IDB WORKING PAPER SERIES Nº IDB-WP-01115

Local labor markets and higher education mismatch

What is the role of public and private institutions?

Elena Arias Ortiz
Mary Kaltenberg
Cristian Jara-Figueroa
Ivan Bornacelly
Dominik Hartmann

Inter-American Development Bank
Social Sector, Education Division

April 2020

Local labor markets and higher education mismatch

What is the role of public and private institutions?

Elena Arias Ortiz
Inter-American Development Bank
Mary Kaltenberg
Brandeis University and UNU-MERIT, Maastricht University
Cristian Jara-Figueroa
MIT Media Lab, Massachusetts Institute of Technology
Ivan Bornacelly
Paris School of Economics
Dominik Hartmann
Federal University of Santa Catarina, Fraunhofer IMW, and University of Hohenheim

Inter-American Development Bank
Social Sector, Education Division

April 2020

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Local labor markets and higher education mismatch: what is the role of public and private institutions? / Elena Arias Ortiz, Mary Kaltenberg, Cristian Jara-Figueroa, Ivan Bornacelly, Dominik Hartmann.
p. cm. — (IDB Working Paper Series ; 1115)
Includes bibliographic references.
1. Labor market-Brazil. 2. Skilled labor-Supply and demand-Brazil. 3. Education, Higher-Brazil. 4. Public universities and colleges-Curricula-Brazil. 5. Private universities and colleges-Curricula-Brazil. I. Arias Ortiz, Elena. II. Kaltenberg, Mary. III. Jara-Figueroa, Cristian. IV. Bornacelly, Ivan. V. Hartmann, Dominik. VI. Inter-American Development Bank. Education Division. VII. Series.
IDB-WP-1115

http://www.iadb.org

Local labor markets and higher education mismatch: What is the role of public and private institutions?[1]

Elena Arias Ortiz
Inter-American Development Bank

Mary Kaltenberg
Brandeis University and UNU-MERIT, Maastricht University

Cristian Jara-Figueroa
MIT Media Lab, Massachusetts Institute of Technology

Ivan Bornacelly
Paris School of Economics

Dominik Hartmann
Federal University of Santa Catarina, Fraunhofer IMW, and University of Hohenheim

Abstract

We analyze mismatches between local labor markets and higher education enrollments in 22 educational fields across 137 mesoregions in Brazil between 2009 and 2016 using a two-stage regression model. Local labor markets have a significant effect on enrollment rates by field in higher education institutions and need to be considered by educational policies to prevent skill mismatches. Moreover, a large share of public universities, urban population and low unemployment rates are associated with low mismatches. Public universities address skills shortages, particularly in poor regions and natural sciences, better than private universities that contribute to oversupply, e.g. in business or law.

Key words: local labor markets, higher education, skills mismatch, public and private institutions

JEL: R12, I23, H52, J24

[1] The project would not have been materialized without the support and funding from the Inter-American Development Bank. The views expressed in this paper are those of the authors and should not be attributed to the Inter-American Development Bank.

I. Introduction

Access to advanced and specialized skills can foster innovation, productivity and economic growth (e.g. Freeman, 2010; Giuri et al., 2007; Toivanen and Väänänen, 2011). Yet, for skills to increase productivity, there needs to be a high level of coordination between the demand of skills in local labor markets and the supply of training provided by higher education. We analyze a large dataset from Brazil to reveal how effective are public and private education institutions at addressing local labor market skill shortages and mismatches.

When the coordination between skill supply and demand fails, it can have negative consequences for individuals, employers, and the economy (McGuiness et al. 2017, Somers et al. 2016). Given the importance and complexity of the match between the demand of labor markets for particular skills and the supply of skills from local education institutions, policymakers have tried a vast array of strategies, from completely public institutions to a totally privatized system, plus everything in between.

Some researchers suggest that private institutions are quicker or more efficient at addressing market demands, and have higher levels of efficiency and accountability (Wilkinson and Yussof, 2005). However, others argue that private institutions tend to offer programs with high private returns but less social returns, and that they are likely to offer professional subjects with low capital intensity (Assad et al., 2017 and Wilkinson and Yussof, 2005). Instead, public universities may deliberately consider the social returns of education when selection majors to teach, as well as address bottlenecks in relatively expensive fields, such as medicine and engineering. Indeed, courses in these critical fields are relatively scarce and to a large extent supplied by public universities, especially in developing economies.

Over the last 20 years, most countries in Latin America saw a dramatic increase in education enrollments in both public and private education institutions. Several employers' surveys in Latin America, however, show that a lack of appropriate or relevant skills is still a major barrier for development and a significant explanatory factor of low growth in labor productivity (Busso et al., 2012, IADB, 2017). Nonetheless, empirical research on different types of skills shortages across regions, as well as on the ability of either private or public education institutions to address these shortages is largely missing. Research on local differences is almost completely lacking; an essential shortcoming considering the large differences between productive structures and local labor markets in developing economies.

We contribute to this gap in literature by analyzing a large dataset on higher education enrollments in 22 educational fields across 137 mesoregions in Brazil between 2009 and 2016. We apply a two-stage regression model to identify relative skill shortages and strengths of the mesoregions and the factors that condition skill mismatches. Our results indicate that skill shortages tend to be more pronounced in less developed regions. We also find that public universities are more effective than private universities in promoting enrollments in natural sciences and addressing the skill needs of less developed regions.

Our results for the case of Brazil are relevant for several reasons. First, as many countries around the world, Brazil's employers consistently report that they cannot find the skilled workers they need (WEF, 2017). Second, as in many other countries in Latin America and the developing world, Brazil experienced a major expansion in higher education over the last 20 years: the number of

institutions almost doubled, the number of programs offered increased from 15,000 in 2002 to almost 35,000 in 2016, and the number of students enrolled in 2016 is almost 10 times the enrollment at the beginning of the 2000s. Third, similarly to other countries in Latin America, the increase in supply has been driven by the expansion of private universities, in particular, for-profit universities that are supported by the assumption that private universities are better able to match skill requirements in the private sector relative to public institutions (Sampaio, 2015). Fourth, the extent to which the private and public universities reduce or increase skills mismatches across regions in Brazil has not been previously empirically studied.

It must be noted that in the education literature, the concept of skills mismatches is often used at the individual level to refer to the degree to which workers possess the right skills or education for their current job (Cedefop, 2010; Leuven and Oosterbeek, 2011). In this paper, we do not focus on individual skill or field of education mismatch per se. Instead, we study a key factor that precedes individual skill mismatch: the match between local educational offer and local economic activities. We use information about the course offering and field enrollment in higher education as a measure of skill supply rather than qualification-at-equilibrium. Furthermore, we evaluate the heterogeneity of labor market and school provision and access across regions, while existing studies have only focused on the national level which mask idiosyncratic differences between local regional markets (Somers et al. 2016, Arias Ortiz et al., 2017). Occupational structure and skill needs differ significantly across regions within a country, especially one that is as large as Brazil (Baer, 1964; Azzoni 2001; Azzoni and Servo, 2002; Furtado, 2009; Simoes and Freitas, 2014; Freitas and Paiva, 2015, da Silva, 2017). Recent work has shown that regional differences in employment density affect the extent of skills mismatches; urban areas with a higher employment density tend to have lower levels of skills mismatches (Berlingieri, 2018). Thus, analyzing education trends and skills mismatches only at the national level may omit the substantial differences in the specific skills mismatches across regions and perpetuate structural development problems associated with less developed regions, especially as it relates to these regions' ability to generate, attract, and retain skills. This can perpetuate regional inequality and limit growth of less developed regions, as human and economic development tend to co-evolve with each other (Ranis et al., 2000; Hartmann, 2014)

The remainder of this article is structured as follows. Section 2 provides a literature review on the association between local labor markets and enrollment rates. In section 3, we describe the Brazilian system of higher education and its recent evolution, as well as the regional differences in the local labor markets. Section 4 introduces the data and methods employed in this study. Section 5 presents the results of a zero inflated regression models and discusses the relative shortages and strength of the 137 mesoregions in 22 educational fields. Section 6 discusses the results from a policy perspective and provides concluding remarks.

II. Literature review

a. The student enrollment decisions and skills mismatch

Local labor markets arguably condition students' enrollment decisions in higher education. Research has shown that individuals decide whether to invest in higher education based on

pecuniary and non-pecuniary factors, such as the cost of higher education and other alternative educational opportunities, changes in population, current wages and income levels, anticipated future earnings, as well as labor market conditions (Becker 1990; Clotfelter 1992; Leslie and Brinkman 1987, Candia et al., 2018). While the majority of research on enrollment demand focuses on how students respond to changes in the price of education, the ‘‘non-price’’ factors of enrollment, such as opportunity costs, demographic changes, and labor market conditions are less extensively studied (Betts and McFarland 1995; Kienzl, Alfonso and Melguizo 2007), in particular in Latin American countries. There is, however, a growing literature on how different aggregate labor market measures (e.g. unemployment rate, wages, etc.) influence enrollment rates (Betts and McFarland 1995, Altonji et al. 2015). For instance, it has been shown that business cycles’ influences on human capital investment includes higher education enrollment (Betts and McFarland 1995, Hershbein 2012), higher education completion (Dynarski 2008, Kahn 2010), and graduate school attendance (Bedard and Herman 2008, Johnson 2013). Additional research has illustrated the significant effect of economic conditions on the choice of specific careers, such as engineering (Freeman, 1975) and investment banking (Oyer 2008). Moreover, economic conditions, such as graduating in times of recessions have been found to explain career losses and mismatches (Liu et al., 2016). These negative effects of recessions, though, depend on the type of industries and enrollment fields; negative effects of recessions on wages and career losses tend to be less pronounced in public sectors, such as education and health, than in private sectors (Liu et al., 2016). But while there is evidence that higher education enrollments seem to respond to labor market conditions, research exploring the precise links on the local level and mismatches between occupational structures and educational enrollment are scarce (Berkes et al., 2018, Candia et al., 2018).

b. Public and private provision of higher education

The provision of education courses in a region can be expected to be associated—at least to some extent—with the local occupational structure. For instance, if a large percentage of people in a region works in agricultural activities, it can be expected that a relatively large number of courses associated with agriculture are offered. Or if many companies in a region are specialized in information and communication technologies, we expect to see many courses in computer sciences.

