Skill Wage Gap in Brazil: 1980-2000

Tiago Freire*

Department of Economics, National University of Singapore

May 13, 2011

Abstract

It is generally accepted that migration will lead an increase in income. However the question is how will income be distributed across individuals in society? If migrants have lower education levels, when compared to current urban workers, then the inflow of migrants will increase the skill wage gap in urban areas. Previous work on this topic has focused on international migration in developed countries. To the best of our knowledge this is the first study to look at the impact of rural-urban migration on city wages. Our results contribute to the evaluation of regional policies, as recent research has found that regional policies can lead to an increase in the number of rural to urban migrants.

We use data the Brazilian Census for 1980 to 2000 to estimate the elasticity of substitution between high and low skill workers. We instrument for the change in the ratio of high to low skill workers, with rural-urban migrants (driven by rainfall shocks in rural areas). Finally, in our simulations we show that migration lead to a 8.5% decrease in the wage gap between high and low skill workers, in Brazil, between 1991 and 2000.

Keywords: Labor Markets; Labor Migration; Rural Urban; Wages.

JEL Classification Numbers: J31, J61, O15, O18, R12.

*Address: Department of Economics, National University of Singapore, 1 Arts Link, AS2 #04-39, Singapore 117570, telephone: +65-65-168910, e-mail: *tiago_freire@nus.edu.sg*. The author is grateful to Rachel Friedberg, Andrew Foster and Sriniketh Nagavarapu and the participants of the Applied Microeconomics Seminar at Brown University for advice and suggestions.

1 Introduction

A substantial amount of resources is spent by both developed and developing countries on regional policies in an effort to reduce attenuate regional inequalities. For instance, regional policies account for over 30% of the EU budget between 2000-06(Puga [2002]). In Brazil, 9.3% of federal government revenues of 1992 came from the Northeast, while expeditures in this region where 14.7% of of the federal budget (Baer [2008]). While a substantial amount of effort has been put towards trying to understand the effectiveness of regional policies in reducing inequalities, little effort has been put forward in attempting to understand the consequences of regional inequalities. For instance, according to Lall et al. [2006], there are few studies looking at the impact of internal migration on urban areas.

We apply a "general equilibrium" approach similar to Card [2001], Borjas [2003] and Ottaviano and Peri [2007] to estimate the elasticity of substitution between high and low skill workers, and simulate the impact of the relative increased of high skilled people on the high to low skill wage gap. In order to solve the endogeneity problem we use weather shocks in rural areas and changes in transportation costs from rural to urban areas, as an instrument for changes in employment of both high and low skill workers. Previous work, such as Card [2001], Card [2009] and Basu [2010], try to control for the endogeneity of migration flows by using historical patterns of migration between origin and destination. However this may not be enough. First, if the historical pattern used is part of a trend then the instrument is not valid. Second, it is still possible that wage shocks in a particular destination (or city) can lead to a larger migration inflow, biasing estimates.

We find an elasticity of substitution between high and low skill workers of 0.54, much smaller that reported by Katz and Murphy [1992], Borjas [2003] and Ottaviano and Peri [2008]. We use this result to simulate the impact of the relative increase of high skill labor on urban wages, between 1991 and 2000 and find that it is responsible for a 8.5% decrease in the skill wage gap.

The paper is organized as follows. In section 2 we introduced the data, and look at the peculiarities of Brazilian Census. In section 3 we look at the empirical framework and results and in section 4 we present the simulation of the impact of migration on the wage gap. In section 5 we conclude.

2 Wages, Employment and Rural-Urban migration in Brazil: 1980 to 2000

We use the 1980, 1991 and 2000 Brazilian Census. The 1980 Census is a 25% sample of the population, while the 1991 and 2000 Census sample 10% of the population in municipalities with a predicted population of more than 15,000 and a sample of 20% in the remaining municipalities. We use the 1980 Census solely to provide a baseline for hourly wages and employment in each agglomeration.

In Brazil, municipalities are the lowest level of Government, according to the 1988 Constitution, with autonomy and its own sources of revenue, which allow an increasing number of public services to be provided. Municipalities have changed considerably over time, increasing from 3951 in 1970 to 5501 in 2000 as pointed out by Reis et al. [2007]. Therefore, to compare municipalities across time we use Minimal Comparable Areas (or MCAs), developed by IPEA to compare municipalities between 1970 and 2000. In figure 15 we have a map of the 3661 MCAs used in our work.

We define cities as agglomerations defined in da Mata et al. [2005]. These agglomerations are based on 1991 MCAs and can be used for comparison of data from 1970 to 2000 The advantage of using this definition is that it has characteristics similar to the U.S. Metropolitan Statistical Areas. Furthermore, IBGE's (Brazilian Statistics Bureau) alternative definition of an urban area, has urban areas in all municipalities in Brazil, even those deep in the Amazon, as pointed out by Camarano and Abramovay [1997]. In figure 15 we have a map of the 123 agglomerations (composed of 447 MCAs) for Brazil.

As you can see from figure 1 and 2, the average education level in Brazil is low. In particular, note from table 1, that men in rural areas have the lowest level of education, with only 3.3 years of education in 1980, increasing to 5.4 years in 2000, while women in urban areas are the best educated group, with 5.4 years of education, while in 1980, increasing to 8.7 years in 2000 (despite the fact that compulsory education now comprises 9 years of schooling). The problem we face is determining what is the appropriate definition of high and low skill, since depending on our definition, creates an implicit assumption that workers within each category (high and low skill) are perfect substitutes. While Card [2007] defines skill using information on occupation within industrial sector, a more common definition, used by Katz and Murphy [1992], Borjas [2006], Ottaviano and Peri [2008] and Peri [2007], relies on years of education. In these papers, low skill is often defined as individuals who graduated from high school or less, while high skill includes people with at least some college. Because the average level of education is so low in Brazil and compulsory education consists of 9 years of education, we define low skill as people having less than 9 years of education (people who didn't finish compulsory education), and high skill as people having 9 or more years of education (people who have complete compulsory education). We will conduct robustness checks, where we use 7 and 12 years of education.

Due to the low level of education in Brazil several individuals join the job market at an early age. In fact, IBGE considers people over 10 years of age as active in the workforce. Also, since 1998, workers in the private sector workers can retire at 60, and women at 55 (Medici [2004]). Therefore for our analysis we consider active workers as people from 15 to 55 years of age (inclusive), who report having labor income and number of hours worked and are not students. From table 1 you can see a small increase in the number of male workers (driven primarily by the increase in population, rather than the increase in labor participation). The number of female workers on the other hand has increased significantly, driven mainly by an increase in labor supply (see Freire [2010]). We will therefore focus on male workers, but also report our results for a sample pooling both men and women.

Since between 1980 and 2000 Brazil went through a period of hyper inflation and changed currency twice, 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. The average number of hours worked is 41 hours in both rural and urban areas in 1991 and increased to 43 hours in 2000 (hours worked not presented). Real hourly wages are not surprisingly higher (double) in urban areas when compared with rural areas, and have increased slowly from 1991 to 2000 (relatively more in rural areas). Hourly wages for each group are presented in table 1.

In figure 3 and 4 we can see a negative correlation between the ratio of (high to low skill) wage and the ratio of (high to low skill) workers for 123 cities in 1991 and 2000. As the ratio of high to low skill workers increases, the ratio of high to low skill wage decreases, consistent with our hypothesis. In particular, in the next section we show how an increase in the relative supply of a particular type of labor leads to a decrease in the relative wage of that group.

As we pointed out before, the problem with estimating this relationship directly is that employment, and in particular, labor participation of each group is also driven by wages. We therefore use rural urban migrants as an exogenous supply shock to estimate the relationship between employment and wages. As the number of high skilled people living in a city increases, the number of high skilled workers increases. In figure 5 and 6 we have a positive correlation between the ratio of (high to low skill) workers and the ratio of (high to low skill) migrants.

We define migrant status based on where people lived 5 years prior to the date of the Census, in a similar fashion to Card [2001]. Rural-Urban migrants are people who moved to an agglomeration in the past 5 years. In figure 7 and 8 note that most migrants come from municipalities close to cities in 1991 and 2000 (consistent with our regressions). Furthermore, in figure 9 and 10 we can see that most migrants went to São Paulo in 1991 and 2000, though they represent a larger proportion of the population in smaller cities, as you can see from figures 11 and 12. We use data on where people were living 5 years before in the 1991 and 2000 Census, to rebuild population and migration patterns from rural areas between 1986 to 1990 and 1995 to 1999.

We use data on precipitation from the Inter-Governmental Panel on Climate Change (IPCC) (see Mitchell et al. [2002] for a complete description of the data set), from 1985 to 1990 and 1994 to 1999 to estimate the impact of weather shocks in rural areas to out migration. The geographical pattern of average monthly rainfall over 100 years (1900 to 2000) over the 3661 MCAs is shown in figure 13.

We also use data from Castro [2002] on transportation costs from each municipality to São Paulo. Castro [2002] constructs an index that measures the benefits of improvements in highway infrastructure from 1970-1995 as the change in equivalent paved road distance from each municipality to the state capital of Sõ Paulo, accounting for the construction of the network as well as the difference in vehicle operating costs between earth/gravel and paved roads. We argue that improvements in the transportation network (reduction in transportation costs) allowed more people to migrate from rural to urban areas.

The basic statistics of rural areas are in table 2. Notice the importance of farming in rural areas, for low skill workers. since most people working in rural areas are farmers, it is likely that their income (and therefore their decision to migrate) will be affected by weather shocks.

The 1991 Census also provides information on where each individual was living in 1981 and time people are living in the current municipality. We use this information to determine historical patterns of migration, in particular, how people moved from rural to urban areas between 1981 and 1985. We can see from figure 14, that high and low skill people don't choose the same destinations, with a large amount of variation across cities. We will show that people move to cities which are closer to their municipality of origin.

