year, birth province and EPH fixed effects. See figure notes for information.

[21] 180 days is also the difference between the 10th and the 90th percentile of strike exposure among the individuals included in our sample.

Figure 3.3: Correlation between teacher strikes and student outcomes

(a) Years of Education **(b)** Earnings



Notes: The figure is a binned scatter plot. The horizontal axis shows the days of teacher strikes during primary education, which varies at birth year- birth province level. The vertical axis of Panel A contains the average years of education and Panel B the average labor income for each birth year- birth province-survey year cell, after controlling for province, cohort and survey year fixed effects. Data is grouped on 20 intervals of equal number of observations according to days of exposure to teacher strikes. Each point correspond to the group average of the variable in the vertical axes. 180 days of teacher strikes is equivalent to a full year of primary school and the difference between the 10th and the 90th percentile of teacher strike exposure among the individuals included in our sample.

3.4.3 Local Labor Market Controls

One of the main threats to our research design is the possibility that teacher strikes are driven by local labor market conditions. If that were the case, the effects we identify do not represent the effect of exposure to teacher strikes during primary school holding all else constant, but rather the effect of teacher strikes *and* local labor market conditions during primary school.

To minimize this identification threat we include two variables in our estimating equation that control for variation in local labor market conditions during primary school across provinces and time. First, we collect data on public administration strikes by province and year from CTI (the occupation with the largest number of strikes during our analysis period

after teachers).[22] By controlling for public administration strikes during primary school, we exploit variation in teacher strikes net of any general province-specific events and conditions that fuel labor conflict. Second, we collect data on province-specific GDP.[23] We average the province-specific GDP during the seven years of primary school for each birth year -birth province cell.

These controls significantly reduces the risk that our results are driven by local labor market conditions; such factors have to be uncorrelated with province-specific GDP and public administration strikes but correlated with teacher strikes and independently affect the outcomes that we examine. One way in which this could happen is if there are province-specific public school conditions that trigger teacher strikes that are unrelated to our local labor market controls and fixed effects. If, for example, poor material conditions or low salaries trigger strikes that effectively changed these variables at the province level, these changes will not be captured by our fixed effects and labor market controls. If this was the case, it is hard to assume that changes in these conditions are not biasing the results, as they could have caused lower outcomes for exposed cohorts even if strikes had not happened.

To ensure that province-specific public school factors are not driving our results, we have obtained data on teacher wages from 1996 through 2009 from the Ministry of Education. Provided that the relationship between teacher strikes and teacher wages during this time period is informative of that same relationship during the period 1977-1998, these data help lower this concern.

[22]Public administration strikes make up more than 25 percent of all labor strikes in Argentina (Chiappe, 2011; Etchemendy, 2013).

[23]This data comes from Mirabella de Sant (2002), who estimates province GDP using residential electricity consumption.

3.4.4 Short-run Outcomes

To examine if the adverse long-run effects of strike exposure that we identify are present immediately after the children have been exposed to strikes, or if the effects develop over time, we complement our main analysis with an analysis on the effect of exposure to teacher strikes on outcomes of students who have just finished primary school.[24] The data that we use for this analysis come from the 2003-2015 EPH waves for children between 12 and 17 years old. We concentrate on educational outcomes since most of these individuals have not yet entered the labor market. These outcomes are: the likelihood of having attended primary school, the probability of attending public school, years of education, the likelihood that the main activity is home production, and the likelihood of being enrolled in secondary school. Unfortunately, we do not have access to any test score data that could provide further evidence on the direct effect of teacher strikes on human capital accumulation.

Though this analysis is informative for better understanding the channels through which our identified long-run effects operate, it is important to note that this sample is different from our main analysis sample, and that these individuals were exposed to teacher strikes in a time period much different from our main analysis sample. Some caution is therefore necessary when comparing the results from the two analyses.[25]

[24] Due to educational reforms during the past two decades, grade 7 became a part of secondary education in 2002, and mandatory education was extended from 7 to 12 years in 1998. In this section the treatment variable is still defined as the days of strike while students were in primary school, which is now when the children were between 6 and 11 years old.

[25] Two major differences are the increase in enrollment in private primary schooling and the large uptake in secondary enrollment rate (compulsory after 1998). Both factors attenuate the effects of public teacher strikes during primary schooling.

3.5 Empirical Methodology

We exploit cross-cohort variation in exposure to teacher strikes during primary school within and across provinces in a dose-response difference-in-difference framework. Specifically, we estimate models of the following form:

$$Y_{pct} = \beta_0 + \beta_1 TS\_Exposure_{pc} + \gamma X_{pc} + \emptyset_t + \vartheta_c + \varphi_p + \delta T_c + \theta T_p + \varepsilon_{pct}, \tag{3.1}$$

where Y_{pct} is an outcome for respondents born in province p, in birth cohort c and observed in EPH year t. Regressions are weighted by the number of observations in each birth province - birth year - calendar year cell. The variable of interest is $TS\_Exposure$ and measures the number of days (in tens) that the cohort was exposed to strikes during primary school. Standard errors are clustered on the birth province level.[26]

Equation (3.1) also includes province (φ_p), birth cohort (ϑ_c) and calendar year ($\emptyset_t$) fixed effects as well as a province-specific linear time trend (θT_p) and a cohort-specific linear time trend (δT_c). θT_p absorbs any trend in Y over time within a province, and δT_c absorbs any trend in Y over time within a birth cohort. Equation (3.1) further contains a vector of province-specific covariates (X_{pc}) that control for average socioeconomic and demographic characteristics of the province while the cohort was in primary school.[27]

In addition to using equation (3.1) as defined above, we estimate models that substitute

[26] As we only have 25 effective clusters, we also estimate our results using cluster bootstrap with asymptotic refinement (wild cluster bootstrap) as discussed in Cameron and Miller (2015). Appendix Table C.4 shows that our results are robust to this adjustment.

[27] In results not shown, we have also estimated this equation using number of strikes, rather than number of days of strikes, as our measure of treatment intensity. The results obtained from this alternative specification are consistent with the results presented in this paper: the number of strikes exposed to during primary school is associated with negative educational attainment and labor market effects. We further find substantial heterogeneity when using this alternative measure: the negative effects are driven exclusively by strikes that lasted for more than two days. That the effects are dependent on the length of the strikes is consistent with Baker (2013).

the time trends for birth province-by-calendar year and birth year-by-calendar year fixed effects. The province-by-calendar year fixed effects control for variation in Y that is common across birth cohorts within a province in a given year (e.g. province-specific macroeconomic shocks) and the birth year-by-calendar year fixed effects control for any systematic difference across birth years that may be correlated with exposure to teacher strikes and the outcomes of interest. Though more flexible than equation (3.1), this is a very demanding specification, in particular bearing in mind our relatively low number of observations. Because our results are robust to both of these models, we consider equation (3.1) to be our preferred specification. Results using the alternative model are shown in the online appendix.[28]

The unit of observation is a birth province — birth year — calendar year, and the identifying variation stems from cross-cohort variation in exposure to teacher strikes during primary school within and across provinces. There are two main assumptions underlying our estimation strategy. First, that there are no shocks (or other policies) contemporaneous with teacher strikes that differentially affect the different cohorts. The most serious threat to this identification assumption is that strikes may be caused by political events or economic conditions that also vary at the birth province – birth year level and independently affect the outcomes of interest. This would bias our results and lead to invalid inference. To

[28]We also perform our analysis using an instrumental variable approach in which we instrument teacher strikes with public administration strikes. This estimation strategy relies on a set of assumptions that are distinct from our preferred cross-cohort difference-in-difference method: that exposure to public administration strikes must be a good predictor of exposure to teacher strikes and that, conditional on the covariates and fixed effects included in the model, exposure to public administration strikes cannot have an independent effect on the outcomes of interest. The most serious threat to the exclusion restriction is that public administration strikes may have an effect on student outcomes that does not operate through exposure to teacher strikes (which is why we have included exposure to public administration strikes as a control variable in equation (3.1)). However, given the rich set of fixed effects as well as the control for province-specific GDP that we include in our model, this is unlikely. Our main results are robust to this alternative approach. The main take-away from this exercise is that – even if we cannot ascertain the validity of the assumptions underlying either one of our two estimation methods – the fact that our results are insensitive to which of these methods we use significantly limit the sources of bias that can invalidate our results. The reason is that the two methods rely on completely different sets of assumptions. Results from the instrumental variable approach are available upon request.

limit this identification threat we control for public administration strikes as well as average province-specific GDP during primary. These controls significantly reduce the risk that our point estimates are driven by local conditions or secular shocks; such shocks would have to be uncorrelated with provincial GDP and public administration strikes but correlated with teacher strikes and have an independent effect on the outcomes that we examine (and survive the fixed effects and linear time trends). Further, to the best of our knowledge, there are no other relevant policies that occurred concurrently with these strikes that are correlated both with variation in teacher strikes across provinces and the outcomes that we examine.

The second assumption underlying our analysis is that the timing of teacher strikes must be uncorrelated with prior trends in outcomes across birth cohorts within each province. The conventional method for examining the validity of this assumption is to estimate event-study models that non-parametrically trace out pre-treatment relative trends as well as time varying treatment effects. Our research design does not lend itself well to this approach, and we rely on two alternative methods for illustrating that the timing of teacher strikes is uncorrelated with prior trends in outcomes across birth cohorts within each province.

First, we incorporate province-specific linear time trends to show that our results are not driven by trends in outcomes across birth cohorts within each province. Second, we reassign the treatment variable for birth cohort c to birth cohort c-7, such that the measure of exposure to teacher strikes is the number of days (in tens of days) of primary school strikes that took place while the individuals were 13 – 19 years old. As these individuals have already completed primary school they should be unaffected by these strikes, and the coefficient on *TS_Exposure* should not be statistically or economically significant.[29]

[29] It should be noted that 13-19 year olds were exposed to teacher strikes as well. To the extent that teacher strikes are correlated across years within provinces, this model may produce economically and statistically significant results. This makes any null results obtained through this falsification test even more powerful in terms of supporting our identifying assumptions.

3.6 Results

3.6.1 Long-term Effects of Teacher Strikes

i. Educational attainment

Panel A of Table 3.2 presents baseline estimates of the effect of teacher strikes on educational attainment stratified by gender. Each cell in the table comes from a separate estimation of equation (3.1).[30] The results in Panel A provide strong evidence of adverse education effects associated with teacher strikes. Specifically, ten days of strikes (0.79 percent of primary school) increases the number of both males and females that do not graduate from high school by 30 out of every 1,000, and reduces the number of years of education by approximately 0.025. These effects represent a decline of 0.5 and 0.2 percent relative to the respective means, which are shown directly below the estimates in the table. With respect to tertiary education, the results suggest that ten days of strikes leads to an increase in the number of males that do not complete college by 30 for every 1,000, but that it does not have an impact on females. Though indicative of heterogeneous gender effects, it is important to note that these point estimates are not statistically significantly different from each other.

The average individual in our sample experienced 88 days of strikes during primary school. This suggests that for males (females) the average cohort in our sample suffered adverse education effects with respect to the proportion of people obtaining a high school diploma, a college degree and years of education equivalent to 4.75, 12.76, and 2.02 (3.69, 2.82, a 2.02) percent respectively, relative to the means.[31] Taken together, these results

[30] Replacing the cohort-specific and province-specific linear time trends with birth province-by-EPH survey year and birth year-by-EPH survey year fixed effects does not affect the results (Appendix Table C.3).

[31] This rescaling assumes linear treatment effects. Given the suggestive evidence in Figure 3.3 this is not

suggest that teacher strikes not only have adverse short-term education effects (reduction in the proportion that obtain a high school diploma), but that these effects persist as individuals move through the various stages of the education system (proportion that obtain a college degree and the average number of years of education).[32] This is an important finding that has not been documented before.

an unreasonable assumption.

[32]In section 6.4 we study the effect of teacher strikes on contemporaneous educational outcomes for children aged 12-17, something that we cannot do for our main analysis sample due to data limitations. This auxiliary analysis reveals negative educational effects consistent with the results for older cohorts discussed in this section.

Table 3.2: Effect of Strike Exposure on Individual Outcomes

Panel A. Educational Attainment

	High School Diploma		College Degree		Years of Schooling	
	Males	Females	Males	Females	Males	Females
Strike Exposure	-0.003**	-0.003**	-0.002***	-0.001	-0.026***	-0.022***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)	(0.006)
% Effect	-0.54%	-0.42%	-1.45%	-0.32%	-0.23%	-0.18%

Panel B. Employment

	Unemployed		Not in Labor Force		Home Production	
	Males	Females	Males	Females	Males	Females
Strike Exposure	0.001**	0.001*	-0.000	0.001	0.001	0.003***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
% Effect	1.91%	1.37%	-0.97%	0.22%	1.30%	0.82%

Panel C. Wages and Earnings

	Log Earnings		Log Wages		Total Earnings	
	Males	Females	Males	Females	Males	Females
Strike Exposure	-0.002*	-0.002	-0.003**	-0.002	-1.704	-1.906***
	(0.001)	(0.001)	(0.001)	(0.0013)	(1.351)	(0.624)
% Effect	-	-	-	-	-0.23%	-0.51%

Panel D. Occupational Quality and Work Hours

	Occupational Sorting		Total Hours		Informal	
	Males	Females	Males	Females	Males	Females
Strike Exposure	-0.002***	-0.000	-0.011	-0.026	-0.000	0.002
	(0.000)	(0.000)	(0.042)	(0.050)	(0.001)	(0.001)
% Effect	-0.85%	-0.11%	-0.03%	-0.12%	-0.06%	0.48%

Notes: Authors' estimation of equation (3.1) using 2003-2015 EPH data on 30-to 40 year old respondents. The unit of observation is a birth priovince - birth year - EPH year, and the sample consists of 2460 observations. Regressions include birth province, birth year and EPH survey year fixed effects as well as local GDP and exposure to public administration strikes. Regressions further include cohort-specific and a province-specific linear time trends. Regressions are weighted by the number of individual observations used to calculate the averages for each birth year-birth province- EPH year cell. The coefficient measures the effect of being exposed to ten additional days of teacher strikes in primary school on the respective outcomes. Standard errors are clustered at the birth province level. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

ii. Employment, labor force participation & home production

Pre-existing research has documented a strong positive relationship between educational attainment and later-in-life labor market opportunities (Ashenfelter *et al.*, 1999; Card, 2011; Harmon *et al.*, 2003; Heckman *et al.*, 2006).[33] This suggests that teacher strikes may also affect students' labor market outcomes. Panel B of Table 3.2 examines this question in detail, showing gender-specific estimates for the proportion of individuals who are unemployed, not in the labor force and whose main activity is home production. Looking across the panel, there is clear evidence that teacher strikes lead to an increase in the proportion that is unemployed: ten days of teacher strikes lead to an increase in the proportion of unemployed individuals by one percentage point. This effect is present among both males and females, and represents an effect of approximately 1.4 percent relative to the mean.

Teacher strikes also increase the proportion of people whose main activity is home production, though this effect is only present among females.[34] In terms of effect size, the point estimate suggest that ten days of teacher strikes induce 30 out of every 1,000 females to move from either working or studying to home production. The male estimate is smaller but not statistically significantly different from the female estimate.

With respect to labor force participation, our results show that there is no statistically significant effect of teacher strikes on the extensive margin of employment. However, once we control for province-specific linear birth year trends in Section 6.3 (Panel F of Table 3.6), we do find significant adverse effects of teacher strike on labor force participation among women. Our inability to detect this effect in our baseline table may be due to secular shifts

[33] However, it is not necessarily the case that adverse educational effects carry over to the labor market (Böhlmark *et al.*, 2017).

