Estimation of the Elasticity of Substitution Between Skilled and Unskilled Men in Agglomerations Between 1980 and 2000 in Brazil and the Simulation of the Impact of Migration on the Skill Wage Gap

Estimativa da Elasticidade de Substituição Entre Homens Qualificados e Não Qualificados em Aglomerações Urbanas, Entre 1980 e 2000, no Brasil, e Simulação do Impacte da Migração na Diferença Salarial dos Trabalhadores Qualificados

Tiago Freire[1]
tiago@tiagofreire.com
Faculty of Business, Government and Law, University of Canberra
Roberts Capital Advisors, LLC

Abstract/ Resumo

We use a production function approach to determine the impact on the wage skill premium of increased labor supply of different skills levels. In particular, we use data from the Brazilian population censuses for 1980 to 2000, for which there is consistent data on cities, to estimate the elasticity of substitution between high and low skill workers. We address the endogeneity of labor supply in urban areas by constructing an instrumental variable using the rural-urban migration response to exogenous shocks, such as rainfall shocks and changes in transport costs in rural areas. We find an elasticity of substitution between high and low skill workers of 0.56, so that high and low skill workers are gross complements and not gross substitutes as has been found previously in the international migration literature. Our simulations show that migration between 1991 and 2000 had only a small impact on the wage skill gap, specifically, a 1.1% decrease in the wage gap between high and low skill workers in Brazil. This explains only 5% of the changes in relative wages between high and low skill

Neste estudo estimamos parâmetros de uma função de produção para determinar o impacto de alterações da oferta laboral nas diferenças salariais entre trabalhadores qualificados e não qualificados. Nomeadamente, utilizamos dados dos Censos populacionais brasileiros entres 1980 e 2000,, para estimar a elasticidade de substituição entre trabalhadores qualificados e não qualificados. Resolvemos o problemas da endogeneidade da oferta laboral nas cidades construindo uma variável instrumental utilizando o impacto de choques exogenos nas zonas rurais, como precipitação e redução nos custos de transporte, nas migrações rurais-urbanas. Ao contrario do que foi encontrado anteriormente em trabalhos sobre migrações internacionais encontramos uma elasticidade de substituição entre trabalhadores qualificados e não qualificados de 0.56, que indica que os trabalhadores com diferentes níveis de qualificação foram complementares e não substitutos. As nossas simulações mostram que a migrações rurais urbanas entre 1991 e 2000 no Brasil reduziram as diferenças salariais entre trabalhadores qualificados e não quali-

[1]The author is grateful to Rachel Friedberg, Andrew Foster, Sriniketh Nagavarapu, Haoming Liu, and other participants of the Applied Microeconomics seminar at Brown University, and at the National University of Singapore for advice and suggestions. I would also like to thank participants at the North America Regional Science Council Conference; European Regional Science Association Congress; Western Economic Association International Biannual Pacific Rim Conference; Centre for Research Analysis of Migration and NORFACE Conference on Migration, Economic Change, and Social Challenges; Portuguese Economic Journal Meeting; and Congresso da Associacao Portuguesa de Desenvolvimento Regional. Furthermore, I would like to thank the European Regional Science Association jury for granting me the Epainos award for best paper by a young researcher.

Our simulations show that migration between 1991 and 2000 had only a small impact on the wage skill gap, specifically, a 1.1% decrease in the wage gap between high and low skill workers in Brazil. This explains only 5% of the changes in relative wages between high and low skill workers seen in the data during this time period. Although this impact is small, it is 2.5 times larger than what we would have found using the elasticity of substitution between high and low skill workers of the international migration literature.

Keywords: Labor markets; Rural–Urban migration; Wages

JEL Codes: R12, J31, J61, O15, O18

ficados em apenas 1.1%. Isto representa apenas 5% da redução do diferencial salarial entre trabalhadores qualificados e não qualificados entre 1991 e 2000. Embora o impacto seja pequeno é 2.5 vezes maior do valor que obteriamos se tivessemos utilizado estimativas da elasticidade de substituição entre trabalhadores qualificados e não qualificados da literatura das migrações internacionais.

Palavras Chave: Mercado de trabalho, Migração rural-urbana, salário

Código JEL: R12, J31, J61, O15, O18

1. INTRODUCTION

In the 1990s, there was a large inflow of immigrants to the US, with the number of foreign born individuals increasing on average 0.85 percentage points (pp) each year between 1990 and 2006 (Ottaviano and Peri, 2008). This compares to the UK, where the number of foreign-born individuals increased on average 0.25 pp a year between 1995 and 2005 (Manacorda et al., 2012). The internal migration flows in developing countries have been just as large. For instance, in China, the percentage of rural migrants living in urban areas increased on average 0.8 pp a year between 2000 and 2010 (Hanewinkel and Oltmer, 2012), while in Brazil, the proportion of migrants from rural areas living in cities increased on average 0.74 pp a year between 1995 and 2000 (author's own calculations). While there are several studies on the impact of immigration on wages in countries of destination, there are fewer studies on the impact of internal migration on wages in regions and cities of destination. In this study, we estimate the elasticity of substitution between high and low skill male workers between the ages of 15 and 55, in agglomeration in Brazil, using exogenous shocks to internal migration, to simulate the impact n the skill wage gap in 123 Brazilian agglomerations.

As pointed out by Borjas (1999), initial studies on the impact of immigration on wages focused on the levels of wages. However, as pointed out by Ottaviano and Peri (2008; 2012) and Peri (2011; 2012), these results are biased as they ignore changes in capital in the receiving communities, as well as the fact that immigrants have different characteristics compared with local residents. For instance, if the average immigrant has more education than the average local resident, then the increased supply of high skill workers would decrease the high skill wage relative to the wage of low skill workers. Therefore, as pointed out by Borjas et al. (2011), recent research has focused on estimating the elasticity of substitution between high and low skill workers and the elasticity of substitution between local residents and immigrants using a production function approach. While evidence for the elasticity of substitution between immigrants and local residents is mixed (Borjas et al., 2011), most studies find an elasticity of substitution between high and low skill workers of between 1.4 and 2.4 (Ottaviano and Peri, 2008). Research on internal migration focuses on wage levels (for recent examples, see Boustan et al., 2010; Lu and Song, 2006; Meng and Zhang, 2001). Therefore, to the best of our knowledge, ours is the first study that attempts to use a production function approach to estimate the impact of internal migration on receiving communities. This is important, first, because we do not know whether estimates for the elasticity of substitution between high and low skill workers in developed countries are the same as that for developing countries, as the industrial composition is different in developing countries and the elasticity of substitution combines both within and across industries (Acemoglu, 2002). Second, it is important to understand the impact of changes in the relative supply of skilled workers on the wage skill premium in developing countries because this tends to be larger in countries with smaller supplies of skilled labor (Gropello and Sakellariou, 2010).

Another major contribution of this study is that it addresses the issue of endogeneity of migration flows. For instance, Borjas (2003), Manacorda et al. (2012), Ottaviano and Peri (2008; 2012), and Peri (2012) all assume that immigration flows to the US and the UK are exogenous, and therefore, are not driven by changes in wages in the US and the UK, respectively. Other studies, such as Card (2001), Peri (2011), and Wozniak and Murray (2012), attempt to address the problem of endogeneity of migration flows by using a supply–push instrumental variable (SPIV) approach. In this approach, the decision to migrate is separated into two factors: (i) the decision to leave the place of origin; and (ii) the decision of which destination country to select. In this approach, it is assumed that the decision to leave a country of origin is exogenous to what is happening in the country of destination, and the endogeneity of the migration flows comes only from the decision of where to go. Therefore, by using historical patterns of migration settlement, under the assumption that migrants are not forward looking, the SPIV is a valid instrument. In this study, we use a similar approach, which does not require the two aforementioned assumptions. We estimate the response of rural migrant outflows from rural areas to exogenous weather shocks and changes in transport costs, and then estimate how distance determines historical patterns of settlement in urban areas. Using these two estimated values, we then construct an instrumental variable that is orthogonal to changes in conditions in urban areas.