Which mechanism explains the coordination between the local occupational structure and the fields offered by higher education institutions? Students might prefer educational fields that are associated to local economic activities. Potentially, the local government might be interested in developing specific fields as part of their industrial policy. Additionally, local companies can promote specific educational fields either by directly investing in education or by offering higher wages or better working conditions to graduates of specific majors (Autor, 2003).

Moreover, is there a difference between public or private providers of higher education in the way they adapt their supply to the private sector’s skills demand? The most frequently used arguments in favor of greater private provision are related to its efficiency, accountability, diversity of choice, increased higher education access, as well as reduced financial burden for the government (Patrinos, 1990, World Bank, 1994). Another advantage that researchers and policymakers often

highlight is that private institutions may respond faster and more efficiently to market demands (Wilkinson, 2005). As private institutions are financed mainly by tuition fees, they have more incentives to respond to labor markets signals and teach the appropriate skills that attract good students and build a good reputation (Assaad et al. 2017). In fact, government regulations in Brazil were modified in recent years to allow for private providers of higher education, expecting to improve the coordination between local labor markets and the education sector. So far, there is little empirical evidence that these types of policies are effective.

c. Higher education in Brazil

Brazil is home to the largest higher education system in Latin America with 2,368 institutions (87% private) and over 7.8 million enrolled students (Robles and Bhandari, 2017). In the last ten years, Brazil had the largest increase in the number of higher education institutions and programs in Latin America, and higher education enrollment more than doubled (World Bank 2017). Returns to higher education in Brazil are very high by international standards - the skill premium (defined as the ratio of the average wages for higher education graduates versus high school graduates) is 2.77 in Brazil, as compared to 1.67 in the United States. This large skill premium can be partially attributed by the fact that only 11% of the working-age population have a higher education degree in contrast to 42% in the United States. Given the high skill premium, higher education enrollment is likely to continue to grow (Ferreyra et al., 2017).

Several factors account for Brazil's growth in higher education enrollment, related to both demand and supply dynamics. On the demand side, the reduction of inequality, the increase of income, and the rise of the middle class have made higher education more accessible and equitable (Ferreira et al, 2006; Torche and Ribeiro, 2010; Souza and Lamounier, 2010; Costa Ribeiro, 2012). Moreover, there has been an increase in the completion rate of secondary school (an increase of 20% over the last decade (IADB, 2018)), and a higher demand of high skilled workers (Andrade, 2012).

On the supply side, the expansion in access to higher education has been driven by two trends - a huge expansion of the private education institutions, and the increase of educational supply outside of state capitals. Public universities accept only a limited number of students, and entry is determined by highly competitive exams. Students who do not have high-quality secondary education are rarely successful, unless they attend expensive preparatory programs (McCowan, 2007). To fill this gap, demand has been partially met by private universities. In Brazil, private universities do not need to adhere to a fully not-for-profit model. In fact, the number of programs offered by private universities has increased mainly due to for-profit universities. As Sampaio (2015) documented, this privatization of higher education is associated with a change of Brazilian legislation in 1997 that recognized the education service performed by for-profit institutions to meet the unsatisfied demand.

Tertiary education institutions in Brazil have some institutional autonomy to decide which degree programs to offer and which segment of the population to cater to (OECD, 2015). Thus, regional characteristics such as the population size, GDP per capita, percent of urban population, the occupational structure or being the home of the state capital, can influence the type of educational fields that are supplied and demanded in a region. McGuiness et al. (2017) showed that

enrollment in Brazil tends to be concentrated in degree programs such as business and social sciences because they are much cheaper to be organized or expanded. Moreover, the recent expansion was characterized by the creation of new smaller private institutions, typically in big cities. The limited money invested in these new education institutions or programs are probably not enough to establish more expensive programs in natural sciences.

Despite this quantitative growth in the supply of higher education, evidence suggests that higher education programs in Brazil lack the flexibility to adapt to the needs of the labor market as higher education institutions have a weak connection with the productive sector. In general, technological research and development is carried out in universities with little connection with the productive sector (Japan Bank for International Cooperation, 2015). As a consequence, enrollment and graduation shows a strong concentration in areas such as business, law, and education.

III. Methodology and Data

a. Data

We use two datasets: 1) Brazilian administrative data from the Annual Report of Social Information (RAIS); and 2) the Brazilian Higher Education Census (INEP), collected by the National Institute of Studies and Educational Research (INEP). RAIS collects information from all tax registered Brazilian firms (e.g. productive sector and location of firms) and every worker (level of education, salary, occupation) in each of these firms. INEP collects data from higher education institutions (HEIs) that offer undergraduate programs and specific sequence training and contains annual data on program enrollments by field and institution. We combine these two data sets to understand the mismatch between the supply of educational programs and labor markets at the regional level between 2010 - 2016.

We use information about 137 Brazilian mesoregions. Brazilian mesoregions are groupings of 558 microregions, which are groupings of 3500 municipalities. We chose mesoregions as our spatial level of analysis because mesoregions capture clustered labor market pools and spatial interactions substantially better than federal states (which often comprise multiple different labor markets), but are still large enough to justify the establishment of universities offering programs in most of the 22 higher education fields. Similar to metropolitan statistical units in the US, mesoregions are designed to capture clustered labor market pools (https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=22269).

RAIS reports employment information using the Brazilian Occupation Classification (CBO), but does not provide information on the links between occupations and educational fields. This is necessary, though, to estimate the mismatches between the occupational structure and the local labor markets. Therefore, we use the O*NET classification scheme from the US to create a crosswalk with the Brazilian data. The Occupational Information Network (O*NET) is a database has been widely used in the literature about the skill content of occupations (e.g. Aedo et al., 2013). O*NET characterizes different occupations in the US economy in terms of worker-oriented factors and job-oriented factors.

We create a crosswalk by linking CBO occupations, the International Standard Classification of Occupations of ILO (ISCO), and the Standard Occupational Classification from the US. Then we

match CBO occupations to the educational fields that are required or advisable to perform each occupation according to (O*NET) classification. For this purpose, we create a crosswalk between education fields in the Brazilian census of higher education institutions (CES), the International Standard Classification of Education (ISCED), and the Classification of Instructional Programs of the US (see Figure X for the crosswalk structure in the online appendix).

Of course, this approach has its limitations, as the skill requirements of an occupation in the USA and Brazil may differ. Due to the lack of data for several countries, this approach has been applied in previous work for a variety of countries (Aedo et al (2013). In addition, evidence suggests that these types of crosswalks are an acceptable proxy. Ospino (2018) shows that education and occupation correspondence of the classification of European Skills, Competences, Qualifications and Occupations (ESCO) is strongly correlated with the same crosswalk from O*NET in terms of level of education and field of studies. Moreover, skill content of occupations in Chile, Colombia, and Mexico are likely to converge towards the ones defined by the US (Dicarlo et al, 2016; Messina, Oviedo and Pica, 2016).

b. Regression model

To what extent do education programs meet the local labor market demand? To answer this question, we model the relationship between education programs and local labor markets by looking at new enrollments (supply of skills) and hired workforce of regions (demand of skills) at the regional level with a two-stage model.

In the first stage, we estimate determinants of enrollment using information about education (total number of students enrolled, applicants, seats offered, and number of programs or careers offered) and labor market characteristics (regional occupational demand, average wages and education requirements for an occupation, migration related to that occupation). Deviations from the trend can be interpreted as mismatches. The second stage uses the unexplained part (i.e. residual) of our first stage regression as the dependent variable. The residual represents the variance that is unexplained by education and labor market characteristics. The second stage regression then estimates what role regional characteristics plays on enrollment, such as unemployment rates, wealth, urban concentration, regional population, private/public education and occupation specialization. This two-stage procedure allows us to estimate the relative strength of a field of education in a region in the first stage, and then analyze if there are general trends at the regional level that explain enrollment in the second stage. This model allows us to reveal general relationships between our variables and enrollment outcomes, however it does not provide a detailed understanding of causality. We cannot identify if the presence of private for-profit universities causes a mismatch, but we can understand if there is a general relationship between the two variables.

The first stage regression predicts new enrollment for a field in a region that can be categorized into two influencing factors: educational supply and labor market conditions. The variables consider students' decision-making, university budgetary and resource constraints, and demand of a field. These variables predict enrollment using a count model, a negative binomial regression, for each field to understand each region's relative educational shortages and strengths.

Our first stage negative binomial count model is as follows:

$$E(Enr_{fry}|x_i,\alpha) = \frac{\rho(Enr_{fry}+\alpha^{-1})}{\rho(Enr_{fry}+1)\rho(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1}+x_i}\right)^{\alpha^{-1}}\left(\frac{x_i}{\alpha^{-1}+x_i}\right)^{Enr_{fry}}$$

Where $x_i = \beta_1 \ln(\sum_{o\epsilon f} F_{ory}) + \beta_2 \ln(\sum_{o\in f} L_{ory}) \quad + \beta_7 \ln Course_{fry} + \beta_8 \ln App_{fry} + \beta_9 \ln Seats_{fry} + \ \beta_{10} \ln Pop_{ry} + \delta_y + \delta_r + \varepsilon_{fry}$

Our dependent variable, Enr_{fry}, is the number of people enrolled in that field, f, in region, r at year, y. Education variables include total enrollment ($\ln E_{fry}$, size effect, natural log), number of applications ($lnApp$, proxy for the demand of the program in natural log), number of programs offered ($\ln Course_{fry}$ proxy of the variety of subject offered in natural log), and number of seats offered ($\ln Seats$, proxy for the supply of that field in natural log). Each of these variables is related to a field, f, in region, r, for year, y.