[34] In our sample, 6 percent are still enrolled in an educational institution and 83 percent of those are enrolled at a university.

in labor market opportunities that occurred for women over the cohorts we consider (Blau and Kahn, 2013; Gasparini and Marchioni, 2015). The effect that we identify in Section 6.3 suggests that exposure to 10 days of strikes reduces female labor force participation by 0.14 percent relative to the mean shown in Appendix Table C.2.

iii. Earnings & wages

The adverse employment and education effects identified in Panels A and B of Table 3.2 suggest that teacher strikes may negatively impact earnings and wages as well. This is examined in Panel C of Table 3.2 with respect to log earnings, log wages and the level of earnings.[35] Looking across the columns in Panel C, the results show interesting effect heterogeneity across genders: while strikes cause a statistically significant reduction in wages and earnings among men conditional on employment, it causes a statistically significant decline in the likelihood of receiving positive earnings among women and therefore shifts the female wage distribution to the left. However, in terms of effects sizes the gender-specific estimates are not statistically significantly different from each other. In terms of interpretation, the estimates indicate that 10 days of strikes lead to a reduction in earnings by 0.2 percent (log-specification), in wages by 0.3 percent, and in earnings by USD 1.8 (level-specification).[36] Scaling the point estimates to account for the average level of exposure to teacher strikes during our analysis period suggests that the average cohort in our sample suffered adverse effects of 1.76, 2.64 and 3.26 percent, respectively.[37]

Another way to interpret our income estimates is to aggregate them up to the country level and consider the total effect on the Argentinian economy. While such back-of-the-

[35] We include the level of earnings (expressed in 2005 PPP dollars) in addition to the log of earnings as individuals with zero earnings automatically are eliminated from the log specification.

[36] The identified effect on the level of employment income is equivalent to 0.34 percent relative to the mean.

[37] These numbers are based on the averages of the gender-specific effects.

envelope calculations must be cautiously interpreted due to the many factors that cannot be taken into account when performing this exercise, it is informative for understanding the potential magnitude of the effect. Using the estimated impact on labor earnings and the average treatment exposure to strikes, the aggregate earnings loss induced by teacher strikes amounts to USD 3265 million.[38] This is equivalent to the cost of raising the average annual employment income of all primary school teachers in Argentina by 87.1 percent.[39] In terms of policy implications, this suggests that it may be worth raising teacher wages if this will prevent them from going on strike.

The point estimates in Panel C of Table 3.2 suggest that the rate of return to education in Argentina is between 5.4 and 9.2 percent.[40] This range is consistent with pre-existing estimates in Argentina of 7-12.5 percent (Kugler and Psacharopoulos, 1989; Galiani and Sanguinetti, 2003; Gasparini *et al.*, 2011a; Montenegro and Patrinos, 2014). However, it is important to note that school missed due to sporadic school closures is fundamentally different than less schooling because one leaves school at an earlier age. In one case, curriculum and learning is repeatedly interrupted and in the other it is not. As such, human capital accumulation might be very different and hence the estimated "return" to years of schooling may not be fully comparable.[41] It is not clear how much one would want to extrapolate

[38]This estimation comes from the sum of the effects on males and females. For males, we multiply the total labor income for the country from EPH 2015 by the percent of employed workers that is male, multiplied by the percent effect for males identified in column 3 of Panel C in Table 3.2, which is scaled by the average treatment exposure. For females, we multiply the total labor income by the percent of employed workers that is female, multiplied by the percent effect for females from column 6 of Panel C in Table 3.2, scaled by the average treatment exposure. This result should be interpreted as the annual labor earnings loss for the entire economy if all adults were exposed to the level of teacher strikes that we find in our sample.

[39]Teachers labor earnings are approximately USD13.000 a year, and there were 289,812 primary school teachers in 2014.

[40]This number is obtained by multiplying the average of the estimated effect on wages for males and on total earnings for females by 18, as the school year consists of 180 instructional days.

[41]For example, conventional education economics suggests that the return to education consists of two components – a human capital component and a signaling component (Lange and Topel, 2006). While both a reduction in formal schooling and teacher strikes may negatively affect human capital accumulation, only a reduction in formal schooling – and not teacher strikes – will likely affect the signaling value of education.

from our estimates about returns to schooling that is more general than the impact of disrupted education. Nevertheless, this type of comparison helps anchor our estimates and put the effects in relation to more known education interventions.

The wage and earnings results in Table 3.2 may conceal important heterogeneous effects across the earnings and wage distributions. We explore this possibility in Table 3.3 with respect to total earnings (Panel A) and log wages (Panel B). The results in Panel B demonstrate that strikes affect all but the tails of both the male and the female distributions. The magnitude of the effect is relatively constant across the different deciles for males and it is smaller for the lower deciles in the case of females. Taken together, these results indicates that the people in the left tail of the wage distribution would have done equally poorly without teacher strikes, and that the people in the right tail of the wage distribution would have done equally well without teacher strikes, while the rest of the individuals would have done better. With respect to total earnings (Panel A), our results are again very similar across men and women, though the magnitudes of the effects are slightly larger for the 40th-90th decile of the female subsample.

To better understand the pattern of these wage and earnings results, Panel C shows results from a similar heterogeneity analysis with respect to educational attainment. These results largely mirror the wage and earning results in the sense that there are statistically significant and adverse effects across all but the bottom and top deciles of the distribution, and that the magnitude of the effects across the different deciles is relatively constant.

Table 3.3: Heterogeneous effects of strike exposure on wages, earnings and educational attainment

Panel A. Total Earnings

	10th	20th	30th	40th	50th	60th	70th	80th	90th
Panel A: Males									
Strike Exposure	0.434	-0.203	-1.368	-1.106	-1.352	-2.128*	-3.120**	-2.393	-5.173
	(1.802)	(1.060)	(0.907)	(0.923)	(0.958)	(1.042)	(1.413)	(1.888)	(3.820)
% Effect	0.25%	-0.06%	-0.31%	-0.20%	-0.21%	-0.28%	-0.35%	-0.22%	-0.38%
Panel B: Females									
Strike Exposure	0.094	0.293	-0.072	-2.041	-2.456	-2.328	-3.770**	-3.126**	-2.353
	(0.070)	(0.316)	(0.954)	(1.688)	(1.763)	(1.714)	(1.342)	(1.480)	(2.285)
% Effect	0.83%	0.98%	-0.10%	-1.39%	-0.98%	-0.63%	-0.75%	-0.47%	-0.26%

Panel B. Log Wages

	10th	20th	30th	40th	50th	60th	70th	80th	90th
Panel A: Males									
Strike Exposure	0.002	-0.003*	-0.004***	-0.004***	-0.004***	-0.005***	-0.005***	-0.004	-0.004
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.008)
Panel B: Females									
Strike Exposure	0.001	0.001	-0.000	-0.002	-0.003*	-0.004**	-0.003*	-0.003**	-0.002
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)

Panel C. Years of Education

	10th	20th	30th	40th	50th	60th	70th	80th	90th
Panel A: Males									
Strike Exposure	-0.012	-0.025**	-0.041*	-0.030*	-0.025*	-0.029**	-0.036***	-0.043***	-0.023*
	(0.008)	(0.010)	(0.020)	(0.016)	(0.013)	(0.010)	(0.009)	(0.013)	(0.012)
% Effect	-0.17%	-0.31%	-0.44%	-0.29%	-0.22%	-0.24%	-0.28%	-0.30%	-0.15%
Panel B: Females									
Strike Exposure	-0.021***	-0.017	-0.012	-0.041**	-0.038***	-0.054***	-0.034**	-0.029**	-0.007
	(0.007)	(0.018)	(0.026)	(0.017)	(0.008)	(0.009)	(0.015)	(0.011)	(0.007)
% Effect	-0.29%	-0.20%	-0.12%	-0.38%	-0.32%	-0.42%	-0.25%	-0.20%	0.04%

Authors' estimation of equation (3.1) using 2003-2015 EPH data on 30-to 40 year old respondents. The unit of observation is a birth priovince - birth year - EPH year, and the sample consists of 2460 observations. Regressions include birth province, birth year and EPH survey year fixed effects as well as local GDP and exposure to public administration strikes. Regressions further include cohort-specific and a province-specific linear time trends. All outcomes are expressed in 2005 PPP dollars. The % effect is dropped for log wage given that the point estimate is already interpreted as a percentage change. Regressions are weighted by the number of individual observations used to calculate the averages for each birth year-birth province- EPH year cell. The coefficient measures the effect of being exposed to ten additional days of teacher strikes in primary school on the respective outcomes. Standard errors are clustered at the birth province level. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

iv. Occupational quality, informal employment & hours worked

In addition to the extensive margin employment effects that we identify above, the adverse effect of strikes on earnings could be driven by a reduction in work hours or worse employment conditions. This is examined in Panel D of Table 3.2, where we look at occupational sorting, hours worked and the proportion that work in the informal sector.

The results suggest that being exposed to 10 days of strikes during primary school has no effect on hours worked, but does have a large negative effect on occupational sorting among men.[42] This effect is not present among women, and we can reject the null hypothesis that there is no difference in effect size across genders. With respect to the average male who was exposed to 88 days of teacher strikes during primary school, the occupational sorting effect represents an effect of 1.32 percent relative to the sample mean in Appendix Table C.2.

With respect to the likelihood of working in the informal sector, we find a precise null effect among males but a sizable effect among females. Although the female coefficient falls just outside of being statistically significant at conventional levels, in alternative specifications where we include province-specific linear cohort trends (Panel F of Table 3.6) or where we replace the province- and cohort-specific time trends with province-by-survey year and birth cohort-by-survey year fixed effects, we find statistically significant effects. We cautiously interpret this as indicative of an effect of strike exposure on the likelihood of working in the informal sector among females. For the average female in our sample who was exposed to the 88 days of teacher strikes during primary school, the increase in the likelihood of working in the informal sector represents an effect of 4.2 percent relative to the mean.

vi. Socioeconomic & intergenerational effects of teacher strikes

[42]The results are robust to alternative measures of occupational quality, such as average wage or years of education in one's occupation.

There is a large literature documenting a strong positive relationship between an individual's education- and labor market outcomes and his/her socioeconomic position (Finer and Zolna, 2014). Teacher strikes may therefore also impact outcomes such as the likelihood of being married, the probability of being the head of the household, the number of children, the educational attainment of the partner, and household per capita income.[43] Table 3.4 explores this question, showing results from estimation of equation (3.1) for each of these outcomes.

Table 3.4 shows evidence of a negative effect on the probability of being household head among females but not males. Relative to the mean, ten days of strikes leads to a 0.19 percent reduction in the likelihood of being household head. That we find effects among females but not males could be due to the heterogeneous effects identified in Section 6.1: while teacher strikes cause males to sort into lower skill occupations, it cause females to move toward home production, potentially lowering their bargaining position in the household. However, evidence of such gender heterogeneity is relatively weak as the male point estimate is not statistically significantly different from the female estimate.

Table 3.4 further shows that strikes affect the characteristics of the partners of the individuals that are exposed to teacher strikes. Specifically, the results show that the partners of females that were exposed to strikes are less educated, such that females experience a marriage downgrading with respect to partner skill: being exposed to the average level of strikes during primary school leads to a decline in the years of education of females' partners by 4.7 percent relative to the sample mean. We do not find a significant effect among males.

Finally, the point estimates in Table 3.4 also show that strikes affect per capita family income: the average individual in our sample is exposed to 88 days of teacher strikes, and this

[43] Given the structure of the EPH, we can only identify children of the head, or the spouse of the head, of the household.

Table 3.4: Effect of Strike Exposure on Socioeconomic Outcomes

	Head of Household		Married		Number of Kids	
	Males	Females	Males	Females	Males	Females
Strike Exposure	-0.002	-0.001*	-0.001	-0.000	0.003	0.004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)
% Effect	-0.20%	-0.12%	-0.07%	-0.06%	-0.23%	-0.51%
	Age of older kid		**Per Capita Family Income**		**Years of Schooling of Partner**	
	Males	Females	Males	Females	Males	Females
Strike Exposure	0.008	0.020	-0.005**	-0.005**	-0.007	-0.037***
	(0.010)	(0.019)	(0.002)	(0.002)	(0.010)	(0.010)
% Effect	0.07%	0.16%	-	-	-0.10%	-0.53%

Notes: Authors' estimation of equation (3.1) using 2003-2015 EPH data on 30-to 40 year old respondents. The unit of observation is a birth priovince - birth year - EPH year, and the sample consists of 2460 observations. Regressions include birth province, birth year and EPH survey year fixed effects as well as local GDP and exposure to public administration strikes. Regressions further include cohort-specific and a province-specific linear time trends. The educational attainment of the partner is defined for heads of households or spouses to heads of households. Regressions are weighted by the number of individual observations used to calculate the averages for each birth year-birth province- EPH year cell. The coefficient measures the effect of being exposed to ten additional days of teacher strikes in primary school on the respective outcomes. Standard errors are clustered at the birth province level. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

is associated with a decline in per capita household income by around 4.4 percent relative to the sample mean. The effect is not statistically significantly different across genders.[44]

Given that teacher strikes not only have adverse effects on student education and labor market outcomes, but also influences the marriage market and family planning decisions, there may be intergenerational effects associated with strikes. This question is explored in Table 3.5, using the intergenerational outcome variables discussed in Section 4 as dependent variables (the probability of not being delayed at school and the educational gap). Across the table, there is evidence of adverse intergenerational education effects among females but not males. This is consistent with the heterogeneous treatment effects identified in Section

[44]The point estimate on per capita family income encompass the effect on labor earnings of exposed individuals, on labor earnings of exposed individual's partner, and on household's composition.

Table 3.5: Intergenerational Treatment Effects

	Not Delayed at School		Gap in Years of Education	
	Males	Females	Males	Females
Strike Exposure	0.001	-0.003***	0.002	-0.007***
	(0.002)	(0.001)	(0.003)	(0.003)
% Effect	0.16%	-0.43%	-0.48%	1.45%

Notes: Authors' estimation of equation (3.1) using 2003-2015 EPH data on 30-to 40 year old respondents. The unit of observation is a birth priovince - birth year - EPH year, and the sample consists of 2460 observations. Regressions include birth province, birth year and EPH survey year fixed effects as well as local GDP and exposure to public administration strikes. Regressions further include cohort-specific and a province-specific linear time trends. Not being delayed at school is a dummy variable that takes the value of one if the age of the child minus years of education plus 6 is greater than zero. The educational gap is defined by years of schooling plus 6 minus age. Regressions are weighted by the number of individual observations used to calculate the averages for each birth year-birth province- EPH year cell. The coefficient measures the effect of being exposed to ten additional days of teacher strikes in primary school on the respective outcomes. Standard errors are clustered at the birth province level. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

6.1. In terms of magnitude, being exposed to ten days of teacher strikes during primary school leads to a 0.43 percent increase in the probability that the child is delayed at school relative to the mean (and to an increase in the education gap of 1.45 percent relative to the mean).

The above discussion shows that teacher strikes not only impact educational attainment and labor market outcomes, but also family planning decisions and the educational outcomes of the affected individuals' children. These results have not been documented before, and additional research that examines these questions should be encouraged.

3.6.2 Heterogeneous Treatment Effects

A large literature has documented that human capital accumulates over time, such that human capital obtained at one point in time facilitates further skill attainment later in life (e.g. Heckman *et al.* (2006)). Therefore, early childhood investments are often argued to yield higher returns than education investments that target older children.[45] With respect to the current analysis, this suggests that exposure to teacher strikes in early grades may have larger adverse effects on long-run educational and labor market outcomes.