We begin by developing a model of labor demand with differentiated labor similar to Ottaviano and Peri (2008; 2012), and Peri (2011) to estimate the elasticity of substitution between high and low skill workers from the relationship between their relative wages and their relative labor supply. Due to data limitations explained in Section 3, we focus on the 1980–2000 Brazilian population censuses to estimate the elasticity of substitution between high and low skill workers in urban areas. Because the supply of high and low skill workers is endogenous, as workers from rural areas may depend on wages, our ordinary least square (OLS) estimates of the elasticity of substitution are biased towards zero. Therefore, we use migration flows as an instrument for change in the number of workers in a city. However, migration flows are themselves endogenous. To address this issue, we use information on rainfall, changes in transport costs, and distance between destination and origin as exogenous sources of variation that explain migration flows from rural to urban areas, and therefore, are instrumental variables for migration flows. As expected, our instrumental variable estimates are smaller than our OLS estimates and we find an elasticity of substitution between high and low skill workers of between 0.56 and 0.58, which is significantly smaller than that found in the migration literature for the UK and US (Manacorda et al., 2012, and Ottaviano and Peri, 2008, respectively). This finding implies that high and low skill workers are gross complements in production in Brazil. Our results hold when we use a standard SPIV with a series of robustness checks. These results suggest that we cannot use elasticity of substitution, as the international migration literature does, to estimate the impact of changes in the relative supply of high skill workers in developing countries, for instance, because of internal migration. In particular, the simulated impact of rural–urban migration on urban wages in Brazil is one third smaller when we use the elasticity of substitution. Despite this, the impact of migration flows on the wage skill premium is small (1.1%) and explains only 5% of the variation in the data. This is because migrants have similar observable characteristics to local urban residents.

Our results have important implications for public policy. Substantial resources are spent each year in both developed and developing countries to reduce inequality. Part of this effort is undertaken via regional policies, in which funds are allocated to attenuate regional inequalities. For example, in Brazil, 9.3% of federal government revenues in 1992 came from the northeast, while expenditure in this region comprised 14.7% of the federal budget (Baer, 2013). Regional funds target investment in transport infrastructure in order to increase agglomeration economies, thus attracting more workers in disadvantaged areas to urban areas and promoting economic growth (see Puga, 2002). While regional economic theory suggests that the migration of workers from poorer rural areas to richer urban regions would decrease interregional inequality, international migration theory suggests that the inflow of migrants could actually increase intraregional inequality within cities (Borjas, 2003). This suggests that there might be a trade-off between the two effects. Therefore, it is important to measure the impact of internal migration in areas of destination when designing regional policies.

The rest of the paper is organized as follows. Section 2 presents the theoretical model on which our work is based. In Section 3, we describe the data and the events that occurred during this time period. In Section 4, we look at the results and describe our instrumental variable approach in detail. In Section 5, we present our simulation results of the impact of migration on the wage gap. Finally, Section 6 concludes.

2. THEORETICAL FRAMEWORK

We use a factor approach to evaluate how changes in factor endowment lead to changes in relative wages. This method was used initially in trade literature and is currently used to evaluate the impact of international migration (see, for instance, Borjas, 2006; Card, 2001; Card, 2009a; Ottaviano and Peri, 2008, 2012; Peri, 2011). The model we develop here is similar to that presented in Freire (2010) and Ottaviano and Peri (2008; 2012).

We start by assuming that production in each city can be represented by a production function with constant returns to capital and labor, as follows.

$$Y=AK^{1-\alpha}N^{\alpha} \qquad (1)$$

where Y is the amount of output in a city in a given year, K is the amount of capital in a city in a particular year, N is the amount of labor (migrant and non-migrant) in a city in a given year, A is the technology used in a city in a given year (or other factors that are city specific and explain agglomeration economies), and α is the output elasticity of labor.

Furthermore, we assume that labor is not homogeneous. In particular, we assume that low skill labor is an imperfect substitute for high skill labor by combining both within a constant elasticity of substitution production function, nested within Equation 1, as follows.

$$N=\left[a_H N_H^{\frac{\sigma_{EDU}-1}{\sigma_{EDU}}}+(1-a_H)N_L^{\frac{\sigma_{EDU}-1}{\sigma_{EDU}}}\right]^{\frac{\sigma_{EDU}}{\sigma_{EDU}-1}} \qquad (2)$$

where N_H and N_L are the number of high and low skill workers, respectively, in a city in a given year, and a_H is time-varying technology parameter that captures the relative productivity of the two education groups. These factors do not depend on the relative supply of each type of worker but may change with time. A higher a_H implies that a worker with extra high skills is relatively more productive than one with extra low skills in the production of goods in a city, when the initial number of high and low skill workers in the city is the same. Finally, σ_{EDU} is the elasticity of substitution between high and low skill workers. If $\sigma_{EDU}>1$, then high and low skill workers are gross substitutes in production, while if o≤ $\sigma_{EDU}1<$, then high and low skill workers are gross complements in production.

Because we assume that the production function in a city exhibits constant returns to scale, the returns to each factor are equal to their marginal productivity. Therefore, high skill wages are given by the following.

$$N=\left[a_H N_H^{\frac{\sigma_{EDU}-1}{\sigma_{EDU}}}+(1-a_H)N_L^{\frac{\sigma_{EDU}-1}{\sigma_{EDU}}}\right]^{\frac{\sigma_{EDU}}{\sigma_{EDU}-1}} \qquad (3)$$

In addition, low skill wages are given by the following.

$$w_L=\alpha AK^{1-\alpha}[N]^{\alpha-1}(1-a_{Hst})N_L^{-\frac{1}{\sigma_{EDU}}} \qquad (4)$$

Several studies on the impact of migration on wages focus on how the proportion of foreign workers affects wages in the destination. As pointed out by Ottaviano and Peri (2008; 2012), wages in a particular city depend on capital and technology in that city. We can control for city specific factors that affect both groups symmetrically (technology, A, and capital, K) by focusing on the ratio of high to low skill wages, which can be interpreted as a proxy for income inequality in a particular city.

$$\frac{w_H}{w_L}=\frac{a_H}{(1-a_H)}\left[\frac{N_H}{N_L}\right]^{-\frac{1}{\sigma_{EDU}}} \qquad (5)$$

This is represented in log form as follows.

$$\ln\left(\frac{w_H}{w_L}\right) = \ln\left[\frac{a_H}{(1-a_H)}\right] - \frac{1}{\sigma_{EDU}}\ln\left[\frac{N_H}{N_L}\right] \quad (6)$$

Equation 6 provides a linear relationship between, on one hand, the ratio of high to low skill wages, and, on the other, the ratio of high to low skill employment. We can use this relationship to obtain an estimate of the elasticity of substitution between high and low skill workers, σ_{EDU}. Once we have established the value of σ_{EDU}, we can use it to estimate the impact of migration on the ratio of high to low skill wages, as in Card (2009b), Ottaviano and Peri (2008; 2012), and Peri (2011).

Previous studies use national data to estimate σ_{EDU} and control for changes in the specific productivity of high skill workers (a_H) and low skill workers ($(1-a_H)$) using time trends.

However, we use a panel of cities in Brazil and, therefore, estimate the regression in first difference, controlling for city fixed effects, as follows.

$$\Delta\ln\left(\frac{w_H}{w_L}\right) = \Delta\ln\left[\frac{a_H}{(1-a_H)}\right] - \frac{1}{\sigma_{EDU}}\Delta\ln\left[\frac{N_H}{N_L}\right] \quad (7)$$

We include a constant in our first difference regressions to control for changes in the specific productivity of high skill workers (a_H) and low skill workers ($(1-a_H)$) over time.

3. DATA

The lowest level of government in Brazil comprises municipalities. Their borders have changed significantly over time, and their numbers have increased from 3,951 in 1970 to 5,501 in 2000, as pointed out by Reis et al. (2009). Therefore, to compare municipalities across time, we merge municipalities into minimal comparable areas (MCAs), following the method developed by Brazil's Institute for Applied Economic Research (IPEA) for 1970–2000. Figure 1 shows a map of the 3,661 MCAs used in our study. We use the term municipality to refer to an MCA. We use the same definition of cities as that proposed by IPEA and the Brazilian Institute of Statistics (IBGE) and defined in Vieira et al. (2012) and used in Mata et al. (2007). These cities or agglomerations are based on 1991 MCAs and can be used to compare data at the city level from 1970–2000. The advantage of using this definition is that it has characteristics similar to the US metropolitan statistical areas. The Brazilian Institute of Geography and Statistics (IBGE) also uses an alternative definition of city, which includes an urban area in every municipality in Brazil, even those deep in the Amazon, as pointed out by Camarano and Abramovay (1997); hence, the official urbanization rate was 80% in 2000, up from 55% in 1980. Figure 1 shows a map of Brazil's 123 agglomerations (composed of 447 MCAs), where 57% of the population lived in 2000, up from 46% in 1980 (Mata et al., 2007). We use micro data from the 1980, 1991 and 2000 censuses[1] but the 2010 census, because our definition of city cannot be used to compare city level results between 1980 and 2010[2].