Labor market variables include, the average wages of occupation o related to field f in region r, for the previous year($\sum_{o\epsilon f}\frac{1}{N}W_{ory-1}$), and the average education level of occupation o related to field f in region r ($\sum_{o\epsilon f}\frac{1}{N}Edu_{ory}$). The choice to continue one's education has an opportunity cost, and thus, we include the average number of years of schooling by occupation. When applying for programs, individuals evaluate the current labor market conditions, including their potential earnings related to that field, which is why we included the lagged average wage for occupations related to a field. Furthermore, individuals may want to enter fields where the demand of a particular occupation is high in terms of its absolute size. Thus, we include a variable that sums all of the associated fields that are related to occupation, o, to calculate the total number of people employed in that region for that occupation $\ln(\sum_{o\in f} L_{ory})$. A nice feature of this variable is that the coefficients for the relevant occupation for an education field will tell us how important that occupation is for that particular field as there are many occupations related to one particular field of study.

Students may choose to study in a particular region that is specialized in a certain field, but after they graduate, that region may not have enough demand to employ all of the new graduates, and thus, students may migrate to where there are employment opportunities for that field. To proxy this effect, we include a migration variable which is defined as the total number of people leaving that region, r, in occupation, o, summed across the associated fields, f, with occupation, o, $ln(\sum_{o\in f} F_{ory})$. We also include a measure of relative regional occupation specialization, which is the number of people employed in an occupation related to field f in region r divided by the national employment in an occupation related to that field ($\frac{\sum_{o\in f} L_{fry})}{\sum_{o\in f} L_{ry})}$). Last, we include the population of a region who are above 18, and fixed effects by region and year.

We retain the residuals (ε_{fry}) from the first stage regression and use them as the dependent variable in the second stage to identify the relative shortages and strengths of regions in certain fields – this is the variable that determines skill mismatches for each field. The absolute value of the residual from this first regression is the general matching ability of a region's educational system to meet the local labor market demand, with 0 representing a perfect match. For ease of interpretation, we always use the absolute value, but run separate regressions for the negative and positive residuals, as well.

The second stage regression focuses on how well the educational system within regions is able to match local labor market demand, and is as follows:

$$E(\left|\widehat{\varepsilon_{fry}}\right|) = \gamma_1 \text{lnPop}_{ry} + \gamma_2 \ln GDPpc_{ry} + \gamma_3 Urban_{ry} + \gamma_4 Unemp_{ry} + \gamma_6 Capital_r + \gamma_7 \frac{\sum_{o\in f} L_{ry}}{\sum_{o\in f} L_y}$$
$$+ \gamma_8 Edu_{ry} + \gamma_9 Public_{ry} + \gamma_{10} Profit_{ry} + \gamma_8 Edudiff_{ry} + \delta_y + \delta_r + \tau_{fry}$$

Our dependent variable is the residual enrollment for region r in field f in year y. Highly populated regions with higher rates of urbanization as well as wealthier regions are likely to have more enrollment and programs, therefore we include the regional GDP per capita ($\ln GDPpc_{ry}$), the share of urban population ($\gamma_3 Urban_{ry}$), and whether that meso-region is the state capital (binary, $\gamma_6 Capital_r$). Regions may have more field specialized programs if there is a high share of people employed in an occupation in that region as compared to the rest of the country, which motivates the inclusion of the regional share of people employed in an occupation ($\gamma_7 \left(\frac{\sum_{o\in f} L_{ry}}{\sum_{o\in f} L_y}\right)$).

There are two variables of considerable importance for our analysis, public vs. private university status and regional skill mismatch. We include the percent of seats that are in a public university ($\gamma_9 Public_{ry}$) or for-profit university ($\gamma_{10} Profit_{ry}$) to analyze whether regions with a high share of private providers experience fewer distortions in enrollment by field relative to the share of occupations in the local labor market. Note that universities can be classified as a non-profit private, for-profit private, or public university. Therefore, the percent of seats in a public university plus the percent of seats in a for-profit university, plus the percent of seats in a non-profit university adds up to one. This is why in our analysis we only estimate whether for-profit private universities better match labor market demands compared to public universities. As such, an increase in non-profit private seats will be reflected as a decrease in the two variables we have added in our analysis. We measure regional skill mismatch with the difference between the required education that is needed to perform an occupation (according to the US data base for occupations O*NET) and the actual level of workers in those occupations ($\gamma_8 Edudiff_{ry}$). We average these differences by sector for each region. In addition, we include other regional controls: the population between 18-29 in the region, the average level of education in the region, and the regional unemployment rate. We also include fixed effects by year and region.

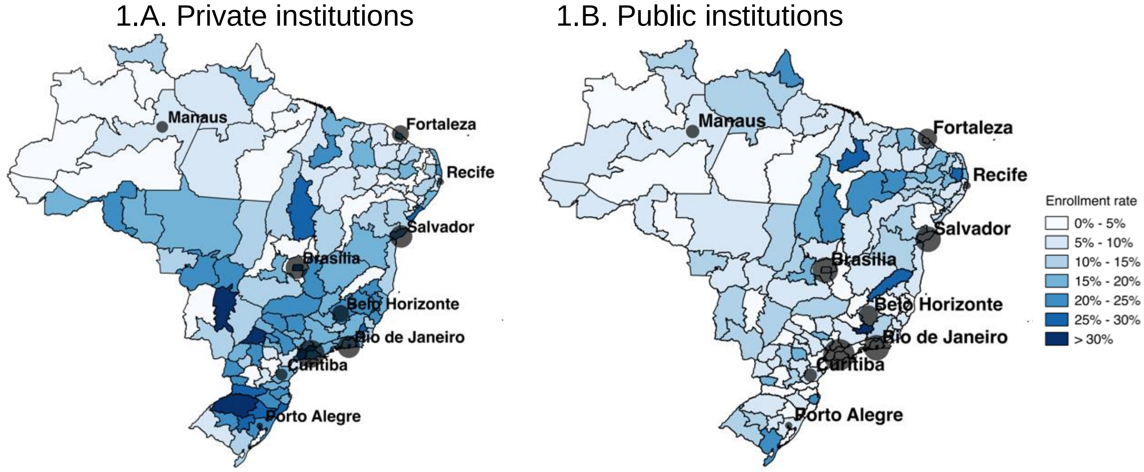
IV. Results

a. Educational programs and enrollment rates

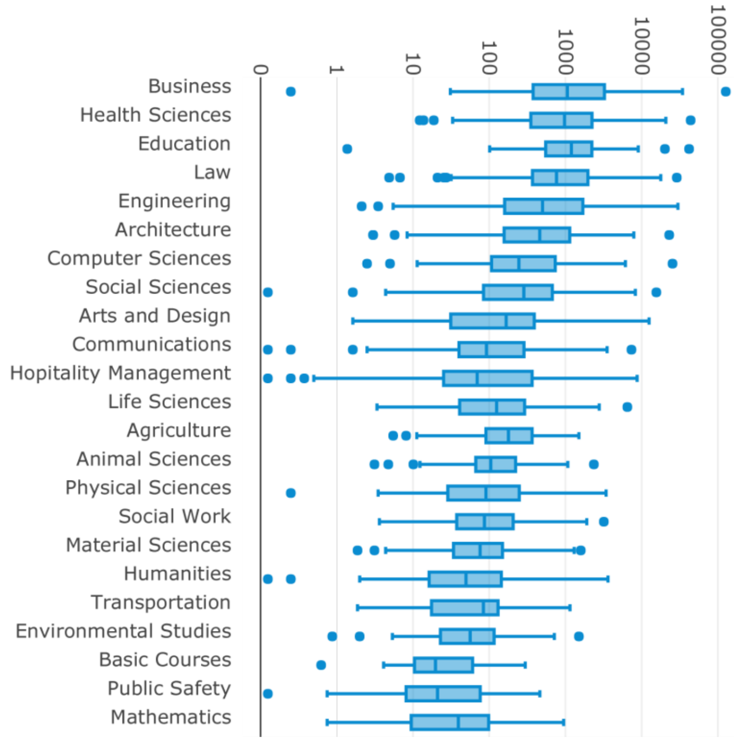
Since the early 2000s, the private provision of higher education rapidly increased. Between 2002 and 2016 about 700 new institutions and 13,000 new degree programs were created, mainly driven by new for-profit and private non-profit institutions (See Table A1 and A2 in the appendix). The growth in Brazilian institutions for higher education has been attributed to a new "dual model" in which private expansion focuses on the provision of higher education course, arguably responding better to local labor market needs, and where public universities focus on research (De Oliveira Barbosa, 2014). It is noteworthy that Sampaio (2000) documented that before 1990, private sector institutions were established in richer and more urbanized regions of the country, while the public sector universities, at the federal and state level, were in charge of the higher education provision in the less-developed regions more distant from the large industrial centers. Figure 1-A and 1-B shows significant differences in the enrollment rates in public and private institutions across mesoregions in Brazil in 2012. Figure A1 in the appendix illustrates that the growth of degree programs from private institutions has been concentrated in some (mainly urban) mesoregions in southeast and southern Brazil, the variation in new degree programs of public university degrees is more evenly spread. Figure 1-C shows boxplots about distribution of enrollments in 22 educational fields across the mesoregion. Business, health sciences, education, and law are the top fields by enrollment, arguably because there is a large demand of these fields and also because they require less capital to offer.

Figure 1. Enrollments rate across mesoregions in Brazil

1.A. Private institutions 1.B. Public institutions



1.C. Total enrollments in different fields across mesoregions.



b. Drivers of enrollment rates across mesoregions

Next, we analyze the effects of total educational supply and demand, local labor market conditions, and other regional characteristics on differences in enrollments. Table 1 shows the regression result for the nine largest educational fields. To simplify our results, we focus on a subsection of educational fields and provide detail estimates in the appendix. As expected, the total number of enrollments, applications, and available seats offered in a region have a positive and significant effect on the total enrollment rates by field across regions. The differences in the size of the effect of total enrollments on each educational field, i.e. the regression coefficients, are relatively small.

We find significant effects of local labor market conditions on enrollment across fields. However, these effects are highly heterogeneous and heavily depend on the type of educational field and occupation. Our regression takes these heterogeneous effects into account. The average wage of an occupation related to higher field in a region has a strong positive and significant effect on differences in the enrollment rates across regions in the educational fields of mathematics, computer sciences, agriculture, engineering, and physical sciences. The average education level of an occupation related to field in a region has a small positive and significant effect on enrollment in health sciences, education, law, life sciences, and transportation. Yet, average wage and average education levels do not have a significant effect on differences across enrollment rates in the other educational fields.