Appendix Table C.5 shows the effect of exposure to teacher strikes on the long-term education and labor market outcomes of students based on whether they were exposed to strikes in grades 1 through 4 or in grades 5 through 7. Across the columns in Appendix Table C.5, there is suggestive evidence that teacher strikes in early grades have noticeably larger adverse effects than strikes in later grades. However, these differences are generally not statistically significant. Only for two outcomes we find that the effect of teacher strikes in early school grades is statistically significantly different from the effect of teacher strikes in later school grades: years of education and total earnings for females.

3.6.3 Robustness & Sensitivity Analysis

The results obtained from our preferred specification support the idea that teacher strikes have adverse effects on long-term educational attainment and labor market outcomes. In this section, we explore evidence on whether these results are driven by other policies, trends or events that are not accounted for by the controls in equation (3.1).

[45] This argument is also based on research that finds young children to be more receptive to learning. See Phillips *et al.* (2000).

In Panel A and Panel B of Table 3.6 we exclude the city of Buenos Aires and the province and city of Buenos Aires, respectively. These geographic areas differ slightly from the rest of Argentina with respect to their institutions and legislation, and the purpose of this exercise is to ensure that our results are not exclusively driven by these geographic areas. The results are robust to the exclusion of these regions.

In Panel C we estimate equation (3.1) without the five provinces that have the highest cross-province mobility rates.[46] The point estimates produced for this subsample of provinces are not statistically significantly different from our baseline results. This demonstrates that our results are robust to accounting for cross-province mobility.

Panel D eliminates pre-2010 EPH survey years to ensure that our results are robust to a balanced panel of age observations. Despite a dramatic loss of observations (recall that our baseline analysis relies on the 2003-2015 EPH waves), the point estimates are not statistically significantly different from our baseline results when imposing this restriction.

Panel E displays results from estimation of equation (3.1) when we have reassigned treatment for birth cohort c to birth cohort c-7. These cohorts are very close in age and are likely exposed to similar province-specific macroeconomic environments. However, the c-7 cohorts have already completed primary school when the documented teacher strikes took place, and if our baseline estimates successfully isolate the effect of teacher strikes on student outcomes, we should not find any statistically effects among these cohorts. None of the point estimates are statistically significant; the results are consistent with the assumption that the strikes are uncorrelated with trends in outcomes across birth cohorts within each province.

Panel F shows results for our preferred specification when province-specific linear birth

[46] Chaco, Corrientes, Misiones, Rio Negro and Santa Cruz . See Appendix Table C..

Table 3.6: Robustness and Sensitivity Checks

	Years of Schooling	Occupational Sorting	Log wage	Total Earnings	Unemployed	Home Production
Panel A: Excluding city of Bs.As.						
Male	-0.0262***	-0.0015***	-0.0032**	-1.7039	0.0008**	0.0009
	(0.0064)	(0.0004)	(0.0012)	(1.3505)	(0.0003)	(0.0005)
Female	-0.0217***	-0.0003	-0.0019*	-1.9064***	0.0009*	0.0027***
	(0.0062)	(0.0004)	(0.0010)	(0.6236)	(0.0004)	(0.0007)
Panel B: Excluding province and city of Bs.As.						
Male	-0.0235***	-0.0012***	-0.0034**	-0.6997	0.0005*	0.0007
	(0.0069)	(0.0004)	(0.0016)	(1.1583)	(0.0003)	(0.0006)
Female	-0.0227***	-0.0004	-0.0021*	-2.1205***	0.0008	0.0027***
	(0.0070)	(0.0004)	(0.0011)	(0.6666)	(0.0005)	(0.0008)
Panel C: Excluding provinces with high migration						
Male	-0.0234***	-0.0013***	-0.0030**	-1.5504	0.0009***	0.0011*
	(0.0055)	(0.0004)	(0.0014)	(1.4377)	(0.0003)	(0.0006)
Female	-0.0228***	-0.0003	-0.0026**	-2.0903***	0.0007*	0.0028***
	(0.0060)	(0.0004)	(0.0011)	(0.6843)	(0.0004)	(0.0008)
Panel D: Balanced panel (survey year greater than 2010)						
Male	-0.0216***	-0.0015***	-0.0023**	-1.8006	0.0008**	0.0012**
	(0.0074)	(0.0004)	(0.0010)	(1.1935)	(0.0004)	(0.0004)
Female	-0.0203***	-0.0001	-0.0016	-1.3777	0.0009	0.0033***
	(0.0068)	(0.0006)	(0.0014)	(0.8659)	(0.0007)	(0.0010)
Panel E: Reassigning treatment from cohort c to cohort c+7						
Male	-0.0061	0.0006	0.0022	-1.7665	-0.0003	0.0002
	(0.0129)	(0.0004)	(0.0013)	(1.2947)	(0.0002)	(0.0005)
Female	-0.0132	-0.0011	0.0007	0.0649	-0.0002	0.0002
	(0.0112)	(0.0007)	(0.0020)	(1.3080)	(0.0005)	(0.0010)
Panel F: Including province-specific linear cohort trends						
Male	-0.0192*	-0.0017***	-0.0045**	-3.9414*	0.0007*	0.0008
	(0.0094)	(0.0005)	(0.0020)	(1.9962)	(0.0004)	(0.0006)
Female	-0.0119	0.0002	-0.0020*	-2.9745***	0.0014**	0.0037***
	(0.0090)	(0.0006)	(0.0011)	(0.7630)	(0.0005)	(0.0008)
Panel G: Eliminating cohorts expose to >200 days of strikes (top 1%)						
Male	-0.0262***	-0.0015***	-0.0032**	-1.7039	0.0008**	0.0009
	(0.0064)	(0.0004)	(0.0012)	(1.3505)	(0.0003)	(0.0005)
Female	-0.0217***	-0.0003	-0.0019*	-1.9064***	0.0009*	0.0027***
	(0.0062)	(0.0004)	(0.0010)	(0.6236)	(0.0004)	(0.0007)

Notes: Authors' estimation using 2003-2015 EPH data on 30-to 40 year old respondents. Each column estimates the authors' preferred version of equation (3.1) unless otherwise specified. Panel A exclude the City of Buenos Aires (CABA). Panel B excludes both CABA and the province of Buenos Aires. Panel C excludes the five provinces with the highest cross-province mobility rates (Chaco, Corrientes, Misiones, Rio Negro and Santa Cruz). Panel D eliminates pre-2010 EPH survey years to obtain a balance panel. Panel E shows results from the falsification test where we have reassigned the treatment variable for cohort c to cohort c+7. Panel F incorporates province-specific linear birth year trends to the estimation of equation (3.1). Panel G drops the top 1 percent of the teacher strike exposure distribution. Standard errors are clustered at the birth-province level. *** indicates significance at the 1% level, ** indicates significance at the 5% level, and * indicates significance at the 10% level.

year trends have been included. These results help us to further examine if our empirical research design has successfully managed to isolate the effect of teacher strikes on student outcomes, or if the coefficient estimates simply are driven by trends in outcomes across birth cohorts within each province. The results from this exercise are not statistically significantly different from our baseline estimates.

One of the main threats to valid inference in our paper, despite the inclusion of fixed effects and demographic controls, is that our results are simply picking up differences in outcomes caused by province-specific variation in macroeconomic performance across time. To explore this question, we use post-2003 EPH data (data on local labor markets do not exist before 2003) to examine the relationship between teacher strikes and local labor market conditions. Provided that the relationship between teacher strikes and local labor markets after 2003 is informative of that same relationship during the period 1977-1998, this analysis reveal if the results simply are picking up differences in outcomes caused by province-specific variation in macroeconomic performance over time.

The results from this exercise are shown in Appendix Table C.7. In Column (1) we show the correlation between teacher strikes and the unemployment rate, the average hourly wages and the average per capita family income. In Column (2) we add days of public administration strikes, calendar year and province fixed effects as well as province-specific time trends.[47] Our main finding is that, once we control for public administration strikes, province-specific time trends and province and year fixed effects, there is no significant relation between the local labor market climate and teacher strikes. These results are consistent with the idea that our results are not simply driven by province-specific variation in macroeconomic performance across time.

[47]The results are robust to the inclusion of the 30th and the 70th percentiles of the per capita family income (intended to capture any effect of a change in the distribution of per capita family income). Results are available upon request.

Even if our results are not driven by province-specific variation in macroeconomic performance over time, it could still be the case that province-specific public school conditions are driving the results (e.g. poor material conditions or low wages). If such school conditions are correlated with teacher strikes but not subsumed by our fixed effects, time trends or local labor market controls, they may contaminate our effects as these conditions could have led to lower outcomes of exposed cohorts even if strikes did not happen. To explore this possibility, we look at the relationship between teacher strikes and teacher wages.

The results from this exercise are shown in Appendix Table C.8. In Column (1) we show the correlation between teacher wages and strikes. In Column (2) we add controls for public administration strikes as well as calendar year and province fixed effects. Our main finding is that there is no significant relationship between teacher wage and teacher strikes. These results demonstrate that our identified results are not simply driven by province-specific public school conditions across time.[48]

3.6.4 Short-run effects

In this section we analyze the effect of exposure to teacher strikes on outcomes of students who have just finished primary school.[49] The purpose of this exercise is to examine if the strike effects occur immediately after the children have been exposed, or if they develop over

[48]To further explore this, Appendix Figure A-1 show the proportional change in real wages and teacher strikes in each province for the entire period 1996-2009. We find no association between changes in wages and the number of teacher strikes during this period.

[49]Due to educational reforms during the past two decades, grade 7 became a part of secondary education in 2002, and mandatory education was extended from 7 to 12 years in 1998. In this section the treatment variable is still defined as the days of strike while students were in primary school, which is now when the children were between 6 and 11 years old.

time. We focus on children between 12 and 17 years old when performing this analysis.[50] We concentrate on educational outcomes since most of these individuals have not yet entered the labor market. These outcomes are: the likelihood of having attended primary school, the probability of attending a public institution, years of education, the likelihood that the main activity is home production, and the likelihood of being enrolled in school. We perform this analysis on the individual level to control for household characteristics.[51]

Table 3.7 displays the results for each one of the outcome variables using two different specifications. Column (1) incorporates the same controls as in our preferred specification.[52] Column (2) incorporates additional local labor market controls (the unemployment rate and the average wage in each province-year) and family characteristics.[53]

With respect to females, the results in Table 3.7 show that there is a decline in public education enrollment of 0.93 percent relative to the sample mean. This represents 5.3 percent relative to the mean when we scale the coefficient to account for the average level of strikes among these individuals (57 days). We also find an increase in the likelihood of home production by 3.45 percent, and a decrease the probability of being enrolled by 4.02 percent, relative to the means. For males, exposure to 10 days of strikes reduces the years of education by 0.29 percent relative to the mean. These results indicate that the negative education effects of teacher strikes are visible immediately after the students finish primary school. That the short-run effects are smaller than the long-run effects estimated in Table 3.2 is consistent with the total effect of teacher strikes during primary school becoming more

[50] We exclude birth cohorts 1986-1990 as the educational reform was taking place at a different rate in each province Alzúa *et al.* (2015).

[51] These results are robust to estimation at the aggregate level used in our main analysis. These results are available upon request.

[52] Except for GDP at the province level for which there is not reliable data available in recent years.

[53] 4 dummies for province-specific quartiles of per capita family income and 5 dummies for the maximum educational level of the head of the household: primary education or less, incomplete secondary, complete secondary, incomplete tertiary, and complete tertiary

noticeable when all educational decisions (including college) have taken place.

Table 3.7: Short-Term Effects of Strike Exposure (12-17 Year Olds)

	Public Education		Years of Education		Home Production		Not Enrolled	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Panel A: Males								
Strike Exposure	-0.0025	-0.0014	-0.0235*	-0.0243*	0.0018	0.0018	0.0031	0.0031
	(0.0030)	(0.0028)	(0.0130)	(0.0119)	(0.0016)	(0.0015)	(0.0023)	(0.0023)
% Effect	-0.32%	-0.18%	-0.29%	-0.31%	2.96%	2.96%	3.89%	3.89%
Panel B: Females								
Strike Exposure	-0.0074*	*-0.0060	-0.0051	-0.0063	0.0021*	0.0022*	0.0032**	0.0032**
	(0.0041)	(0.0035)	(0.0102)	(0.0106)	(0.0012)	(0.0012)	(0.0014)	(0.0015)
% Effect	-0.93%	-0.75%	-0.05%	-0.07%	3.45%	3.61%	4.02%	4.02%
Controlling for wage and unemployment		X		X		X		X
Controlling for household characteristics		X		X		X		X

Notes: Authors' estimation of equation (1) using 2003-2015 EPH data on 12 to 17 year old respondents. Column (1) show results using individual-level data and the same controls as in our preferred baseline specification. The model used to produce the results in Column (2) incorporates local labor market variables that may influence the wealth of the family: the unemployment rate and the average wage in each province. The model underlying the results in Column (2) further includes 4 dummies of province-specific quartiles of per capita family income and 5 dummies for the maximum educational level of the head or spouse of the household (primary education or less, incomplete secondary, complete secondary, incomplete tertiary, and complete tertiary). Public education is a dummy variable equal to one if attending a public school. Home production is a dummy that equals 1 if the respondent is neither working nor studying. Standard errors are clustered at the birth province level. The coefficients are interpret as the effect of being exposed to teacher strikes for ten extra days during primary school. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

In Section 2.3 we note that there may be heterogeneous treatment effects of teacher strikes with respect to the socioeconomic characteristics of the student's parents: wealthy parents can afford to move their children to private institutions if they believe the strikes hurt their children, and more educated parents are more likely to be capable to replace lost instructional days with home schooling. Even though we do not have information on parental wealth and educational attainment for the individuals included in our main analysis, we can examine this for children that are between 12-17 years old. In Appendix Table C.9, we estimate the effect of strikes by per capita family income and maximum years of education of the head of the household. Consistent with our predictions, we find clear evidence that the most affected students are those from the most socioeconomically disadvantaged households.

3.7 Discussion and Conclusion

Teacher industrial action is a prevalent feature of public education systems across the globe. Despite a large theoretical literature on labor strikes and a reignited debate over the role of teachers' unions in education, there is a lack of empirical research that credibly evaluates the effect of teacher strikes on student outcomes. This paper contributes to the literature by providing a detailed analysis of the effect of teacher strikes during primary school on long-run education and labor market outcomes. This is not a full analysis of the benefits and costs associated with teacher strikes, but rather a partial equilibrium analysis that uses teacher strikes to measure the effect of school disruptions on student long-term outcomes. Nevertheless, the success of education policies ultimately depends on how they impact the long-run outcomes of students, and identifying the net effect of teacher strikes on student outcomes is therefore of great independent importance.

Our results identify adverse long-run educational and labor market effects for both males and females. For males, we find that exposure to teacher strikes during primary school leads to a reduction in educational attainment, an increase in the likelihood of being unemployed, occupational downgrading, and has adverse effects on both labor market earnings and hourly wages. For females, teacher strikes reduce educational attainment in a way similar to that of men. We find a reduction in the level of earnings among females as well, generated by a reduction on labor force participation towards home production. By looking at 12-17 years old, we demonstrate that the negative educational effects are visible immediately after children have finished primary school, and that these effects are concentrated among children from the most vulnerable households.

Our analysis reveals that strikes affect individuals on other socioeconomic dimensions as well. Specifically, individuals exposed to teacher strikes have less educated partners, lower per capita family income, and are less likely to be the head of the household. While our results suggest that males and females may be affected somewhat differently by teacher strikes, the difference in our gender-specific estimates are often not statistically significant.