Finally, we use data from the Agricultural Census of 1985 and 1995 on agricultural area (hectares), as a control for the importance of agriculture in each municipality.

3 Empirical Framework

We use the factor approach to evaluate how changes in factor endowment lead to changes in relative wages. This method was initially used in the international trade literature and is currently used to evaluate the impact of international migration (see for instance in Card [2001], Borjas [2006], Card [2007] and Peri [2007]).

We start by assuming that production in each city can be represented by a production function with constant returns to capital and labor. Therefore we have

$$Y = AK^{1-\alpha}N^{\alpha} \tag{1}$$

Where Y is the amount of output in a city in a given year, K and N is the amount of capital and labor (respectively) in a city in a given year, A is the technology (or other factors that are city specific and explain agglomeration economies) used in a city in a given year, and α is the elasticity of substitution between capital and 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 them within a constant elasticity of substitution (CES) production function, nested within equation 1. Labor N is therefore given by:

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

Where N_H and N_L are the number of high and low skill workers (respectively) in a city in a given year, a_H and $(1 - a_H)$ are productivity of high and low skill (respectively) workers. A higher a_H implies that high skilled labor is relatively more productive than low skill labor in the production of goods in the city. Finally, σ_{EDU} is the elasticity of substitution between high and low skill workers.

Our objective is to try to estimate the impact of a change in local labor supply that resulted from migration, on local wages. Since we have assumed 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:

$$w_H = \alpha A K^{1-\alpha} [N]^{\alpha-1} (a_{Hst}) N_H^{-\frac{1}{\sigma_{EDU}}} \tag{3}$$

And low skill wages are given by:

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

We can control for city specific factors affecting both groups symmetrically (technology, A and capital, K) by focusing on ratio of wages of high to low skill.

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

Or in logs:

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

Which gives us the relationship between high to low skill wages and high to low skill employment. More importantly, we can use this equation to obtain an estimate for the elasticity of substitution between high and low skill workers, σ_{EDU}.

Our results for equation 6 are in column 1 and 2 (controlling for $\frac{a_H}{(1-a_H)}$ with a trend) of table 3 with a sample consisting of men only. When we control for time trend, allowing the ratio $\frac{a_H}{(1-a_H)}$ to vary across time, our estimates of σ_{EDU} (3.81) are not far from what we observe in the literature (between 2.4 to 3.2 reported by a survey in Ottaviano and Peri [2008]).

One problem with estimating equation 6, is that productivity factors may vary by industry, and since industry composition varies by city, we expect our estimates to be biased. Therefore control for city fixed effects by looking at the first difference (difference between Census waves). We will estimate:

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

Uur estimates for equation 7 are in column 3 and 4 of table 3. As soon as we control for city fixed effects our estimates become significantly smaller than those in the literature. In particular, we find an elasticity of substitution between high and low skill (σ_{EDU}) of 1 (or 0.86 when we allow different time trends across cities).

As we explained before, it is possible (even likely) that changes in wages can lead to changes in labor supply (reverse causality). Therefore we instrument for changes in the ratio of high to low skill employment with migration flows. We argue that changes in the relative availability of high versus low skill labor (driven by migration) affect how much high and low skill workers firms decide to hire.

We build an exogenous migration shock to urban areas by breaking down the decision to migrate from rural areas and where migrants go.

3.1 How Many People Migrate From Rural Areas

In its simplest specification, the number of people who migrate out of a rural location, is a function of wages in that area and transportation costs. Therefore we use weather variation in current year and the previous year as a proxy for changes in rural wages and changes in transportation costs to São Paulo to measure improvements in the transport network. Our basic regression is then:

$$\ln Migrants_{i,rural,t} = \alpha_0 + \alpha_1 \ln N_{i,rural,(t-10)} + \alpha_2 Rain_{rural,t} + \alpha_3 Transp_{rural,t-5} + \alpha_4 X_i + e_{i,rural,t} \quad (8)$$

Where $N_{i,rural,(t-10)}$ is the (lagged) number of people living in a rural area, $rural$, in the previous Census year, $(t-10)$, with high or low education, i; $Rain_{rural,t}$ is the (log) rainfall in a rural area, $rural$, in year t (since the timing of the drought and its impact is uncertain we also check if lagged rainfall shocks impact migration), and $Transp_{rural,t}$ is the (log) cost of transportation to São Paulo, from a rural area, $rural$ in 1985 and 1990, $t-5$, while X_i is a set controls for characteristics of the rural area of origin (these include log agricultural area, year dummies, municipality fixed effects, and controls for skill group).

Our results are in table 7 for a sample of men only. In columns 1 through 4 of table 7 we run the regression 8 for each group (high and low skill, men and women). Since we include rural municipalities fixed effects, we can interpret our coefficients as response to shocks (deviations from average across the periods). We find that reductions in the transportation costs lead to more out migration of high skilled people (and surprisingly not low skilled people), in particular, for men only, a 10% decrease in transportation costs, increases the number of high skilled migrants by 3%. Furthermore, rainfall shocks affect migration of low skilled people (and not high skilled people), in particular, for men only, a 1 s.d. decrease in rainfall leads to an increase in migration of 5%. In the drought area of north east, the impact of a drought is different. This region, as pointed out by Baer [2008], receives government aid in years of drought (often misused). Therefore during good years, when there is no government transfers, out migration increases, of both low and high skill. Finally notice the large number of zeros (up to the 17% in the men only sample, not reported in the table). This implies that our estimates, using least squares, are biased towards zero.

In the next section we try to explain how migrants decide where to migrate to.

3.2 Where do Migrants Go

The previous literature shows the importance of distance and networks in the decision of migrants of where to migrate. While we don't have a good measure for networks, we do measure distance between origin and destination for migrations between 1981 and 1985 (previous to the period of analysis in the previous section).

To understand how destination depends on distance from origin, we construct a cross section of migrants from 1981 to 1985, for 123 urban destinations, from 3207 rural origins, for each of the groups (high and low skill, men and women). We run the following regression

$$\frac{Migrants_{i,c,rural,80-85}}{\sum_c Migrants_{i,c,rural,80-85}} = \beta_0 + \beta_1 Distance_{urban,rural} + \epsilon_{i,c,rural,80-85} \quad (9)$$

The dependent variable $\frac{Migrants_{i,j,c,rural,80-85}}{\sum_c Migrants_{i,c,rural,80-85}}$ is the share of migrants from a rural area moving into an urban area, c, for each group (high and low skill, i).

Our results are in table 5. As you can see from column 1 and 2, migrants are more likely to move to cities which are closer, and distance matters less for high skilled people and more for women (high and low skill). In particular, a city that is 10% closer, receives 0.2 percentage points more migrants. In columns 3 and 4 we check whether this result can be driven by the supply of men or women, high or low skill, by controlling by how many people are living in the rural and urban area. Distance is still statistically significant for every group.

In the next section we combine these two estimates (how many people migrate from rural areas and where do migrants go) to construct an instrument.

3.3 Instrumental Variable

After determining how people respond to droughts and changes in transportation costs in rural areas, and where people decide to go, we can build an exogenous migration shock for each city (exogenous to changes in cities) in the following way:

$$\left(\Delta \hat{N_{ij,s,91}}\right)^{mig} = \sum_{t=86}^{91} \sum_{rural} \frac{\hat{Migrants_{ij,rural,s,85}}}{\sum_s Migrants_{ij,rural,s,85}} \hat{Migrants_{ij,rural,st}} \qquad (10)$$

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 relative wages. Therefore we use the ratio of high to low skill migrants to instrument for the change in the ratio of high to low skill workers.

The results for the first stage are in table 6. In both column 1 and 2 (with city specific trends) we can see a positive and statistically significant coefficient (with a corresponding F-statistic above 10).

Our results from the second stage are in columns 5 and 6 of table 3. Our estimates of σ_{EDU} are even smaller than OLS (consistent with the problem of reverse causality). In particular, the elasticity of substitution of high to low skill workers is only 0.53 (close to 50% lower than our OLS estimates), even smaller than most results in the literature.

3.4 Robustness Checks

We first run a weighted regression, where weights are city population. 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 table 7 and 8 (for the first stage of the IV regressions). Most of the results in table 7 and 8 are similar we obtained in the previous section.

We also worry about the role of São Paulo in Brazil. As we explained before, São Paulo receives a disproportional share of rural-urban migrants in Brazil. Furthermore, it is disproportionally large than other cities in Brazil. Therefore we run our regressions excluding the city of São Paulo. Or results are in table 9 and 10. Again the results for σ_{EDU} are similar, except for column 5, where in the first stage the instrument is not statistically significant and we obtain an insignificant coefficient in the second stage as well.

There is also the concern about the public sector. If a significant proportion of workers in our cities are public employees then our assumption of how the job market works might be a problem, as the public sector offers other benefits not reflected in wages. We try to see how big a problem this might be by dropping the Federal capital, Brasilia (under the assumption that this will be the city with the largest number of public employees). Our results are in table 11 and 12, and are not significantly different from our previous results.

Next we redefine our sample to include both men and women (pooled). The results for the OLS regression of equation 7 are reported in table 13, the results for equation 8 are reported in table 7, the results for equation 9 are reported in table 15, the first stage of the IV regression are reported in table 16 and the second stage of the IV regression are in columns 5 and 6 of table 13. All the results are consistent with our previous findings. The results for the elasticity of substitution between high and low skill, σ_{EDU}, are even smaller (0.32 and 0.4).

We also redefine our definition of high skill as workers with 12 or more years of education and low skill as workers with less than 12 years of education (similar to previous studies using US data). The results for the OLS regression of equation 7 are reported in table 17, the results for equation 8 are reported in table 7, the results for equation 9 are reported in table 19, the first stage of the IV regression are reported in table 20 and the second stage of the IV regression are in columns 5 and 6 of table 17. All the results are consistent with our previous findings. The results for the elasticity of substitution between high and low skill, σ_{EDU}, are also smaller than our previous findings (0.33 and 0.5), similar to our previous results but still inconsistent with the previous literature.