These results indicate that the occupational structure, i.e. type and number of occupations present in a region, does affect enrollment rates in different educational fields across regions. Yet, as

expected, each educational enrollment is associated with different types of occupations. While the presence of some occupations, such as customer service or other service-related technicians, in a region has a strong, positive, and significant effect on enrollments in business, other occupations, such as a high number of workers in mining and quarrying is associated with relatively lower numbers of enrollments in business, but rather higher enrollments in transportation and material sciences.

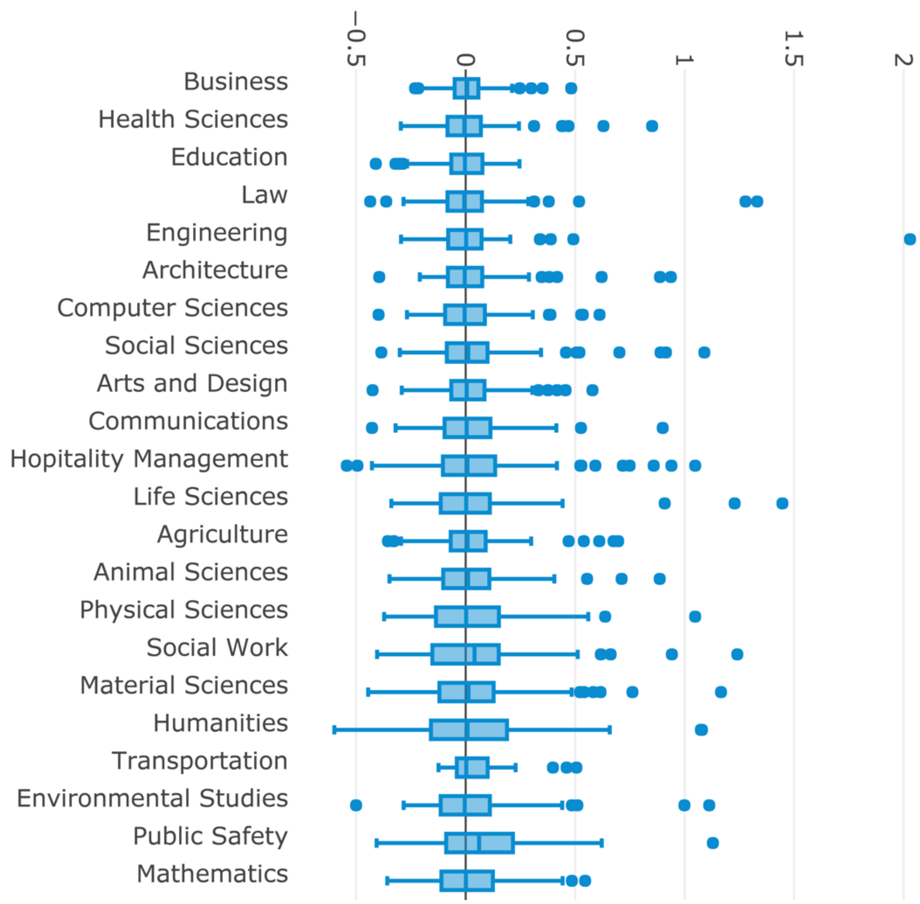
Table 1. The effects of total educational supply and demand, local labor market conditions, and other regional characteristics on differences in enrollment rates across regions

	Business	Health Sciences	Education	Law	Engineering	Architecture	Social Sciences	Computer Sciences	Agriculture
N. Courses (ln)	0.071***	0.075*	-0.096***	-0.015	0.070**	-0.028	0.045	0.074**	0.057**
	(0.024)	(0.038)	(0.027)	(0.030)	(0.029)	(0.039)	(0.034)	(0.032)	-(0.027)
N Apps. (ln)	0.091***	0.040***	0.089***	0.052***	0.075***	0.021	0.100***	0.093***	0.043***
	(0.013)	(0.014)	(0.014)	(0.012)	(0.011)	(0.014)	(0.013)	(0.011)	(0.010)
N. Seats (ln)	0.180***	0.249***	0.240***	0.281***	0.296***	0.547***	0.294***	0.210***	0.387***
	(0.021)	(0.027)	(0.021)	(0.022)	(0.023)	(0.024)	(0.025)	(0.021)	(0.022)
Pop <18 yrs.	0.056	-0.017	-0.023	0.024	0.046	0.002	0.021	-0.122***	0.064
	(0.053)	(0.052)	(0.047)	(0.029)	(0.045)	(0.051)	(0.047)	(0.036)	(0.046)
N Enroll. (ln)	0.581***	0.578***	0.694***	0.641***	0.492***	0.360***	0.442***	0.584***	0.475***
	(0.027)	(0.031)	(0.029)	(0.026)	(0.024)	(0.024)	(0.026)	(0.027)	(0.021)
Av. Edu.	0.072	0.153**	0.075***	0.039**	-0.051	0.044	-0.101	0.037	0.019
	(0.045)	(0.059)	(0.027)	(0.020)	(0.050)	(0.037)	(0.109)	(0.036)	(0.022)
Av. Wage (t-1)	0.038	0.165**	-0.027	0.03	-0.083	-0.097	0.06	-0.017	-0.078*
	(0.065)	(0.069)	(0.055)	(0.023)	(0.069)	(0.087)	(0.075)	(0.039)	(0.044)
Occ. Spec.	0.396	0.688	2.793***	0.065	1.351**	1.545*	0.434	1.415***	0.618
	(0.510)	(0.686)	(0.981)	(0.347)	(0.625)	(0.804)	(0.464)	(0.339)	(1.338)
Constant	-2.168***	-2.894***	-0.853	-1.412***	0.576	-0.622	-1.056	-0.185	-0.68
	(0.623)	(0.785)	(0.569)	(0.325)	(0.623)	(0.733)	(1.270)	(0.454)	(0.447)
Observations	907	873	949	850	765	722	764	861	833
Log Likelihood	-6,233.13	-5,952.60	-6,658.89	-5,592.68	-4,766.79	-4,425.88	-4,457.33	-4,872.64	-4,343.32
	27.932***	18.415***	19.330***	26.245***	26.584***	23.959***	19.728***	21.619***	24.382***
theta	(1.427)	(0.941)	(0.920)	(1.400)	(1.547)	(1.397)	(1.149)	(1.204)	(1.418)
Akaike Inf. Crit.	12,558.27	11,977.20	13,385.77	11,225.35	9,645.57	8,943.75	8,978.67	9,809.27	8,774.64

c. The relative mismatches between local labor markets and higher education enrollments

The predicted values from the first stage regression allow us to estimate the relative strengths and shortages of student enrollment for each region and educational fields. Figure 4 shows the distribution of the residuals. More ubiquitous fields, such as business, law or education tend to have a relative lower spread of residuals, then less ubiquitous fields, such as life sciences, humanities, or mathematics.

Figure 2. Boxplots on the distribution of shortages and relative strengths of the 137 mesoregions in the 22 education fields.



However, there is great variety of relative shortages and mismatches across regions. Several relatively poor regions, such as the Norte Maranhense, Centro-Sul Baiano, and Mesorregião Ocidental do Tocantins have a large number of educational fields with enrollment shortages, while most regions in the South, South-East, and Centro-Oeste of Brazil have lower number of shortages (see Figure 3-A).

Our analysis also considers the local needs for different educational fields by the local labor market. Naturally, regions in the tropical rainforests in Northern Brazil require different types and

quantities of educational fields to match their needs than semi-arid regions in the Northeast or the more industrialized areas in the South and South-East. The residuals from the first stage regressions help identify specific shortages and relative strengths. Figure 3-B to 3-E illustrates that particular Northeastern regions have relative enrollment shortages in several types of education fields, such as business, health sciences, education, and law. It is noteworthy that there are significant differences across mesoregions within federal states.

Figure 3-A. Number of under-represented fields for each region.

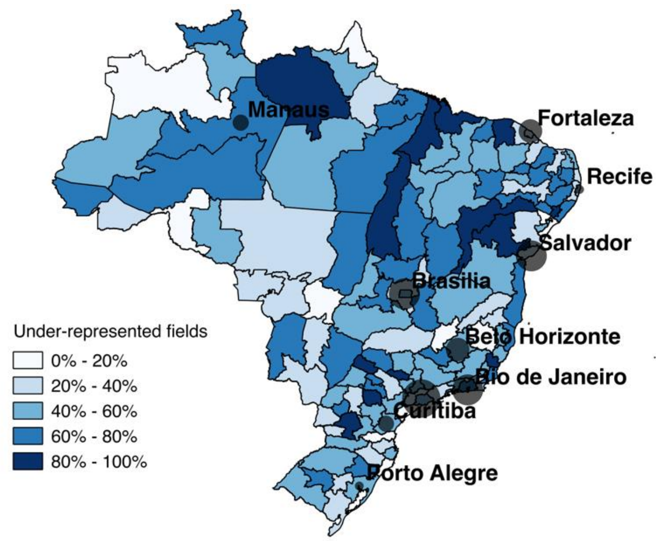
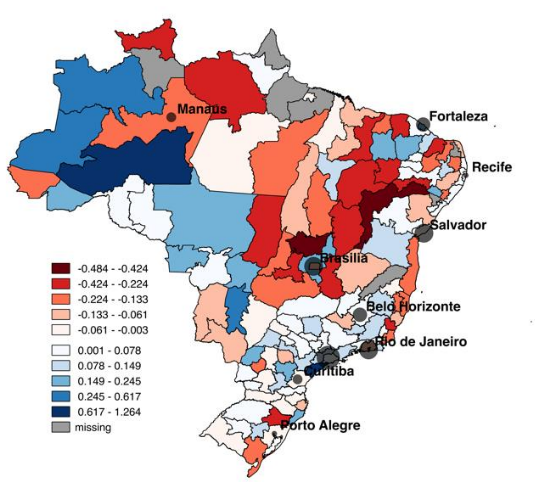
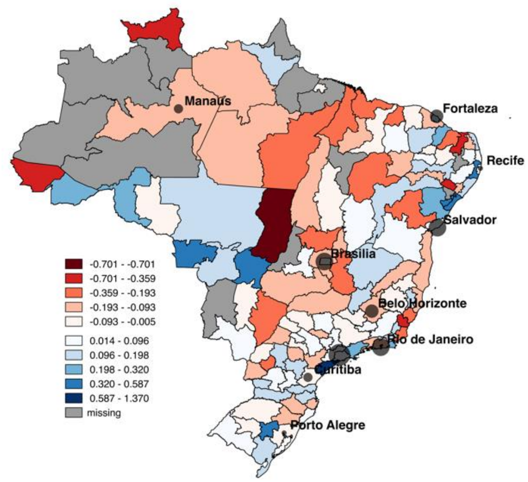


Figure 3.B-E Relative shortages and strengths in the fields A. Business, B. Health Sciences, C. Education, and D. Law. Blue color indicates relative strengths, red colors (relative) shortages.

