Department of Economics- FEA/USP

Rainfall, Internal Migration and Local Labor Markets in Brazil

RAPHAEL CORBI
TIAGO FERRAZ

WORKING PAPER SERIES Nº 2018-24

DEPARTMENT OF ECONOMICS, FEA-USP
WORKING PAPER Nº 2018-24

Rainfall, Internal Migration and Local labor Markets in Brazil

Raphael Corbi (rcorbi@usp.br)

Tiago Ferraz (tiago.ferraz@usp.br)

Abstract:

We investigate the labor market impacts of weather-induced internal migration in Brazil between 1987 and 2010. We instrument the number of migrants at the destination municipalities using a two-step approach. First, we exploit the variation of out-migration flows from the Brazilian Semiarid, driven by deviations from historical average rainfall, to predict the number of internal migrants leaving their hometowns. Then, we distribute this predicted flow according to the preexisting support network in each destination based on the migrant's region of origin. Our results indicate that increasing in-migration rate by 1*p.p*. reduces native employment by 0.3*p.p*., mostly in the formal sector, decreases wages in the informal sector by 0.2% and deepens earnings inequality.

Keywords: Migration; labor supply; wage effects; rainfall

JEL Codes: J21; J22; J61; R23

Rainfall, Internal Migration and Local Labor Markets in Brazil

Raphael Corbi* and Tiago Ferraz†

Preliminar version: December 17, 2018

Abstract

We investigate the labor market impacts of weather-induced internal migration in Brazil between 1987 and 2010. We instrument the number of migrants at the destination municipalities using a two-step approach. First, we exploit the variation of out-migration flows from the Brazilian Semiarid, driven by deviations from historical average rainfall, to predict the number of internal migrants leaving their hometowns. Then, we distribute this predicted flow according to the preexisting support network in each destination based on the migrant's region of origin. Our results indicate that increasing in-migration rate by $1p.p.$ reduces native employment by $0.3p.p.$, mostly in the formal sector, decreases wages in the informal sector by 0.2% and deepens earnings inequality.

Keywords: Migration, labor supply, wage effects, rainfall

JEL Codes: J21, J22, J61, R23

*University of São Paulo, Department of Economics, email: rcorbi@usp.br
†University of São Paulo, Department of Economics, email: tiago.ferraz@usp.br

1 Introduction

Internal migration is a powerful tool to fight poverty and, therefore, plays a key role in reshaping the economic geography in developing countries. Individuals moving into a higher productivity area can have access to a more dynamic labor market, better educational and health systems as well as a wide range of public services that could improve their living conditions.

Migration can also be seen as a coping strategy for climate change. Although it is hard to precisely estimate its economic impacts, there is a growing concern that climate change could harm vulnerable populations in rural areas of developing countries (Skoufias et al., 2011). In such scenario, migration could help to attenuate the longer term welfare impacts of changes in environmental conditions (Chein and Assunção, 2008). At the same time, what happens to the regions where people are moving into is unclear. Some individuals may suffer because of increasing competition in the labor market, while others will benefit from a growing demand for goods and services, rendering the net effect a question to be empirically answered.

Our paper evaluates the impact of weather-induced internal migration from the Brazilian Semiarid region on employment and wages across destination municipalities, during the period 1987-2010. We follow a two-step approach to conduct our empirical analysis. First, we exploit exogenous rainfall shocks at the origin region to predict the number of individuals leaving their hometowns. Then, we use the pre-existent migrant network to allocate them in the destination areas. Thus, we can use this predicted flow as instrumental variable for the observed in-migration. [1]

Our results indicate that increasing the share of migrants by one percentage point reduces native employment by 0.3%, mostly in the formal sector, and decreases wages in the informal sector by 0.2%. This result is consistent with some degree of wage stickiness in the formal sector. Because those markets are more regulated and nominal wages can't be reduced, adjustment to shocks occurs at the employment margin, while the opposite happens in the informal sector. In addition, we show that internal migration worsens earnings inequality. Individuals

[1] For a similar strategy, see Boustan et al. (2010); Imbert et al. (2016).

below the 25th percentile of the wage distribution are negatively affected while those above the 85th percentile are benefited by migration shocks.

This paper is related to the literature on the economic impacts of internal migration on labor markets, like Boustan et al. (2010) who focus on the Great Depression period to study the effects of internal migration in the US. Imbert et al. (2016) use Chinese data to assess the impact of internal migration on manufacturing growth. More closely related to our work is Kleemans and Magruder (2017) who analyze the Indonesian labor market and show that internal migration reduces employment in the formal sector and wages in the informal sector.

The reasons for focusing on the Brazilian Semiarid are twofold. First, over this period more than 7 million Semiarid's residents (almost one third of the region's current population) decided to pursue a better life elsewhere. One should expect that such a large movement would have some impact on labor markets in destination areas. Second, although earlier works have already discussed the effects of immigration on labor markets, most of them focused on international migration to developed countries. Less attention had been paid to the impacts of internal migration. Such distinction is important because the number of individuals moving within countries is much bigger than international migrants and also because labor markets in developing countries usually are structurally different from those in richer economies. In a developing country, a regulated formal labor market usually coexists with a more competitive informal sector. One exception is the already cited paper from Kleemans and Magruder (2017), but in their case the focus is on the average effect on employment and wages.

Hence, our main contribution to this literature is to discuss how internal migration affects not only the average employment and wages, but also earnings inequality in a two-sector economy.

This paper is organized as follows. In the next section, we first present some background information on the Brazilian Semiarid region and labor markets, then we discuss the data used in our empirical analysis. In section 3, we explain the empirical strategy and the identifying assumptions we make. In section 4, we present and analyze the main results. Finally, we show some robustness checks in Section 5 and present our conclusions in Section 6.

2 Background and Data

In this section, we first describe the economic background and weather conditions at the Semiarid region and the functioning of local labor markets in Brazil, in an effort to contextualize our analysis.

We then discuss the main sources of data regarding labor market outcomes, migration flows and weather, and present some descriptive statistics.