Finally, we check whether the in-migration of rural workers has lead to out-migration from urban areas. For the number of log rural-urban migrants we proxy it with our instrument, defined in equation 10, to test this possibility. Our results are reported in table 21. In the first three columns the coefficient for the number of rural-urban migrants has either the wrong sign (negative) or is statistically insignificant. Once we control for city fixed effects, we obtain a positive and statistical significant coefficient, evidence of crowding out of urban migrants. However, if we run our regression for each group separately the coefficient on the number of rural-urban migrants becomes insignificant again.

4 Simulation Results

Finally we can use our estimates to simulate the effect of the the relative increase in the number of high skilled workers on the wage gap. In particular, we can use equation 7, (assuming no change in $\frac{a_H}{(1-a_H)}$ over time), our results from column 6 in talbe 3 and the actual change in the ratio of high to low skill workers, due to rural-urban migration, reported in table 1, to simulate the impact on the wage gap between high and low skill workers.

From table 1, we can deduce that the wage gap between high and low skill men decreased by 23% between 1991 and 2000. The percent of high skill workers increased from 26% in 1980 to 46% in 2000. Using our best estimate elasticity of substitution between men and women from column 6 of table 3 of 0.54, we find a reduction in the wage gap of 8.6%, significantly smaller than the 23% reduction that occurred between 1991 and 2000. Our results explain 63% of the reduction in the wage gap between high and low skill. Even if we use the estimates of the elasticity of substitution between high and low skill from Ottaviano and Peri [2008], the simulated effect of the proportion of high skilled workers would be even smaller than the actual effect.

Our results clearly indicate that, thought the relative increase in high skill workers led to a decrease in the wage gap (and therefore wage inequality), migration contributed to this trend.

5 Conclusion

We looked at the impact of internal migration on the wage gap between high skilled workers and low skill workers in Brazil between 1991 and 2000. To the best of our knowledge this is the first paper to focus on internal migration. We assume that workers with less than 9 years of education (low skilled) are imperfect substitutes for workers with 9 years of education or more (high skilled) in the production of goods. We estimated the elasticity of substitution. Our least square estimates were significantly smaller than those in the literature (0.99 to 0.86 in the sample containing only men, while Ottaviano and Peri [2008] reports estimates between 1.54 and 2.27).

Unlike the previous literature we also address the endogeneity problem (not only are wages affected by the number of workers, but labor supply is also affected by wages), First we use rainfall shocks and changes in transportation costs to estimate how people decide to out migrate from rural areas. We find that a 1 s.d. decrease in rainfall leads to an increase in low skill migration of 5%, and a 10% decrease in transportation costs, increases the number of high skilled migrants in 3% (for a sample consisting of only men). Next we determine how people decide where to migrate to, in particular how migrants respond to distance. A city that is 10% closer, receives 0.2 percentage points more migrants.

We are then able to build an exogenous migration shock to each city. Our instrumental variable estimates of the elasticity of substitution between high and low skill workers (0.54) are smaller than our ordinary least squares estimates (0.99 and 0.86), consistent with our assumption of reverse causality. This suggests that previous simulations of the impact of increased in the proportion of high skilled workers on the high to low skill wage gap are downward biased (up to to twice as large). Our results are consistent across different specifications.

Finally we simulated the impact of the increase in the number of high skilled worker on the urban high to low skill wage gap and found that rural-urban migration accounts for 63% in the reduction in the high to low skill wage gap.

References

W. Baer. *The Brazilian Economy: Growth and Development*. Rienner, 6th edition, 2008.

S. Basu. *Wage Gap Profiles of a New Group of Asian Immigrants: Effects of Larger Inflows*. PhD thesis, University of Rochester, 2010.

G. Borjas. The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *Quarterly Journal of Economics*, 118(4):1335–74, 2003.

G. Borjas. Native internal migration and the labor market impact of immigration. *The Journal of Human Resources*, 41(2):221–258, 2006.

A. A. Camarano and R. Abramovay. xodo rural, envelhecimento e masculinizao no brasil: Panorama dos ltimos cincoenta anos. In *Anais do 1 Encontro Nacional Sobre Migrao*, 1997.

D. Card. Immigration inflows, native outflows, and the local market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64, January 2001.

D. Card. How immigration affect u.s. cities. CReAM Discussion Paper Series, 2007. CDP No 11/07.

D. Card. Immigration and inequality. *American Economic Review*, 99(2):1–21, 2009.

N. Castro. Custos de transporte e producao agricola no brasil: 1970-1996. Texto para Discussao - NEMESIS LXVI, 2002.

C. Corseuil and M. Foguel. Uma sugesto de deflatores para rendas obtidas a partir de algumas pesquisas do ibge. IPEA Working Paper no. 897, July 2002.

D. da Mata, U. Deichmann, J. Henderson, S. Lall, and H. Wang. The determinants of city growth in brazil. NBER Working Paper, August 2005. w11585.

T. Freire. *Essays in Applied Economics*. PhD thesis, Brown University, June 2010.

L. Katz and K. Murphy. Changes in the wage structure 1963-87: Supply and demand factors. *Quarterly Journal of Economics*, 107(1):35–78, February 1992.

S. Lall, H. Selod, and Z. Shalizi. Rural-urban migration in developing countries: A survey of theoretical predictions and empirical findings. World Bank Working Paper, May 2006. WPS 3915.

A. Medici. The political economy of reform in brazil's civil servant pension scheme. PROGRAM ON THE GLOBAL DEMOGRAPHY OF AGING Working Paper, September 2004.

T. Mitchell, M. Hulme, and M. New. Climate data for political areas. *Area*, 34:109–112, January 2002.

G. Ottaviano and G. Peri. Rethinking the effects of immigration on wages:. University of California, Davis, November 2007.

G. Ottaviano and G. Peri. Immigration and national wages: Clarifying the theory and the empirics. University of California, Davis, July 2008.

G. Peri. Immigrants' complementarities and native wages: Evidence from california. University of California, Davis, August 2007.

D. Puga. European regional policies in light of recent location theories. *Journal of Economic Geography*, 2:373–406, 2002.

E. Reis, M. Pimentel, and A. I. Alvarenga. reas mnimas comparveis para os perodos intercensitrios de 1872 a 2000. IPEA Document, August 2007.

6 Appendix 1 - Figures

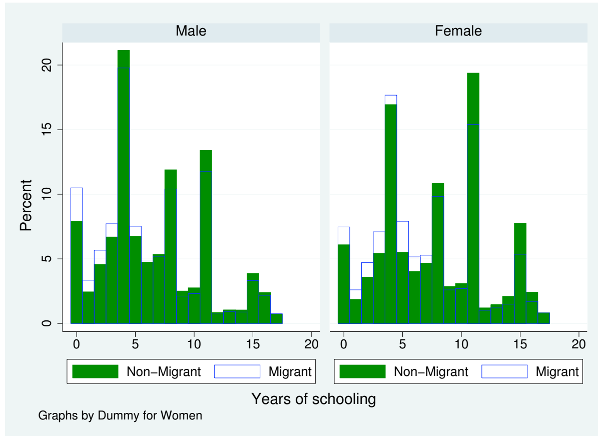

Figure 1: Education of Active Urban Population (green) and Rural-Urban Migrants (blue line) for 1991. Notice that there is a larger percentage of rural-urban migrants in the low education categories, than urban dwellers. We consider high skill people as having 9 or more years, while low skill are people having less than 9 years of education.

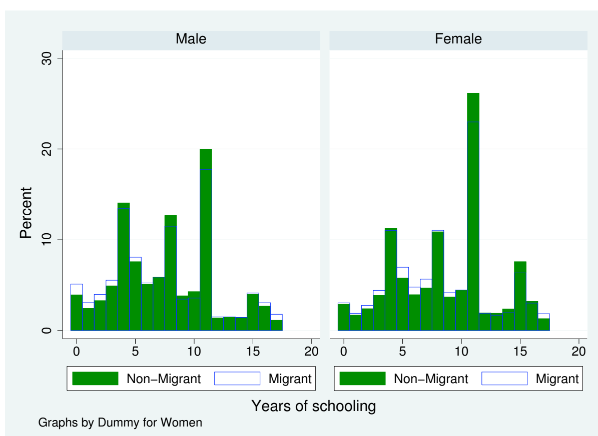

Figure 2: Education of Active Urban Population (green) and Rural-Urban Migrants (blue line) for 2000. Notice that there is a larger percentage of rural-urban migrants in the low education categories, than urban dwellers. We consider high skill people as having 9 or more years, while low skill are people having less than 9 years of education.

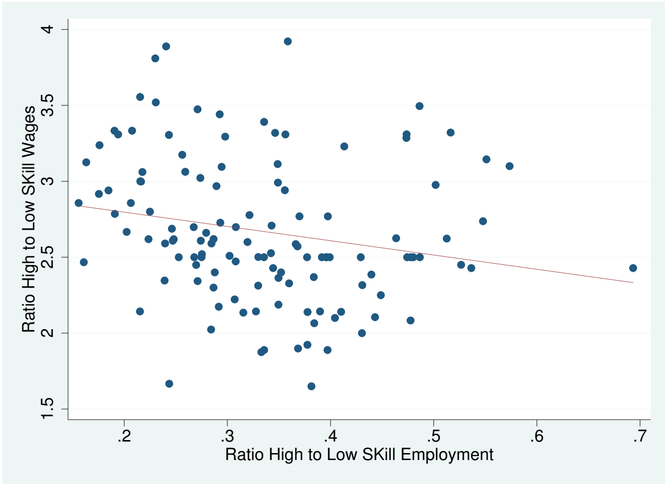

Figure 3: Negative relationship between the ratio of high to low skill hourly wage (vertical axis) and the ratio of high to low skill workers for each of the 123 agglomerations in 1991 (Men only).