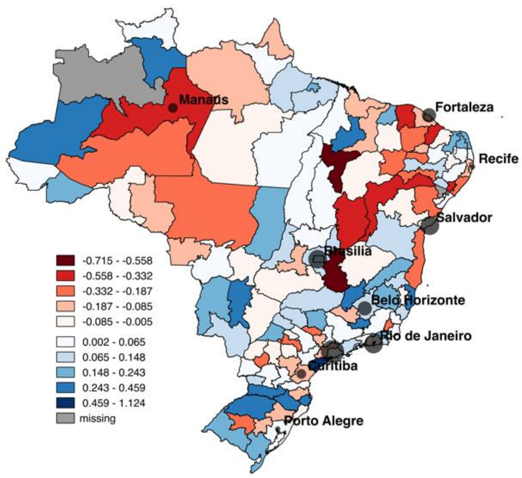
B. Business



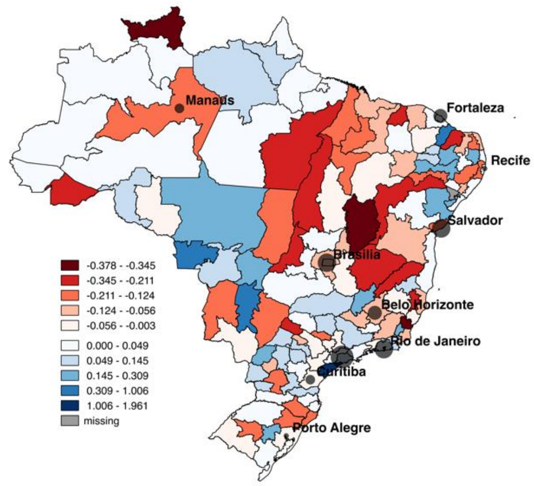
C. Health



D. Education



E. Law



d. Factors explaining the mismatches

In the second stage regression, we analyze factors that explain mismatches, which are proxied by the absolute value of the regression residual—i.e. the unexplained part of the student enrollments—from the first stage regression. A value of zero indicates a perfect match between education supply and labor market demand, and larger residuals correspond to a larger mismatch. As explanatory factors we scrutinize the role of public and private universities as well as a set of regional characteristics, such as the urban population, the average level of education, and the regional unemployment rate.

We estimate six models: (1) ordinary least squares (OLS) and (2) random effect models (RE) for all mismatches as well as separate regressions for over- and underrepresented enrollments (models 3-6). We chose a random effect model in addition to an OLS model to capture the percentage of the variance coming from between regions as compared to within regions. Moreover, we applied a Hausman test to ensure that a random effects model can be used over a fixed effects model.

Table 2. Second Stage Regression Results: Regional Characteristics on Estimated Mismatch

	(1) OLS	(2) RE	(3) OLS Pos	(4) RE Pos	(5) OLS Neg	(6) RE Neg
lnPop 18-29	-0.0424***	-0.0458***	-0.0585***	-0.0651***	-0.0280***	-0.0257***
	(0.005)	(0.006)	(0.008)	(0.010)	(0.005)	(0.005)
lnGDPpc	-0.00771	-0.0119	-0.0207	-0.0364*	0.00589	0.00670
	(0.009)	(0.010)	(0.013)	(0.019)	(0.010)	(0.011)
Edu	0.0912***	0.0880***	0.117**	0.0847	0.0698**	0.0751***
	(0.033)	(0.034)	(0.052)	(0.065)	(0.029)	(0.027)
Req. Edu Diff	-0.115***	-0.122***	-0.143**	-0.130*	-0.0892***	-0.0968***
	(0.035)	(0.038)	(0.055)	(0.071)	(0.032)	(0.030)
Public %	-0.0435*	-0.0322	-0.0512	-0.0737*	-0.0428*	-0.0259
	(0.022)	(0.022)	(0.035)	(0.041)	(0.023)	(0.023)
Profit %	0.0197	0.0120	0.0109	-0.0153	0.0285	0.0383**
	(0.020)	(0.019)	(0.030)	(0.033)	(0.018)	(0.017)
Urban Pop %	-0.119***	-0.109**	-0.104*	-0.0762	-0.144***	-0.133***
	(0.039)	(0.047)	(0.062)	(0.087)	(0.045)	(0.044)
Unemp. Rate	0.00316**	0.00328**	0.00320	0.00376	0.00330**	0.00286*
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
Capital City	-0.0149	-0.0154	-0.0145	-0.0123	-0.0175*	-0.0185*
	(0.010)	(0.011)	(0.015)	(0.017)	(0.010)	(0.010)
Occ Spec.	0.166	0.241	0.405**	0.525*	-0.0939	-0.126
	(0.134)	(0.164)	(0.194)	(0.268)	(0.129)	(0.125)
Constant	0.275	0.349	0.352	0.712	0.199	0.115
	(0.251)	(0.258)	(0.397)	(0.514)	(0.227)	(0.214)
N	13506	13506	6320	6320	7186	7186
Within R2		0.0246		0.0424		0.0204
Between R2		0.196		0.117		0.212
Overall R2		0.0991		0.0995		0.124
Adj. R2	0.0976		0.0967		0.120	
RMSE	0.183	0.164	0.223	0.183	0.134	0.120
N Regions	137	137	137	137	137	137

Note - Standard errors in parentheses, Robust-clustered standard errors * p < 0.10, ** p < 0.05, *** p < 0.01

Our results provide several takeaways. First, population matters: the percentage of population in university age (18-29 years old), and the percentage of urban population are negative and significantly associated with skills mismatches. A one percentage point increase in the young adult population is associated with a 4.5% decrease in the skills mismatch. Urban regions also have significant, but lower negative effect, with a 1 precentage point increase being associated with a 1.09% decrease in mismatch. Second, regions that have higher rates of unemployment are not as efficient at matching local labor market demand – a 1 percentage point increase in unemployment rate is associated with an increase in mismatch by .3%. Third, public universities tend to be better in matching labor market demands than for-profit private universities. A 1% point increase in the percent of seats offered to public schools is associated with a 3.2% decrease in mismatch. The statistical significance disappears, though, when using a Random Effect model because the magnitude of effect decreases.

Fourth, surprisingly, the wealth of a region does not impact enrollment rates. This is likely due to the fact that the supply side of education is already taken into account in the first stage regression. Thus, the wealth of a region does not seem to explain enrollment rates beyond the existing supply of education within a region. Furthermore, the average educational attainment by region is included in this regression, and education and individual wealth are highly correlated, this can further explain why the wealth of a region may not be significant here. Fifth, the required education difference has a large and negative effect. A difference in one level of education is associated with a 12.2% decrease in mismatch.

To better understand these effects, we run separate regressions for negative or positive residuals, i.e. for under- and overrepresented enrollments (models 3-6). Regions with higher population and urban density tend to be better at matching labor market needs. However, regions with a capital city are better at reducing underestimated mismatch, but do not have a significant effect on reducing overestimated mismatch. Regions that face higher unemployment rates tend to have more underestimated mismatch. This suggests that regions with high unemployment are unable to efficiently match local demand of jobs. Regions with a high occupation specialization tend to overestimate mismatching – a 10% rise in specialization leads to a 4% increase in overestimated mismatch. It seems that regions may recognize their comparative advantage, and as a consequence invest more into programs where they know they regionally specialize. Investigating the positive and negative effects of this specialization is a promising question for future research, yet beyond the scope of this paper.

Moreover, several variables lead to underestimating labor market needs. The required education difference measures the difference between the years of schooling needed to do a particular job compared to the actual years of schooling that job holders have obtained. A one-year increase in required education difference is associated with a 9.6% decrease in underestimated labor mismatch and with a 13% decrease in overestimated labor mismatch. In contrast, an increase in education is significantly associated to higher mismatch across our regressions. While this result may seem counterintuitive, it suggests that even if the average educational level increases this will not directly result in a reduction in labor market mismatches. This is because the private sector is unable to efficiently match the needs of the labor market needs, and therefore, increases in any type of education doesn't necessarily translate to effective matching.

Regarding the role of public and private school enrollment, it seems that public universities tend to reduce mismatch, while private education institutes tend to increase it. It must be noted, though, that the results are not statistically robust across all regression models. A 10% increase in public university seats offered leads to a 0.73% decrease in underestimated labor mismatch (statistically significant), while a 10% increase in private university seats offered leads to a 0.15% increase in underestimated labor mismatch (not statistically significant). In regard to overestimated labor mismatch, a 10% increase in public university seats is associated with a decrease in .25% (not statistically significant), while a 10% increase in private university seats is associated with an increase in .38% mismatch (statistically significant). In other words, public universities are better able to match the labor market needs, while private schools have an opposite effect and may even contribute to regional skills mismatches.

V. Discussion and conclusions

We used a two-stage regression model in this article to analyze whether local labor market conditions are significant predictors of enrollment rates and to identify factors that drive relative mismatches.

The results of the first stage regressions show that local labor markets significantly affect the differences in enrollment rates across regions and fields. Using the residuals from the first stage regression, we could identify trends that contribute to relative mismatches. Considering the relatively low total number of university graduates in Brazil, probably most mesoregions could gain from higher total enrollments in most fields. Yet the most urgent bottlenecks can widely differ across regions. While one region may suffer from enrollment shortages in architecture and material sciences, another region from a lack of enrollments in health sciences.