As Figure 2 shows, Brazil's average education level is low but increasing, with average years of education for men rising from 6.2 in 1991 to 7.64 in 2000, based on our own calculations. This is despite the fact that there were eight years of compulsory education until 2006, and nine thereafter. Between 1980 and 2000 there were improvements in educational levels, but these where restricted to the number of people achieving compulsory education, with little progress made at higher levels of education (Souza, 2001). As in Borjas (2006), Katz and Murphy (1992), Ottaviano and Peri (2008), and Peri (2011) we use education to define skills. High skill workers are those with nine or more years of education and low-skill workers have less than nine years of education (that is, completed compulsory education or less). In our robustness checks, we use a definition of high skill workers as those with 12 years of education or more, and less than 12 years for low skill workers. Because women have been found to have a higher elasticity of labor supply, we focus on men only[3]. To determine hourly income, because of low levels of education, we define active workers as people from 15 to 55 years of age[4], inclusive, who are not tudents and report having labor income and report numbers of hours worked.

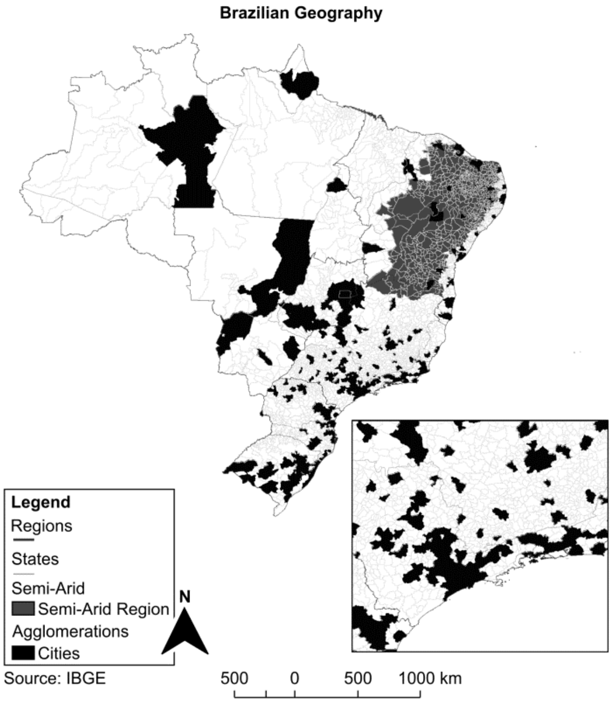
[1] We exclude cases with missing information regarding their education, age, income or place of residence.

[2] We also cannot use the 2010 census because we don't have a complete panel for our variables, in particular a transportation cost variable we use in our instrumental variable approach.

[3] We other robustness checks we did not show, the inclusion of women does not qualitatively change our results.

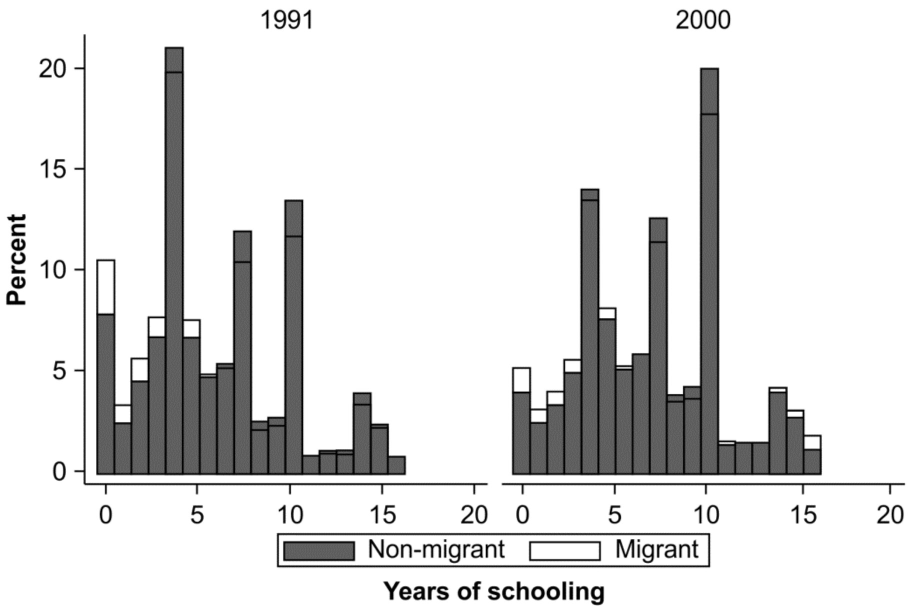
[4] Acording to Baer (2013) the retirment age in the public sector for some public sector workers is 55. Because we need to take into consideration more elaborate labor suply considerations for men above 55, we limit our sample to men in that age.

Figure 1: Map of 123 agglomerations (cities), 28 states, 5 regions, and a semi-arid (chronic drought) region in Brazil.



Note: The enlarged map is for the area around São Paulo

Figure 2: Education of active urban population (gray) and rural–urban migrants (black line) for 1991 and 2000.



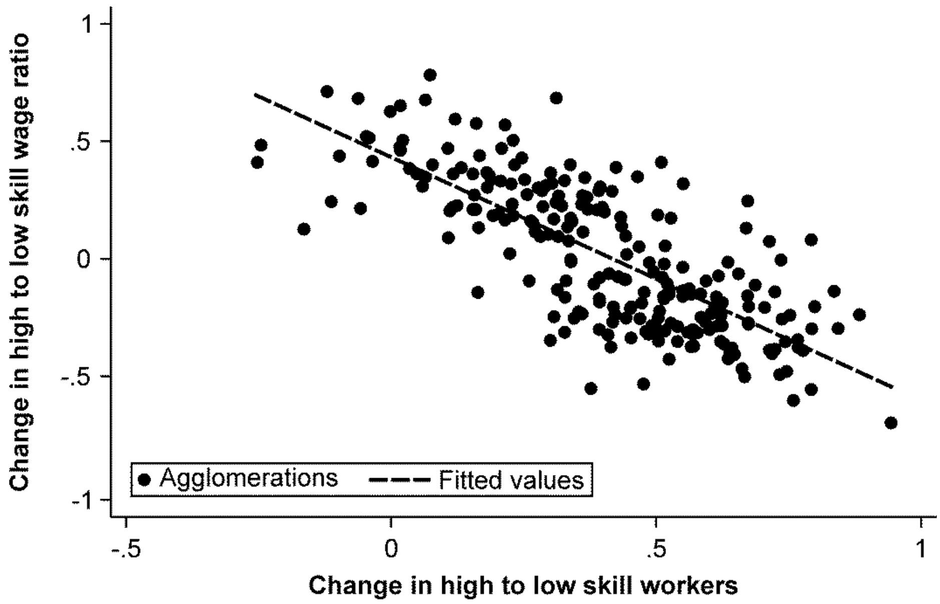
Our basic statistics regarding wages and numbers of workers in urban areas are reported in Table 1. Between 1980 and 2000, Brazil went through a period of hyperinflation and replaced its currency system twice; thus, we convert wages to January 2002 values using deflators provided by Corseuil and Foguel (2002). We focus on median hourly labor income for each group in order to avoid bias from outliers, although our results remain the same as if we had used average hourly wages. The average number of hours worked increased from 41 hours a week in 1991 to 43 hours in 2000, based on our calculations. As Table 1 shows, real hourly wages decreased for high skill workers between 1980 and 2000, while real hourly wages for low skill workers decreased between 1980 and 1991 but recovered partially between 1991 and 2000. Finally, the average number of high skill workers in cities increased slowly between 1980 and 2000, while the average number of low skill workers increased between 1980 and 1991, but decreased slightly between 1991 and 2000. This is consistent with an increase in average educational levels in Brazil in this timeframe. Figure 3 shows a negative relationship between the ratio of high to low skill wages and the ratio of high to low skill workers in cities.

Table 1: Basic statistics (mean and standard deviation) for 123 agglomerations (cities) on wages, employment, and rural to urban migration for 1980, 1991, and 2000 in Brazil.

Basic statistics for 123 agglomerations						
	1980		1991		2000	
	Mean	SD	Mean	SD	Mean	SD
(Median) Hourly wage						
High skill	3.39	0.85	2.84	0.73	2.54	0.54
Low skill	1.76	0.46	1.11	0.37	1.26	0.34
Average employment						
High skill	25,779	82,095	40,636	118,797	64,293	180,778
Low skill	84,311	250,150	100,070	257,366	92,901	218,622
Average migrants						
High skill			1,230	13,694	1,906	21,226
Low skill			3,319	36,962	4,025	44,823

Note: High skill is defined as people with nine years of education or more, and low skill less than nine years of education. Reported median hourly wages are deflated to current values (January 2002). Migrants are defined as people who lived in different municipalities five years before the censuses.