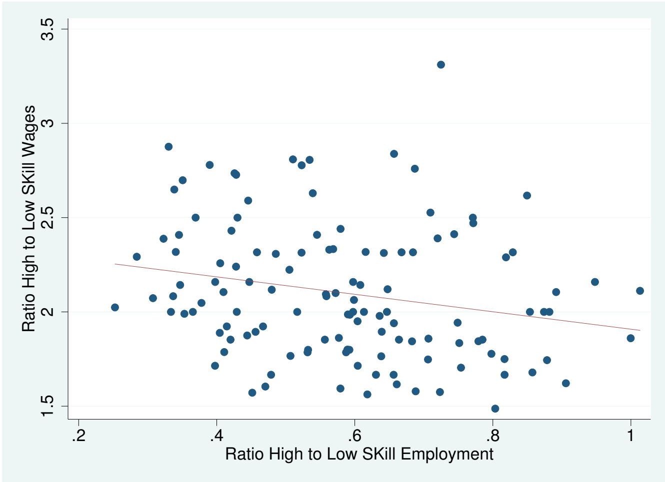

Figure 4: Negative relationship between the ratio of high to low skill hourly wage (vertical axis) and the ratio of high to low skill workers for each of the 123 agglomerations in 2000 (Men only).

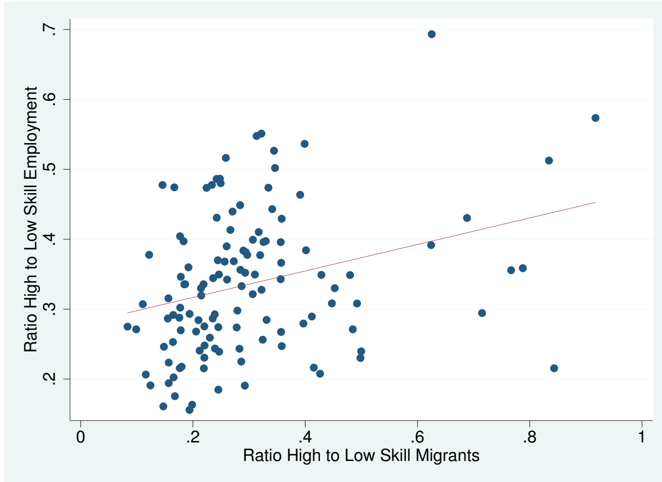

Figure 5: Positive relationship between the ratio of high to low skill migrants (horizontal axis) and the ratio of high to low skill workers for each of the 123 agglomerations in 1991 (Men only).

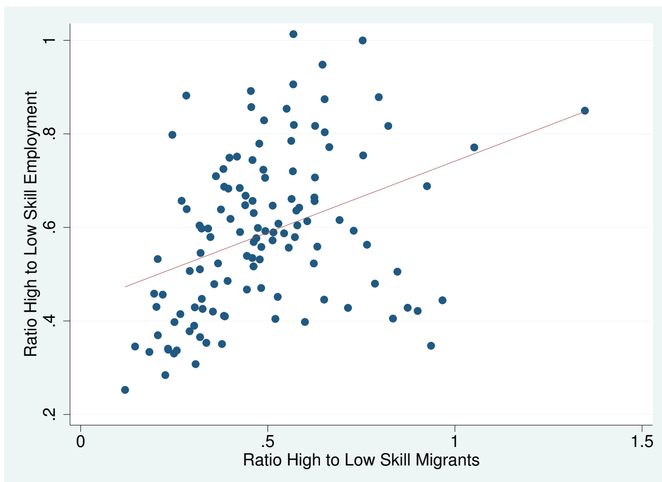

Figure 6: Positive relationship between the ratio of high to low skill migrants (horizontal axis) and the ratio of high to low skill workers for each of the 123 agglomerations in 2000 (Men only).

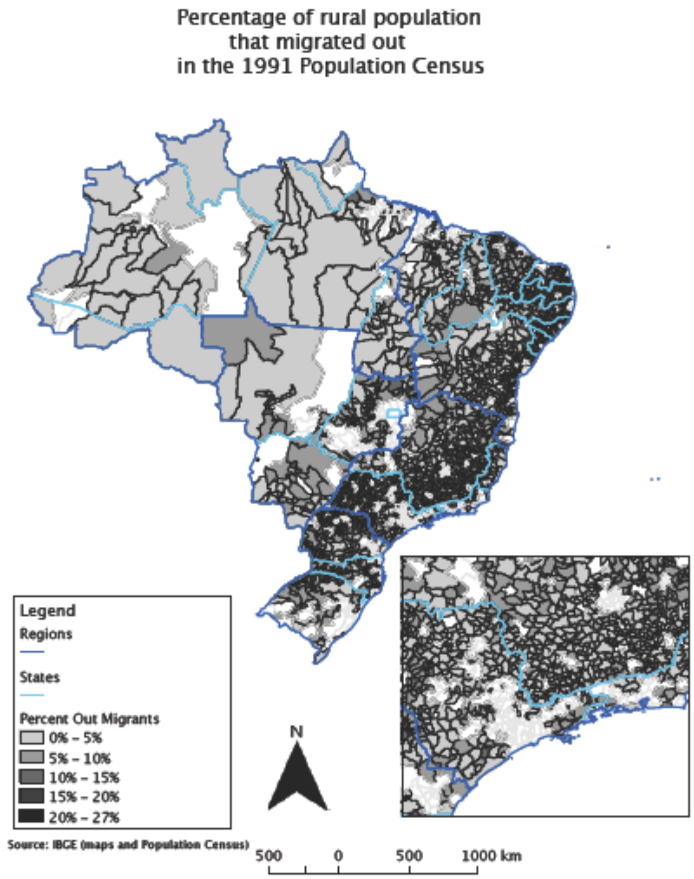

Figure 7: Map of out migration from rural areas (total migrants - summing all groups - as a percentage of total population living in rural area - summing all groups) in 1991. Blow-up map is of São Paulo area. Notice that rural areas surrounding urban areas send a larger percentage of migrants.

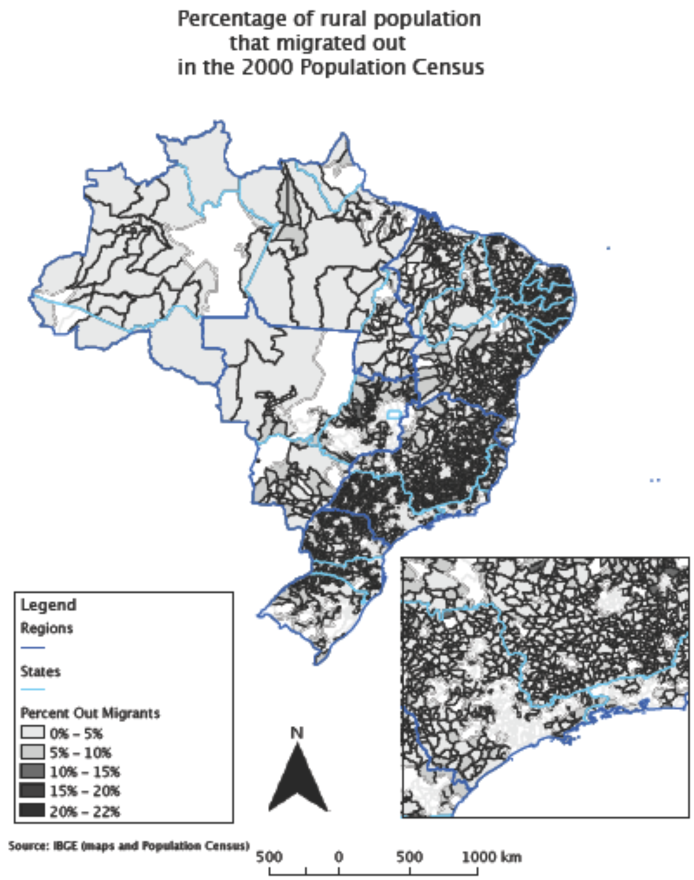

Figure 8: Map of out migration from rural areas (total migrants - summing all groups - as a percentage of total population living in rural area - summing all groups) in 2000. Blow-up map is of São Paulo area. Notice that rural areas surrounding urban areas send a larger percentage of migrants.

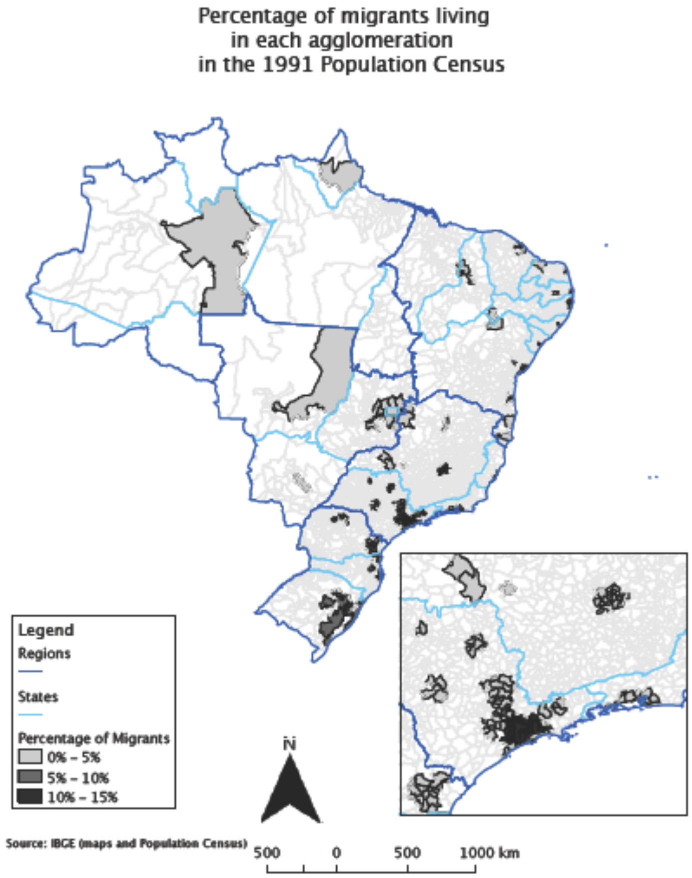

Figure 9: Map of percentage of in migrants to each urban area as percent of all migrants to urban areas in 1991. Blow-up map is of São Paulo area. Notice that São Paulo is the largest recipient of rural-urban migrants.