2.1 Brazilian Semiarid

According to the official definition provided by the Ministry of National Integration, the Brazilian semiarid encompasses 1,133 municipalities distributed by 9 states, covering an area of around 976,000km^2 (roughly 11 percent of the country's territory).[2]. A municipality officially belongs to the semiarid region if at least one of these three criteria holds:

(i) yearly average precipitation below 800 mm (in the period 1961-1990);

(ii) aridity index up to 0.5 (measured by Thornthwaite Index, which combines humidity and aridity for a given area, in the same period);

(iii) has an index of risk of drought above 60% (defined as the share of days under hydric deficit, using the period 1970-1990).

Average historical precipitation in the Semiarid is about 740mm, as opposed to around 1,300 mm for the rest of the country, while average temperature is around 25°C. The rainy season occurs between November and April, with the highest levels of precipitation after February, when the sowing typically starts.

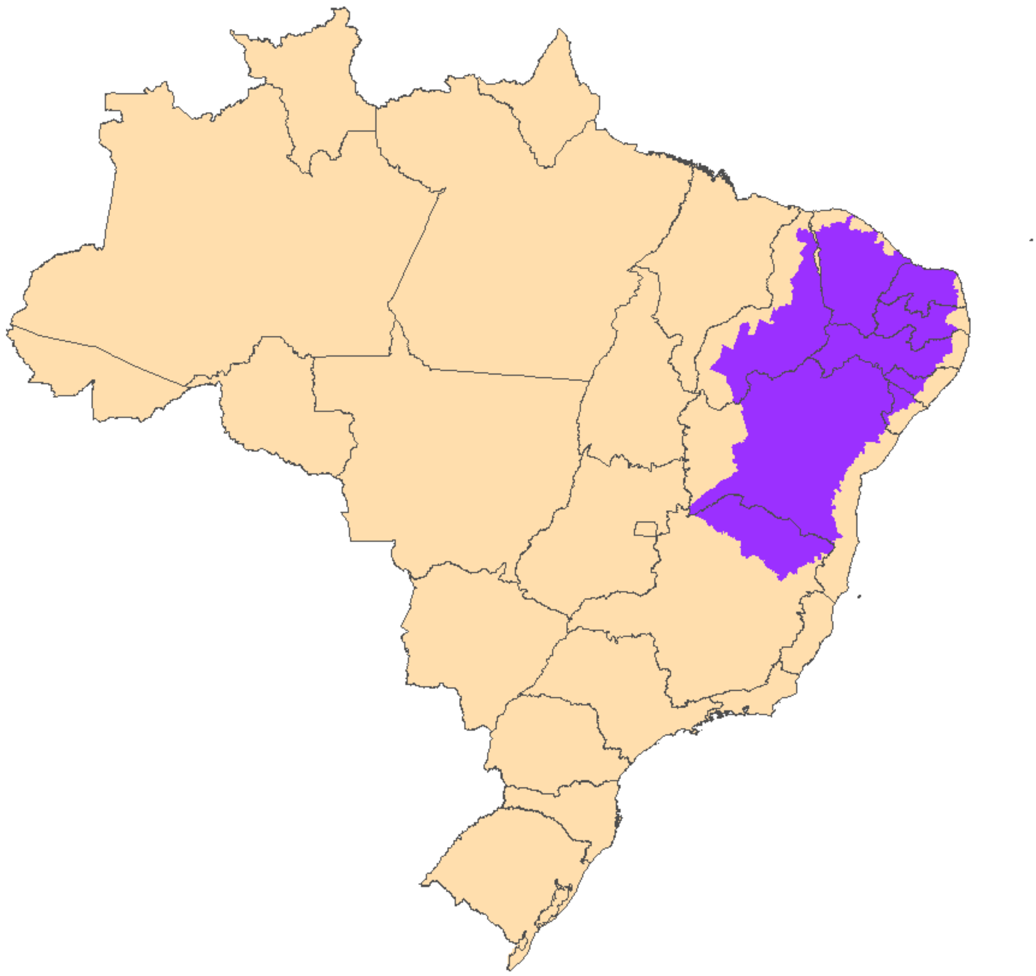
Since Colonial Era, Brazilian Semiarid faced a great number of episodes of severe droughts. Between 1877 and 1879, the event known as The Great Drought took the life of around half million people either by starvation or epidemics (Sousa and Pearson, 2009) and drove away hundreds of thousands more (Greenfield, 1986).

The Semiarid is one of the least developed regions in the country with 80% of the children in households below the poverty line and infant mortality reached

[2] It includes almost all Northeastern states, except for *Maranhão*, plus the northern area of *Minas Gerais*. See Figure 1.

31 per 1000 births in 1996, compared to a national average of 25% and 15 per 1000 births, respectively.[3] Municipalities are typically small (population median is around 20,000) and their economies are based on low productivity subsistence agriculture and cattle raising, both activities highly susceptible to suffer from climate shocks. Average human capital level is relatively low as 60% of the adult population have less than 8 years of schooling, as opposed to a national average of 45%.

Figure 1 – **Brazilian Semiarid**



[3] See Rocha and Soares (2015).

2.2 Local labor markets in Brazil

A common feature of labor markets in developing countries is the existence of a two-sector economy, where firms operating in the formal sector coexist with those in the more unregulated informal sector. When firms hire workers formally they have to comply with several legal obligations, like paying minimum wages or observing safety regulations, among others.

According to the Brazilian law, firms have to register every employee and

2.3 Migration, Labor Market and Weather Data

We used several sources to construct our main dataset. Migration data were extracted from the Brazilian Census (1980, 1991, 2000 and 2010) that provided information regarding the number of years in the destination and the origin municipality for all migrants and allowed us to construct a measure of yearly out-migration from each *origin* municipality in the Semiarid and a measure of in-migration to each *destination* during 1974-2010. There is a gap in our data because the 2000 Census only asked respondents where they lived 5 years before, so for this particular round, we can only track the individuals during the period 1996-2000. Our migration dataset spans over 1974-1991 and 1996-2010. Along this period several municipalities were split into new ones. In order to avoid potential miscoding regarding migration status or municipality of origin, we aggregate our data using the original municipal boundaries as they were in 1970.

We also built a measure of pre-existing networks by associating the share of migrants from each semiarid origin municipality in each destination, using the previous round of the Census. This is especially relevant for our identification strategy, discussed in more detail in the next section, to resolve endogeneity problems that could arise when migrants choose the place where they move to.