Figure 3: Relationship between the growth of the ratio of high to low skill median male hourly wages and growth of the ratio of high to low skill male workers for each of the 123 agglomerations, pooling data for 1991 and 2000.



We define migrants based on where people were living five years prior to the census, as in Card (2001). This information is recorded in the 1991 and 2000 censuses only. If individuals were living outside one of our 123 cities (agglomerations) five years ago, and are currently living in a city, then we count them as a rural–urban migrant. As Table 1 shows, the total number of high and low skill rural–urban migrants increased between 1980 and 2000, although the number of high skill migrants increased faster than low skill migrants. Figure 3 shows a positive relationship between the inflow of rural–urban migrants and the local supplies of high and low skill workers in cities.

4. EMPIRICAL RESULTS

In each row of Table 2 we present different versions of our model for robustness checks that will be discussed in section 4.2. Our baseline model is presented in row 1, while row 2, 3 and 4 we use a different instrument variable, in row 5 we use agglomeration size to weight each observation, while in row 6 we exclude São Paulo from our sample and in row 7 we redefine the skill split off from 9 years to 12 years of education.

Each column of Table 2 has a different specification. We begin, by estimating $-\frac{1}{\sigma_{EDU}}$ in Equation 6 controlling for changes in the relative productivity,$\Delta[ln a_H/(1-a_H)]$ with a trend. Our results, reported in column 1 of Table 2, imply an elasticity of substitution between high and low skill workers, σ_{EDU}, of 3.82, which is statistically different from the 1.4–2.4 reported in the literature survey of Ottaviano and Peri (2008). Because productivity factors may vary by industry, and industry composition varies by city, we include city fixed effects in column 2 to allow differences in worker productivity to vary across cities. Thus, our estimate of σ_{EDU} is now 1.19, which is not statistically different from previous results.

Table 2: Estimates of the inverse of the elasticity of substitution between high and low skill workers, $-\frac{1}{\sigma_{EDU}}$, in Equation 6, for different specifications.

<table>
<tr><td rowspan="4"></td><td colspan="6">Log ratio of high to low skill hourly wages</td></tr>
<tr><td colspan="4">OLS</td><td colspan="2">IV</td></tr>
<tr><td colspan="2">Level</td><td colspan="4">First difference</td></tr>
<tr><td>(1)</td><td>(2)</td><td>(3)</td><td>(4)</td><td>(5)</td><td>(6)</td></tr>
<tr><td>Basic regression</td><td>-0.262
(6.10)**</td><td>-0.838
(14.77)**</td><td>-1.007
(17.32)**</td><td>-1.159
(16.69)**</td><td>-1.730
(9.98)**</td><td>-1.771
(13.13)**</td></tr>
<tr><td>SPIV-1</td><td></td><td></td><td></td><td></td><td>-1.730
(13.88)**</td><td>-1.751
(13.55)**</td></tr>
<tr><td>SPIV-2</td><td></td><td></td><td></td><td></td><td>-3.817
(1.10)</td><td>-1.635
(15.44)**</td></tr>
<tr><td>SPIV-3</td><td></td><td></td><td></td><td></td><td>-1.729
(13.86)**</td><td>-1.752
(13.53)**</td></tr>
<tr><td>Weighted</td><td>-0.177
(2.75)**</td><td>-0.932
(7.98)**</td><td>-1.163
(10.13)**</td><td>-1.407
(6.71)**</td><td>-1.830
(7.93)**</td><td>-2.078
(10.98)**</td></tr>
<tr><td>Excluding São Paulo</td><td>-0.264
(6.03)**</td><td>-0.838
(14.71)**</td><td>-1.006
(17.24)**</td><td>-1.156
(16.63)**</td><td>-1.733
(9.86)**</td><td>-1.770
(13.03)**</td></tr>
<tr><td>12 years of education</td><td>-0.318
(8.17)**</td><td>-0.939
(16.00)**</td><td>-0.992
(15.23)**</td><td>-1.052
(11.81)**</td><td>-1.762
(3.35)**</td><td>-1.814
(11.52)**</td></tr>
<tr><td>Trend</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td></tr>
<tr><td>City (agglomeration) fixed effects</td><td>No</td><td>Yes</td><td>No</td><td>Yes</td><td>No</td><td>Yes</td></tr>
<tr><td>Observations</td><td>360</td><td>360</td><td>237</td><td>237</td><td>237</td><td>237</td></tr>
</table>

Note: High skill workers are individuals with 9 or more years of education except in row 6, where the threshold is moved to 12 or more years of education. Our instrumental variable (IV) is described in Sub-section 4.1, while SPIV-1, SPIV-2, and SPIV-3 are three versions of the supply–push instrumental variable. All regressions include a trend. In first difference regressions, this is done by including a constant. The sample includes men only between the ages of 15 and 55 years. The T and Z-statistics in parentheses are calculated using clustered-robust standard errors. * is significant at the 5% level; ** is significant at the 1% level

We obtain the same results when we estimate the model in first difference as in Equation 7, as shown in column 3, which includes a constant as the trend. Finally, as suggested by Borjas et al. (2011), we allow each city to have a different path of productivity change over time by estimating a random growth model (Wooldridge, 2010) with the inclusion of city fixed effects. Thus, our results in column 4 become statistically different from those in the international migration literature, with an implied σ_{EDU} of 0.86.

As pointed out in Section 3, despite the fact that we are using men only in our sample, the supply of workers still depends on wages, which implies that reverse causality is possible and causes our estimates of $-\frac{1}{\sigma_{EDU}}$ to be biased toward zero, or σ_{EDU} to be biased upward. Furthermore, as suggested by Borjas (2006) and Wozniak and Murray (2012), rural migrants may displace urban dwellers who migrate to other cities, which would also bias our estimates of $-\frac{1}{\sigma_{EDU}}$ toward zero. We address these issues with an instrumental variable approach, as discussed in the next sub-section and show our results in columns 5 and 6 of Table 2.

4.1 Instrumental variable

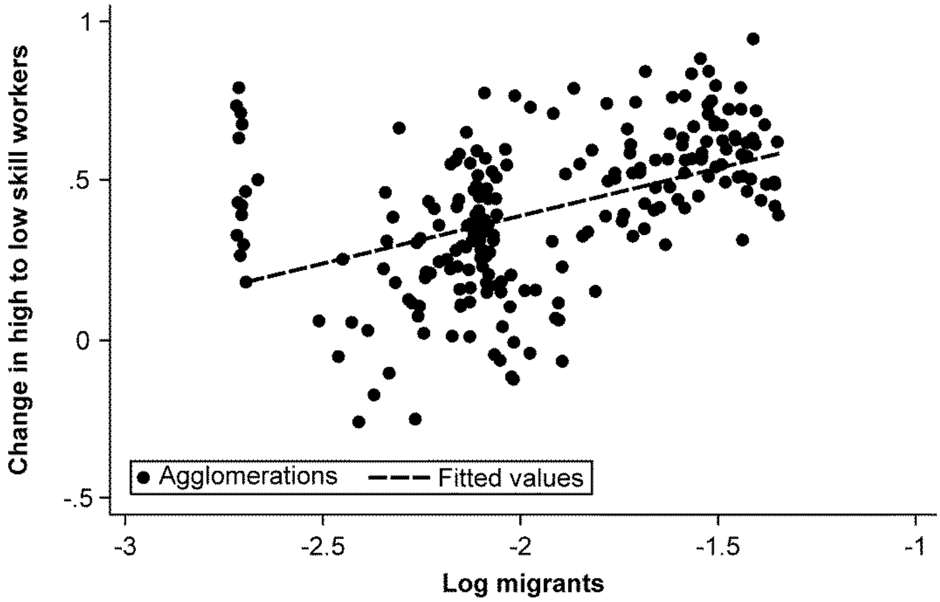
As explained in Section 2, the number of workers in a city includes both migrants and non-migrants. Therefore, we can use migration flows as an instrumental variable for changes in the number of workers in a city. Figure 4 shows a positive relationship between growth of the ratio of high to low skill male workers and the log ratio of high to low skill recent migrants. However, migration is itself an endogenous decision that may depend on wages in urban areas. Therefore, we must find an instrument for the migration flows themselves. The standard solution in the literature is to use a SPIV. This method splits migration flows into two components: (i) the decision to leave a foreign country or rural area; and (ii) the decision of where to go within the country, namely, which city. The validity of using this instrument relies on two assumptions. First, we must assume that the decision to leave a foreign country or rural area is driven only by conditions in the source of origin and not the destination and is, therefore, orthogonal to changes in conditions in the destination. Therefore, the endogeneity of migration flows derives from only the decision of where to move. To address this issue, we use historical patterns of migration so that they are not related to current conditions in the destination. This entails an underlying assumption that migrants are not forward looking (or, at least, forward looking to only some degree). The SPIV is then

$$SPIV_{irt}=\Omega_{ir}\Delta Mig_{it} \tag{8}$$

where ΔMig_{it} is the total inflow of recent migrants of a particular skill level *i* into the country or urban area in a particular year *t*, and Ω_{ir} is the historical pattern of settlement of migrants of a particular skill level *i*, measured by the fraction of migrants who moved to a region or city *r* in a time period before *t*.