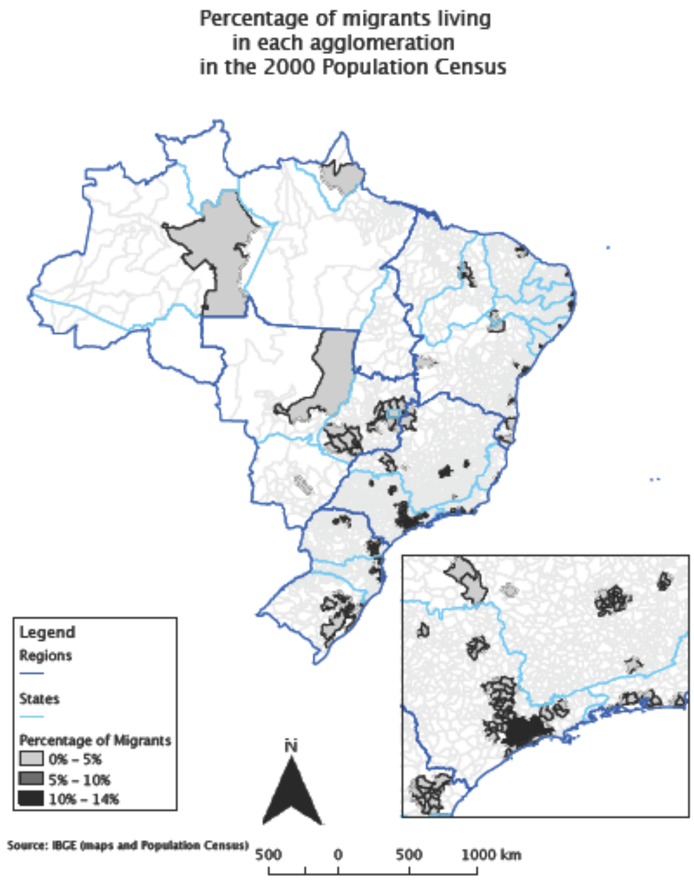

Figure 10: Map of percentage of in migrants to each urban area as percent of all migrants to urban areas in 2000. Blow-up map is of São Paulo area. Notice that São Paulo is the largest recipient of rural-urban migrants.

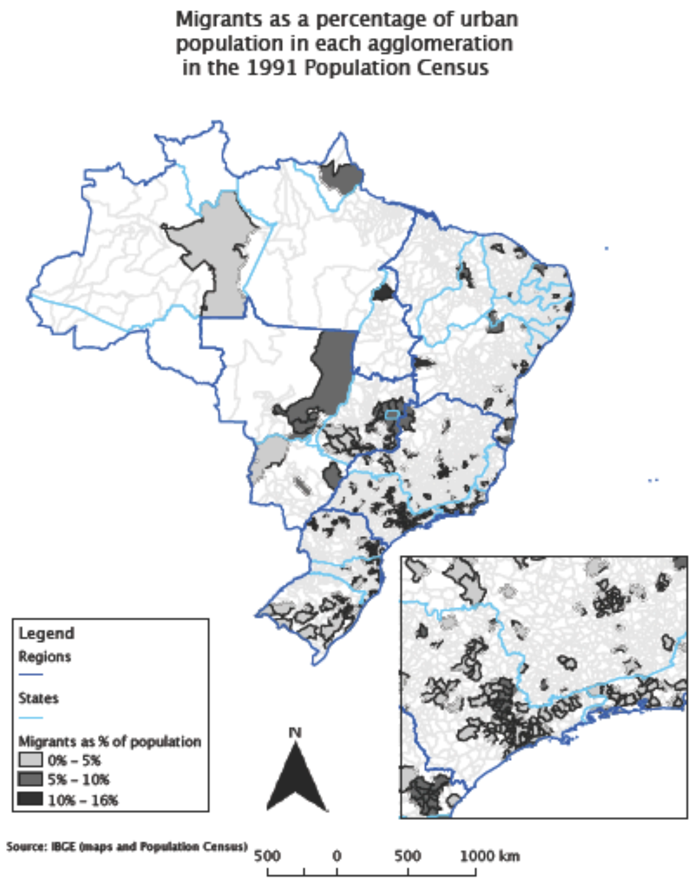

Figure 11: Map of migrants to each urban area as a percentage of total urban population in 1991. Blow-up map is of São Paulo area. Notice that rural-urban migrants represent a larger percentage of urban population in smaller cities (not São Paulo or Rio de Janeiro).

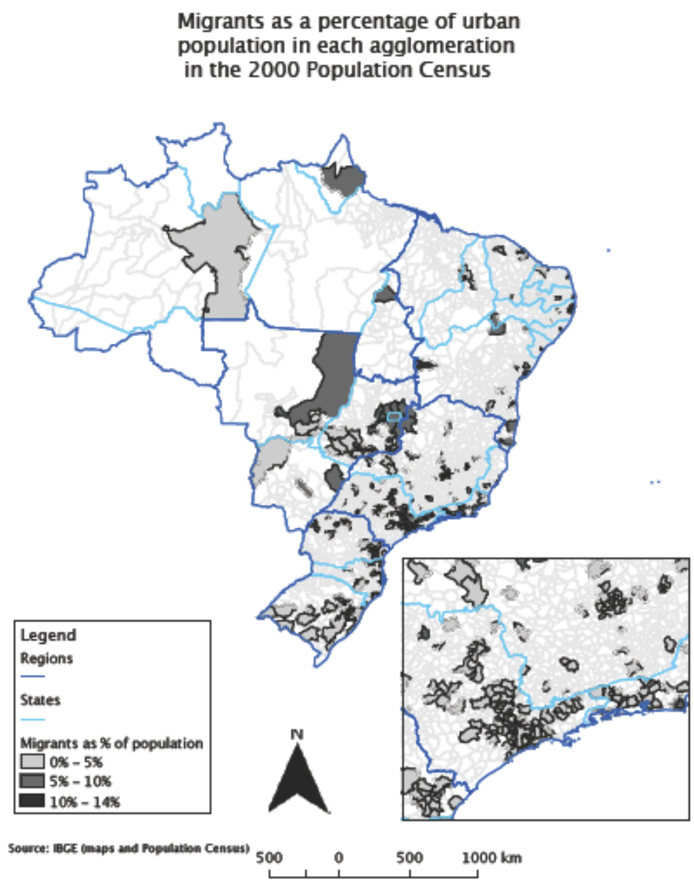

Figure 12: Map of migrants to each urban area as a percentage of total urban population in 2000. Blow-up map is of São Paulo area. Notice that rural-urban migrants represent a larger percentage of urban population in smaller cities (not São Paulo or Rio de Janeiro).

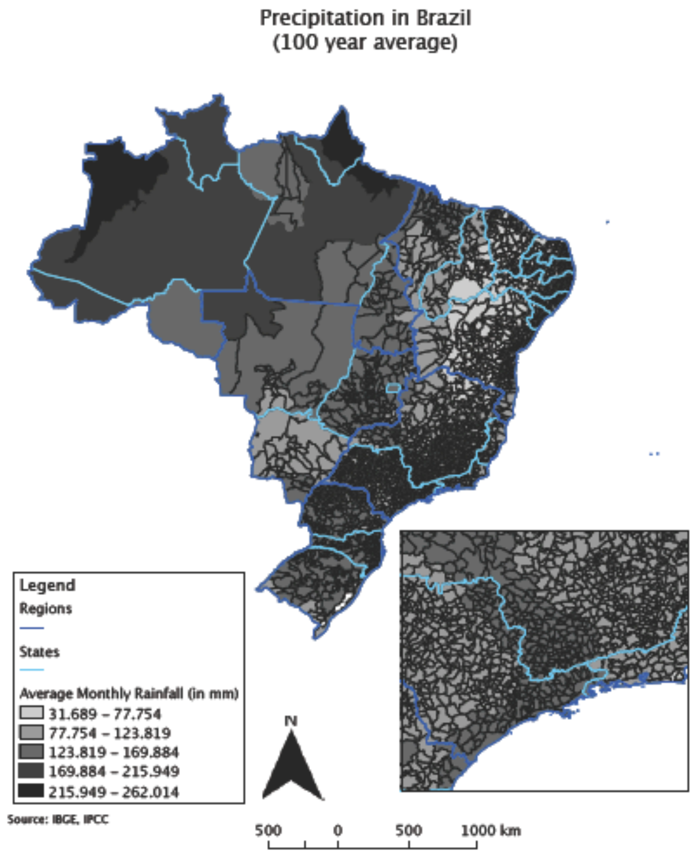

Figure 13: Map of distribution of average monthly rainfall (in mm) in Brazil from 1900 to 2000. Blow-up map is of the São Paulo area. Notice that the semi-arid area has the lowest average monthly rainfall.

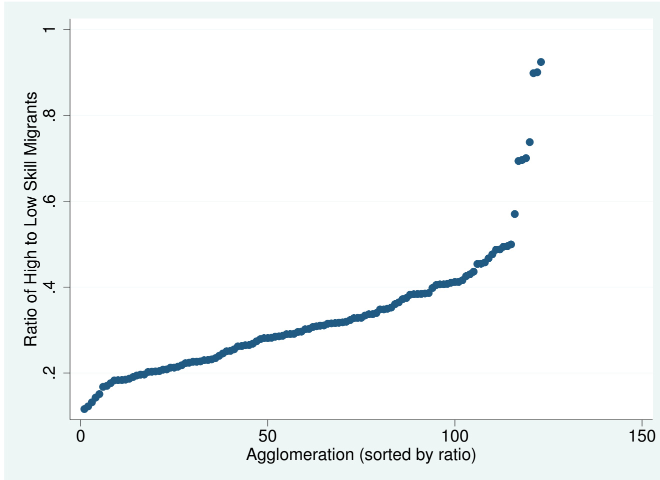

Figure 14: Historical patterns of migration. Plot of the ratio of high to low skill migrants from 1981 to 1985, for 123 agglomerations, sorted from smallest to largest ratio on the horizontal axis.