The prevalence of teacher strikes in Argentina means that the effect of teacher strikes on the economy as a whole is substantial: A back-of-the-envelope calculation suggests an aggregate annual earnings loss of $3265 million. This is equivalent to the cost of raising the average employment income of all primary school teachers in Argentina by 87.1 percent. In terms of policy implications, this suggests that it may be worth raising teacher wages if this will prevent them from going on strike.

Taken together, our results stress the importance of stable labor relations between government and industry and emphasize the necessity of a good bargaining environment that reduces the number of strikes that students are exposed to. Given that the negative effects

that we identify last for years and even generations, both unions and government should make substantial attempts to limit the prevalence of strikes. One policy could be to introduce labor contracts that extend over several years and only allow teachers to strike if a bargaining impasse is reached when renewing these multi-year contracts. This would eliminate sporadic strikes while still allowing teachers to use industrial action as a tool to ensure fair contracts.

APPENDIX A

CHAPTER 1 OF APPENDIX

A.1 Appendix

A.1.1 Proof of Lemma 1

The profit of a competitive firm producing tasks i is:[1]

$$\Pi(i) = p(i)\left(L\alpha_L(i)L(i) + A_M\alpha_M(i)M(i) + A_H\alpha_H(i)H(i)\right) - w_L L(i) - w_M M(i) - w_H H(i),$$

where the price of the task $p(i)$ and the wage of each type of worker are given for the firm. The first order conditions with respect to $L(i)$, $M(i)$, and $H(i)$ imply

$$J(i) > 0 \qquad if \qquad \pi_J(i) = p(i)A_J\alpha_J(i) - w_J \geq 0 \qquad \text{for } J = \{L, M, H\}.$$

Competition assures that $\Pi(i) \leq 0$. Moreover, the Cobb-Douglas production function for the final good implies that production of all task is positive. So $\forall i$ it must be that either $\pi_L(i) = 0$, or $\pi_M(i) = 0$, or $\pi_H(i) = 0$ or two of these conditions has to hold. Unemployment of any type of worker is not possible in this economy because it would imply that wages fall to zero and firms could make a profit in producing task i with that type. Then, a positive amount has to be produce of each task using either low, medium, or high educated workers. Therefore, given the assumption of increasing comparative advantage on i for more educated workers, there must exist two thresholds levels I_L and I_H (where $0 < I_L < I_H < 1$) such that: a) $\pi_L(i) - \pi_M(i) > 0$ and $\pi_L(i) - \pi_H(i) > 0$ for all $i<I_L$, $\pi_L(I_L) - \pi_M(I_L) = 0$; b)

[1]The proof presented here follows the analisys in Acemoglu and Zilbotti (1999).

$\pi_M(i) - \pi_L(i) > 0$ and $\pi_M(i) - \pi_H(i) > 0$ for all $i = (I_L, I_H)$; c) $\pi_M(I_H) - \pi_H(I_H) = 0$; and d) $\pi_H(i) - \pi_L(i) > 0$ $\pi_H(i) - \pi_M(i) > 0$ for all $i > I_H$.

A.1.2 Proof of propositions 1-3

The proof of propositions (1), (2) and (3) proceeds as follows. First, I compute the thresholds' changes that determine the distributions of workers' types to tasks. Then, the effects on wages and the overall distribution of occupations follow directly from the new thresholds.

Changes in task composition within each educational level

The task composition within each educational level is determined by the thresholds levels I_L and I_H. I use comparative statics to estimate the effects of an educational expansion on the thresholds levels.

Proof of $(i.1)$. I follow Acemoglu and Autor (2011) by expressing (1.14) and (1.15) in logs:

$$ln\,(1 - I_H) - ln\,(I_H - I_L) - C_{HM}(I_H) - ln\,h + ln\,m = 0 \tag{A.1}$$

$$ln\,(I_H - I_L) - ln\,(I_L) - C_{ML}(I_L) - ln\,m + ln\,(1 - m - h) = 0. \tag{A.2}$$

Now consider the effect of a change in h by totally differentiating these equations. We thus obtain:

$$\begin{bmatrix} -\frac{1}{(1-I_H)} - \frac{1}{(I_H-I_L)} - C'_{HM}(I_H) & \frac{1}{(I_H-I_L)} \\ \frac{1}{(I_H-I_L)} & -\frac{1}{(I_H-I_L)} - \frac{1}{I_L} - C'_{ML}(I_L) \end{bmatrix} \begin{bmatrix} dI_H \\ dI_L \end{bmatrix} = \begin{bmatrix} \frac{1}{h} \\ \frac{1}{(1-m-h)} \end{bmatrix} dh \quad \text{(A.3)}$$

The determinant of the first matrix, Δ is positive (see Acemoglu and Autor (2011)), therefore by Cramerś rule:

$$\frac{dI_H}{dh} = \frac{\frac{1}{h}\left(-\frac{1}{(I_H-I_L)} - \frac{1}{I_L} - C'_{ML}(I_L)\right) - \frac{1}{(1-m-h)}\frac{1}{(I_H-I_L)}}{\Delta} < 0 \quad \text{(A.4)}$$

$$\frac{dI_L}{dh} = \frac{\frac{1}{(1-m-h)}\left(-\frac{1}{(1-I_H)} - \frac{1}{(I_H-I_L)} - C'_{HM}(I_H)\right) - \frac{1}{h}\frac{1}{(I_H-I_L)}}{\Delta} < 0 \quad \text{(A.5)}$$

The inequality in the first equation is straightforward because all its terms are negative. For the second equation to be positive it is necessary that

$$-C'_{HM}(I_H) > \frac{(1-m-h)}{h}\frac{1}{(I_H-I_L)} + \frac{1}{1-I_H} + \frac{1}{(I_H-I_L)}$$

Which leads to a contradiction since $C'_{HM}(I_H) > 0$ and the RHS is positive. Therefore $\frac{dI_L}{dh} < 0$. ■

An increase in h leads to a reduction in both thresholds, given that higher educated workers start performing medium-level tasks, and medium educated workers are pushed towards lower-level tasks displacing low educated workers. With an increase in h, the average task performed by each educational level diminishes leading to conditional occupational downgrading.

Proof of $(ii.1)$. Following a similar procedure it is possible to write:

$$\frac{dI_L}{dm} = \frac{\left(\frac{1}{m} + \frac{1}{(1-m-h)}\right)\left(-\frac{1}{(1-I_H)} - \frac{1}{(I_H-I_L)} - C'_{HM}(I_H)\right) + \frac{1}{m}\frac{1}{(I_H-I_L)}}{\Delta} < 0 \tag{A.6}$$

$$\frac{dI_H}{dm} = \frac{-\frac{1}{m}\left(-\frac{1}{(I_H-I_L)} - \frac{1}{I_L} - C'_{ML}(I_L)\right) - \left(\frac{1}{m} + \frac{1}{(1-m-h)}\right)\frac{1}{(I_H-I_L)}}{\Delta} > 0 \tag{A.7}$$

The inequality in the first equation is straightforward. For the second equation to be positive it is necessary that:

$$C'_{ML}(I_L) > \frac{m}{(1-m-h)}\frac{1}{(I_H-I_L)} - \frac{1}{I_L}$$

From (A.2) the RHS is < 0, because

$$\frac{m}{(1-m-h)}\frac{1}{(I_H-I_L)} = \frac{1}{I_L}\frac{A_L\alpha_L(I_L)}{A_M\alpha_M(I_L)} < \frac{1}{I_L}.$$

The result $\frac{dI_H}{dm} > 0$ follows by the comparative advantage schedule, given that $C'_{ML}(I_L) > 0$. ■

It is easy to see that with an increase in m the average task performed by low educated workers $(I_L/2)$ and that of high educated workers $((1-I_H)/2)$ will increase. The change in the average task performed by medium educated workers depends on the parameters of the model. Moreover, an increase in m leads to a larger share of tasks performed by medium educated workers, by increasing its participation in low-level and high-level tasks. We are

also interested in $\frac{d(I_H+I_L)}{dm}$ since the average task performed by medium educated workers can be defined as $I_L + (I_H - I_L)/2 = (I_H + I_L)/2$. Depending on which effect is stronger, $\frac{dI_H}{dm}\frac{dI_L}{dm} \lesseqgtr 0$. The average task performed by medium educated workers will increase or decrease depending on whether this inequality is positive or negative. We can write it as follows:

$$\frac{d(I_H + I_L)}{dm} = \frac{-\frac{1}{m}\left(-\frac{1}{I_L} - C'_{ML}(I_L)\right) - \frac{1}{(1-m-h)}\frac{1}{(I_H-I_L)}}{\Delta} + \frac{\left(\frac{1}{m} + \frac{1}{(1-m-h)}\right)\left(-\frac{1}{(1-I_H)} - \frac{1}{(I_H-I_L)} - C'_{HM}(I_H)\right) + \frac{1}{m}\frac{1}{(I_H-I_L)}}{\Delta} \lesseqgtr 0$$

When m and h increase separately, I_L always declines and I_H increases in the first case and decreases in the second case. When m and h increase simultaneously there is a mixture of these effects. As a consequence, I_L will diminish but I_H can increase or decrease depending on what effect is stronger.

Proof of $(iii.3)$. Suppose that $\frac{m'}{h'} < \frac{m}{h}$ and $I'_H > I_H$. By equations (A.5) and (A.6) $I_L < I'_L$. From (1.15), $\frac{m}{h} = c_{HM}(I_H)(I_H - I_L)$, with $c_{HM} = exp(C_{HM})$ and where $c'_{HM}(I_H) > 0$. After the educational expansion $\frac{m'}{h'} = c_{HM}(I'_H)(I'_H - I'_L) > \frac{m}{h}$, which is a contradiction.■

It means that I_H cannot increase unless the relative supply of medium with respect to high educated workers also increases. However, this condition does not work in the other direction.

Changes task composition of the economy

In only a few cases the model provides a sharp prediction of the changes in the task composition of the economy that does not depend on the parameters of the model.

Proof of $(i.2)$ *and* $(ii.2)$. When there is an increase in I_H, some medium educated workers will be added to tasks originally performed only by high educated workers. Therefore, the share of employment in high-level tasks in the economy will increase.

$$E_T = h$$

$$E'_T = h' + \int_{I_H}^{I'_H} \frac{m'}{I'_H - I'_L} di$$

$$\Delta E_T = \Delta h + \int_{I_H}^{I'_H} \frac{m'}{I'_H - I'_L} di > 0 \ \blacksquare$$

This inequality is written for the more general case in which m and h increases. Note that the inequality stills holds for proposition 2 when $\Delta h = 0$.

In the rest of the cases, when I_H declines, it is only possible to define boundaries for the changes in the occupational composition of employment. For example,

$$E_B = l$$

$$E'_B = l' + \int_{I'_L}^{I_L} \frac{m'}{I'_H - I'_L} di$$

$$\Delta E_B = (l' - l) + (I_L - I'_L) \frac{m'}{I'_H - I'_L} > \Delta l.$$

The inequality establishes that when the share of low educated workers is reduced by Δl, the share in bottom-level tasks will decline less than Δl since some medium educated workers will start performing some of these tasks. The extent to which medium educated workers will start performing those tasks depends on the extent of the decline in I_L, which at the same time depends on the comparative advantage schedules and whether the decline in l was due to an increase in m, h or both.

A similar result arises when looking at changes in employment for top-level occupations when I_H declines. Formally,

$$\Delta E_T = (h' - h) - (I_H - I'_H)\frac{h'}{1 - I'_H} < \Delta h.$$

A decline in I_H is only possible if $\Delta h > 0$. The inequality establishes that there is less increase in employment in top-level occupations than the increase in the share of high educated workers. Some high educated workers start performing middle-level tasks. To study the determinants for an increase in employment in top-level tasks, it is possible to state that:

$$\Delta E_T > 0 \Leftrightarrow \frac{\Delta h}{h'} > \frac{(I_H - I'_H)}{(1 - I'_H)}$$

that is, the increase in the supply of high educated workers has to be large with respect to the change in the threshold I_H.

Changes in the wages

We now turn to the case of changes in the wage levels. From the wage equations (1.22),

(1.23) and (1.24) the threshold levels determine the wage levels. After an educational expansion I_L always falls, while I_H can increase or decrease depending on the case. I consider these two cases separately.

Case 1: $\Delta I_L < 0$ *and* $\Delta I_H < 0$

Proof of proposition $(i.3)$.. After an educational expansion, wages for high educated workers can be expressed as:

$$ln\, W'_H = AVP' + I'_H C_{HM}(I'_H) + I'_L C_{ML}(I'_L).$$

Therefore, the wage increase $\Delta ln\, W_H = ln\, W'_H - ln\, W_H$ can be expressed as:

$$\Delta ln\, W_H = \underbrace{\int_{I'_L}^{I_L} C_{ML}(i)\, di + I'_L C_{ML}(I'_L) - I_L C_{ML}(I_L)}_{\Delta I_L \text{ in } W_H \text{ effect} < 0} + \underbrace{\int_{I'_H}^{I_H} C_{HM}(i)\, di + I'_H C_{HM}(I'_H) - I_H C_{HM}(I_H)}_{\Delta I_H \text{ in } W_H \text{ effect} < 0} < 0,$$

where the inequalities for each term comes from the assumption about the comparative advantage schedules. Given the comparative and absolute advantage schedule, $C_{ML}(I_L) > C_{ML}(i) > C_{ML}(I'_L)$ for all $i = (I'_L, I_L)$ and $C_{HM}(I_H) > C_{HM}(i) > C_{HM}(I'_H)$ for all $i = (I'_H, I_H)$. I can establish the following inequalities:

$$\int_{I'_L}^{I_L} C_{ML}(i)\, di < I_L C_{ML}(I_L) - I'_L C_{ML}(I_L) < I_L C_{ML}(I_L) - I'_H C_{ML}(I'_L)$$

$$\int_{I'_H}^{I_H} C_{HM}(i)\, di < I_H C_{HM}(I_H) - I'_H C_{HM}(I_H) < I_H C_{HM}(I_H) - I'_H C_{HM}(I'_H)$$

Therefore, $\Delta ln\, W_H < 0$.

Similarly, define $\Delta ln\, W_L = ln\, W_L - ln\, W'_L$. The change in wages for low educated workers can be expressed as

$$\Delta ln\, W_L = \underbrace{\int_{I'_L}^{I_L} C_{ML}(i)\, di - (1 - I'_L) C_{ML}(I'_L) + (1 - I_L) C_{ML}(I_L)}_{\Delta I_L \text{ in } W_L \text{ effect} > 0} +$$
$$\underbrace{\int_{I'_H}^{I_H} C_{HM}(i)\, di - (1 - I'_H) C_{HM}(I'_H) + (1 - I_H) C_{HM}(I_H)}_{\Delta I_H \text{ in } W_L \text{ effect} > 0} > 0,$$

where the inequalities arise from considering

$$\int_{I'_L}^{I_L} C_{ML}(i) > I_L C_{ML}(I'_L) - I'_L C_{ML}(I'_L)$$

and,

$$\int_{I'_H}^{I_H} C_{HM}(i) > I_H C_{HM}(I'_H) - I'_H C_{HM}(I'_H)$$

then,

$$\begin{aligned}\Delta ln\, W_L &> I_L C_{ML}(I'_L) - I_L C_{ML}(I_L) + C_{ML}(I_L) - C_{ML}(I'_L) \\ &\quad + I_H C_{HM}(I'_H) - I_H C_{HM}(I_H) + C_{HM}(I_H) - C_{HM}(I_H L') \\ &= (1 - I_L)(C_{ML}(I_L) - C_{ML}(I'_L)) + (1 - I_H)(C_{HM}(I_H) - C_{HM}(I'_H)) > 0.\end{aligned}$$

Finally, wages of medium educated workers can be expressed as:

$$\Delta ln\, W_M = \underbrace{\int_{I'_L}^{I_L} C_{ML}(i)\, di + I'_L C_{ML}(I'_L) - I_L C_{ML}(I_L)}_{\Delta I_L \text{ in } W_M \text{ effect} < 0} +$$

$$\underbrace{\int_{I'_H}^{I_H} C_{HM}(i)\, di - (1 - I'_H) C_{HM}(I'_H) + (1 - I_H) C_{HM}(I_H)}_{\Delta I_H \text{ in } W_M \text{ effect} > 0} \lesseqgtr 0,$$

where the results follow from the inequalities in opposite directions derived above for $\Delta ln\, W_H$ and $\Delta ln\, W_L$.