Weather data were retrieved from the Climatic Research Unit at University of East Anglia (Harris et al., 2014). The CRU Time Series provides worldwide monthly gridded data of precipitation and temperature, at the $0.5° \times 0.5°$ level ($0.5°$ is around 56km on the equator). We construct municipality-level monthly precipitation and temperature measures based on grid-level raw data as the weighted average of the municipality grid's four nodes using linear distances to the centroid as weights.[4]

[4] This methodology follows that used by Rocha and Soares (2015).

Then we define the rainfall shocks as monthly rainfall deviations from the historical mean. More specifically

$$Rainfall_{oy} = \ln\left(\sum_{\tau=1}^{12} r_{o\tau y}\right) - \ln(\bar{r}_o) \tag{1}$$

where $r_{o\tau y}$ is the rainfall in municipality of origin o in month τ of year y, and $\bar{r}_o$ is the municipality's historical average precipitation for the same months. Historical averages are calculated over the period from 1927 to 2010. Temperature shocks are computed in a similar way, using average yearly temperatures instead of the summation.

Table 1 describes municipality-level data for origin (Panel A) and destination (Panel B) municipalities. Semiarid's areas show lower levels of rainfall and slightly higher temperatures than destination municipalities. On average, 2.0% of Semiarid's population move to larger cities within the country every year.

We also used Census data to collect the labor market outcomes. We restricted the sample to individuals between 18 and 65 years old. The outcomes we use are the log of wages and several dummies indicating whether the individual is employed; whether the job is in the formal sector, the informal sector or self-employment; and some combinations of these dummies with indicators for low/high skilled and age below/above 30 years old.

To avoid biases from changes in demographics characteristics we residualized the outcome variables by running a regression on age dummies, five education levels (no education, elementary incomplete, high school incomplete, high school graduated and college graduated), five race categories (white, black, Asian, mulatto and indigenous), gender interacted with marital status and dummies for the 27 Brazilian states for each Census separately and took the average of these residuals at the municipality-by-Census round level.

Table 2 describes the individual data. All the regressions in our analysis use only native individuals at the destination areas, but we include statistics on migrants from the Semiarid in this table just for comparison. In our sample, 56% of individuals are employed, 21% have a formal job, 19% have an informal job and 16% are self-employed. The average monthly wage[5] is R$ 955.73. Compared to natives in our sample, migrants are younger, less educated and more likely to hold a formal job, although they receive lower wages (even when compared

[5] Wages are measured in R$ 2010.

to natives with the same skill level).[6] Their average monthly wage is R$ 651.50, roughly two thirds of the monthly wage for natives.

3 Empirical Strategy and Identifying Assumptions

In this section, we first describe the empirical framework that allows us to (i) isolate the observed variation in out-migration induced by exogenous rainfall shocks, and (ii) the in-migration flows into destination municipalities predicted by the pre-existing migrant networks. We then discuss and present supportive evidence on the validity of this procedure that is key to isolate the causal effect of in-migration on labor market outcomes for native workers.

3.1 Empirical Framework. Rainfall-induced Migration.

We specify a regression model of labor market outcomes of native individuals as a function of internal migration flows. Specifically we assume that

$$\Delta y_{dt} = \alpha + \beta m_{dt} + \epsilon_{dt} \tag{2}$$

where y_{dt} is a vector of labor outcomes at destination municipality d in census year t, m_{dt} is the municipality-level within-census cumulated migrant flow and ϵ_{dt} is the error term. By differencing the outcome variables we can account for time-invariant unobserved municipality-level characteristics that could be correlated with in-migration flows, but the error term may include unobserved time-varying confounders which would potentially bias OLS estimates. In particular, migrants could choose a specific destination municipality due to higher wages or job prospects.

In order to account for this endogeneity problem we follow a two-step procedure to construct the instrument for cumulated migration in the destination. First we project m_{oy}, the out-migration rate from origin municipality o in year y, onto weather shocks in the previous year:

$$m_{oy} = \alpha + \beta' Z_{oy-1} + \phi_o + \delta_y + \varepsilon_{oy} \tag{3}$$

[6] We define individuals as low skilled if they have not completed elementary school.

where Z is a vector of rainfall and temperature shocks at the origin municipality o in the previous year, ϕ_o is a municipality fixed effect, δ_y is a time fixed effect and ε_{oy} is a random error term. For each year the predicted number of migrants who leave their hometowns is obtained by multiplying this predicted rate by the total population in the previous Census:

$$\widehat{M}_{oy} = \widehat{m}_{oy} \times pop_{oy} \tag{4}$$

In the second step we use the pre-existent network of migrants from the origin o to municipality d in order to distribute them throughout the destination areas. An example may clarify this point. Suppose that 10% of the migrants from *Campina Grande* are living in *São Paulo* in 1991 and we predict that 1,000 people will leave *Campina Grande* in 1996. Then, we allocate 100 people as in-migrants to *São Paulo* in 1996.[7] This allocation procedure is similar to those implemented by Munshi (2003) and Kleemans and Magruder (2017). It is important to highlight that we are using two different time notations, because predicted out-migration is calculated for each year y while labor market outcomes are available only in census years t. Hence we aggregate the number of migrants entering each destination municipality d in year y into the five[8] years period t

$$\widetilde{M_{dt}} = \sum_{o} \sum_{y \in t} \widehat{M_{oy}} \times network_{od,t-1} \tag{5}$$

Finally we obtain the instrument for in-migration rate $\widetilde{m_{dt}}$ by dividing the number of migrants predicted to enter each destination $\widetilde{M_{dt}}$ by the municipality's labor force in the previous Census, which is completely predetermined, and plug it into our baseline specification 2. Hence the specification

$$\Delta y_{dt} = \alpha + \beta \widetilde{m_{dt}} + \epsilon_{dt} \tag{6}$$

can be thought as a reduced-form relationship that associates labor market outcomes and the cumulated predicted migrant flow at the destination.

[7] For each decade we use the predetermined share of migrants as reported in the previous Census to allocate them. We tested another specification with a constant network using only the 1980 census. Our first stage remains robust, but updating the network provides narrow intervals for the estimates.

[8] We cumulated the migration data into five years in order to make all periods comparable because in the 2000 Census we can only track individuals up to the previous five years. We also tried to aggregate the data for ten years before each Census, except for 2000 which is not possible. The results do not change very much.