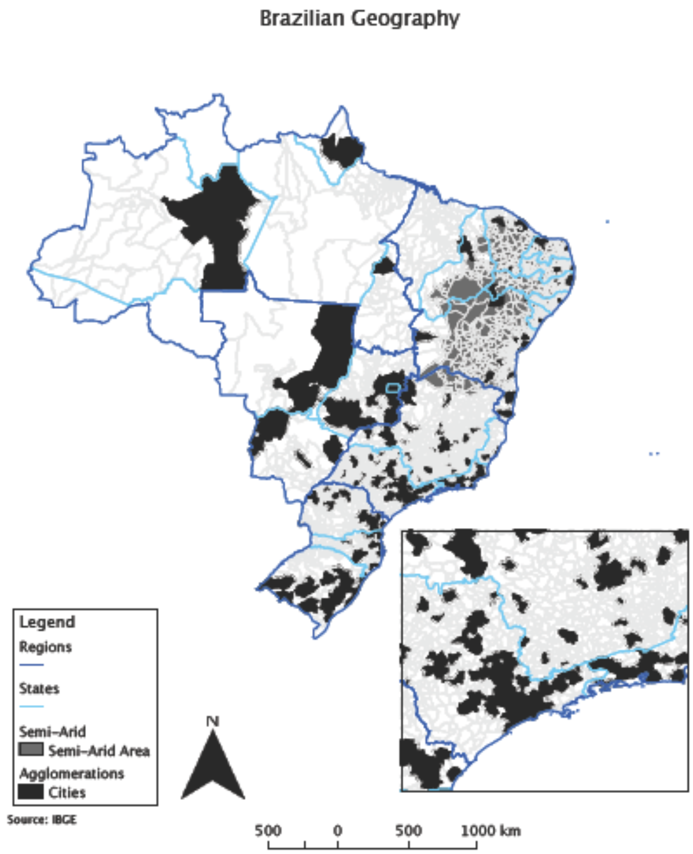

Figure 15: Map of 123 Agglomerations (Cities), 28 States, 5 Regions and Semi-Arid (chronic drought) Region in Brazil. Blow-up map is of area around São Paulo.

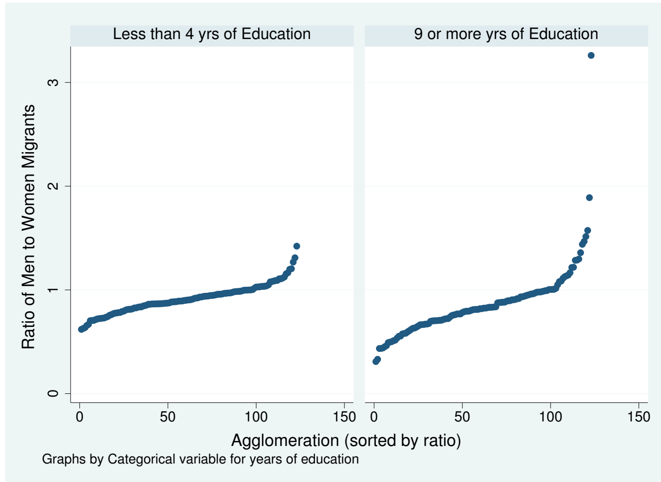

Figure 16: Historical patterns of migration. Plot of the ratio of men to women migrants by skill (high and low) from 1981 to 1985, for 123 agglomerations, sorted from smallest to largest ratio on the horizontal axis.

7 Appendix 2 - Tables

Basic Statistics for 123 Agglomerations						
	1980		1991		2000	
	Mean	SD	Mean	SD	Mean	SD
Avg Hourly Wage						
Men High Skill	3.39	0.85	2.84	0.73	2.54	0.54
Women High Skill	1.77	0.43	1.63	0.48	1.62	0.35
Men Low Skill	1.76	0.46	1.11	0.37	1.26	0.34
Women Low Skill	0.93	0.26	0.66	0.23	0.84	0.21
Avg Employment						
Men High Skill	25,779	82,095	40,636	118,797	64,293	180,778
Women High Skill	15,831	50,377	32,193	90,656	56,473	156,021
Men Low Skill	84,311	250,150	100,070	257,366	92,901	218,622
Women Low Skill	35,597	114,132	46,783	127,290	50,409	124,024
Avg Migrants						
Men High Skill			1,230	13,694	1,906	21,226
Women High Skill			1,238	13,789	2,147	23,906
Men Low Skill			3,319	36,962	4,025	44,823
Women Low Skill			3,251	36,200	3,745	41,699

Table 1: Statistics for 123 cities (agglomerations) on wages, employment, rural to urban migration for 1980, 1991 and 2000 in Brazil. High skill is defined as people having 9 years of education or more, and low skill less than 9 years of education. Reported hourly median wages are deflated to current values (January 2002). Migrants are defined as people who lived in a different municipality 5 years before the Census.

	1991				2000			
	Low Skill		High Skill		Low Skill		High Skill	
	Male	Female	Male	Female	Male	Female	Male	Female
Perc. Administrative	5.43%	4.50%	30.84%	29.44%	3.98%	3.60%	20.96%	24.14%
Perc. Technical or Scientific	0.90%	5.39%	16.81%	43.75%	1.34%	4.23%	16.25%	36.78%
Perc. Farming	52.14%	26.25%	7.52%	0.93%	37.43%	13.68%	6.71%	0.77%
Perc. Mining	1.38%	0.12%	0.33%	0.03%	0.66%	0.05%	0.20%	0.01%
Perc. Industry	19.03%	11.45%	12.97%	3.84%	25.38%	10.05%	17.51%	4.57%
Perc. Commerce and Trade	6.64%	8.93%	12.64%	11.22%	8.21%	10.78%	14.38%	15.13%
Perc. Transport	5.11%	0.71%	4.46%	1.31%	6.88%	0.51%	6.12%	1.38%
Perc. Services	3.73%	38.96%	3.38%	5.89%	6.27%	50.04%	5.26%	12.77%
Perc. Domestic Services	0.49%	26.08%	0.14%	2.68%	0.94%	33.76%	0.21%	6.98%
Perc. Security and National Defense	1.04%	0.05%	6.30%	0.35%	0.95%	0.04%	6.16%	0.38%
Perc. Other	4.59%	3.63%	4.75%	3.24%	8.91%	7.03%	6.46%	4.07%
No Migrants	*1403454*	*435290*	*284529*	*179502*	*1125683*	*426150*	*400489*	*296881*
No Non-Migrants	*13660763*	*4338761*	*2288760*	*1991011*	*11255160*	*4524492*	*3630459*	*3451467*
Rainfall	*11.27*	*4.19 (sd)*			*11.14*	*4.34 (sd)*		
Transportation costs to Sao Paulo	*1810.7*	*1436.99 (sd)*			*1549.34*	*1126.06 (sd)*		

Table 2: Basic Statistics for 3214 rural municipalities (MCA - Minimal Comparable Areas), by group: Men/Women and High Education (9 or more years of education) and Low Education (less than 9 years of education) for 1991 and 2000. Rows 1-11 has the percentage of people (of each group) working in each sector (low skill workers concentrate in farming). Rows 12 and 13 has the average number of migrants (people who moved out of a rural area in the last 5 years) and non-migrants for each group. Row 14 have the mean monthly rainfall and standard deviation in brackets in rural areas for 1991 and 2000. Row 15 has the average and standard deviation in brackets of the index of transportation costs from rural area to São Paulo.

OLS Estimate for the Elasticity of Substitution between High and Low Skill Men Only						
	Log Ratio of High to Low Skill Hourly Wages					
	OLS				IV	
	Level		First Difference		First Difference	
Log Ratio of High to Low Skill Workers	-0.088 (4.11)**	-0.262 (6.10)**	-1.0069 (17.32)**	-1.159 (16.69)**	-1.901 (6.67)**	-1.860 (12.37)**
Trend	No	Yes	No	No	No	No
City (agglomeration) Fixed Effects	No	No	No	Yes	No	Yes
Observations	360	360	237	237	237	237
R-squared	0.03	0.15	0.58	0.66		
Number of Cities (agglomerations)				123		123

Table 3: Estimates for the elasticity of substitution between high and low skill workers, using men only. T-statistics in parentheses (robust standard errors, clustered at the city level). * significant at 5%; ** significant at 1% . First difference regressions are regressions on the change in log ratio of high to low skill hourly wage and the change in high to low skill employment. In the first difference regressions the constant can be interpreted as the trend and city dummies as city specific trends.