When $\Delta I_L < 0$ and $\Delta I_H < 0$, W_H increases, W_L decreases, and W_L can increase or decrease depending on the parameters of the model.

■

Case 2: $\Delta I_L < 0$ *and* $\Delta I_H > 0$

Proof of proposition $(ii.3)$.. Similarly, in the case$I_H > 0$ it is possible to express:

$$\Delta ln\, W_H = \underbrace{\int_{I'_L}^{I_L} C_{ML}(i)\, di + I'_L C_{ML}(I'_L) - I_L C_{ML}(I_L)}_{\Delta I_L \text{ in } W_H \text{ effect} < 0} +$$

$$\underbrace{-\int_{I_H}^{I'_H} C_{HM}(i)\, di + I'_H C_{HM}(I'_H) - I_H C_{HM}(I_H)}_{\Delta I_H \text{ in } W_H \text{ effect} > 0} \lesseqgtr 0,$$

$$\Delta ln\, W_M = \underbrace{\int_{I'_L}^{I_L} C_{ML}(i)\, di + +I'_L C_{ML}(I'_L) - I_L C_{ML}(I_L)}_{\Delta I_L \text{ in } W_M \text{ effect} < 0} +$$

$$\underbrace{-\int_{I_H}^{I'_H} C_{HM}(i)\, di - (1 - I'_H) C_{HM}(I'_H) + (1 - I_H) C_{HM}(I_H)}_{\Delta I_H \text{ in } W_M \text{ effect} < 0} < 0,$$

$$\Delta ln\, W_L = \underbrace{\int_{I'_L}^{I_L} C_{ML}(i)\, di - (1 - I'_L)C_{ML}(I'_L) + (1 - I_L)C_{ML}(I_L)}_{\Delta I_L \text{ in } W_L \text{ effect} > 0} +$$

$$\underbrace{- \int_{I_H}^{I'_H} C_{HM}(i)\, di - (1 - I'_H)C_{HM}(I'_H) + (1 - I_H)C_{HM}(I_H)}_{\Delta I_H \text{ in } W_L \text{ effect} < 0} \lesseqgtr 0,$$

■

Changes in relative wages

Relative wages are easy to estimate in the model. Their level depends on the changes in the thresholds and are equal to the difference of the equations stated above (alternatively, they can be obtained directly from equations (1.25)-(1.27)) .

Proof of $(i.4)$ *and* $(ii.4)$. Relative wages can be expressed as follows.

$$ln\, W'_H - ln\, W'_M - (ln\, W_H - ln\, W_M) = C_{HM}(I'_H) - C_{HM}(I_H) \begin{cases} > 0, & \text{if } I'_H > I_H. \\ < 0, & \text{if } I'_H < I_H. \end{cases}$$

$$ln\, W'_M - ln\, W'_L - (ln\, W_M - ln\, W_L) = C_{ML}(I'_L) - C_{ML}(I'_L) < 0$$

$$lnW'_H - lnW'_L - (lnW_H - lnW_L) = C_{HM}(I'_H) - C_{HM}(I_H) + C_{ML}(I'_L) - C_{ML}(I'_L) \begin{cases} \lesseqgtr 0, & \text{if } I'_H > I_H. \\ < 0, & \text{if } I'_H < I_H. \end{cases}$$

■

I find that in any educational expansion the wage gap between medium and low educated workers always fall according to the model. This is the case because medium educated workers will always start performing bottom-level tasks when there is an educational expansion (decline in I_L), and the productivity differential of medium with respect to low educated workers always decline, and this is what defines the wage gaps.

APPENDIX B

CHAPTER 2 OF APPENDIX

B.1 Oaxaca Decomposition of Average Wages

This section evaluates the relationship between changes between 1995 and 2014 in average wages of each educational group and their occupational composition of employment. Wages for a given educational group may change in part because workers switched to occupations with a different wage, and in part because there was an increase or decrease in wages of occupations they were already performing. The Oaxaca decomposition disentangles the importance of these two effects on changes in average wages.

The changes in log wages between 1995 and 2014 can be decomposed by the method proposed by Blinder (1973) and Oaxaca (1973), and the role of occupations is obtained following Ferreira *et al.* (2016). Let's W_{it}^s be the wage of individual i from the educational group s in time t and X be a vector of occupational dummies and other personal characteristics (gender and age group dummies). Then we can express:

$$log(W_{it}^s) = \beta_t^s X_{it}^s + \epsilon_{it}^s$$

Consider $t = 1995, 2015$, it is possible to write a model with the two periods pooled together as follows:

$$log(W_i^s) = \beta^s X_i^s + \epsilon_i^s$$

The difference in average log wages between 1995 and 2014 can be written as:

$$E(log(W^s_{i2014})) - E(log(W^s_{i1995})) = E(X^s_{i2014})'(\beta^s_{2014} - \beta^s) + E(X^s_{i1995})'(\beta^s - \beta^s_{1995}) + (E(X^s_{i2014}) - E(X^s_{i1995}))'\beta$$

Let $\bar{j}$ denote sample average of variable j, $\hat{\beta}$ denote ordinary last squares estimates of parameter β, 1 and 2 be 1995 and 2014 respectively, and w be $log(W)$. We can write the sample estimate of the above expression as:

$$(\bar{w}^s_2 - \bar{w}^s_1) = \underbrace{\bar{X}^s_2(\hat{\beta}^s_2 - \hat{\beta}^s) + \bar{X}^s_1(\hat{\beta}^s - \hat{\beta}^s_1)}_{\hat{\Delta}^s_p\text{: pay structure effect}} + \underbrace{(\bar{X}^s_2 - \bar{X}^s_1)\hat{\beta}}_{\hat{\Delta}^s_c\text{:composition effect}} .$$

where $\hat{\Delta}^s_p$ and $\hat{\Delta}^s_c$ are, respectively, the estimate of the pay structure effect and the composition effect for group s. The first term, the composition effect, reflects the changes in average wages of the group s due to changes in the return to covariates Xs. The second term, the composition effect, estimates the change in average wages that is due to changes in the distribution of covariates Xs.

We are interested in studying how changes in returns and composition of occupations are related to changes in average wages. Let X_j be composed by 82 occupational dummies from ISCO-88-3 digit level classification, and X_g be other covariates. Given the linearity of both the pay structure and the composition effect, it is easy to get the part of the effect driven by occupations. It is possible to write:

$$\hat{\Delta}^s_p = \underbrace{\sum_{j=1}^{j=82} \bar{X}^s_{2,j}(\hat{\beta}^s_{2,j} - \hat{\beta}^s_j) + \bar{X}^s_{1,j}(\hat{\beta}^s_j - \hat{\beta}^s_{1,j})}_{\textbf{Occupations}} + \underbrace{\sum_g \bar{X}^s_{2,g}(\hat{\beta}^s_{2,g} - \hat{\beta}^s_g) + \bar{X}^s_{1,g}(\hat{\beta}_g{}^s - \hat{\beta}^s_{1,g})}_{\textbf{Other covariates}}$$

$$\hat{\Delta}^s_c = \underbrace{\sum_{j=1}^{j=82} (\bar{X}^s_{2,j} - \bar{X}^s_{1,j})\hat{\beta}_j}_{\textbf{Occupations}} + \underbrace{\sum_g (\bar{X}^s_{2,g} - \bar{X}^s_{1,g})\hat{\beta}_g}_{\textbf{Other covariates}}$$

Table B.1 shows the results of the decomposition. The decomposition is performed for the entire workforce and for each educational group separately. In column (1) the vector

of personal characteristics X only incorporates occupational dummies. In Column (2) X also contains dummies for combinations of gender, 4 age groups, and 6 levels of education, to account for changes in these characteristics for the entire workforce and within each educational group.

The results of the decomposition are stated as follows. For the entire workforce, the average log wage increased 0.382 between 1995 and 2014, mainly due to a general raise in wages within each occupation. The composition effect due to a change in occupations is small (only explains 11 percent of the increase in wages), while most of the increase in wages comes from the pay structure effect. This is consistent with Section 2.3 where we stated that the occupational composition of the Brazilian economy did not change much during the period 1995-2014. The increase in average wages is mainly due to a generalized increase in wages across all occupations and a decline in the occupational premiums, reflected in the negative effects of occupational premiums in the pay-structure effect and the large positive change in the constant of the regression. These results are robust to incorporate other covariates into the decomposition (column (2)).

The decomposition exercise shows heterogeneous results for workers with different educational level. For low educated workers, the average wages increased due to small but negative composition effect and a large and positive pay structure effect. This indicates that they occupational composition deteriorated (they are more concentrated in low wage in 2014 when compared to 1995), but the wage level had a generalized increase (reflected in the constant) that more than compensate for the occupational downgrading. Note that the sign of the occupation term in the pay structure effect is negative, indicating that the occupational premiums (with respect to the constant) declined. For medium educated workers, the composition effect was large and negative, and it was not compensated by a small

and positive pay structure effect. It reflects that the occupational composition of medium educated workers deteriorated, and even when there was an increase in wages within each occupation it was not enough to compensate for the occupational downgrading. In the case of high educated workers, both the composition and the pay structure effect are negative, indicating that not only there was an occupational downgrading during this period but also that wages within each occupation decline as well. The fact that the occupation term in the pay structure effect is positive shows that the occupational premiums increase (with respect to the constant that largely diminished). This is because wages fall for high educated workers in low wage occupations. The results for each educational level practically do not change when other covariates are incorporated into the decomposition (column (2)).

A better intuition of these results can be obtained from Figure B.1. The figure displays a locally weighted smoothing regression of the changes in log wages between 1995 and 2004 on the ranking of occupations for all workers in panel (a) and for workers with different educational level in panel (b). From the first panel, it is clear that wages increased more in low wage occupations and occupational premiums declined. While the smoothing regression in the second panel shows that wages increased for all occupations in the case of low educated workers, only increased in occupations with a low ranking for medium educated workers, and wages declined for almost all occupations for high educated workers (except those at the top), and this decline was more pronounced in low ranked occupations.

In summary, this section decomposes the changes in wages during the period 1995-2014 between an occupational composition effect (changes in the occupational structure of employment) and a pay structure effect (changes in wages within occupations). We find that for all workers the composition effect is small, consistent with small changes in the occupational composition of the economy; while the increase in wages is driven by higher wages at each oc-

cupation, especially those that started with a lower wage. We also find that all educational levels experienced a negative occupational composition effect, consistent with conditional occupation downgrading. For low educated workers wages within each occupation largely increase, compensating for the negative composition effect, and average wages increased. For medium educated workers, wages increase only in a small number of occupations and it was not enough to compensate for the composition effect, so average wages decreased. For high educated workers, wages declined in each occupation, reinforcing the negative composition effect, and average wages fell.

These results are in line with to those of the model presented on Section 1.2 for the case that I_H declines, as was the case in the calibrated estimation in Section 2.6.1.[1] In the model, an educational expansion generates a deterioration in occupations conditional on education, which is larger for medium educated workers; this is exactly what the composition effect indicates in the data. The model predicts a large increase in the price of occupations where low educated workers are employed, an mild change in prices of occupations performed by medium educated workers (can increase or decrease, but always between the changes in prices of low and high educated workers), and a sharp decline in the prices of occupations performed by high educated workers. These predictions are consistent with the heterogeneous pay structure effect that is observed in the data.

[1] In this discussion, I assume the tasks in the model can be directly interpreted as occupations. See Section 2.2 for a detail discussion on occupations being interpreted as tasks

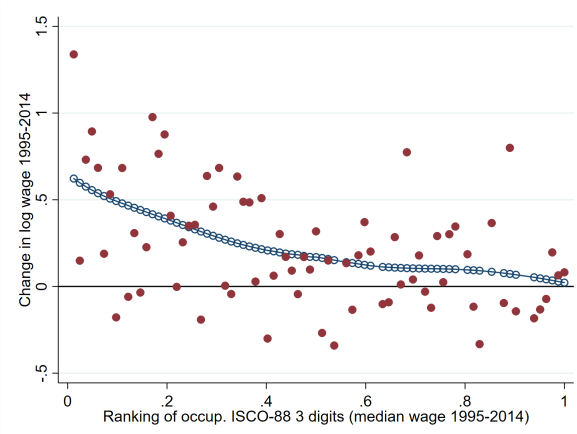
Table B.1: Oaxaca decomposition of changes in mean wages

	All workers		Low educated		Medium Educated		Highly Educated	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Total change in mean ln wages	0.382	0.382	0.333	0.333	-0.038	-0.038	-0.151	-0.151
Decomposition								
Composition effect	0.041	0.200	-0.045	0.016	-0.123	-0.084	-0.055	-0.049
Occupations	0.041	0.023	-0.045	-0.031	-0.123	-0.117	-0.055	-0.052
Gender x edu x age		0.178		0.048		0.032		0.003
Pay structure effect	0.341	0.181	0.379	0.317	0.084	0.046	-0.096	-0.101
Occupations	-0.568	-0.363	-0.487	-0.338	-0.585	-0.616	0.400	0.442
Gender x edu x age		0.302		-0.399		0.169		0.068
Constant	0.909	0.242	0.865	1.054	0.670	0.493	-0.496	-0.611
Relative importance								
Composition/total	0.11	0.53	-0.14	0.05	3.21	2.21	0.36	0.33
Structure/total	0.89	0.47	1.14	0.95	-2.21	-1.21	0.64	0.67
Occupations composition/total	0.11	0.06	-0.14	-0.09	3.21	3.06	0.36	0.35
Occupations structure/total	0.89	-0.32	1.14	2.15	-2.21	3.21	0.64	1.12

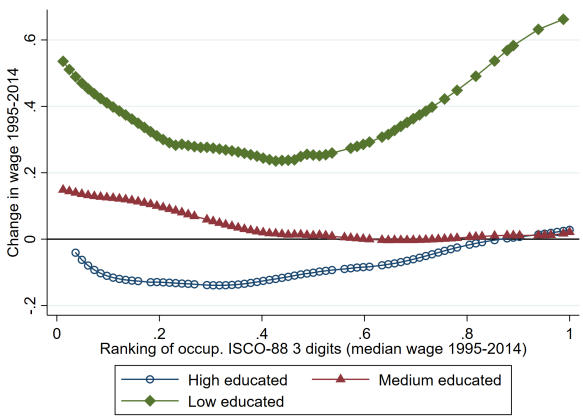
Note: The table presents the Oaxaca decomposition of average log wages between a composition effect and a pay structure effect, which are in turn decompose into the effect of occupations and those of other covariates. Column (1) only considers occupational dummies according to ISCO-88-3 digit level. Column (2) incorporates group dummies for combinations of gender, 4 age groups and 6 levels of education (GenderxEduxAge).