3.2 Identifying Assumptions.

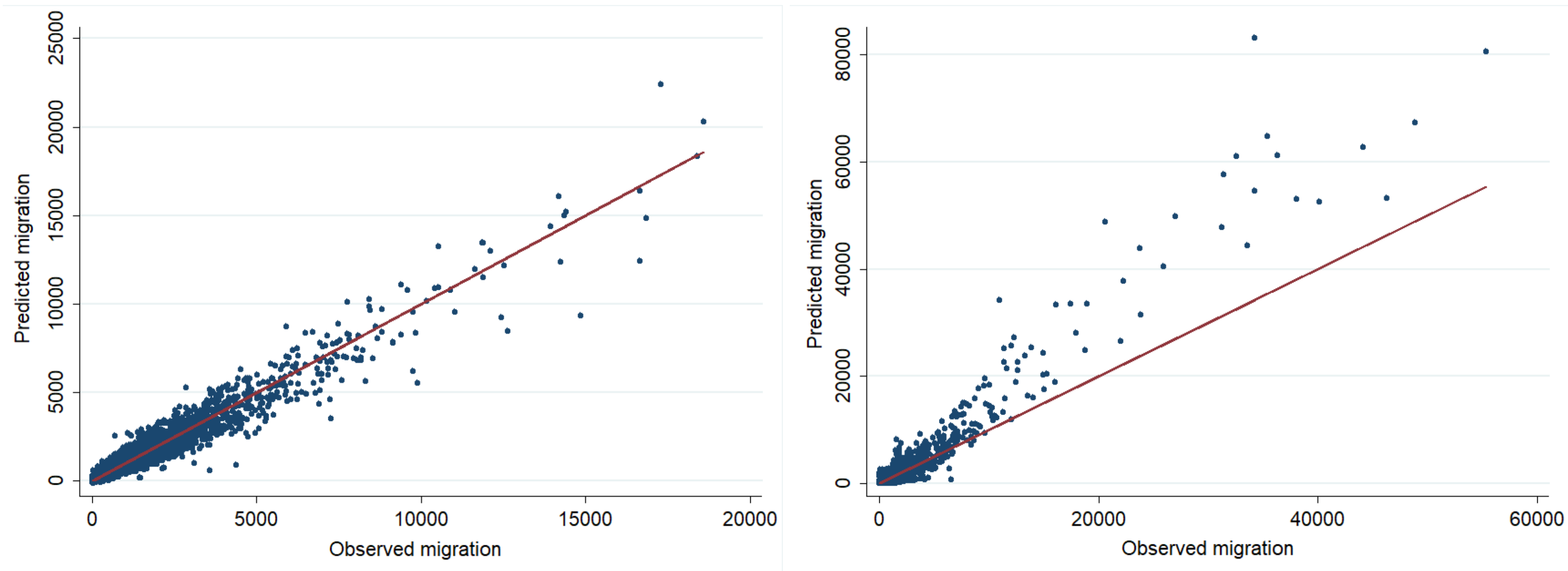
We use semiarid municipality-level data to estimate variations of specification 3 and report the results in Table 3. Columns (2)-(4) include additional lags of rainfall shocks. All specifications include municipality and year fixed effects and the standard errors are clustered at the municipality level.

As expected, rainfall shocks in the previous year are negatively correlated with out-migration indicating that Semiarid's inhabitants are induced to leave the region during drought periods. Coefficient estimates are remarkably stable across specifications and adding more lags in specifications (2) and (3) do not change the results. More important to our identification, we include as control rainfall and temperature shocks one year forward to ensure that our instrument is not contaminated by serial correlation in the weather measures. The coefficient on $rainfall_{y+1}$ reported in column (4) is insignificant and small in magnitude, while the coefficient for $rainfall_{y-1}$ is a little higher and remains significant. Our preferred specification is (1) as it yields a F-statistic of joint significance of 47.84.[9] Our estimates indicate that a municipality in the Semiarid where monthly rainfall is 10% below historical average will experience an increase of $2p.p.$ in the out-migration rate.

In the second step, we distribute the predicted out-migration shock using the pre-existent network of migrants from origin municipality o to destination d. A *sine qua non* requirement implicit in our empirical framework is that both predicted out- and in-migration rates, $\widetilde{m_{oy}}$ and $\widetilde{m_{dt}}$ respectively, are strongly correlated with their observed counterparts. Figure 2 plots observed and predict out-migration flows from origin municipalities on Panel (a) and show that most observations float around the 45^o line. A similar picture arises as we focus on in-migration on Panel (b).

9 Staiger and Stock (1997) suggests that the F-statistic should be greater than 10 when using one endogenous variable.

Figure 2 – **Predicted vs observed migration**



(a) Out-migration from Semiarid

(b) In-migration into Non-Semiarid

Table 4 shows the first stage relationship between observed and predicted in-migration rates. We include year and state fixed effects and the first stage relationship remains virtually unchanged. The point estimates are almost the same and the F-statistic is sufficiently high in all specifications. In columns (4)-(6) we show the first stage for each year, separately. Our preferred specification is column (2), which uses only year fixed effects and from now on we will use it in all the second stage regressions.

4 Migration Flows and Labor Market Outcomes

Now we turn our attention to the labor markets and try to answer the question: how does internal migration impacts wages and employment prospects of native workers? We begin by reporting estimates of the reduced-form relationship that associates the probability of being employed and predicted in-migration rate as in Equation 6.

If firms could perfectly replace native workers for migrants, we should expect negative coefficients in the reduced form regressions. Table 5 shows that our prediction holds. In columns (1)-(4) the dependent variable is the change in residual proportions of employment, while in columns (5)-(6) the outcomes are the residualized log-wages, in each sector. Increases in the predicted migration rate are associated with decreases on employment and wages.

In Table 6 we report the 2SLS-IV estimates. Columns (1)-(4) show the effect at the employment margin. Increasing the share of migrants by $1p.p.$ reduces the probability of being employed by $0.3p.p.$ About half of this effect comes from natives in the formal sector. For every 100 migrants from the Semiarid entering a destination municipality, 17 native workers in the formal sector lose their jobs. The effect on employment among informal workers and self-employed individuals is much smaller and estimated with less precision.

Columns (5)-(8) report the effect of an in-migration shock on wages. There is no significant average effect when we consider all native workers, although for those employed in the informal sector wages are reduced by 0.2%, but this coefficient is only significant at the 5% level.