OLS Estimate for Impact of Change in Transportation Costs and Rainfall on Log Number of Migrants from Rural Areas				
	Log Migrants			
	Low Skill		High Skill	
	Men	Women	Men	Women
Log Lag Natives	0.8248	0.8579	0.1032	0.1054
	(11.77)**	(13.00)**	(3.39)**	(3.06)**
Log Agricultural Area (in He)	-0.0585	-0.0617	0.0234	-0.0643
	(1.97)*	(1.76)	(0.57)	(1.70)
Change in Transportation Cost to SP	0.000042	0.000029	-0.0000902	-0.00017
	(1.40)	(0.94)	(2.18)*	(4.09)**
Average Monthly Rainfall	-0.0063	-0.0074	-0.0075	-0.00046
	(2.54)*	(3.03)**	(1.75)	(0.11)
Previous year's Avg Monthly Rainfall	-0.0027	-0.0043	-0.0045	-0.0056
	(0.99)	(1.59)	(1.02)	(1.29)
Average Monthly Rainfall in Semi-Arid Area	0.0149	0.0075	0.0625	0.0469
	(2.64)**	(1.33)	(5.90)**	(4.68)**
Last year's Average Monthly Rainfall in Semi-Arid Area	0.0277	0.0158	0.0591	0.0644
	(4.43)**	(2.76)**	(5.48)**	(6.14)**
Observations	25712	25712	25594	25700
Dummy for Women	No	No	No	No
Dummy for High Skilled	No	No	No	No
Dummy for High Skilled Women	No	No	No	No
MCA Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Number of AMC7000	3214	3214	3210	3214
R-squared	0.07	0.08	0.08	0.11

Table 4: Estimates of the impact of weather shocks and reductions in transportation costs on out migration from rural areas (3214 rural MCAs - minimal comparable areas) for 1987 to 1991 and 1996 to 2000, by group: high and low Skill (men and women separated). T-statistics in parentheses (robust standard errors, clustered at the MCA level). * significant at 5%; ** significant at 1%.

Estimates for Distance as an Explanation for Migration Location Decision				
	Percent Rural-Urban Migrants			
Log Distance	-0.0212 (93.17)**	-0.02196 (87.38)**	-0.0233 (91.99)**	-0.0224 (89.35)**
Log Distance (for Women)		-0.0012 (8.37)**	-0.0015 (9.91)**	-0.0013 (8.97)**
Log Distance (for High Skilled)		0.0033 (13.23)**	0.0063 (21.45)**	0.00401 (14.95)**
Log Distance (for High Skilled Women)		-0.00105 (3.71)**	-0.0018 (6.11)**	-0.0012 (4.14)**
(Log) People living in rural area* (log) people living in urban area			0.0011 (94.45)**	0.0012 (35.16)**
Log number of people living in rural area				-0.0091 (30.00)**
Log number of people living in urban area				0.00094 (4.04)**
Dummy for High Skill	Yes	Yes	Yes	Yes
Dummy for Women	Yes	Yes	Yes	Yes
Dummy for High Skilled Women	Yes	Yes	Yes	Yes
Municipality (MCA) Fixed Effects	Yes	Yes	Yes	Yes
Observations	1581288	1581288	1578213	1578213
Number of Rural Municipalities (MCA)	3214	3214	3214	3214
R-squared	0.05	0.05	0.07	0.08

Table 5: Estimates of distance as an explanation of choice in destination (123 cities/agglomeration) from rural areas (3214 MCAs) for each group: High and Low Skill (for the sample of men and women separated). T-statistics in parentheses (robust standard errors, clustered at the rural municipality (MCA - minimal comparable area)). * significant at 5%; ** significant at 1%.

First Stage of IV Ratio of Workers and Migration		
	Change in Log Ratio of High to Low Skill Workers	
	Men Only	
Log Ratio of High to Low Skill Migrants	0.268 (4.52)**	0.723 (8.88)**
City (agglomeration) Fixed Effects	No	Yes
Observations	237	237
R-squared	0.12	0.46
Number of Cities (agglomerations)		123

Table 6: Estimates for the impact of ratio of high to low skill migrants on the ratio of high to low skill employment (first stage of instrumental variable regressions). Sample includes men only. T-statistics in parentheses (robust standard errors, clustered at the city level). * significant at 5%; ** significant at 1%.

OLS Estimate for the Elasticity of Substitution between High and Low Skill Men Only (weighted)						
	Log Ratio of High to Low Skill Hourly Wages					
	OLS				IV	
	Level		First Difference		First Difference	
Log Ratio of High to Low Skill Workers	-0.017 (0.57)	-0.177 (2.75)**	-1.163 (10.13)**	-1.407 (6.71)**	-1.748 (5.25)**	-2.166 (9.72)**
Trend	No	Yes	No	No	No	No
City (agglomeration) Fixed Effects	No	No	No	Yes	No	Yes
Observations	360	360	237	237	237	237
R-squared	0.00	0.06	0.56	0.71		
Number of Cities (agglomerations)				123		123

Table 7: Estimates for the elasticity of substitution between high and low skill workers, using men only, with each observation weighted by city size. T-statistics in parentheses (robust standard errors, clustered at the city level). * significant at 5%; ** significant at 1% . First difference regressions are regressions on the change in log ratio of high to low skill hourly wage and the change in high to low skill employment. In the first difference regressions the constant can be interpreted as the trend and city dummies as city specific trends.

First Stage of IV Ratio of Workers and Migration		
	Change in Log Ratio of High to Low Skill Workers	
	Men Only, Weighted	
Log Ratio of High to Low Skill Migrants	0.28	0.61
	(3.82)**	(5.71)**
City (agglomeration) Fixed Effects	No	Yes
Observations	237	237
R-squared	0.21	0.64
Number of Cities (agglomerations)		123

Table 8: Estimates for the impact of ratio of high to low skill migrants on the ratio of high to low skill employment (first stage of instrumental variable regressions). Sample includes men only, weighting each regression by city size. T-statistics in parentheses (robust standard errors, clustered at the city level). * significant at 5%; ** significant at 1%.

OLS Estimate for the Elasticity of Substitution between High and Low Skill						
Men Only, excluding Sao Paulo						
	Log Ratio of High to Low Skill Hourly Wages					
	OLS				IV	
	Level		First Difference		First Difference	
Log Ratio of High to Low Skill Workers	-0.088 (4.08)**	-0.264 (6.03)**	-1.0056 (17.24)**	-1.156 (11.53)**	-1.910 (6.59)**	-1.859 (11.45)**
Trend	No	Yes	No	No	No	No
City (agglomeration) Fixed Effects	No	No	No	Yes	No	Yes
Observations	360	360	237	237	237	237
R-squared	0.03	0.15	0.58	0.69		
Number of Cities (agglomerations)				123		123

Table 9: Estimates for the elasticity of substitution between high and low skill workers, using men only and excluding the city of São Paulo. T-statistics in parentheses (robust standard errors, clustered at the city level). * significant at 5%; ** significant at 1% . First difference regressions are regressions on the change in log ratio of high to low skill hourly wage and the change in high to low skill employment. In the first difference regressions the constant can be interpreted as the trend and city dummies as city specific trends.

First Stage of IV Ratio of Workers and Migration		
	Change in Log Ratio of High to Low Skill Workers	
	Men Only, excluding São Paulo	
Log Ratio of High to Low Skill Migrants	0.266 (4.48)**	0.723 (6.11)**
City (agglomeration) Fixed Effects	No	Yes
Observations	237	237
R-squared	0.12	0.58
Number of Cities (agglomerations)		123

Table 10: Estimates for the impact of ratio of high to low skill migrants on the ratio of high to low skill employment (first stage of instrumental variable regressions). Sample includes men only and excludes the city of São Paulo. T-statistics in parentheses (robust standard errors, clustered at the city level). * significant at 5%; ** significant at 1%.

OLS Estimate for the Elasticity of Substitution between High and Low Skill						
Men only, excluding Brasilia						
	Log Ratio of High to Low Skill Hourly Wages					
	OLS				IV	
	Level		First Difference		First Difference	
Log Ratio of High to Low Skill Workers	-0.087	-0.25999	-1.0087	-1.161	-1.866	-1.841
	(4.03)**	(6.05)**	(17.33)**	(11.58)**	(6.79)**	(11.65)**
Trend	No	Yes	No	No	No	No
City (agglomeration) Fixed Effects	No	No	No	Yes	No	Yes
Observations	360	360	237	237	237	237
R-squared	0.03	0.15	0.58	0.70		
Number of Cities (agglomerations)				123		123

Table 11: Estimates for the elasticity of substitution between high and low skill workers, using men only and excluding the city of Brasilia. T-statistics in parentheses (robust standard errors, clustered at the city level). * significant at 5%; ** significant at 1% . First difference regressions are regressions on the change in log ratio of high to low skill hourly wage and the change in high to low skill employment. In the first difference regressions the constant can be interpreted as the trend and city dummies as city specific trends.

First Stage of IV Ratio of Workers and Migration		
	Change in Log Ratio of High to Low Skill Workers	
	Men Only, excluding Brasilia	
Log Ratio of High to Low Skill Migrants	0.278 (4.55)**	0.733 (6.21)**
City (agglomeration) Fixed Effects	No	Yes
Observations	237	237
R-squared	0.12	0.59
Number of Cities (agglomerations)		123

Table 12: Estimates for the impact of ratio of high to low skill migrants on the ratio of high to low skill employment (first stage of instrumental variable regressions). Sample includes only men and excludes the city of Brasilia. T-statistics in parentheses (robust standard errors, clustered at the city level). * significant at 5%; ** significant at 1%.

OLS Estimate for the Elasticity of Substitution between High and Low Skill Both Men and Women						
	Log Ratio of High to Low Skill Hourly Wages					
	OLS				IV	
	Level		First Difference		First Difference	
Log Ratio of High to Low Skill Workers	0.0187 (1.07)	-0.1399 (3.09)**	-0.968 (15.60)**	-1.179 (14.73)**	-3.13 (4.12)**	-2.49 (8.87)**
Trend	No	Yes	No	No	No	No
City (agglomeration) Fixed Effects	No	No	No	Yes	No	Yes
Observations	360	360	237	237	237	237
R-squared	0	0.07	0.45	0.54		
Number of Cities (agglomerations)				123		123

Table 13: Estimates for the elasticity of substitution between high and low skill workers, pooling both men and women. T-statistics in parentheses (robust standard errors, clustered at the city level). * significant at 5%; ** significant at 1% . First difference regressions are regressions on the change in log ratio of high to low skill hourly wage and the change in high to low skill employment. In the first difference regressions the constant can be interpreted as the trend and city dummies as city specific trends.