Figure B.1: Changes in log wages by occupation. Period 1995-2014

(a) All Workers



(b) By Educational Level



Note: The figure plots a locally weighted smoothing regression of the changes in log wages between 1995 and 2004. Occupations are ranked as in Figure 2.2. Source: Author's calculation based on PNAD 1995 and 2004.

B.2 Additional tables

Table B.2: Correlation between the ranking of occupations using different samples

	1995-2015 (reweigthed)	1995-2015 (no union nor public)	1995-2015	1995	2004	2014
1995-2015 (re-weighted)	1.000					
1995-2015 (no union nor public)	0.980	1.000				
1995-2015	0.999	0.980	1.000			
1995	0.944	0.944	0.941	1.000		
2004	0.990	0.960	0.990	0.914	1.000	
2014	0.978	0.952	0.980	0.896	0.976	1.000

Notes: The table shows the correlation between the ranking of occupations classified by ISCO-88 3 digit (a total of 82 occupations) based on their median wage for different samples. The sample 1995-2015 (re-weighted) refers to the ranking based on their median wage using all household surveys over the period 1995-2014 with re-weights to account for changes in demographic characteristics of employed workers. The sample 1995-2015 (no union nor public) drop workers that belong to a union or that work in the public sector. The sample 1995-2015 also considers the entire period but it does not account for changes in demographic characteristics. Samples 1995, 2004, and 2014 only take into account household survey in each of those years.

Table B.3: Changes in the occupational structure of employment in Brazil between 1995 and 2014

	Average	Total workforce		Low educated		Medium educated		High educated	
	wage	**1995**	**1995-2014**	**1995**	**1995-2014**	**1995**	**1995-2014**	**1995**	**1995-2014**
Occupations ISCO-88 1 digit									
Professionals	6.3	5.1	5.3	0.6	0.2	3.9	-2.1	33.9	3.2
Legislators and managers	5.5	7.9	-2.8	4.4	-2.5	12.5	-8.8	20.4	-10.1
Military	4.5	2.1	-1.2	1.3	-1.2	4.5	-3.5	2.5	-0.7
Technicians	3.3	7.8	-0.2	3.4	-2.0	17.1	-8.9	17.3	-3.8
Clerks	2.4	8.2	3.6	2.9	0.1	21.9	-7.0	15.6	2.1
Plant and machine operators	2.2	8.1	-0.9	10.5	-1.0	4.3	4.4	0.7	1.3
Craft and related trades	2.0	20.5	-2.1	25.8	1.8	12.5	7.6	2.7	2.0
Service workers and salers	1.8	24.2	6.7	28.8	9.0	19.0	17.8	5.8	6.0
Agriculture and elementary	1.6	16.1	-8.2	22.2	-4.3	4.2	0.5	1.2	0.1
Total		**100.0**	0.0	**100.0**	**0.0**	**100.0**	**0.0**	**100.0**	**0.0**

Notes: I construct a ranking of occupations classified by ISCO-88 3 digit (a total of 82 occupations) based on their median wage over the period 1995-2014.Three categories ISCO-88 3 digit: 82 occupations are divided into 3 groups of an equal number of occupations according to their average wage. Bottom-third refers to the 27 occupations with lower average wage, medium-third are the next 27, and the remaining 28 are classified as top-third occupations. Workers are classified into low (less than secondary), medium (some secondary), and high educated (some university).

Table B.4: Labor market effects of the Brazilian educational expansion under different functional forms

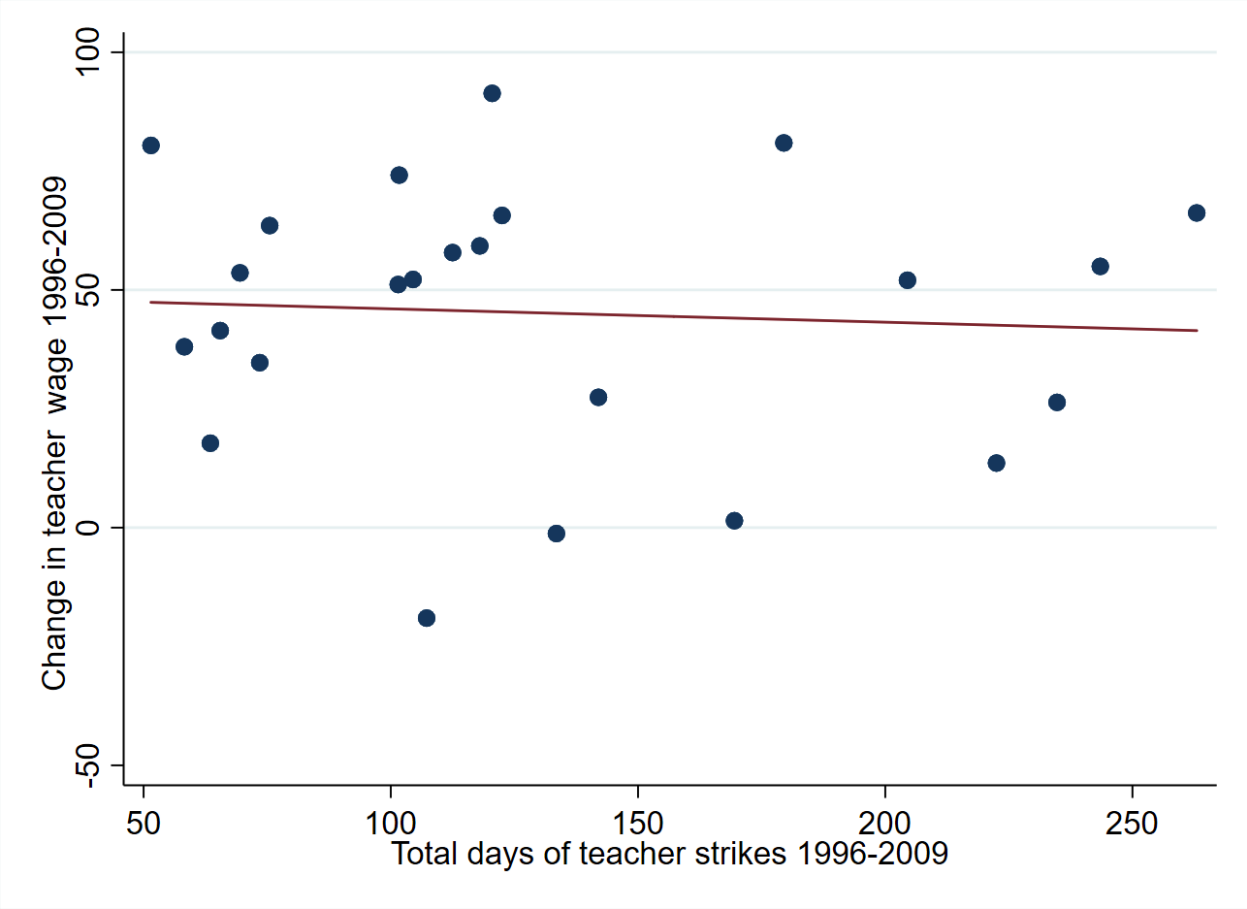
	Brazil 1995-2014	Model: different functional forms $C_{HM}(i)$ & $C_{ML}(i)$ equal to:			
		i	$(i+1)^2$	$i+1$	i^2
Panel A: Thresholds					
Initial I_L		0.424	0.424	0.424	0.424
Initial I_H		0.664	0.664	0.664	0.664
Final I_L		0.257	0.222	0.195	0.300
Final I_H		0.58	0.559	0.534	0.610
Panel B: Changes in occup. composition of employment					
Bottom-third	-0.011	-0.072	-0.034	-0.004	-0.124
Middle-third	-0.005	-0.004	-0.033	-0.054	0.033
Top-third	0.015	0.076	0.067	0.058	0.091
Panel C: Changes in mean occup. ranking					
Low educated	-0.037	-0.083	-0.101	-0.114	-0.062
Medium educated	-0.129	-0.126	-0.154	-0.180	-0.089
High Educated	-0.058	-0.042	-0.053	-0.065	-0.027
Panel D: % Changes in wages					
Low educated	37.3	46.9	29.6	16.7	65.9
Medium educated	-10.5	-10.9	-4.6	-2.1	-17.0
High Educated	-21.6	-22.2	-16.3	-9.7	-30.0
Panel E: Changes in wage gaps					
W_H/W_M	-12.4	-12.7	-12.3	-7.8	-15.7
W_M/W_L	-34.8	-39.3	-26.4	-16.1	-50.0
W_H/W_L	-42.9	-47.1	-35.5	-22.6	-57.8

Notes: Column 1 shows the changes observed in the data from 1995 to 2014. Columns 2-5 show the results of the educational expansion in the model using different fucntional forms for the comparative advantage curves across tasks (i) following Section 2.5. For comparison, column 2 reproduces the preferred estimation from Table 2.6.1.

APPENDIX C

CHAPTER 3 OF APPENDIX

Figure C.1: Correlation between teacher strikes and teacher wages



Notes: The figure is a binned scatter plot. The horizontal axis shows the days of teacher strikes between 1996 and 2009, which varies at province level. The vertical axis depicts the change in real teacher wage over the same period for each of the provinces in Argentina.

Table C.1: Cross-province mobility of 13 year olds

Province	Fraction Non-movers
Buenos Aires	0.979
Catamarca	0.963
Chaco	**0.855**
Chubut	0.930
Ciudad Bs.As.	0.999
Cordoba	0.947
Corrientes	**0.850**
Entre Rios	0.905
Formosa	0.942
Jujuy	0.932
La Pampa	0.952
La Rioja	0.968
Mendoza	0.947
Misiones	**0.836**
Neuquen	0.979
Rio Negro	**0.715**
Salta	0.943
San Juan	0.949
San Luis	0.945
Santa Cruz	**0.835**
Santa Fe	0.975
Sgo del Estero	0.942
T. del Fuego	0.943
Tucuman	0.952

Notes: Authors' tabulations from 2003-2015 EPH data on 13 year old respondents. The table shows the fraction of 13 year olds during 2003-2015 that live in the same province they were born. Bold numbers represents provinces with fraction of non-movers higher than 0.9.

Table C.2: Dependant variable means

	Male	**Female**
Panel A: Educational Attainment		
Secondary Education Completed	0.559	0.620
Years of Education	11.178	11.731
Tertiery Education Completed	0.166	0.248
Panel B: Employment		
Unemployment	0.042	0.066
Not in Labor Force	0.041	0.312
Home Production	0.069	0.329
Informal Sector	0.309	0.354
Hours Worked	42.265	21.239
Occupational Sorting	0.177	0.284
Panel C: Wage and Earnings		
Log Total Earnings	6.489	6.123
Total Earnings	731.8	372.3
Log Wage	1.255	1.257
Panel D: Other Socioeconomic Outcomes		
Head of Household or Spouse	0.743	0.801
Married	0.716	0.688
Number of Children	1.353	1.671
Log Per Capita Family Income	6.791	6.650
Years of Schooling of Partner	11.732	10.357
Age of older kid	11.331	12.315
Panel D: Intergenerational Outcomes		
Not Delayed at School	0.728	0.714
Gap in Years of Education	-0.462	-0.503

Notes: Authors' tabulations from 2003-2015 EPH data on 30-40 years old respondents from 1971-1985 cohorts. Home production is defined as neither working nor studying. Informality is defined as the share of employed workers that are salaried employee in a small firm (less than 5 employees), or works as self-employed without a university degree, or is a family worker with zero earnings. Occupational sorting is evaluated by constructing an index of occupation quality based on the proportion of workers in each occupation with more than a high school degree. Not being delayed at school is defined as a dummy variable takes the value of one if the age of the child minus years of education plus 6 is greater than zero, and it takes the value of zero otherwise. The educational gap defined by years of schooling plus 6 minus age.

Table C.3: Effect of Strike Exposure on Individual Outcomes; two-dimensional fixed effects

Panel A. Educational Attainment

	High School Diploma		College Degree		Years of Schooling	
	Males	Females	Males	Females	Males	Females
Strike Exposure	-0.003**	-0.003**	-0.003***	-0.001	-0.026***	-0.024***
	(0.001)	(0.001)	(0.001)	(0.006)	(0.007)	(0.006)
% Effect	-0.52%	-0.47%	-1.39%	-0.40%	-0.23%	-0.20%

Panel B. Employment

	Unemployed		Not in Labor Force		Home Production	
	Males	Females	Males	Females	Males	Females
Strike Exposure	0.001**	0.001*	-0.000	0.001	0.001	0.003***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
% Effect	1.67%	1.22%	-0.73%	0.29%	1.30%	0.85%

Panel C. Wages and Earnings

	Log Earnings		Log Wages		Total Earnings	
	Males	Females	Males	Females	Males	Females
Strike Exposure	-0.002*	-0.002	-0.003***	-0.002*	-1.653	-1.772**
	(0.001)	(0.001)	(0.001)	(0.001)	(1.264)	(0.663)
% Effect	-	-	-	-	-0.23%	-0.48%

Panel D. Occupational Quality and Work Hours

	Occupational Sorting		Total Hours		Informal	
	Males	Females	Males	Females	Males	Females
Strike Exposure	-0.001***	-0.000	-0.010	-0.019	-0.000	0.002**
	(0.000)	(0.000)	(0.043)	(0.058)	(0.001)	(0.001)
% Effect	-0.79%	-0.14%	-0.02%	-0.09%	-0.03%	0.54%

Notes: Authors' estimation of equation (3.1) using 2003-2015 EPH data on 30-to 40 year old respondents. The unit of observation is a birth priovince - birth year - EPH year, and the sample consists of 2460 observations. Regressions include birth province, birth year and EPH survey year fixed effects as well as local GDP and exposure to public administration strikes. Regressions further include birth provice by EPH survey year and birth year by EPH survey year fixed effects. Regressions are weighted by the number of individual observations used to calculate the averages for each birth year-birth province- EPH year cell. The coefficient measures the effect of being exposed to ten additional days of teacher strikes in primary school on the respective outcomes. Standard errors are clustered at the birth province level. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.4: P-values from Wild Cluster Bootstrap Standard Errors Method

	Years of Education	Occupational Sorting	Log Wage	Total Earnings	Unemployment	Home Production
Panel A: Males						
Strike Exposure	-0.026**	-0.002**	-0.003**	-1.704	0.001**	0.001
P-Value from Wild Cluster Bootstrap Standard Error Method	0.029	0.016	0.032	0.275	0.045	0.134
Panel A: Females						
Strike Exposure	-0.022**	-0.0000	-0.002	-1.906*	0.001	0.003**
P-Value from Wild Cluster Bootstrap Standard Error Method	0.044	0.682	0.119	0.057	0.143	0.022

Notes: Authors' estimation of equation (3.1) using 2003-2015 EPH data on 30 to 40 year old respondents. The unit of observation is a birth province - birth year - EPH year, and the sample consists of 2460 observations. Regressions include birth province, birth year and EPH survey year fixed effects as well as local GDP and exposure to public administration strikes. Regressions further include cohort-specific and a province-specific linear time trends. Regressions are weighted by the number of individual observations used to calculate the averages for each birth year-birth province- EPH year cell. The coefficient measures the effect of being exposed to ten additional days of teacher strikes in primary school on the respective outcomes. P-value is estimated following the wild cluster method with Rademacher 2 point distribution following Cameron and Miller (2015). The bootstrap uses 999 replications. The p-values show the probability of observing the given coefficient value under the null hypothesis of no effect. To facilitate interpretation of the results, stars (*) have been used after the coefficient estimate to indicate which level the coefficient estimate was significant at when the standard errors were clustered at the birth province level. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.5: Heterogeneous Treatment Effects of Strike Exposure by School Grade