The results of the second stage regressions illustrate that a high level of urban population, low unemployment, and a high share of public universities are associated with significantly lower mismatches. Instead, private universities do not contribute to a better matching and there is no empirical evidence that they are more efficient than the public sector. This is arguably the case, because private institutions tend to move into regions and educational fields that are already well supplied, such as business or law studies. In contrast, public universities in Brazil are deliberately designed to address particular skills demands and shortages of the local economy in more expensive educational fields, such as medicine or engineering, or in less developed regions.

Several limitations must be considered. First, we estimate relative and not absolute shortages and mismatches. These relative shortages are highly likely to be associated with absolute shortages, yet applied policies will require additional research, i.e. case studies on the precise absolute needs of each mesoregion. Secondly, our estimates are not causal, but rather associations and provide general relationship between these variables. Third, some fields might require further disaggregation. For instance, strengths in the broad category health may hide a shortage in doctors and relative oversupply of nurses or other health related studies.

Nonetheless, our study provides important insights on the effects of local labor market conditions on educational enrollment rates across regions; and the presented methods can also be applied to other countries. Identification of skills mismatches on the national level may ignore substantial

differences on the local level. In this regard, the recent increase in private universities in Brazil, facilitated by new regulations, may not have achieved its original goal of ensuring a more efficient matching of the skills demand with the local skills needs. In contrast, in some poor region the relative oversupply of some courses, such as business studies, may have even increased, while relative shortages are still not addressed. Subsequent research may need to consider that private universities may cater the local demand in terms of students' enrollment decisions (e.g. in Brazil for law and business), but not necessarily the local economies main skills shortages (e.g. in engineering, medicine, or informatics).

In sum, our results imply that educational policies need to take the differences in local labor markets demand into account, or they may perpetuate regional mismatches. Finally, it cannot simply be assumed that private institutions perform better than public universities in addressing local needs. Arguably, incentives should be given to both public and private institutions to deliver concrete results in terms of alignment with each region's most significant skills shortages.

References

Almeida, R., Corseuil, C. and Poole, J. (2017). "The Impact of Digital Technologies on Routine Tasks: Do Labor Policies Matter?", *World Bank Policy Research Working Paper 8187*.

Altonji, J., Arcidiacono, P., and Maurel, A. (2015), Chapter 7 – "The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects", Editor(s): Eric A. Hanushek, Stephen Machin, Ludger Woessmann, *Handbook of the Economics of Education*, Elsevier, Volume 5, 2016, Pages 305-396.

Andrade, C. (2012). *"Acesso ao ensino superior no Brasil: equidade e desigualdade social. A evolução¸o dos últimos 15 anos."* Ensino Superior. Campinas: Unicamp, Agosto.

Arias Ortiz, Bornacelly and Elacqua (2017), "The Evolution of tasks, skills and occupations in Brazil Between 2003 and 2017", *mimeo*, Inter-American Development Bank.

Assaad, R. C. Krafft and D. Salehi-Isfahani (2018), "Does the type of higher education affect labor market outcomes? Evidence from Egypt and Jordan", *Higher Education*, 75:945-995.

Autor, D., F. Levy, and R. Murnane (2003), "The Skill content of recent Technological Change: An Empirical Exploration", The Quarterly Journal of Economics, 118 (4): 1279-1333.

Azzoni, C. R. (2001). Economic growth and regional income inequality in Brazil. The annals of regional science, 35(1), 133-152.

Azzoni, C. R., & Servo, L. M. (2002). Education, cost of living and regional wage inequality in Brazil. Papers in regional science, 81(2), 157-175.

Baer, W. (1964). Regional inequality and economic growth in Brazil. Economic Development and Cultural Change, 12(3), 268-285.

Becker, G., Murphy K. and Tamura, R. (1990), "Human Capital, Fertility, and Economic Growth", *Journal of Political Economy*, 98 (3).

Berkes, E., Mohnen, P. and Taska, B. (2018). The Consequences of Initial Skill Mismatch for College Graduates: Evidence from Online Job Postings

Berlingieri, F. (2017). Local labor market size and qualification mismatch. ZEW-Centre for European Economic Research Discussion Paper, (17-055).

Betts J. and L. McFarland (1995), "Safe Port in a Storm: The Impact of Labor Market Conditions on Community College Enrollments", Journal of Human Resources, 1995, (30)4, 741-765.

Busso, M. Bassi, M., Urzúa, S. and Vargas, J. (2012), *Disconnected: Skills, Education, and Employment in Latin America*, Inter-American Development Bank, Washington DC.

Busso,M. Cristia, J., Hincapié, D., Messina, J. and Ripani, L. (2017), *Learning better: public policy for skills development Inter-American Development Bank*, Washington DC.

Candia, C., Encarnação, S., & Pinheiro, F. L. (2018). The Higher Education Space, Connecting Degree Programs from Individuals' Choices. arXiv preprint arXiv:1810.03676.

Cedefop (2010). *The skill matching challenge: analyzing skill mismatch and policy implications*. Luxembourg: Publications Office.

Clotfelter, C. (1992), *Who Benefits from the Nonprofit Sector?,* University of Chicago Press.

Costa, R. (2012). *“Quatro décadas de mobilidade social no Brasil”.* Rio de Janeiro, Dados, 55, 614-678.

Dicarlo, E., Salvatore Lo Bello, S., Monroy-Taborda, A., Oviedo, M., Sanchez-Puerta, L., and Santos, I. (2016). “The Skill Content of Occupations across Low- and Middle-Income Countries: Evidence from Harmonized Data.” IZA Discussion Paper, no. 10224.

Ferraz, D., Mariano, E., Rebelatto, D., & Hartmann, D. (2019). Linking Human Development and the Financial Responsibility of Regions: Combined Index proposals using methods from Data Envelopment Analysis. *Available at SSRN 3401374*.

Ferreira, H., Leite, P., Litchfield, J., & Ulissea, G. (2006). “*Ascensão o e queda da desigualdade de renda no Brasil. Económica.”* Rio de Janeiro, 8(1), 147169.

Marta Ferreyra, M., Avitabile, C., Botero Álvarez, J., Haimovich Paz, F., & Urzúa, S. (2017). At a Crossroads: Higher Education in Latin America and the Caribbean. Directions in Development. Washington, DC: World Bank.

Freitas, E. E., & Paiva, E. A. (2015). Diversificação e sofisticação das exportações: uma aplicação do product space aos dados do Brasil. Revista Econômica do Nordeste, 46(3), 79-98.

Furtado C. (2009). O Nordeste e a Saga da Sudene. Rio de Janeiro, Contraponto.

Hartmann, D. (2014). Economic complexity and human development: How economic diversification and social networks affect human agency and welfare. London: Routledge

Leslie, L. L., & Brinkman, P. T. (1987). Student price response in higher education. *Journal of Higher Education*, 58, 181-204.

Leuven, E., & Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. In Handbook of the Economics of Education (Vol. 4, pp. 283-326). Elsevier.

Liu, K., Salvanes, K. G., & Sørensen, E. Ø. (2016). Good skills in bad times: Cyclical skill mismatch and the long-term effects of graduating in a recession. European Economic Review, 84, 3-17.

McCowan, T. (2007), Expansion without equity: An analysis of current policy on access to higher education in Brazil. *Higher Education*. 53: 579.

McGuinness, S., Pouliakas, K. and Redmond. P. (2017), *How Useful is the Concept of Skills Mismatch?,* International Labour organization, Geneva.

Messina, J., Oviedo, A., and G. Pica (2016). *“Job Polarization in Latin America.”* Unpublished Manuscript.

Moreira R., 2014. A formação espacial brasileira: contribuição crítica aos fundamentos espaciais da geografia do Brasil. Rio de Janeiro, Consequência, 2nd ed.

OECD (2016), Boosting Productivity and Inclusive Growth in Latin America, Paris.

Ospino, C, (2018), *“Ocupaciones laborales: Clasificaciones, taxonomías y ontologías para los mercados laborales del siglo XXI”,* Nota técnica, IDB-TN-1513. Banco Inter-Americano de

Desarrollo.

Patrinos, H. (1990), “The privatization of higher education in Colombia, effects on quality and Equity”, *Higher Education*, 20(2), 161-173.

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

Ranis, Stewart & Ramirez, 2000, “Economic Growth and Human Development”, World Development, 28 (2), 197-219.

Robles, C. and Bhandari, R. (2017), Higher Education and Student Mobility A Capacity Building Pilot Study in Brazil, IIE Center for Academic Mobility Research and Impact.

Silva, S. A. D. (2017). Regional Inequalities in Brazil: Divergent Readings on Their Origin and Public Policy Design. EchoGéo, (41).

Simões, R. F., & Freitas, E. E. (2014). Urban attributes and regional differences in productivity: Evidence from the external economics of Brazilian micro-regions from 2000-2010. *Journal of Economic & Financial Studies*, *2*(01), 27-39.

Souza, A., & Lamounier, B. (2010). A classe média brasileira: ambições, valores e projetos de sociedade. Rio de Janeiro: Elsevier. Brasília, DF: CNI.

Sampaio, H., (2015), *"Higher Education in Brazil: Stratification in the Privatization of Enrollment" In Mitigating Inequality: Higher Education Research, Policy, and Practice in an Era of Massification and Stratification.";* 53-8.

Somers, M., Cabus, S., Groot M. & Maassen van den Brink H. (2016), “Horizontal Mismatch between Employment and the Field of Education: Evidence from a Systematic Literature Review”, *TIER Working Paper Series 16/02.*

Torche, F. and Ribeiro, C. (2010), “Pathways of change in social mobility: Industrialization, education and growing fluidity in Brazil”. Research in social stratification and mobility 28, 291-307.

Wilkinson, R. and I. Yussof (2005), “Public and Private Provision of Higher Education in Malaysia: A comparative Analysis”, *Higher Education*, 50 (3), 361-386.