In our approach, we drop these two assumptions and instead estimate the response of ΔMig_{it} to exogenous shocks in rural areas and let predetermined characteristics of rural areas determine Ω_{ir}.

Figure 4: Relationship between the growth of the ratio of high to low skill male workers and the log ratio of high to low skill recent migrants for each of the 123 agglomerations, pooling data for 1991 and 2000.



4.1.1 How many people migrate from rural areas

The 1991 and 2000 Brazilian population censuses include information on not only where people were living five years prior to the censuses, but also when people moved to their current areas of residence. We use this information to build out-migration flows for each rural area for each year in 1986–1990 and 1995–1999 for male individuals between the ages of 15 and 55 years. In Table 3, we show the average number of out-migrants each year from rural municipalities reported in the 1991 and 2000 censuses. There is a large percentage of observations with zero migrants. Therefore, we estimate a Tobit model with fixed effects using non-linear methods, as suggested by Honore (1992).

In its simplest specification, the decision to migrate depends on wages in both rural and urban areas and transport costs in the destination and origin. As Table 3 shows, more than 37% of low skill men living in rural areas worked in farming in both the 1991 and 2000 censuses. It is likely that their income from agriculture depends on the weather. We use data on precipitation from the Intergovernmental Panel on Climate Change (IPCC) (for a complete description of the dataset, see Mitchell et al., 2002,) in 1985–1990 and 1994–1999 to estimate the impact on out-migration of weather shocks in rural areas.

Table 3 shows that the number of high skill people living in rural areas increased 59% between 1991 and 2000, which is the same rate of growth in urban areas (Table 1) and is much faster than the growth rate of low skill people living in rural areas (15%). However, the rate at which high skill individuals migrated remained the same at 48% (calculated using Table 1). Furthermore, the average cost of moving dropped 14% between 1991 and 2000, as measured by the index of transport costs from rural municipalities to São Paulo constructed by Castro (2002)[5]. Beine et al. (2008), Docquier and Rapoport (2004), and Vidal (1998) argue that migration due to reduced transport costs leads to more individuals investing in education in the areas of origin Therefore, more high skill individuals become available in rural areas, and a larger percentage migrate out of rural areas. In addition, we point out that it is unlikely that people are leaving rural areas to acquire higher levels of education in urban areas because the number of university students in Brazil stayed relatively stable over the time period of our study (see Souza 2001). In particular, we show in Sub-section 4.2 that our results are the same when we redefine high skill workers as those with 12 or more years of education.

Our basic regression is then

$$\operatorname{Ln} Migrants_{i,rural,t} = \delta_0 + \delta_1 \ln N_{i,rural,(t-10)}$$

$$+\delta_2 Rain_{rural,t} + \delta_3 \Delta \ln Transp_{rural,t-5} + \delta_4 X_{rural,t} +$$

$$v_{i,rural,t} \qquad (9)$$

where $N_{i,rural,(t-10)}$ is the (lagged) number of men living in a rural area, *rural*, in the previous census year, $(t-10)$, with high or low skill, *i*; $Rain_{rural,t}$ is the (log) rainfall in *rural* in year *t* (because the timing of the drought and its impact is uncertain, we also check if lagged rainfall shocks impact on migration); $\Delta \ln Transp_{rural,t-5}$ is the growth rate of transport costs to São Paulo from *rural* in *t* (the relevant years between 1970 and 1985 and between 1985 and 1990); while $X_{rural,t}$ is a set control for characteristics of the rural area of origin, which include log agricultural area (in hectares) taken from the agricultural censuses of 1985 and 1995, year dummies, municipality fixed effects, and dummies for the skill groups. Our results in Table 4 are for our sample of men between the ages of 15 and 55 years. Because we include rural municipalities fixed effects, we can interpret our coefficients as responses to shocks (deviations from the average across periods). We find that rainfall shocks affect migration of low skill men only; in particular, a 1 standard deviation decrease in rainfall leads to an increase in out-migration of low skill men of 2.4%. In the drought area of the northeast, the impact of a drought is different. This region, as pointed out by Baer (2013), receives government aid in years of drought, although the aid is often misused. Therefore, during good years, when there are no government transfers, out-migration increases for both low and high skill men. Furthermore, we find that reductions in

[5] Castro (2002) construct an index, centered around 1,000, that measures the benefits of improvements in highway infrastructure in 1970–1995 as the change in equivalent paved road distance from each municipality to São Paulo city. It is based on a linear programming exercise that takes into account the design of the road network as well as the difference in vehicle operating costs between earth/gravel and paved roads.

transport costs lead to higher out-migration of high skill people; in particular, a 10% decrease in transport costs increases the number of high skill migrants by 2.3%.

Table 3: Basic statistics (mean and standard deviation) for 3,214 rural municipalities (minimal comparable areas), for high skill (9 or more years of education) and low skill (less than 9 years of education) for 1991 and 2000.

	1991		**2000**	
	Less than 9 years of education	**9 or more years of education**	**Less than 9 years of education**	**9 or more years of education**
No. of migrants	*144.38 (426.08)*	*27.56 (230.54)*	*128.33 (398.90)*	*42.61 (273.59)*
No. of non-migrants	*4,317.12 (19,581.85)*	*532.58 (8,158.37)*	*5,018.67 (17,728.05)*	*847.06 (12,835.33)*
Fraction of observations that are zero	*2.90%*	*36.67%*	*2.93%*	*24.70%*
Occupation				
Administrative	5.43%	30.84%	3.98%	20.96%
Technical or scientific	0.9%	16.81%	1.34%	16.25%
Farming	52.14%	7.52%	37.43%	6.71%
Mining	1.38%	0.33%	0.66%	0.2%
Industry	19.03%	12.97%	25.38%	17.51%
Commerce and trade	6.64%	12.64%	8.21%	14.38%
Transport	5.11%	4.46%	6.88%	6.12%
Services	3.73%	3.38%	6.27%	5.26%
Domestic services	0.49%	0.14%	0.94%	0.21%
Security and national defense	1.04%	6.3%	0.95%	6.16%
Other	4.59%	4.75%	8.91%	6.46%
Average agricultural area (in hectares)	*107,457.5 (471,670.6)*		*96,362.07 (453,752.7)*	
Average rainfall	*11.27 (4.192)*		*11.14 (4.346)*	
Average transport costs	*1,811.316 (1,436.84)*		*1,549.82 (1,125.901)*	

Note: Rows 4–14 show the percentage of people of each group working in each sector (low skill workers are concentrated in farming). Rows 15–17 include the basic characteristics of rural municipalities. The sample includes men only between the ages of 15 and 55 years.

Table 4: Estimates of the impact of rainfall shocks and growth in transport costs on out-migration from rural areas (3,214 rural municipalities or minimal comparable areas (MCAs)) for 1987–1991 and 1996–2000 for high and low skill men between the ages of 15 and 55 years.