All the results we found are consistent with having different degrees of nominal wage rigidity between sectors. Firms operating in the formal sector are more exposed to enforcement of labor regulation[10]. Because the law prevents them to reduce nominal wages, adjustment to shocks happens at the employment margin. Firms hiring informal employees have more flexibility to reduce wages, thus is more likely that we observe a decrease in earnings for those workers.

Our results are in line with those reported by Kleemans and Magruder (2017) who analyzed the Indonesian labor market and found that internal migration, induced by weather shocks, reduced formal employment and informal earnings, although in our case the magnitude of the effects on wages are smaller.

4.1 Effects on earnings inequality

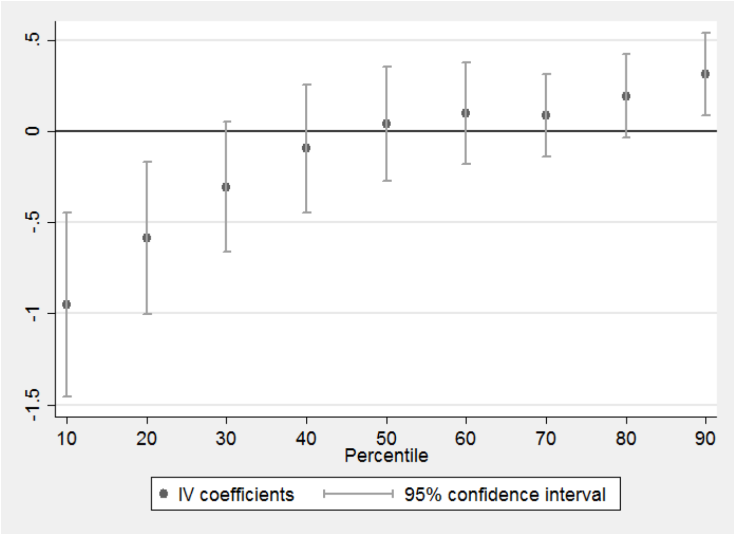
So far we have only discussed the average effect of on labor market outcomes. We address another important question: how internal migration affects earnings inequality? Our approach here is to use our instrumental variable to analyze the impact of migration along the wage distribution for native workers. This estimation strategy was developed by Dustmann et al. (2012), but in that case immigrants were not preassigned to skill groups according to observed characteristics. Their reason for doing this is that immigrants in the U.K. are, on average, more educated than native workers but they downgrade upon arrival. For Brazilian migrants from the Semiarid region this is unlikely to happen. In our sample, native workers are more educated than migrants and are paid higher wages, on average, in every sector.

We present the estimates for each decile of the earnings distribution in Table 7. Column (1) consider all native workers and show that individuals on the bottom of the wage distribution suffer with the competition from migrants while those with higher remuneration in fact perceive an increase in their earnings.

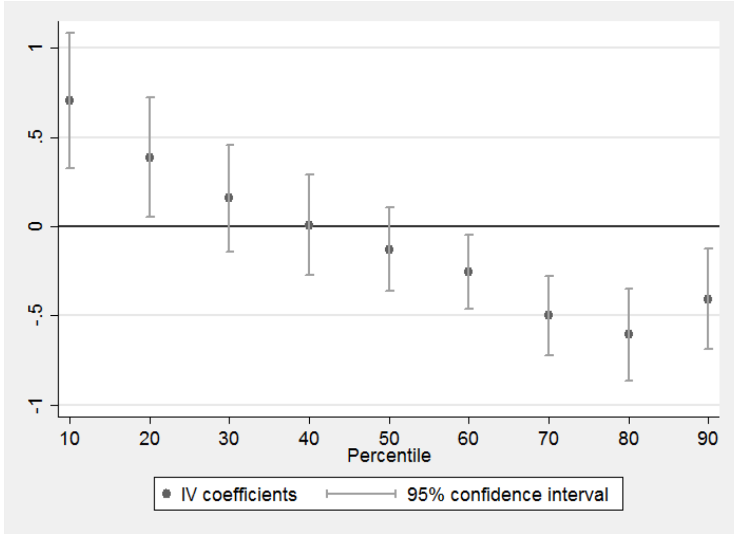
In columns (2)-(4) we break this results by sector. Interestingly the effects run in the opposite direction when we compare native workers in the formal sector with those holding an informal job or self-employed. While earnings inequality is deepening for the latter group, wage dispersion is actually being reduced for individuals employed in the formal sector. Figure 3 helps to illustrate this point.

10 (Almeida and Carneiro, 2012) argue that labor inspectors often focus on formal firms because they are easier to find than informal firms.

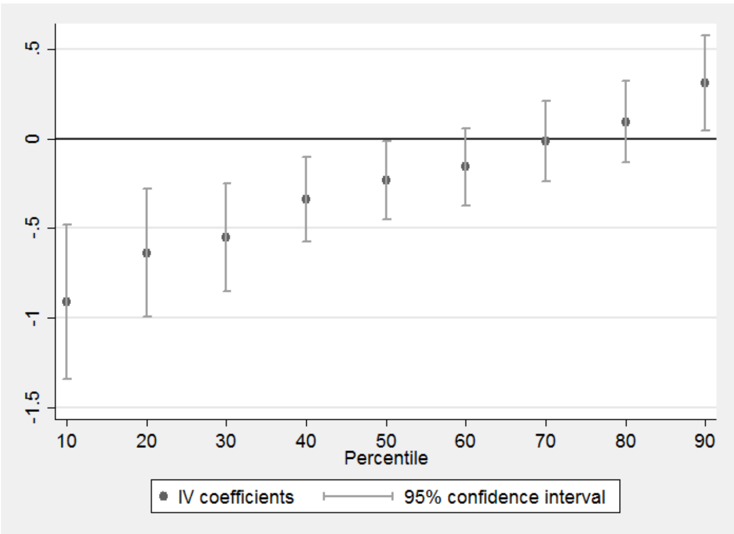
Figure 3 – **Impacts of internal migration across wage distribution**