OLS Estimate for Impact of Change in Transportation Costs and Rainfall on Log Number of Migrants from Rural Areas			
	Log Migrants		
		Men and Women	
	Pooled	Low Skill	High Skill
Log Lag Natives	0.641 (41.90)**	0.838 (12.70)**	0.122 (3.42)**
Log Agricultural Area (in He)	-0.0299 (1.13)	-0.0593 (2.10)*	-0.0142 (0.37)
Change in Transportation Cost to SP	-0.00024 (9.29)**	-0.000032 (1.15)	-0.000147 (3.78)**
Change in Transportation Cost to SP (x Dummy High Skilled People)	0.00027 (11.09)**		
Average Monthly Rainfall	-0.00904 (3.98)**	-0.0069 (3.22)**	-0.0042 (1.13)
Average Monthly Rainfall (x Dummy High Skilled People)	0.01296 (4.26)**		
Previous year's Avg Monthly Rainfall	-0.0061 (2.46)*	-0.0038 (1.67)	-0.0056 (1.37)
Previous year's Avg Monthly Rainfall (x Dummy High Skilled People)	0.0097 (3.10)**		
Average Monthly Rainfall in Semi-Arid Area	0.0292 (5.58)**	0.0110 (2.37)*	0.0602 (6.53)**
Average Monthly Rainfall in Semi-Arid Area (x Dummy High Skilled People)	0.0238 (3.46)**		
Last year's Average Monthly Rainfall in Semi-Arid Area	0.0507 (8.91)**	0.0205 (4.10)**	0.0623 (6.15)**
Last year's Average Monthly Rainfall in Semi-Arid Area (x Dummy High Skilled People)	0.00097 (0.15)		
Observations	51420	25712	25708
Dummy for High Skilled	Yes	No	No
MCA Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Number of AMC7000	3214	3214	3214
R-squared	0.56	0.10	0.12

Table 14: Estimates of the impact of weather shocks and reductions in transportation costs on out migration from rural areas (3214 rural municipalities (MCA - minimal comparable areas) for 1987 to 1991 and 1996 to 2000, by group: high and low Skill (pooling both men and women). T-statistics in parentheses (robust standard errors, clustered at the MCA level). * significant at 5%; ** significant at 1%.

Estimates for Distance as an Explanation for Migration Location Decision		
	Percent Rural-Urban Migrants	
Log Distance	-0.02307 (96.44)**	-0.02346 (98.29)**
Log Distance (for High Skilled)		0.00077 (3.97)**
Dummy for High Skill	Yes	Yes
Dummy for Women	Yes	Yes
Dummy for High Skilled Women	Yes	Yes
Municipality (MCA) Fixed Effects	Yes	Yes
Observations	790644	790644
Number of Rural Municipalities (MCA)	3214	3214
R-squared	0.06	0.06

Table 15: Estimates of distance as an explanation of choice in destination (123 cities/agglomeration) from rural areas (3214 MCAs) for each group: High and Low Skill (for the sample pooling men and women). T-statistics in parentheses (robust standard errors, clustered at the rural municipality (MCA - minimal comparable area)). * significant at 5%; ** significant at 1%.

First Stage of IV Ratio of Workers and Migration		
	Change in Log Ratio of High to Low Skill Workers	
	Men Only	
Log Ratio of High to Low Skill Migrants	0.173 (2.98)**	0.475 (6.50)**
City (agglomeration) Fixed Effects	No	Yes
Observations	237	237
R-squared	0.06	0.33
Number of Cities (agglomerations)		123

Table 16: Estimates for the impact of ratio of high to low skill migrants on the ratio of high to low skill employment (first stage of instrumental variable regressions). Sample pooling both men and women. T-statistics in parentheses (robust standard errors, clustered at the city level). * significant at 5%; ** significant at 1%.

OLS Estimate for the Elasticity of Substitution between High and Low Skill Men Only						
	Log Ratio of High to Low Skill Hourly Wages					
	OLS				IV	
	Level		First Difference		First Difference	
Log Ratio of High to Low Skill Workers	-0.232 (6.07)**	-0.318 (8.17)**	-0.992 (15.23)**	-1.052 (11.81)**	-2.996 (1.31)	-1.988 (7.23)**
Trend	No	Yes	No	No	No	No
City (agglomeration) Fixed Effects	No	No	No	Yes	No	Yes
Observations	360	360	237	237	237	237
R-squared	0.11	0.32	0.61	0.61		
Number of Cities (agglomerations)				123		123

Table 17: Estimates for the elasticity of substitution between high and low skill workers, using men only where the cutoff between high and low skill is 12 years of education. T-statistics in parentheses (robust standard errors, clustered at the city level). * significant at 5%; ** significant at 1% . First difference regressions are regressions on the change in log ratio of high to low skill hourly wage and the change in high to low skill employment. In the first difference regressions the constant can be interpreted as the trend and city dummies as city specific trends.

OLS Estimate for Impact of Change in Transportation Costs and Rainfall on Log Number of Migrants from Rural Areas			
	Log Migrants		
		Men	
	Pooled	Low Skill	High Skill
Log Lag Natives	0.3381 (26.36)**	0.7852 (11.39)**	0.0618 (3.43)**
Log Agricultural Area (in He)	-0.0785 (2.99)**	-0.0442 (1.43)	-0.1044 (2.65)**
Change in Transportation Cost to SP	-0.0003 (10.17)**	-0.0001 (1.84)	-0.0001 (3.07)**
Change in Transportation Cost to SP (x Dummy High Skilled People)	0.00043 (15.04)**		
Average Monthly Rainfall	-0.01197 (4.71)**	-0.0054 (2.27)*	-0.0069 (1.58)
Average Monthly Rainfall (x Dummy High Skilled People)	0.0109 (2.95)**		
Previous year's Avg Monthly Rainfall	-0.0073 (2.80)**	-0.0006 (0.23)	-0.0043 (0.94)
Previous year's Avg Monthly Rainfall (x Dummy High Skilled People)	0.0063 (1.66)		
Average Monthly Rainfall in Semi-Arid Area	0.0287 (5.27)**	0.0189 (3.49)**	0.0401 (4.08)**
Average Monthly Rainfall in Semi-Arid Area (x Dummy High Skilled People)	0.0069 (0.83)		
Last year's Average Monthly Rainfall in Semi-Arid Area	0.0389 (6.78)**	0.0324 (5.42)**	0.0311 (3.06)**
Last year's Average Monthly Rainfall in Semi-Arid Area (x Dummy High Skilled People)	-0.0054 (0.69)		
Observations	48437	25712	22725
Dummy for High Skilled	Yes	No	No
MCA Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Number of AMC7000	3214	3214	2992
R-squared	0.80	0.07	0.01

Table 18: Estimates of the impact of weather shocks and reductions in transportation costs on out migration from rural areas (3214 rural municipalities (MCA - minimal comparable areas) for 1987 to 1991 and 1996 to 2000, by group: high and low Skill (men only where the cutoff between high and low skill is 12 years of education). T-statistics in parentheses (robust standard errors, clustered at the MCA level). * significant at 5%; ** significant at 1%.

Estimates for Distance as an Explanation for Migration Location Decision		
	Percent Rural-Urban Migrants	
Log Distance	-0.0160 (79.76)**	-0.0210 (81.40)**
Log Distance (for High Skilled)		0.0099 (29.32)**
Dummy for High Skill	Yes	Yes
Dummy for Women	Yes	Yes
Dummy for High Skilled Women	Yes	Yes
Municipality (MCA) Fixed Effects	Yes	Yes
Observations	790644	790644
Number of Rural Municipalities (MCA)	3214	3214
R-squared	0.04	0.04

Table 19: Estimates of distance as an explanation of choice in destination (123 cities/agglomeration) from rural areas (3214 MCAs) for each group: High and Low Skill (for the sample of men only where the cutoff between high and low skill is 12 years of education). T-statistics in parentheses (robust standard errors, clustered at the rural municipality (MCA - minimal comparable area)). * significant at 5%; ** significant at 1%.

First Stage of IV Ratio of Workers and Migration		
	Change in Log Ratio of High to Low Skill Workers	
	Men Only	
Log Ratio of High to Low Skill Migrants	0.054 (0.91)	0.716 (4.86)**
City (agglomeration) Fixed Effects	No	Yes
Observations	237	237
R-squared	0.00	0.19
Number of Cities (agglomerations)		123

Table 20: Estimates for the impact of ratio of high to low skill migrants on the ratio of high to low skill employment (first stage of instrumental variable regressions). Sample includes only men and low skilled people are defined as workers with less than 12 years of education and high skill have 12 or more years of education. T-statistics in parentheses (robust standard errors, clustered at the city level). * significant at 5%; ** significant at 1%.

OLS Estimate for Out Migration from Urban areas (Men Only)						
	Log Out Urban Migrants					
	Pooled				Low Skill	High Skill
	(1)	(2)	(3)	(4)	(5)	(6)
Log Rural-Urban Migrants	-0.079	-0.069	0.108	0.410	0.876	0.164
	(4.46)**	(3.86)**	(1.47)	(5.55)**	(1.52)	(0.21)
Log Natives	0.874	0.882	0.887	0.758	0.583	0.888
	(42.09)**	(42.78)**	(44.06)**	(12.09)**	(2.55)*	(6.42)**
Control for year 2000	No	Yes	Yes	Yes	Yes	Yes
Control for High skill	No	No	Yes	Yes	No	No
City (agglomeration) Fixed Effects	No	No	No	Yes	Yes	Yes
Observations	492	492	492	492	246	246
R-squared	0.9	0.91	0.91	0.83	0.26	0.72
Number of Agglomeration ID (123) living in now				123	123	123

Table 21: Estimates of the impact of rural-urban migration on out migration. The number of rural-urban migrants used is obtained from equation 10. Sample includes only men. T-statistics in parentheses (robust standard errors, clustered at the city level). * significant at 5%; ** significant at 1%.