OLS estimate of the impact of changes in transport costs and rainfall shocks on the log number of migrants from rural areas		
	Log migrants	
	Low skill	High skill
Log lag population	0.843 (11.66)**	0.134 (3.11)**
Log agricultural area (in hectares)	-0.064 (2.05)*	0.0602 (1.26)
Growth rate of transport costs to São Paulo	-0.028 (0.50)	-0.226 (2.35)*
Average monthly rainfall	-0.0058 (2.36)*	-0.0098 (1.89)
Previous year's average monthly rainfall	-0.0016 (0.58)	-0.0103 (1.91)
Average monthly rainfall in semi-arid area	0.01605 (2.82)**	0.084 (5.52)**
Last year's average monthly rainfall in semi-arid area	0.0302 (4.81)**	0.075 (4.94)**
Municipality (MCA) fixed effects	Yes	Yes
Year dummies	Yes	Yes
Number of municipalities	3,213	3,209
Observations	25,704	25,586
Censured observations	753	7,944
F-statistic	32.81	55.05
	635.28	

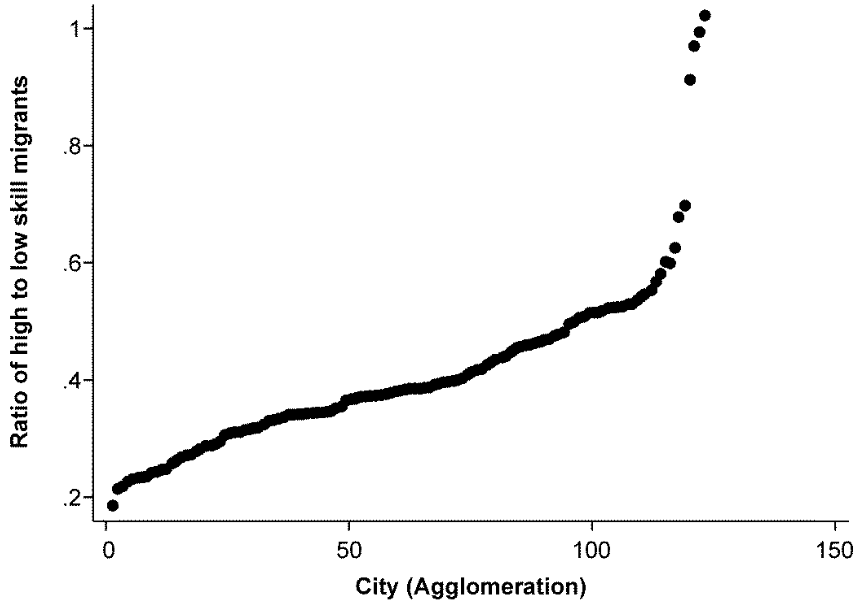
4.1.2 Where migrants go

The 1991 census provides information on where each person was living in 1981 and when he or she moved to the current municipalities. We use this information to determine historical patterns of migration, in particular, how people moved from 3,207 rural municipalities to 123 urban areas between 1981 and 1985. Figure 5 shows that high and low skill people do not choose the same destinations and that there is a large amount of variation across cities. In particular, distance between origin and destination (Our measure of distance is the great circle distance between the center of the municipality (MCA) of origin and the city (agglomeration), using the median radius of the earth) affects the cost of transport, and therefore, the number of migrants deciding to move to a given location. Therefore, we run the following regression.

$$\frac{Migrants_{i,rural,urban}}{\sum_c Migrants_{i,rural,c}} = \eta_0 + \eta_1 Distance_{rural,urban} + \eta_2 W_{i,rural,urban} + \varpi_{i,rural,urban} \qquad (10)$$

The dependent variable $\frac{Migrants_{i,rural,urban}}{\sum_c Migrants_{i,rural,c}}$ is the share of migrants from a rural area moving to an urban area, *c*, for each group (high and low skill, *i*).

Figure 5: Historical patterns of migration.



Note: Plot of the ratio of high to low skill migrants from 1981 to 1985 for 123 agglomerations sorted from smallest to largest ratio on the horizontal axis.

Table 5 shows our results. As columns 1 and 2 indicate, migrants are more likely to move to cities that are closer to the origin, although distance matters less for high skill men. In particular, a city that is 10% closer receives 0.2 percentage points more rural migrants. In columns 3 and 4, we check whether this result is driven by the supply of high or low skill men in areas nearer to the cities of destination by controlling for how many people live in the rural and urban areas. Despite the fact that the supply of high and low skill men matters, distance remains statistically significant for every group.

Table 5: Estimates of the impact of distance on the likelihood of moving from a particular rural area to one of the 123 agglomerations (cities).

Estimates for distance as an explanation for migration location decision				
	Percentage Rural–Urban Migrants			
	(1)	(2)	(3)	(4)
Log distance	-0.02003 (82.36)**	-0.022 (84.80)**	-0.023 (88.55)**	-0.022 (86.27)**
Log distance (for high skill)		0.0033 (13.23)**	0.0062 (21.20)**	0.0041 (14.82)**
(Log) People living in rural area* (log) People living in urban area			0.00103 (80.54)**	0.0012 (30.07)**
Log number of people living in rural area				-0.0094 (25.82)**
Log number of people living in urban area				0.00044 (1.52)
Dummy for high skill	Yes	Yes	Yes	Yes
Municipality (MCA) fixed effects	Yes	Yes	Yes	Yes
Number of rural municipalities (MCAs)	3,214	3,214	3,214	3,214
Observations	790,644	790,644	787,815	787,815
F-statistic	6,782.95	3,730.08	4,064.25	3,900.55
R-squared	0.05	0.05	0.07	0.07

Note: The sample includes men only between the ages of 15 and 55 years. T-statistics in parentheses are calculated using clustered-robust standard errors. * is significant at the 5% level; ** is significant at the 1% level.

4.1.3 Instrumental variable results

After determining how people respond to droughts and changes in transport costs in rural areas, and where people decide to go related to the distance between origin and destination, we can build an exogenous migration shock for each city, *r*, by skill group, *j*, for a particular year, *t*, which is exogenous to changes in cities. Therefore, our instrument is as follows

$$\left[\Delta\widehat{N_{\iota,c,t}}\right]^{mig} = \sum_{j=t-5}^{t}\sum_{rural}\left(\frac{\widehat{Migrants_{\iota,rural,urban}}}{\sum_c Migrants_{\iota,rural,c}}\right)\widehat{Migrants}_{\iota,rural,j} \qquad (11)$$

We argue that an increase in the availability of high versus low skill workers will lead firms to hire more high skill workers, which will affect the relative wages between these two groups. Therefore, we use the ratio of the flow of high to low skill migrants as an instrument for the change in the ratio of high to low skill workers.

Row 1 of Table 6 shows the results for the first stage of our basic regression. In both columns 1 and 2 (with city specific trends), we see a positive and statistically significant coefficient with a corresponding F-statistic well above 10.

Columns 5 and 6 of row 1 in Table 2 show our results from the second stage. As expected, the coefficient is larger in absolute terms, which is consistent with the possibility of reverse causality, and is also consistent with the outflow of urban residents to other urban areas, which implies that our estimates of σ_{EDU} are even smaller than before. In particular, the elasticity of substitution of high to low skill workers, σ_{EDU}, of only 0.56–0.58, is significantly smaller than that previously found in the literature and implies that high and low skill workers are gross complements rather than gross substitutes in production.

4.2 Robustness checks

Because our results are statistically different from those previously found, our first concern is the validity of our instrument. Therefore, we begin by running the same regressions but using the standard SPIV. We construct three versions of this instrument in Equation 8. In the first version, SPIV-1, we pool all rural municipalities as if all rural areas where the same. In the second version, SPIV-2, we construct our instrument for the flow of migrants from each of the rural areas, and then add them up to obtain the inflow into each specific city. The third version, SPIV-3, is similar to SPIV-1, except that we exclude the own city contribution to the migration flows, $\Delta\left[Mig_{it}-Mig_{irt}\right]$, as in Wozniak and Murray (2012). Our results are shown in the first column of rows 2, 3 and 4 of Table 6 (first stage results) and the second column of the same rows (second stage results). When the SPIV is correlated with the endogenous variable, the elasticity of substitution between high and low skill workers, σ_{EDU}, is not statistically different from our basic regression. It ranges between 0.57 and 0.61 for those regressions that do not have weak instruments.

Another potential source of concern regarding our estimates is the larger variation in city size in our sample (see the standard deviation in the number of high and low skill workers in urban areas in Table 1). Furthermore, this period of time saw an increase of urban-urban migration as pointed out by Lima et al (2016). As such, the impact of rural-urban migration could be biased downwards, as pointed out by Borjas (2003). To address both these problems, we run a weighted regression, in which we weight each observation by city size. We can interpret the results from these regressions as the impact of rural–urban migration on the urban system (rather than on the average city). The results are presented in row 5 of Table 6 (first stage results) and in row 5 of Table 2 (OLS and second stage results). Again, our estimates of σ_{EDU} are not statistically different from that found before and range between 0.48 and 0.55.

Another possible reason for the difference between our results and those of other researchers comes from the potential bias from the inclusion of São Paulo in our sample. Based on our calculation using 1991 and 2000 census data, São Paulo receives a disproportionate share of 15% of rural–urban migrants in Brazil, and is a disproportionately large city in Brazil. Furthermore, we use the distance to São Paulo from rural areas to construct our instrument, and thus, it may be correlated with the error term. Therefore, we run our regressions excluding the city of São Paulo and report our results in row 6 of Table 6 (first stage results) and in row 6 of Table 2 (OLS and second stage results). Again, our estimates of σ_{EDU} are not statistically different from our previous findings and range between 0.56 and 0.58.