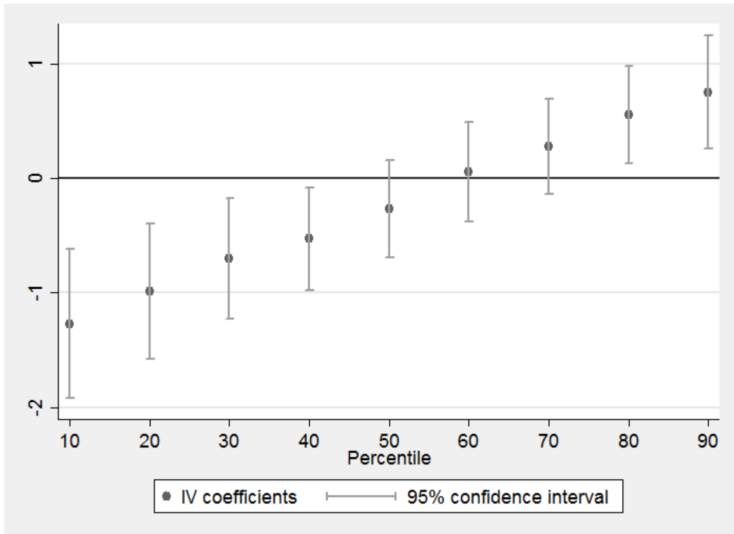
(a) Overall



(b) Formal Sector



(c) Informal Sector



(d) Self Employment

We also estimated the effect of internal migration on other measures of earnings inequality. We report in Table 8 the impacts on the change in the ratio of 90th to 10th and 50th to 10th percentiles and the Gini index of residualized log-wages. By any measure we adopt it seems that increasing migration worsens within-municipality wage inequality.

4.2 Heterogeneous effects

In the analysis presented so far we implicitly assumed that, once we controlled for individuals characteristics, native workers and migrants are perfectly substitutable. We relax this assumption now and consider the possibility that individuals with different skill levels or who belong to different age groups are not affected the same way.

In our data, the best measure we have for skill level is the number of years of schooling. We define as low skilled an individual who have up to 7 years of schooling, which is equivalent to an incomplete elementary education. In our sample, 64% of natives are low skilled in comparison with 72% of migrants from the Semiarid. High skilled individuals are those with 8 years or more of schooling.

One should expect that native workers who are more alike Semiarid's migrants would suffer from competition in the labor market. Table 9 confirms such prediction. Panel A shows the results for natives employed in the formal sector, where the adjustment is mostly at the employment margin. Young and low-skilled workers have a higher probability of losing their jobs. Although there is no effect on average wages, we also see that younger individuals perceive a negative effect while there is a positive impact for those who are more educated.

5 Robustness

In this section we assess the validity of our findings by performing some robustness checks.

First, our identification relies on the assumption that rainfall at origin municipalities affects destination labor markets only through internal migration. This assumption would be violated if, due to a low rainfall, a negative income shock at the origin had reduced trade flows with some of the destination areas, for instance. We believe that this problem is already addressed when we decided to use only destinations outside de Semiarid region. Nevertheless, we also test the consistency of our estimates using only long distance migration. If trade is the main driver for our results, then we should expect this effect to fade out with distance.

In Table 10 we compare our baseline estimates with different thresholds of long distance migration. The first-stage F statistic remains sufficiently high, showing that we still have a good instrument. The decrease on employment caused by one percentage point increase on migration ranges between 0.2 *p.p.* and 0.3 *p.p.*. The effect on wages is not statistically significant, although point estimates are very similar to the baseline.

Another concern we need to address is that weather shocks at the origin could be correlated with rainfall and temperature shocks at the destination municipalities. If the labor market outcomes are affected by weather conditions at the destination, as one should expect they are, then our identification assumption would not be valid. To account for this we run the same specification as in Table 6 including the 5-year average of rainfall and temperature shocks at the destination municipalities. Table 11 confirms that our main findings are mostly unaffected. The negative impact on employment and wages in the informal sector is slightly higher, but our conclusions continue to be valid.

6 Conclusion

In this paper we investigated the labor market impacts of weather-induced internal migration in Brazil. We exploit exogenous variation in the number of migrants entering each destination municipality by following a two-step approach. First, we explore the variation of out-migration flows from the semiarid driven by deviations from historical rainfall averages. Second, we distribute the predicted out-migration flow based on the pre-existing support network in each destination based on the migrant's region of origin. By adding the in-migration flow from each area of origin in each destination, we are able to calculate the predicted flow of migrants into each destination driven by exogenous shocks to rainfall in the origin and the pre-existing support network.

We find that increasing in-migration rate by $1p.p.$ reduces formal employment, especially for low skilled and younger natives, and have no effect on average wages, although it does affect different people in different ways. For those located at the bottom of the wage distribution in the informal sector and self employment, internal migration has a negative but increasing impact; and for natives holding formal jobs the impact is positive for those in the lower tail but decreasing and becomes negative for every decile above the median.

References

Almeida, R. and Carneiro, P. (2012). Enforcement of labor regulation and informality. *American Economic Journal: Applied Economics*, 4(3):64–89.

Boustan, L. P., Fishback, P. V., and Kantor, S. (2010). The effect of internal migration on local labor markets: American cities during the great depression. *Journal of Labor Economics*, 28(4):719–746.

Chein, F. and Assunção, J. (2008). Climate change, agricultural productivity and poverty. In *XXX Meeting of the Brazilian Econometric Society*.

Correia, S. (2016). Linear models with high-dimensional fixed effects: An efficient and feasible estimator. Technical report. Working Paper.

Dustmann, C., Frattini, T., and Preston, I. P. (2012). The effect of immigration along the distribution of wages. *Review of Economic Studies*, 80(1):145–173.

Greenfield, G. M. (1986). Migrant behavior and elite attitudes: Brazil's great drought, 1877-1879. *The Americas*, 43(1):69–85.

Harris, I., Jones, P., Osborn, T., and Lister, D. (2014). Updated high-resolution grids of monthly climatic observations–the cru ts3. 10 dataset. *International Journal of Climatology*, 34(3):623–642.

Imbert, C., Seror, M., Zhang, Y., and Zylberberg, Y. (2016). Migrants and firms: Evidence from china.

Kleemans, M. and Magruder, J. (2017). Labour market responses to immigration: Evidence from internal migration driven by weather shocks. *The Economic Journal*.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the us labor market. *The Quarterly Journal of Economics*, 118(2):549–599.