Appendix

Table A1. Evolution of the Higher Education system in Brazil (2002-2016)

	2002				2016			
	N. of HEI	N. of programs	Enrollment	% of enrollment	N. of HEI	N. of programs	Enrollment	% of enrollment
Total	1,637	14,445	1,451,922		2,407	34,226	10,982,985	
Total in-person programs	1,637	14,399	1,411,208	97.2	2,407	32,602	9,491,655	86.4
By administrative category								
Public	195	5,252	320,354	22.7	296	9,924	1,800,704	19.0
Private	1,442	9,147	1,090,854	77.3	2,111	22,673	7,690,951	81.0
By academic organization								
Unversities	162	8,486	765,454	54.2	197	14,693	6,222,585	65.6
University Centres	77	1,413	195,215	13.8	166	4,469	1,029,585	10.8
Isolated colleges	1,345	4,127	428,665	30.4	2,004	12,127	2,083,702	22.0
Centre of tech. education	53	373	21,874	1.6	40	1,313	155,783	1.6
Total distance programs		46	40,714	2.8		1,624	1,491,330	13.6

Note - Total in person programs and total distance programs sum 100%. Each category sum 100%. (2) HEI means Higher Education Institutions

Table A2. Number of new programs created between 2010-2016 by field of study

Field of study	Total	Public	For profit	Non-for-profit
Engineering and engineering trades	1387	250	767	370
Business and administration	1135	189	992	-46
Architecture and building	851	76	486	289
Health	843	-8	712	139
Life sciences	196	35	108	53
Arts	187	11	98	78
Agriculture, forestry and fishery	139	65	65	9
Computing	118	105	118	-105
Law	92	-1	104	-11
Social and behavioural science	90	-27	132	-15
Security services	89	3	85	1
Veterinary	79	9	46	24
Manufacturing and processing	61	64	20	-23
Physical science	42	49	1	-8
Humanities	38	15	0	23
Social services	35	7	28	0
Personal Services	35	30	34	-29
Mathematics and statistics	19	20	-1	0
Transport services	12	6	9	-3
Enviromental protection	-8	23	17	-48
Journalism and information	-226	-41	-46	-139
Teacgher training and education science	-511	-53	232	-690
Total	4703	827	4007	-131

Note - New programs is the difference between the number of programs offered in 2016 and the number of programs offered in 2010.

Figure A1. Variation in the number of programs (2010-2016)

1.A Private institutions 1.B. Public institutions

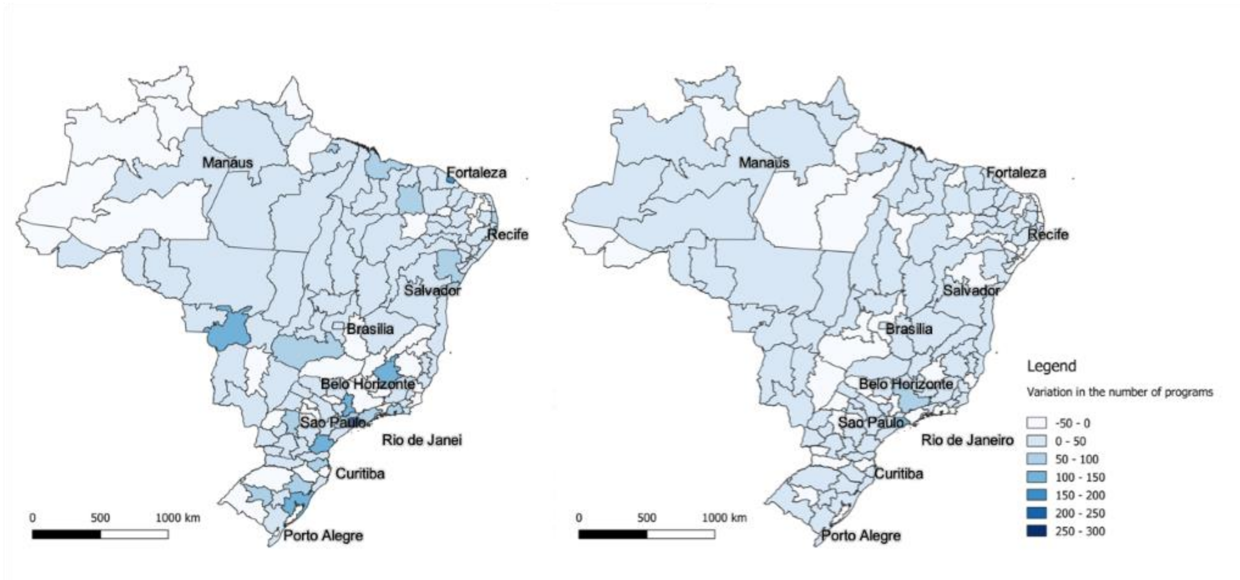


Table A3. First Stage Regression Detailed Results Corresponding to Part 1

The first column is an occupation. Each occupation is connected to a field and the coefficient is the relevance of that coefficient to that field of enrollment. Corresponding to our model, these are the variables that sums all of the associated fields that are related to occupation, o, to calculate the total number of people employed in that region for that occupation $\ln\left(\sum_{o\in f} L_{ory}\right)$.

	Animal Sciences	Trans.	Hospitality Mngmt.	Social Work	Material Sciences	Comm.	Public Safety
Courses	0.195***	0.156*	0.057	-0.066	0.136**	0.052	-0.247*
	(0.060)	(0.093)	(0.054)	(0.049)	(0.057)	(0.048)	(0.130)
Applications	0.032**	0.045	0.107***	0.184***	0.140***	0.051***	0.032
	(0.014)	(0.034)	(0.017)	(0.018)	(0.016)	(0.015)	(0.035)
Seats	0.462***	0.161***	0.157***	0.292***	0.224***	0.177***	0.420***
	(0.029)	(0.061)	(0.029)	(0.031)	(0.031)	(0.024)	(0.064)
Youth Pop.	-0.027	0.718*	0.251***	-0.105	0.051	0.036	-0.266
	(0.043)	(0.401)	(0.064)	(0.077)	(0.095)	(0.045)	(0.255)
Total Enrll.	0.333***	0.583***	0.624***	0.402***	0.432***	0.616***	0.500***
	(0.028)	(0.051)	(0.030)	(0.032)	(0.028)	(0.030)	(0.055)
Avg. Edu	-0.169	0.916***	0.049	0.054	0.038	-0.098*	-0.265*
	(2.025)	(0.244)	(0.054)	(0.049)	(0.055)	(0.057)	(0.145)
Av. Wages (t-1)	0.044	-0.047	0.032	0.005	-0.159	0.022	0.319*
	(0.036)	(0.321)	(0.168)	(0.081)	(0.099)	(0.073)	(0.194)
Occ. Spec.	0.918**	0.258	1.122	4.019**	2.72	0.660*	3.732**
	(0.422)	(3.308)	(0.782)	(1.585)	(1.798)	(0.344)	(1.765)
2010	0.049	0.131	0.077	-0.056	0.014	-0.043	-0.242*
	(0.043)	(0.086)	(0.050)	(0.083)	(0.054)	(0.039)	(0.130)
2011	0.035	0.213**	0.074	-0.066	-0.041	0.011	-0.188
	(0.044)	(0.096)	(0.053)	(0.082)	(0.054)	(0.040)	(0.132)
2012	0.160***	-0.039	0.154***	-0.025	-0.073	0.119***	-0.091
	(0.043)	(0.101)	(0.055)	(0.084)	(0.056)	(0.039)	(0.141)
2013	0.122***	-0.046	0.081	-0.111	-0.092	0.128***	-0.189
	(0.044)	(0.119)	(0.062)	(0.086)	(0.059)	(0.044)	(0.155)
2014	0.116***	-0.298**	0.086	-0.193**	-0.215***	0.075*	-0.217
	(0.045)	(0.121)	(0.064)	(0.088)	(0.061)	(0.043)	(0.147)
2015	0.057	-0.535***	0.09	-0.369***	-0.318***	0.014	-0.227
	(0.045)	(0.139)	(0.065)	(0.090)	(0.063)	(0.044)	(0.151)

Table A3. First Stage Regression Detailed Results Corresponding to Part 1

	Animal Sci	Transp.	Hosp. Mngmt.	Social Work	Mat. Sci.	Comms	Public Safety
Emp. Bio. Health	0.134*		0.545***	-0.13	-0.03	-0.053	
	(0.081)		(0.124)	(0.149)	(0.154)	(0.085)	
Migr. Bio. Health	-0.174**		-0.448***	0.089	-0.06	-0.022	
	(0.073)		(0.097)	(0.108)	(0.116)	(0.064)	
Emp. Teacher	0.085**		-0.006	0.032	-0.023	-0.053*	-0.087
	(0.037)		(0.043)	(0.046)	(0.052)	(0.031)	(0.109)
Migr. Admin.		0.22	0.13	0.065			0.409
		(0.372)	(0.122)	(0.135)			(0.388)
Emp. Soc. Sci.				-0.177			-0.283
				(0.149)			(0.446)
Migr. Soc. Sci.				-0.045			0.028
				(0.139)			(0.466)
Emp. Comm/Arts				0.107		0.258***	
				(0.106)		(0.097)	
Migr. Comm/Arts				-0.001		-0.064	
				(0.087)		(0.081)	
Emp. Sci. Tech.					-0.136		0.368
					(0.155)		(0.494)
Migr. Sci. Tech.					0.08		-0.291
					(0.148)		(0.496)
Emp. Bio. Health Techn			0.287*	0.016	-0.032		
			(0.158)	(0.188)	(0.199)		
Migr. Bio. Health Techn			0.205	-0.227	0.058		
			(0.145)	(0.174)	(0.182)		
Emp. Mngrs		-1.332*	-0.128		0.626**	-0.28	0.549
		(0.686)	(0.216)		(0.256)	(0.205)	(0.574)
Migr. Mngrs		2.111***	-0.02		0.345	0.253	-0.835
		(0.634)	(0.194)		(0.233)	(0.189)	(0.537)
Emp. Police					0.041		-0.039
					(0.030)		(0.090)
Migr. Police					0.013		0.075
					(0.027)		(0.080)
Emp. Scientists		0.058			0.139	0.036	-0.109
		(0.312)			(0.098)	(0.078)	(0.320)
Migr. Scientists		0.136			-0.099	-0.058	-0.541*
		(0.351)			(0.103)	(0.075)	(0.308)
Emp. Transp. Techn.		-0.642***					
		(0.196)					
Migr. Transp. Techn.		0.724***					
		(0.167)					