We use nine years of education as the split between high and low skill workers because of Brazil's low levels of education. However, previous studies have used 12 years of education as the distinction. Therefore, we redefine high skill workers as those with 12 or more years of education and low skill workers as those with less than 12 years of education. Our results are reported in row 7 of Table 6 (first stage results) and row 7 of Table 2 (OLS and second stage results). Again, our estimates of σ_{EDU} are not statistically different from our previous findings and range between 0.57 and 0.58.

Finally, we conduct other robustness checks to address issues that are specific to Brazil but do not report these results as there is no comparison to be made in the literature. A major difference between developed and developing countries is the existence of large informal markets in developing countries. This may affect our estimates in two ways. First, wages will not reflect full compensation in the formal sector; however, if there are compensating wage differentials, they do reflect full compensation in the informal sector. Therefore, changes in the size of each sector across time would bias our estimates. Furthermore, if there is market segmentation and wages in the informal sector do not compensate individuals for the absence of other benefits, then workers in the formal and informal sectors are imperfect substitutes in production, and thus, our production function specification needs to include this. We rerun the regressions in the formal and informal sector that we obtain in the second stage of our instrumental variable approach with city dummies; this gives an estimate of σ_{EDU} of 0.45 for the formal sector and 0.82 for the informal sector, both of which are not statistically different from our previous results.

Another major difference between Brazil and countries in the existing literature is the difference between public and private sector labor incomes, which are not reflected in wages.

For instance, public sector employees benefit from more stable jobs and better pension schemes (Emilio et al., 2012). Nevertheless, previous research has found that returns to education are higher in the public sector (Tannen, 1991) and that workers gain on average 3.9% more in wages when they transfer from the private to the public sector (Emilio et al., 2012). Therefore, we run the same regressions without workers in the public sector[6]. Our estimate of σ_{EDU} in the second stage of our instrumental variable approach with city dummies, of 0.37, is not statistically different from our previous results.

Table 6: First stage estimates of the instrument on the log of recent migrants for different specifications described in Sub-sections 4.1 and 4.2.

First stage of IV - Ratio of workers and migration		
	Change in log ratio of high to low skill workers	
	(1)	(2)
Basic regression	0.296 (43.45)	0.533 (126.69)
SPIV-1	0.542 (144.29)	0.541 (137.20)
SPIV-2	0.0093 (1.02)	0.628 (143.32)
SPIV-3	0.542 (144.00)	0.541 (65.12)
Weighted	0.292 (31.30)	0.444 (55.48)
Excluding SP	0.294 (42.68)	0.533 (124.79)
12 years of Education	0.078 (3.74)	0.426 (34.86)
Trend	Yes	Yes
City fixed effects	No	Yes
Year dummies	No	No
Observations	237	237

Note: High skill is defined as individuals with 9 or more years of education, except in row 6, where the threshold is moved to 12 or more years of education. SPIV-1 and SPIV-2 are two versions of the standard supply–push instrumental variable. The sample includes men only between the ages of 15 and 55 years. F-statistics are in parentheses from clustered-robust standard errors.

Finally, Borjas et al. (2011) argue that trends do not correctly control for changes of specific productivity over time. Therefore, we run the same regression, including year dummies, to control for changes in productivity over time. While our point estimates remain unchanged, our second stage instrumental variable coefficients are no longer statistically significant.

5. SIMULATION RESULTS

Table 1 shows that the wage gap between high and low skill men decreased by 23% between 1991 and 2000. In this section, we show how increases in the relative supply of high to low skill workers in urban areas, driven by rural migration, affect this wage skill premium. The ratio of high to low skill workers increased from 40% in 1991 to 69% in 2000 in urban areas, while the ratio of high to low skill rural migrants increased from 37% in 1991 to 47% in 2000 (Table 1), despite the fact that in rural areas the ratio of high to low skill non-migrant only increased from 12% in 1991 to 16% in 2000 (Table 3).

We use our estimates of σ_{EDU} to simulate the impact on the wage skill premium from the relative increase in the number of high skill male workers due to rural migration. Following Borjas (1999), we use Equation 6 and redefine $N_{it} = Loc_{it}(1 + m_{it})$, where

[6] We exclude only workers employed directly by the government because we cannot identify state owned enterprises. However, the number of these state owned enterprises started to decline after 1979 with successive privatization initiatives (Baer, 2013), so that total employment in state owned enterprises declined to 10% of the total workforce in 2008 (OECD, 2011).

$m_{it} = \frac{Mig_{it}}{Loc_{it}}$ is the ratio of the stock of migrants of skill i at year t to the number of non-migrants local workers, Loc , of skill i in the same year t.

$$\ln\left(\frac{w_H}{w_L}\right) = \ln\left[\frac{a_H}{(1-a_H)}\right] - \frac{1}{\sigma_{EDU}}\ln\left[\frac{Loc_{Ht}}{Loc_{Lt}}\right] - \frac{1}{\sigma_{EDU}}[\ln(1+m_{Ht}) + \ln(1+m_{Lt})] \quad (12)$$

If we assume as in Borjas (1999) that there are no changes in $\frac{a_H}{(1-a_H)}$ over time and that the numbers of high and low skill non-migrants and high and low skill workers remain constant over time, then using the approximation $\ln(1+m_{it}) \approx m_{it}$, we can obtain the first difference as follows.

$$\Delta\ln\left(\frac{w_H}{w_L}\right) = -\frac{1}{\sigma_{EDU}}\ln[\Delta m_{Ht} - \Delta m_{Lt}] \quad (13)$$

This allows us to simulate the contribution of rural–urban migration on the changes in relative wages. Because we assume that the number of non-migrants did not change, then $\Delta mit{=}\Delta Migit/Loci{,}t{-}10$, where ΔMig_{it} is the flow of rural–urban migrants into the city between years, which is reported in Table 1, and $Loci{,}t{-}1$ is the number of local male workers in the city from the previous census year in the same table. Surprisingly, $\Delta m_{Ht} - \Delta m_{Lt}$ is only 0.6%, which implies that migrants have very similar observable characteristics to local non-migrant city residents.

Using our initial estimate of the elasticity of substitution between high and low skill workers, σ_{EDU}, of 0.56, we find that the inflow of rural migrants into cities in Brazil leads to only a 1.1% reduction in the wage gap between high and low skill workers, which explains only 5% of the variation in the data.

By comparison, if we were to use the estimate of σ_{EDU} from the international migration literature, that is, 1.5 (Ottaviano and Peri, 2008), then the inflow of rural migrants into urban cities leads to only a 0.4% reduction in the high to low skill wage gap, which explains 1.8% of the variation in the data. Therefore, use of the international migration literature parameters greatly underestimates the impact of internal migration in the order of more than half.

These results clearly indicate that, although the inflow of rural migrants into cities in Brazil reduces income inequality, which is consistent with the findings of Borjas et al. (1992) and Ottaviano and Peri (2008), this inflow is not the main driver of the changes in the wage skill premium in Brazilian cities.

Perhaps more importantly, use of the elasticity of substitution between high and low skill workers from the international migration literature would bias our results downward even further.

6. CONCLUSION

Internal migration flows, particularly in developing countries, are just as large as international migration flows. While there is an extensive body of literature on the impact of international migration on receiving communities, the literature on internal migration is much smaller, in particular regarding the impact on wage skill premiums. Analysis of this issue is important as there is much debate about how to decrease inequality within public policy. In particular, regional policies that attempt to reduce interregional migration often target investment in transport infrastructure in order to attract workers from surrounding areas and promote agglomeration economies. This increased internal movement of workers can lead to increased intraregional inequalities, which must be taken into consideration when designing regional policies.

This study showed that rural–urban migration contributed to reducing the wage skill premium, a measure of income inequality, by 1.1% in Brazilian cities between 1991 and 2000. However, this explained only 5% of the actual variation in the data, suggesting that rural–urban migration is not the main driver of the decline of the wage skill premium. This is because, in Brazil, rural–urban migrants have similar observable characteristics to local urban residents and despite the fact that the elasticity of substitution between high and low skill workers in Brazil is up to a third smaller than that found in the international migration literature. This implies that changes in the relative supply of high and low skills workers have a much larger impact on the wage skill gap in Brazil than in the US or UK.