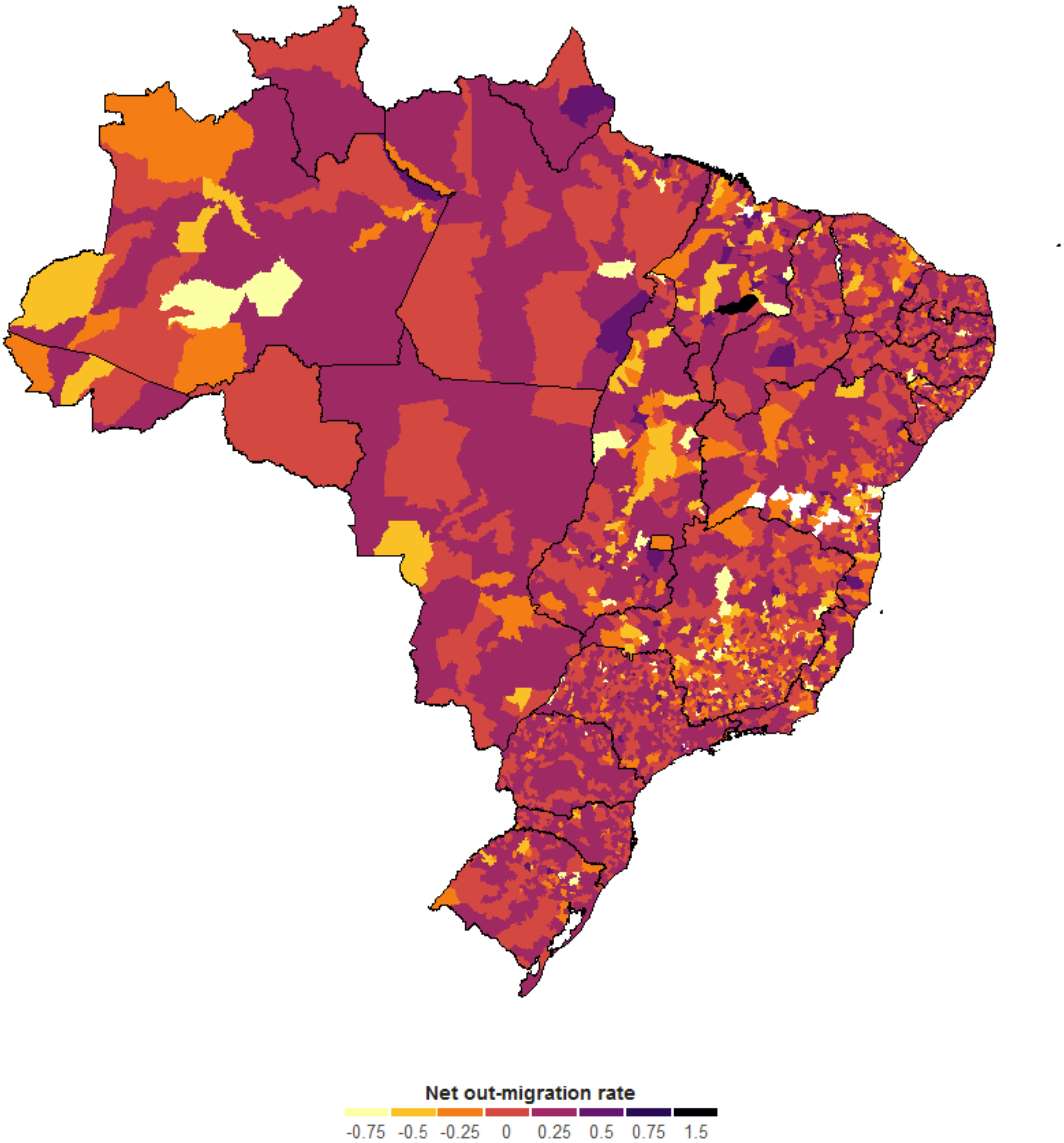
Rocha, R. and Soares, R. R. (2015). Water scarcity and birth outcomes in the brazilian semiarid. *Journal of Development Economics*, 112:72–91.

Skoufias, E., Vinha, K., and Conroy, H. V. (2011). *The impacts of climate variability on welfare in rural Mexico*. The World Bank.

Sousa, A. Q. and Pearson, R. (2009). Drought, smallpox, and emergence of leishmania braziliensis in northeastern brazil. *Emerging infectious diseases*, 15(6):916.

Staiger, D. and Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, 65(3):557–586.

Figure 4 – **Internal Migration - Brazil: 1970-2010 - Share of 1970 Population**



Source: Census Microdata obtained from IBGE.

Table 1 – **Summary statistics: Weather and migration data**

Panel A: Semiarid Region

Variables	Mean	Std. Dev.	Min	Max	Observations	Municipalities
Out-Migration	463.45	834.38	1.52	18,598	26,894	826
Predicted Out-Migration	467.59	831.1	-80.22	22,960	26,323	826
Population	23,522.57	40,249.98	0.00	832,850	26,894	826
Rainfall (level)	784.00	306.63	0.00	2,684.62	20,989	826
Temperature (level)	25.32	1.32	21.74	28.99	26,324	826

Panel B: Non- Semiarid Region

Variables	Mean	Std. Dev.	Min	Max	Observations	Municipalities
Rainfall (level)	1,343.96	627.60	0.00	5,057.09	69,911	2815
Temperature (level)	22.56	2.59	15.62	28.96	61,555	2815
In-Migration from Semiarid	66.94	744.23	0.00	55,356	69,911	2815
Predicted In-Migration from Semiarid	76.94	1,160.16	-0.34	83,704	68,912	2815
Population	44,192.15	240,032.86	0.00	10,435,546	69,911	2815
Labor Force	21,757.73	126,576.48	254.69	5,578,407	69,896	2,815

Notes: Rainfall is measured in mm. Temperature is measured in degrees Celsius. Population and labor force are those observed in the previous Census.