	Years of Education	Occupational Sorting	Log Wage	Total Earnings	Unemploy.	Home Production
Panel A: Males						
Strike Exposure 1-4 grade	-0.0295***	-0.0015**	-0.0039**	-1.5495	0.0011**	0.0010*
	(0.0099)	(0.0006)	(0.0015)	(1.4588)	(0.0004)	(0.0006)
Strike Exposure 5-7 grade	-0.0240***	-0.0014***	-0.0027	-1.8088	0.0006	0.0008
	(0.0073)	(0.0004)	(0.0021)	(1.5825)	(0.0004)	(0.0006)
Panel B: Females						
Strike Exposure 1-4 grade	-0.0355***	-0.0011*	-0.0035**	-3.2691***	0.0013*	0.0027**
	(0.0074)	(0.0006)	(0.0013)	(0.9841)	(0.0007)	(0.0013)
Strike Exposure 5-7 grade	-0.0126*	0.0003	-0.0008	-0.9976	0.0006	0.0026***
	(0.0065)	(0.0006)	(0.0012)	(0.8975)	(0.0005)	(0.0008)

Notes: Authors' estimation of their preferred version of equation (3.1) using 2003-2015 EPH data on 30-to 40 year old respondents (controlling for birth province, birth year and EPH survey year fixed effects as well as local GDP and exposure to public administration strikes and including a cohort-specific and a province-specific linear time trend). The treatment variable has been split into 2: teacher strikes that occur in grades 1-4 and those that took place in grades 5-7. Regressions are weighted by the number of individual observations used to calculate the averages for each birth year-birth province-year. The coefficient is interpret as the effect of being exposed to teacher strikes for ten extra days during primary school. Standard errors are clustered at the birth province level. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.6: Effect of controlling for non-teacher strikes and GDP

	Years of Education	Occupational Sorting	Log Wage	Total Earnings	Unemploy.	Home Production
Panel A: Without controls for PA strikes and GDP						
i. Male						
Strike Exposure	-0.0233***	-0.0015***	-0.0034***	-2.1796*	0.0008***	0.0006
	(0.0064)	(0.0004)	(0.0010)	(1.1480)	(0.0003)	(0.0004)
ii. Female						
Strike Exposure	-0.0176***	-0.0003	-0.0020***	-2.5964***	0.0010**	0.0029***
	(0.0053)	(0.0003)	(0.0007)	(0.6296)	(0.0004)	(0.0007)
Panel B: With controls for PA strikes and GDP						
i. Male						
Strike Exposure	-0.0262***	-0.0015***	-0.0032**	-1.7039	0.0008**	0.0009
	(0.0064)	(0.0004)	(0.0012)	(1.3505)	(0.0003)	(0.0005)
PA Strike Exposure	0.0004	-0.0005	-0.0014	-2.0821	-0.0001	-0.0010
(0.0123)	(0.0009)	(0.0035)	(2.3253)	(0.0004)	(0.0008)	
GDP	-1.4222***	-0.0355	-0.0345	-6.9421	-0.0020	0.0132
	(0.3645)	(0.0271)	(0.0871)	(67.6629)	(0.0184)	(0.0323)
ii. Female						
Strike Exposure	-0.0217***	-0.0003	-0.0019*	-1.9064***	0.0009*	0.0027***
	(0.0062)	(0.0004)	(0.0010)	(0.6236)	(0.0004)	(0.0007)
PA Strike Exposure	0.0121	-0.0005	-0.0012	-3.3382**	0.0002	0.0009
	(0.0110)	(0.0009)	(0.0020)	(1.4787)	(0.0014)	(0.0014)
GDP	-0.7139	-0.0662*	-0.0531	-74.3703	-0.0049	0.0406
	(0.4904)	(0.0365)	(0.0468)	(62.6525)	(0.0328)	(0.0546)

Notes: Authors' estimation of equation (3.1) using 2003-2015 EPH data on 30-to 40 year old respondents. Panel A excludes controls for public administration strikes and province-specific GDP. Panel B includes these controls, both defined at the time the cohorts were in primary school. Regressions are weighted by the number of individual observations used to calculate the averages for each birth year-birth province-year. The coefficient is interpret as the effect of being exposed to teacher strikes for ten extra days during primary school. Standard errors are clustered at the birth province level. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.7: Effect of local labor market conditions on teacher strikes

	Teacher Strikes	
	(1)	(2)
Unemployment rate	0.6355**	1.1255
	(0.2591)	(0.9366)
Average wage	0.3605	-1.8366
	(0.6432)	(5.0689)
Average per capita income	0.0016*	-0.0072
	(0.0009)	(0.0061)
Public administration strike exposure		X
Province FE		X
Year FE		X
Province-specific time trends		X
R-squared	0.047	0.407

Notes: Authors' estimation of equation (3.1) using 2003-2015 EPH data and strike data from CTI. The unemployment rate, average wages and average per capita family income describe the labor market conditions for each birth province-calender year cell. Column (1) regress the days of teacher strikes during the period 2003-2015 only on labor market conditions. Column (2) adds days of strikes in public administration, calendar year and province fixed effects and province-specific time trends. Regressions are weighted by the number of individual observations used to calculate the averages for province-year. Robust standard errors in parenthesis. The coefficient is interpret as the effect of local labor market conditions to days of teacher strikes. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.8: Effect of teacher wages on teacher strikes

	Teacher Strikes	
	(1)	(2)
Teacher wage year t	0.0102	-0.0119
	(0.0126)	(0.0130)
Teacher wage year t-1	0.0103	0.0229
	(0.0133)	(0.0212)
Teacher wage year t+1	-0.0177	0.0272
	(0.0114)	(0.0174)
Public administration strike exposure		X
Province FE		X
Year FE		X
Province-specific time trends		X
R-squared	0.018	0.569

Notes: Authors' estimation of equation (3.1) using 1996-2009 data on teacher wages from the Ministry of Education in Argentina and strike data from CTI. The wages correspond to the wages of primary school teachers with 10 years of experience in each province-calendar year. Both columns include province and calendar year fixed effects. Standard errors are clustered at the province level. The coefficient is interpreted as the effect of strike exposure on teacher wages. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.9: Heterogeneous Treatment Effects on Short-Term Outcomes (12-17 Year Olds)

	Male				*Female*			
	Public School	Years of Schooling	Home Production	Not Enrolled	Public School	Years of Schooling	Home Production	Not Enrolled
Panel A: Stratification by Parental Education								
At most primary education	-0.0073**	-0.0513***	0.0030	0.0043	-0.0087**	-0.0256	0.0011	0.0023
	(0.0034)	(0.0151)	(0.0019)	(0.0032)	(0.0038)	(0.0202)	(0.0025)	(0.0027)
Some secondary education	-0.0039	-0.0307*	0.0019	0.0038	-0.0089**	-0.0077	0.0029*	0.0041**
	(0.0030)	(0.0160)	(0.0018)	(0.0030)	(0.0034)	(0.0122)	(0.0016)	(0.0019)
Secondary education	-0.0011	-0.0181	0.0020	0.0032	-0.0037	-0.0101	0.0017	0.0028*
	(0.0029)	(0.0114)	(0.0014)	(0.0020)	(0.0034)	(0.0136)	(0.0011)	(0.0014)
Some tertiary education	0.0019	-0.0176*	0.0013	0.0019	-0.0027	0.0032	0.0024*	0.0031**
	(0.0034)	(0.0097)	(0.0019)	(0.0024)	(0.0034)	(0.0101)	(0.0013)	(0.0014)
Tertiary education	0.0009	-0.0087	0.0012	0.0017	-0.0051	0.0067	0.0012	0.0022**
	(0.0042)	(0.0130)	(0.0013)	(0.0016)	(0.0060)	(0.0083)	(0.0009)	(0.0010)
Panel B: Stratification by Family Income								
First quartile	-0.0059	-0.0339**	0.0021	0.0041	-0.0110**	-0.0220	0.0033*	0.0050**
	(0.0044)	(0.0144)	(0.0017)	(0.0031)	(0.0052)	(0.0165)	(0.0017)	(0.0020)
Second quartile	-0.0022	-0.0353**	0.0029	0.0047*	-0.0077*	-0.0069	0.0025	0.0033*
	(0.0022)	(0.0136)	(0.0018)	(0.0024)	(0.0039)	(0.0097)	(0.0015)	(0.0017)
Third quartile	0.0024	-0.0179	0.0011	0.0021	-0.0020	-0.0003	0.0017	0.0024*
	(0.0019)	(0.0162)	(0.0017)	(0.0025)	(0.0021)	(0.0101)	(0.0011)	(0.0012)
Fourth quartile	-0.0010	-0.0080	0.0009	0.0012	-0.0052	0.0027	0.0012	0.0023**
	(0.0049)	(0.0089)	(0.0012)	(0.0016)	(0.0059)	(0.0090)	(0.0008)	(0.0010)

Notes: Authors' estimation of equation (3.1) using 2003-2015 EPH data on 12 to 17 year old respondents. The results are based on individual-level regressions and the underlying model contains the same controls as that in column (2) of Table 17. Panel A interact the treatment variable with 5 dummies for the maximum educational level of the head or spouse of the household (primary education or less, incomplete secondary, complete secondary, incomplete tertiary, and complete tertiary). Panel B interacts the treatment variables with 4 dummies of province-specific quartiles of per capita family income. Standard errors are clustered at the birth province level.*** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.10: Effect of Strikes During Secondary School

	Years of Education	Occupational Sorting	Log Wage	Total Earnings	Unemploy.	Home Production
Panel A: Exposure to strikes during primary and secondary school						
i. Male						
Strike Exposure	-0.0013	-0.0004	-0.0002	-0.8551	0.0010**	0.0017**
	(0.0138)	(0.0007)	(0.0029)	(2.1016)	(0.0004)	(0.0006)
ii. Female						
Strike Exposure	-0.0083	0.0005	0.0017	-0.8612	0.0011**	0.0008
	(0.0069)	(0.0006)	(0.0018)	(1.1780)	(0.0005)	(0.0009)
Panel B: Exposure to strikes by educational level						
i. Male						
Primary Exposure	-0.0157	-0.0012*	-0.0023	-2.0373	0.0013***	0.0019***
	(0.0119)	(0.0006)	(0.0025)	(1.4239)	(0.0004)	(0.0006)
Secondary Exposure	0.0131	0.0003	0.0021	0.2899	0.0007	0.0013*
	(0.0130)	(0.0007)	(0.0029)	(2.9300)	(0.0005)	(0.0008)
ii. Female						
Primary Exposure	-0.0188**	0.0003	0.0001	-1.9758*	0.0015***	0.0024**
	(0.0076)	(0.0006)	(0.0018)	(1.1483)	(0.0005)	(0.0010)
Secondary Exposure	0.0020	0.0006	0.0033	-0.2375	0.0007	-0.0004
	(0.0093)	(0.0007)	(0.0019)	(1.0948)	(0.0005)	(0.0010)

Notes: Authors' estimation of equation (3.1) using 2003-2015 EPH data on 30-to 40 year old respondents. Panel A defines strike exposure as all the teacher strikes that took place during the years cohorts were supposed to attend primary and secondary school (age 6 to 17). Panel B differentiates exposure to teacher strikes between primary (age 6 to 12) and secondary (13 to 17). Regressions are weighted by the number of individual observations used to calculate the averages for each birth year-birth province-year. The coefficient is interpret as the effect of being exposed to teacher strikes for ten extra days. Standard errors are clustered at the birth province level. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table C.11: Differential Effect of Exposure to Long Lasting Strikes within an School Year

	Years of Education	Occupational Sorting	Log Wage	Total Earnings	Unemploy.	Home Production
i. Male						
Strike Exposure	-0.0301***	-0.0016***	-0.0020*	-1.6521	0.0008**	0.0010*
	(0.0061)	(0.0005)	(0.0010)	(1.0521)	(0.0003)	(0.0005)
Dummy of Exposure to 26-44 Strikes	-0.1222*	-0.0033	-0.0073	-12.7880	0.0032	0.0028
	(0.0691)	(0.0045)	(0.0177)	(12.4007)	(0.0033)	(0.0044)
Dummy of Exposure to More than 45 Strikes	-0.0287	0.0004	-0.0344	-13.4280	0.0017	-0.0001
	(0.1226)	(0.0080)	(0.0348)	(18.9655)	(0.0051)	(0.0057)
ii. Female						
Strike Exposure	-0.0259***	0.0001	-0.0005	-1.7583***	0.0005	0.0029***
	(0.0075)	(0.0006)	(0.0009)	(0.4546)	(0.0005)	(0.0007)
Dummy of Exposure to 26-44 Strikes	0.0371	0.0098	0.0032	24.1319*	-0.0048	-0.0200*
	(0.0814)	(0.0061)	(0.0163)	(12.1390)	(0.0047)	(0.0097)
Dummy of Exposure to more than 45 Strikes	0.1308	0.0011	-0.0274	19.7115	0.0043	-0.0249
	(0.1585)	(0.0111)	(0.0193)	(22.1678)	(0.0085)	(0.0179)

Notes: Authors' estimation using 2003-2015 EPH data on 30-to 40 year old respondents. Each column estimates equation (3.1) adding two dummies to account for differential effect of strike exposure to long lasting strikes within one school year. These variables are constructed by first identifying the maximum number of strikes in a school year for each birth province-birth year cell. We create three mutually exclusive categories of strike maximum strike exposure if the maximum was lower than 25 days within one year (42.9% of the sample), between 26 and 45 days (45.9 percent), or more than 45 days (11.1 percent). Regressions are weighted by the number of individual observations used to calculate the averages for each birth year-birth province-year. The coefficient on teacher strikes exposure is interpret as the effect of being exposed to teacher strikes for ten extra days during primary school. Standard errors are clustered at the birth-province level. *** indicates significance at the 1% level, ** indicates significance at the 5% level, and * indicates significance at the 10% level.

BIBLIOGRAPHY

Acemoglu, D. (1998). Why do new technologies complement skills? directed technical change and wage inequality. *The Quarterly Journal of Economics*, **113** (4), 1055–1089.

— (2002). Directed technical change. *The Review of Economic Studies*, **69** (4), 781–809.

— and Angrist, J. (2000). How large are human-capital externalities? evidence from compulsory schooling laws. *NBER macroeconomics annual*, **15**, 9–59.

— and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In D. Card and O. Ashenfelter (eds.), *Handbook of Labor Economics*, vol. 4, Elsevier, pp. 1043 – 1171.

— and Restrepo, P. (2018). *Artificial Intelligence, Automation and Work*. Tech. Rep. 24196, National Bureau of Economic Research.

— and Zilbotti, F. (1999). *Productivity Differences*. Working Paper 6879, National Bureau of Economic Research.

Adão, R. (2015). Worker heterogeneity, wage inequality, and international trade: Theory and evidence from brazil. *Mimeo*.

Alvarez, J. (2017). *The Agricultural Wage Gap: Evidence from Brazilian Micro-data*. Tech. rep., Mimeo.

—, Benguria, F., Engbom, N. and Moser, C. (2017). *Firms and the decline in earnings inequality in Brazil*. Tech. rep., Columbia Business School Research Paper No. 17-47.

Alzúa, M. L., Gasparini, L. and Haimovich, F. (2015). Educational reform and labor market outcomes: the case of argentina's ley federal de educacion. *Journal of Applied*

Economics, **18** (1).

Anderson, D. M. (2014). In school and out of trouble? the minimum dropout age and juvenile crime. *Review of Economics and Statistics*, **96** (2), 318–331.