The fact that we found that high and low skill workers are gross complements in production $0 \leq \sigma_{EDU} < 1$, implies that, unlike what has previously been found for the US by Kiley (1999), Acemoglu (2002), Klump et al. (2007), and Leon-Ledesma et al. (2010), skill biased technological progress over this time period is one of the drivers in the reduction in the wage skill premium. As pointed out by Acemoglu (2002), an improvement in the relative productivity of skilled workers increases the demand for both high and low skill workers so that wages for low skill workers increase by more than those of high skill workers, reducing the wage skill premium. As a result, recent public policies investing in human capital can potentially have more benefits in Brazil than usually believed. In particular, such policies can not only increase income of low skill workers but also benefit high skill workers, while also reducing income inequality.

A further contribution of this study is that it addressed the potential endogeneity of migration flows by using exogenous push factors from rural areas (rainfall shocks and reduction in transport costs) to construct instrument variables for migration flows into urban areas. Our results were consistent with those using the standard SPIV technique of the migration literature. Given the fact that each technique relies on different assumptions for identification, it is particularly reassuring for our results that we addressed potential endogeneity problems correctly.

REFERENCES

Acemoglu, Daron. 2002. Technical change, inequality, and the labor market. J. Econ. Lit. 40(1), 7 ¬ 72.

Baer, Werner. 2013. The Brazilian Economy: Growth and Development, seventh ed. Westport, CT: Lynne Rienner Publishers

Baer, Werner. 2013. The Brazilian Economy:Growth and Development, seventh ed.Westport, CT: Lynne Rienner Publishers

Beine, Michel, Docquier, Frederic, Rapoport, Hillel. 2008. Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers. Econ. J. 118 631¬ 652.

Borjas, George. 1999. Handbook of Labor Economics Vol. 3. Elsevier B.V. Chap. The economic analysis of immigration, 1697–1760.

Borjas, George. 2003. The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. Q. J. Econ. 118(4), 1335–74.

Borjas, George J. 2006. Native internal migration and the labor market impact of immigration. J. Hum. Resour. 41(2), 221–258.

Borjas, George J., Bronars, Stephen G., Trejo, Stephen J. 1992. Self-selection and internal migration in the United States. J. Urban Econ. 32(2), 159–185.

Borjas, George J., Grogger, Jeffrey, Hanson, Gordon H. 2011. Substitution between immigrants, natives, and skill groups. National Bureau of Economic Research Working Paper no.17461.

Boustan, Leah Platt, Fishback, Price V., Kantor, Shawn. 2010. The effect of internal migration on local labor markets: American cities during the Great Depression. J. Labor Econ. 28(4), 719¬746.

Camarano, Ana Amélia, Abramovay, Ricardo. 1997. Êxodo rural, envelhecimento e masculinização no Brasil: Panorama dos ùtimos cincoenta anos. Anais do 1° Encontro Nacional Sobre Migração.

Card, David. 2001. Immigration inflows, native outflows, and the local market impacts of higher immigration. J. Labor Econ. 19(1), 22–64.

Card, David. 2009a. How immigration affects US cities, in: Inman, Robert. P (Ed.), Making Cities Work: Prospects and Policies for Urban America, Princeton University Press, pp. 158–200.

Card, David. 2009b. Immigration and Inequality. Am. Econ. Rev. 99(2), 1–21.

Castro, Newton. 2002. Custos de Transporte e Produção Agrícola no Brasil: 1970–1996. Agricultura em São Paulo, 49, 87–109.

Corseuil, Carlos, Foguel, Miguel. 2002. Uma sugestão de deflatores para rendas obtidas a partir de algumas pesquisas do IBGE. Instituto de Pesquisa Economica Aplicada Working Paper no. 897.

Docquier, Frederic, Rapoport, Hillel. 2004. Skilled migration: the perspective of developing countries. World Bank Policy Research Working Paper Series, no. 3382.

Emilio, Daulins, Ponczek, Vladimir, Botelho, Fernando. 2012. Evaluating the wage differential between public and private sectors in Brazil. Revista de Economia Politica, 32(1), 72–86.

Freire, Tiago. 2010. Essays in Applied Economics. Ph.D. thesis, Brown University.

Gropello, Emanuela, Sakellariou, Chris. 2010. Industry and skill wage premiums in East Asia. World Bank Policy Research Working Paper Series, no. 5379.

Hanewinkel, Vera, Oltmer, Jochen. 2012. Internal Migration in China: Opportunity or Trap? Policy Brief 19, Institute for Migration Research and Intercultural Studies (IMIS) of the University of Osnabrück.

Honore, Bo E. 1992. Trimmed lad and least squares estimation of truncated and censored regression models with fixed effects. Econometrica, 60(3), 533–565.

Katz, Lawrence, Murphy, Kevin. 1992. Changes in the wage structure 1963–87: Supply and demand factors. Q. J. Econ. 107(1), 35–78.

Kiley, Michael T. 1999. The supply of skilled labour and skill-biased technological progress. Econ. J. 109(458), 708–724.

Klump, Rainer, McAdam, Peter, Willman, Alpo. 2007. Factor substitution and factor-augmenting technical progress in the United States: A normalized supply-side system approach. Rev. Econ. Stat. 89(1), 183–192.

Leon-Ledesma, Miguel A., McAdam, Peter, Willman, Alpo. 2010. Identifying the elasticity of substitution with biased technical change. Am. Econ. Rev. 100(4), 1330–1357.

Lima, A.C.C., Simões, R. and Hermeto, A.M., 2016. Desenvolvimento regional, hierarquia urbana e condição de migração individual no Brasil entre 1980 e 2010. EURE (Santiago), 42(127),.29-54.

Lu, Zhigang, Song, Shunfeng. 2006. Rural–urban migration and wage determination: The case of Tianjin, China. China Econ. Rev. 17(3), 337–345.

Manacorda, Marco, Manning, Alan, Wadsworth, Jonathan. 2012. The impact of immigration on the structure of wages: Theory and evidence from Britain. J. Eur. Econ. Assoc. 10(1), 120–151.

Mata, Daniel, Deichmann, Uwe, Henderson, J. Vernon, Lall, Somik V., Wang, Hyoung G. 2007. Determinants of city growth in Brazil. J. Urban Econ. 62(2), 252–272.

Meng, Xin, Zhang, Junsen. 2001. The two-tier labor market in urban China: Occupational segregation and wage differentials between urban residents and rural migrants in Shanghai. J. Comp. Econ. 29(3), 485–504.

Mitchell, Thimothy, Hulme, Mike, New, Mark. 2002. Climate data for political areas. Area, 34, 109–112.

Organization of Economic Cooperation and Development (OECD). 2011. Government at a Glance 2011. OECD Publishing.

Ottaviano, Gianmarco, Peri, Giovanni. 2012. Rethinking the effect of immigration on wages. J. Eur. Econ. Assoc. 10(1), 152–197.

Ottaviano, Gianmarco I.P., Peri, Giovanni. 2008. Immigration and national wages: Clarifying the theory and the empirics, National Bureau of Economic Research Working Paper no. 14188.

Peri, Giovanni. 2011. Rethinking the area approach: Immigrants and the labor market in California. J. Intern. Econ. 84(1), 1–14.

Peri, Giovanni. 2012. The effect of immigration on productivity: Evidence from US states. Rev. Econ. and Stat. 94(1), 348–358.

Puga, Diego. 2002. European regional policies in light of recent location theories. J. Econ. Geogr. 2, 373–406.

Reis, Eustáquio, Pimentel, Márcia, Alvarenga, Ana I. 2009. Áreas mínimas comparáveis para os períodos intercensitários de 1872 a 2000. Rio de Janeiro, Instituto de Pesquisa Economica Aplicada-Diretoria de Estudos e Politicas Macroeconomicas.

Souza, Paulo Renato. 2001. Education and development in Brazil, 1995–2000. CEPAL Review, 73, 65–80.

Tannen, Michael B. 1991. New estimates of the returns to schooling in Brazil. Econ. Educ. Rev. 10(2), 123–135.

Vidal, Jean-Pierre. 1998. The effect of emigration on human capital formation. J. Popul. Econ., 11(4), 589–600.

Vieira, A.D.Á., Nizzola, L.J., Kallas, L.M.E., Brito, M.F., Schvasberg, B. and de Faria, R.S., 2012. Estudos recentes sobre a rede urbana brasileira: diferenças e complementaridades. Revista Brasileira de Estudos Urbanos e Regionais, 13(2), .55-70.

Wooldridge, Jeffrey M. 2010. Econometric Analysis of Cross Section and Panel Data. MIT press.

Wozniak, Abigail, Murray, Thomas J. 2012. Timing is everything: Short-run population impacts of immigration in US cities. J. Urban Econ. 72(1), 60–78.