Table A3. First Stage Regression Detailed Results Corresponding to Part 1

	Animal Sci.	Transp.	Hosp. Mngmt.	Social Work	Mat. Sci.	Comms	Public Safety
Emp. Admin.		-0.617*	-0.144	0.12			0.227
		(0.340)	(0.120)	(0.13)			(0.33)
Emp. Serv. Techn.			0.146*		0.061	0.079	0.225
			(0.09)		(0.10)	(0.08)	(0.31)
Migr. Serv. Techn.			-0.061		0.011	-0.045	0.189
			(0.08)		(0.09)	(0.07)	(0.30)
Emp. Other Techn.							-0.118
							(0.27)
Migr. Other Techn.							0.213
							(0.27)
Emp. Services		-0.242	-0.425**	0.446**	0.057		-0.043
		(0.48)	(0.19)	(0.19)	(0.23)		(0.56)
Migr. Services		-0.588	0.306	-0.161	0.093		0.312
		(0.44)	(0.19)	(0.17)	(0.21)		(0.53)
Emp. Sales		-0.348					
		(0.78)					
Migr. Sales		-0.271					
		(0.64)					
Emp. Farmers					-0.013		
					(0.03)		
Migr. Farmers					0.041*		
					(0.02)		
Emp. Agric.					0.034		
					(0.07)		
Migr. Agric.					-0.034		
					(0.06)		
Emp. Fishermern					-0.011		
					(0.03)		
Migr. Fishermern					-0.013		
					(0.03)		
Emp. Mining		-0.053			-0.074		
		(0.22)			(0.10)		
Migr. Mining		0.296			0.047		
		(0.18)			(0.09)		
Emp. General		-0.578					
		(0.45)					
Migr. General		0.38					
		(0.41)					
Emp. Textile					0.18		
					(0.12)		
Migr. Textile					-0.189		
					(0.119)		

Table A3. First Stage Regression Detailed Results Corresponding to Part 1

	Animal Sci.	Transp.	Hosp. Mngmt.	Social Work	Mat. Sci.	Comms	Public Safety
Emp. Wood					0.005		
					(0.17)		
Migr. Wood					-0.007		
					(0.16)		
Emp. Rel. Food Manuf.					-0.009		
					(0.10)		
Migr. Rel. Food Manuf.					0.028		
					(0.09)		
Constant	1.2	-6.886	-2.199^{*}	-0.591	1.531	-0.062	0.757
	(18.22)	(4.22)	(1.30)	(0.94)	(1.05)	(0.75)	(2.28)
Observations	583	122	546	630	604	576	155
Log Likelihood	-2937.40	-555.67	-2900.9	-3318.22	-3018.55	-2926.19	-759.99
	16.025^{***}	33.725^{***}	14.731^{***}	9.262^{***}	10.702^{***}	20.540^{***}	8.415^{***}
theta	(1.076)	(5.884)	(1.046)	(0.576)	(0.686)	(1.514)	(1.038)
Akaike Inf. Crit.	5910.814	1173.346	5857.92	6692.449	6129.118	5904.383	1587.993

Note - $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table A4. First Stage Regression Detailed Results Corresponding to Part 2

	Envir. Studies	Arts & Design	Humanities	Math	Life Sciences	Phys. Sci.
Courses	0.132**	0.139***	-0.034	0.149***	-0.03	0.021
	(0.05)	(0.04)	(0.05)	(0.05)	(0.04)	(0.04)
Applications	0.085***	0.068***	0.076***	0.082***	0.099***	0.064***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)
Seats	0.175***	0.160***	0.296***	0.413***	0.310***	0.493***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)
Youth Pop.	0.142**	-0.103	-0.094	-0.081	0.044	0.089
	(0.07)	(0.10)	(0.07)	(0.09)	(0.05)	(0.07)
Total Enrll.	0.520***	0.508***	0.544***	0.378***	0.483***	0.365***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)
Avg. Edu	-0.057	-0.029	0.042	-0.038	-0.006	-0.047
	(0.05)	(0.06)	(0.06)	(0.05)	(0.11)	(0.07)
Av. Wages (t-1)	0.187**	-0.065	0.098	-0.022	0.163***	0.166**
	(0.08)	(0.08)	(0.08)	(0.08)	(0.06)	(0.08)
Occ. Spec.	1.902*	1.800***	0.566	-0.091	1.085	1.302
	(0.99)	(0.53)	(0.92)	(0.74)	(0.80)	(0.89)
2010	0.035	0.046	0.026	0.158*	-0.131***	0.031
	(0.07)	(0.04)	(0.09)	(0.09)	(0.04)	(0.05)
2011	0.012	0.059	0.012	0.191**	-0.137***	-0.037
	(0.06)	(0.04)	(0.09)	(0.09)	(0.04)	(0.06)
2012	-0.05	0.156***	0.062	0.210**	-0.098**	0.011
	(0.07)	(0.05)	(0.09)	(0.09)	(0.04)	(0.06)
2013	-0.128*	0.042	-0.093	0.089	-0.139***	-0.184***
	(0.07)	(0.05)	(0.09)	(0.10)	(0.04)	(0.06)
2014	-0.259***	-0.02	-0.210**	-0.031	-0.151***	-0.270***
	(0.07)	(0.05)	(0.09)	(0.09)	(0.04)	(0.06)
2015	-0.344***	-0.047	-0.192**	-0.076	-0.262***	-0.298***
	(0.07)	(0.05)	(0.09)	(0.09)	(0.04)	(0.06)

Table A4. First Stage Regression Detailed Results Corresponding to Part 2

	Envir. Studies	Arts & Design	Humanities	Math	Life Sciences	Phys. Sci.
Emp. Mngrs			0.659**	-0.432		0.216
			(0.29)	(0.40)		(0.24)
Migr. Mngrs			-0.735***	0.28		-0.221
			(0.28)	(0.35)		(0.23)
Emp. Police	0.067**		0.109**	0.025	-0.02	0.014
	(0.03)		(0.04)	(0.05)	(0.02)	(0.03)
Migr. Police	-0.003		-0.093**	-0.082**	0.00004	-0.003
	(0.03)		(0.04)	(0.04)	(0.02)	(0.03)
Emp. Bio. Health	-0.235**	0.217**			-0.132	
	(0.11)	(0.11)			(0.10)	
Migr. Bio. Health	0.161*	-0.261***			0.043	
	(0.10)	(0.09)			(0.07)	
Emp. Scientists			0.085	0.073	0.156**	-0.048
			(0.13)	(0.17)	(0.07)	(0.10)
Migr. Scientists			0.017	0.154	-0.140**	0.109
			(0.13)	(0.17)	(0.07)	(0.10)
Emp. Teacher	-0.0001	0.045	0.135**	0.127**	0.063*	0.177***
	(0.05)	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)
Emp. Science.Tech.	-0.118					0.256*
	(0.09)					(0.15)
Migr. Science.Tech.	-0.034					-0.093
	(0.09)					(0.15)
Emp. Social.Science			-0.511**	0.142		
			(0.22)	(0.24)		
Migr. Social.Science			0.450**	-0.243		
			(0.22)	(0.23)		
Emp. Comm/Arts		-0.197*	0.14			
		(0.12)	(0.16)			
Migr. Comm/Arts		0.131	-0.169			
		(0.10)	(0.15)			

Table A4. First Stage Regression Detailed Results Corresponding to Part 2

	Envir. Studies	Arts & Design	Humanities	Math	Life Sciences	Phys. Sci.
Emp. Serv. Techn		0.018				
		(0.09)				
Migr. Serv. Techn		0.039				
		(0.09)				
Emp. Clerks		0.021				
		(0.15)				
Migr. Clerks		0.008				
		(0.15)				
Emp. Services	0.036	-0.039				
	(0.17)	(0.21)				
Migr. Services	0.021	0.353*				
	(0.14)	(0.21)				
Emp. Mining	-0.224**					
	(0.10)					
Migr. Mining	0.159**					
	(0.08)					
Emp. Admin.				-0.144		-0.121
				(0.21)		(0.13)
Migr. Admin.				0.039		-0.203
				(0.23)		(0.14)
Emp. Bio. Hlth. Techn.					0.136	0.127
					(0.13)	(0.14)
Migr. Bio. Hlth. Techn					-0.062	-0.162
					(0.11)	(0.14)
Emp. Other Techn					-0.054	-0.002
					(0.06)	(0.08)
Migr. Other Techn					0.049	0.005
					(0.06)	(0.08)
Emp. Manuf. Process.						0.001
						(0.06)
Migr. Manuf. Process.						-0.026
						(0.07)
Emp. Utility						-0.162**
						(0.08)
Migr. Utility						0.059
						(0.08)

Table A4. First Stage Regression Detailed Results Corresponding to Part 2

	Envir. Studies	Arts & Design	Humanities	Math	Life Sciences	Phys. Sci.
Emp. Sales		-0.188				
		(0.31)				
Migr. Sales		-0.073				
		(0.27)				
Emp. Elect. Manuf.		0.06				
		(0.05)				
Migr. Elect. Manuf.		-0.021				
		(0.05)				
Emp. Manuf. Assmb.		0.066*				
		(0.04)				
Migr. Manuf. Assmb.		-0.092***				
		(0.03)				
Emp. Jewelers		0.054				
		(0.07)				
Migr. Jewelers		-0.031				
		(0.07)				
Emp. Textile		0.195*				
		(0.11)				
Migr. Textile		-0.153				
		(0.10)				
Constant	-1.416	0.976	-0.623	1.714*	-2.316**	-1.631
	(0.90)	(0.85)	(1.00)	(0.95)	(1.18)	(1.02)
Observations	427	443	429	263	692	512
Log Likelihood	-2,013.00	-2,442.51	-2,258.71	-1,189.84	-3,628.36	-2,710.85
theta	13.847*** (1.115)	24.443*** (1.929)	7.878*** (0.578)	17.314*** (1.840)	15.639*** (0.965)	11.593*** (0.803)
Akaike Inf. Crit.	4,078.00	4,957.02	4,569.41	2,431.67	7,308.72	5,489.70

Note: *p<0.1; **p<0.05; ***p<0.01