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. (2013). The" task approach" to labor markets: an overview. *Journal for Labour Market Research*, **46** (3), 185–199.

— (2014). Skills, education, and the rise of earnings inequality among the"other 99 percent". *Science*, **344**, 843–851.

— and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *The American Economic Review*, **103** (5), 1553–1597.

Autor, D. H., Levy, F. and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, **118** (4), 1279–1333.

Baker, M. (2013). Industrial actions in schools: strikes and student achievement. *Canadian Journal of Economics*, **46** (3), 1014–1036.

Barro, R. J. and Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of development economics*, **104**, 184–198.

Barros, R., Mirela de Carvalho, S. F. and Mendonca, R. (2010). *Markets, the State and the Dynamics of Inequality: The Case of Brazil*. Tech. rep., UNDP Discussion Paper.

Beaudry, P., Green, D. A. and Sand, B. M. (2016). The great reversal in the demand for skill and cognitive tasks. *Journal of Labor Economics*, **34** (S1), S199–S247.

Becker, G. S. (1964). Human capital: A theoretical and empirical analysis, with special reference to schooling. *NY: National Bureau of Economic Research*.

Bekman, E., Bound, J. and Machin, S. (1998). Implications of skill-biased technological change: international evidence. *The quarterly journal of economics*, **113** (4), 1245–1279.

Blau, F. D. and Kahn, L. M. (2013). Female labor supply: Why is the united states falling behind? *American Economic Review*, **103** (3), 251–56.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, pp. 436–455.

Böhlmark, A. and Lindquist, M. J. (2006). Life-cycle variations in the association between current and lifetime income: Replication and extension for sweden. *Journal of Labor Economics*, **24** (4), 879–896.

—, Willén, A. *et al.* (2017). *Tipping and the Effects of Segregation*. Tech. rep., IFAU Working Paper 2017:14.

Bowlus, A. J. and Robinson, C. (2012). Human capital prices, productivity, and growth. *The American Economic Review*, **102** (7), 3483–3515.

Bruns, B., Evans, D. and Luque, J. (2011). *Achieving World-Class Education in Brazil: The Next Agenda*. World Bank Publications.

Burstein, A., Morales, E. and Vogel, J. (2016). *Changes in between-group inequality: computers, occupations, and international trade*. Tech. rep.

Cahan, S. and Cohen, N. (1989). Age versus schooling effects on intelligence development. *Child development*, pp. 1239–1249.

— and Davis, D. (1987). A between-grade-levels approach to the investigation of the absolute effects of schooling on achievement. *American Educational Research Journal*, **24** (1), 1–12.

Caldwell, W. E. and Jeffreys, L. M. (1983). The effect of teacher strikes on student achievement: New evidence. *Government Union Review*, **4** (1), 40–58.

Cameron, A. C. and Miller, D. L. (2015). A practitioneś guide to cluster-robust inference. *Journal of Human Resources*, **50** (2), 317–372.

Card, D. (2011). The causal effect of education on earnings. In O. Ashenfelter and D. Card (eds.), *Handbook of labor economics*, vol. 3, Amsterdam:: North-Holland, pp. 1801–1863.

—, Dooley, M. D. and Payne, A. A. (2010). School competition and efficiency with publicly funded catholic schools. *American Economic Journal: Applied Economics*, **2** (4), 150–76.

Cascio, E. U. and Lewis, E. G. (2006). Schooling and the armed forces qualifying test evidence from school-entry laws. *Journal of Human resources*, **41** (2), 294–318.

Chau, N. H. and Kanbur, R. (2018). *Employer Power, Labor Saving Technical Change, and Inequality*. Tech. rep., Centre for Economic Policy Research Discussion Paper DP12925.

Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K. and Rogers, F. H. (2006). Missing in action: teacher and health worker absence in developing countries. *Journal of Economic perspectives*, **20** (1), 91–116.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W. and Yagan, D. (2011). How does your kindergarten classroom affect your earnings? evidence from project star. *The Quarterly Journal of Economics*, **126** (4), 1593–1660.

—, Hendren, N. and Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, **106** (4), 855–902.

Chiappe, M. (2011). *La conflictividad laboral entre los docents publicos argentines 2006-2010*. Tech. rep., Buenos Aires: Direccion de Estudios de Relaciones del Trabajo.

Cingano, F. (2014). Trends in income inequality and its impact on economic growth. *OECD*.

Colasanti, M. (2008). *State Collective Bargaining Policies for Teachers*. Tech. rep., Denver: Edu cation Commission of the States.

Confederacion de Educadores Argentinos (2009). *Historia del movimiento obrero y del sindicalismo en Argentina*. Tech. rep., Confederacion de Educadores Argentinos.

Consejo Tecnico de Inversiones (1977-2014). *Tendencias Económicas y Financieras*. Tech. rep., Consejo Técnico de Inversiones.

Costinot, A. and Vogel, J. (2010). Matching and inequality in the world economy. *Journal of Political Economy*, **118** (4), 747–786.

Cunha, F. and Heckman, J. (2007). The technology of skill formation. *American Economic Review*, **97** (2), 31–47.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, pp. 1–48.

—, Cohodes, S., Jennings, J. and Jencks, C. (2016). School accountability, postsecondary attainment, and earnings. *Review of Economics and Statistics*, **98** (5), 848–862.

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. and Kovak, B. K. (2015). Trade liberalization and the skill premium: A local labor markets approach. *American Economic Review*, **105** (5), 551–57.

Dorn, D., Katz, L. F., Patterson, C., Van Reenen, J. *et al.* (2017). Concentrating on the fall of the labor share. *American Economic Review*, **107** (5), 180–85.

Engbom, N. and Moser, C. (2017). Earnings inequality and the minimum wage: Evidence from brazil. *CESifo Working Paper Series No. 6393*.

Etchemendy, S. (2013). *Conflictividad laboral docente*. Tech. rep., Buenos Aires: Mimeo.

Ferreira, F. H., Firpo, S. P. and Messina, J. (2016). *Understanding Recent Dynamics of Earnings Inequality in Brazil*. Oxford University Press.

Ferreyra, M. M., Avitabile, C., Paz, F. H. *et al.* (2017). *At a Crossroads: Higher Education in Latin America and the Caribbean*. World Bank Publications.

Fields, G. S. (1974). The private demand for education in relation to labour market conditions in less-developed countries. *The Economic Journal*, **84** (336), 906–925.

— (1995). *Educational Expansion and Labor Markets*. Pergamon.

— (2018). The labor market effect of educational expansion in a segmented labor market model. *Mimeo*.

Finer, L. B. and Zolna, M. R. (2014). Shifts in intended and unintended pregnancies in the united states, 2001–2008. *American journal of public health*, **104** (S1), S43–S48.

Fitzpatrick, M. D., Grissmer, D. and Hastedt, S. (2011). What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment. *Economics of Education Review*, **30** (2), 269–279.

Galiani, S. and Sanguinetti, P. (2003). The impact of trade liberalization on wage inequality: evidence from argentina. *Journal of development Economics*, **72** (2), 497–513.

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

—, Jaume, D., Serio, M. and Vázquez, E. (2011b). La segregación entre escuelas públicas y privadas en argentina. reconstruyendo la evidencia. *Desarrollo Económico: Revista de Ciencias Sociales*, pp. 189–219.

— and Marchioni, M. (2015). *Bridging gender gaps? The rise and deceleration of female labor force participation*. CELDLAS-UNLP.

Goldin, C. D. and Katz, L. F. (2009). *The race between education and technology*. Harvard University Press.

Goodman, J. (2014). *Flaking out: Student absences and snow days as disruptions of instructional time*. Tech. rep., National Bureau of Economic Research.

Goos, M., Manning, A. and Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *The American Economic Review*, **104** (8), 2509–2526.

Gormley, W. T. and Gayer, T. (2005). Promoting school readiness in oklahoma an

evaluation of tulsa's pre-k program. *Journal of Human resources*, **40** (3), 533–558.

Haider, S. and Solon, G. (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review*, **96** (4), 1308–1320.

Hansen, B. (2011). *School year length and student performance: Quasi-experimental evidence*. Tech. rep., Mimeo: University of California-Santa Barbara.

Hanushek, E. A. and Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of economic literature*, **46** (3), 607–668.

— and — (2012). Do better schools lead to more growth? cognitive skills, economic outcomes, and causation. *Journal of economic growth*, **17** (4), 267–321.

Harmon, C., Oosterbeek, H. and Walker, I. (2003). The returns to education: Microeconomics. *Journal of economic surveys*, **17** (2), 115–156.

Harris, J. R. and Todaro, M. P. (1970). Migration, unemployment and development: a two-sector analysis. *The American economic review*, **60** (1), 126–142.

Heckman, J. J., Lochner, L. and Taber, C. (1998). Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of economic dynamics*, **1** (1), 1–58.

—, Lochner, L. J. and Todd, P. E. (2006). Chapter 7 earnings functions, rates of return and treatment effects: The mincer equation and beyond. In E. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education*, vol. 1, Elsevier, pp. 307 – 458.

Henry, B., Caspi, A., Moffitt, T. E., Harrington, H. and Silva, P. A. (1999). Staying in school protects boys with poor self-regulation in childhood from later crime: A longitudinal study. *International Journal of Behavioral Development*, **23** (4), 1049–1073.

Jaume, D. (2013). *Un Estudio sobre el Incremento de la segregación escolar en Argentina.* Tech. rep., Documentos de Trabajo del CEDLAS 143.

Johnson, D. R. (2011). Do strikes and work-to-rule campaigns change elementary school assessment results? *Canadian Public Policy*, **37** (4), 479–494.

Katz, L. F., Kearney, M. S. *et al.* (2006). The polarization of the us labor market. *American Economic Review*, **96** (2), 189–194.

— and Murphy, K. M. (1992). Changes in relative wages, 1963–1987: supply and demand factors. *The quarterly journal of economics*, **107** (1), 35–78.

Khanna, G. (2015). Large-scale education reform in general equilibrium: Regression discontinuity evidence from india. *unpublished paper, University of Michigan.*

Kugler, B. and Psacharopoulos, G. (1989). Earnings and education in argentina: an analysis of the 1985 buenos aires household survey. *Economics of Education Review*, **8** (4), 353–365.

Lange, F. and Topel, R. (2006). Chapter 8 the social value of education and human capital. *Handbook of the Economics of Education*, vol. 1, Elsevier, pp. 459 – 509.

Lazear, E. P., Shaw, K. L. and Stanton, C. T. (2016). *Who Gets Hired? The Importance of Finding an Open Slot.* Tech. rep., National Bureau of Economic Research.

Lee, J.-W. and Barro, R. J. (2001). Schooling quality in a cross–section of countries. *Economica*, **68** (272), 465–488.

Leuven, E., Lindahl, M., Oosterbeek, H. and Webbink, D. (2010). Expanding schooling opportunities for 4-year-olds. *Economics of Education Review*, **29** (3), 319–328.

— and Oosterbeek, H. (2011). Overeducation and mismatch in the labor market1. *Handbook of the Economics of Education*, **4**, 283.

Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The manchester school*, **22** (2), 139–191.

Lledo, V. D. (2005). *Tax systems under fiscal adjustment: a dynamic CGE analysis of the Brazilian tax reform*. 5-142, International Monetary Fund.

López-Calva, L. F., Lustig, N., Valderrama, D. *et al.* (2016). *Understanding the Dynamics of Labor Income Inequality in Latin America*. Tech. Rep. No. 1608, Tulane University, Department of Economics.

Lovenheim, M. and Willén, A. (2016). *The long-run effects of teacher collective bargaining*. Tech. rep., CESifo Working Paper Series No. 5977.

Luyten, H. (2006). An empirical assessment of the absolute effect of schooling: regression-discontinuity applied to timss-95. *Oxford Review of Education*, **32** (3), 397–429.

Maloney, W. F. and Molina, C. (2016). *Are automation and trade polarizing developing country labor markets, too?* The World Bank.

Marcotte, D. E. (2007). Schooling and test scores: A mother-natural experiment. *Economics of Education Review*, **26** (5), 629–640.

— and Hemelt, S. W. (2008). Unscheduled school closings and student performance. *Education Finance and Policy*, **3** (3), 316–338.

Meghir, C., Narita, R. and Robin, J.-M. (2015). Wages and informality in developing countries. *American Economic Review*, **105** (4), 1509–46.

Meisels, S. J. and Shonkoff, J. P. (2000). Early childhood intervention: A continuing evolution. *Handbook of early childhood intervention*, **2**, 3–31.

Mejia-Reyes, P. (1999). Classical business cycles in latin america: turning points, asymmetries and international synchronisation. *Estudios Económicos*, pp. 265–297.

Messina, J. and Silva, J. (2017). *Wage Inequality in Latin America: Understanding the Past to Prepare for the Future*. World Bank Publications.

Michele, B. and Dinand, W. (2010). Do teacher strikes harm educational attainment of students? *LABOUR*, **24** (4), 391–406.

Mirabella de Sant, M. (2002). *Diferencias de bienestar entre provincias de Argentina*. Tech. rep., XXXVII Reunión Anual de la AAEP.

Montenegro, C. E. and Patrinos, H. A. (2014). *Comparable estimates of returns to schooling around the world*. The World Bank.

Moretti, E. (2004a). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of econometrics*, **121** (1), 175–212.

— (2004b). Human capital externalities in cities. *Handbook of regional and urban economics*, **4**, 2243–2291.

Murillo, M. V. and Ronconi, L. (2004). Teachersśtrikes in argentina: Partisan alignments and public-sector labor relations. *Studies in Comparative International Development*, **39** (1), 77–98.

Narodowski, M. and Moschetti, M. (2015). The growth of private education in argentina: evidence and explanations. *Compare: A Journal of Comparative and International Education*, **45** (1), 47–69.

Neal, D. A. and Johnson, W. R. (1996). The role of premarket factors in black-white wage differences. *Journal of political Economy*, **104** (5), 869–895.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, pp. 693–709.

OECD (2011). Brazil: Encouraging lessons from a large federal system. *OECD*.

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

Pischke, J.-S. (2007). The impact of length of the school year on student performance and earnings: Evidence from the german short school years. *The Economic Journal*, **117** (523), 1216–1242.

Rauch, J. E. (1993). Productivity gains from geographic concentration of human capital: evidence from the cities. *Journal of urban economics*, **34** (3), 380–400.

Rivkin, S. G. and Schiman, J. C. (2015). Instruction time, classroom quality, and academic achievement. *The Economic Journal*, **125** (588), F425–F448.

Salardi, P. (2014). The evolution of gender and racial occupational segregation across formal and non-formal labor markets in brazil, 1987 to 2006. *Review of Income and Wealth*, pp. 1–22.

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

— (2012). Qualitative mismatches. *Foundations and Trends in Microeconomics*, **8** (1-2), 1–168.

Sims, D. P. (2008). Strategic responses to school accountability measures: It's all in the timing. *Economics of Education Review*, **27** (1), 58–68.

Teulings, C. N. (2005). Comparative advantage, relative wages, and the accumulation of human capital. *Journal of Political Economy*, **113** (2), 425–461.

Thornicroft, K. (1994). Teachers strikes and student achievement: Evidence from ohio. *Journal of Collective Negotiations*, **23** (1), 27–40.

Ulysseas, G. (2014). Firms, informality and development: Theory and evidence from brazil.

Zirkel, P. (1992). The academic effects of teacher strikes. *Journal of Collective Negotiations in the Public Sector*, **21** (2), 123–138.

Zwerling, H. L. (2008). Pennsylvania teachers strikes and academic performance. *Journal of Collective Negotiations*, **32** (2), 151–172.