Table 2 – **Summary statistics: Natives and Semiarid's migrants in destination municipalities**

	All	Natives	Migrants	Diff
Male	0.49 (0.50)	0.49 (0.50)	0.48 (0.50)	0.008***
Age	34.69 (14.80)	34.80 (14.81)	29.53 (13.07)	5.265***
Black	0.07 (0.25)	0.07 (0.25)	0.05 (0.23)	0.012***
Low skilled	0.64 (0.48)	0.64 (0.48)	0.72 (0.45)	-0.077***
High skilled	0.36 (0.48)	0.36 (0.48)	0.28 (0.45)	0.077***
Employed	0.56 (0.50)	0.56 (0.50)	0.57 (0.50)	-0.013***
Formal Sector	0.21 (0.41)	0.21 (0.41)	0.25 (0.43)	-0.038***
Informal Sector	0.19 (0.39)	0.19 (0.39)	0.20 (0.40)	-0.013***
Self Employed	0.16 (0.36)	0.16 (0.37)	0.12 (0.33)	0.038***
Wages	955.73 (1886.29)	963.53 (1900.02)	651.50 (1198.67)	312.03***
Wages Formal Sector	1110.45 (1738.88)	1117.43 (1751.90)	821.68 (1028.93)	295.75***
Wages Informal Sector	534.32 (963.70)	539.16 (973.64)	398.57 (608.07)	140.59***
Wages Self Employed	1067.95 (2562.08)	1075.08 (2572.16)	701.03 (1939.26)	374.05***
Wages Low Skilled	614.42 (1105.24)	618.14 (1111.88)	501.00 (871.28)	117.14***
Wages High Skilled	1346.13 (2438.26)	1352.79 (2449.36)	967.72 (1647.19)	385.07***
Observations	33,457,327	32,787,967	669,360	

Notes: Low skilled indicates individuals with incomplete elementary schooling. Wages are measured in R$ 2010.

Table 3 – **Out-migration response to weather shocks**

	Dependent variable: Out-migration rate			
	(1)	(2)	(3)	(4)
Rainfall$_{y-1}$	-0.002***	-0.002***	-0.002***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
Rainfally$_{y-2}$		-0.002***	-0.002***	
		(0.000)	(0.000)	
Rainfall$_{y-3}$			0.001**	
			(0.000)	
Rainfall$_{y+1}$				-0.001
				(0.000)
F-Statistic	47.84	32.61	25.97	54.65
Observations	26,323	26,323	26,323	26,323
Municipalities	826	826	826	826
R-squared	0.508	0.508	0.509	0.511

Notes: Each observation is a municipality-year. Dependent variable is the number of individuals who left the origin municipality divided by the total population in the previous Census. Rainfall is measured as log-deviation from historical average. All specifications include controls for temperature shocks, municipality and year fixed effects. Standard errors are clustered at the municipality level. *** significant at the 1% level. ** significant at the 5% level. * significant at the 10% level.

Table 4 – **First Stage: Relationship between predicted and observed in-migration rates**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	1991-2010				1991	2000	2010
In-migration	0.658***	0.654***	0.618***	0.647***	0.885***	0.647***	0.578***
	(0.059)	(0.060)	(0.061)	(0.061)	(0.109)	(0.085)	(0.064)
Observations	8,442	8,442	8,442	8,433	2,810	2,811	2,812
Municipalities	2815	2815	2815	2812	2,810	2,811	2,812
Year FE	NO	YES	YES	NO	YES	YES	YES
State FE	NO	NO	YES	NO	YES	YES	YES
State-Year	NO	NO	NO	YES	NO	NO	NO

Notes: Dependent variable is the observed number of migrants from the Semiarid region entering each destination municipality, measured as a fraction of the labor force in the previous Census. The regressor is the predicted number of migrants in each destination municipality distributed by the preexistent network (also measured as a fraction of the predetermined labor force). Standard errors clustered at municipality level in parentheses. *** significant at the 1% level. ** significant at the 5% level. * significant at the 10% level.

Table 5 – **Reduced Form: Impacts of predicted internal migration, by sector**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Employment				Wages			
	Overall	Formal	Informal	Self Employed	Overall	Formal	Informal	Self Employed
In-migration	-0.195***	-0.111***	-0.042**	-0.042*	-0.095	0.021	-0.153**	-0.147
	(0.030)	(0.020)	(0.021)	(0.024)	(0.096)	(0.072)	(0.074)	(0.137)
Observations	8,442	8,442	8,442	8,442	8,442	8,432	8,434	8,421
Municipalities	2815	2815	2815	2815	2815	2815	2815	2815
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Columns (1)-(4) show the estimated coefficients from a OLS regression of the change in residual proportions of employment, in each sector, on the predicted in-migration rate. In columns (5)-(8) dependent variable is the residualized log-wages in each sector. Standard errors clustered at municipality level in parentheses. *** significant at the 1% level. ** significant at the 5% level. * significant at the 10% level.

Table 6 – **Second Stage: Impacts of predicted internal migration, by sector**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Employment				Wages			
	Overall	Formal	Informal	Self Employed	Overall	Formal	Informal	Self Employed
In-migration	-0.298***	-0.170***	-0.064*	-0.064*	-0.146	0.033	-0.233**	-0.224
	(0.057)	(0.032)	(0.033)	(0.039)	(0.143)	(0.112)	(0.113)	(0.206)
Observations	8,442	8,442	8,442	8,442	8,442	8,432	8,434	8,421
Municipalities	2815	2815	2815	2815	2815	2815	2815	2815
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Dependent variables are defined the same way as in Table 5. All regressions use 2SLS estimation. Standard errors clustered at municipality level in parentheses. *** significant at the 1% level. ** significant at the 5% level. * significant at the 10% level.

Table 7 – **Impacts of internal migration on wages across the distribution**

	(1)	(2)	(3)	(4)
	Overall	Formal	Informal	Self Employed
10th percentile	-0.951***	0.706***	-0.912***	-1.271***
	(0.254)	(0.192)	(0.217)	(0.329)
20th percentile	-0.586***	0.388**	-0.638***	-0.990***
	(0.211)	(0.168)	(0.181)	(0.300)
30th percentile	-0.304*	0.157	-0.550***	-0.703***
	(0.181)	(0.150)	(0.153)	(0.265)
40th percentile	-0.093	0.007	-0.339***	-0.528**
	(0.177)	(0.142)	(0.119)	(0.226)
50th percentile	0.040	-0.130	-0.234**	-0.269
	(0.159)	(0.118)	(0.111)	(0.214)
60th percentile	0.100	-0.255**	-0.159	0.057
	(0.141)	(0.104)	(0.108)	(0.218)
70th percentile	0.087	-0.501***	-0.013	0.277
	(0.115)	(0.113)	(0.114)	(0.208)
80th percentile	0.193*	-0.608***	0.094	0.553**
	(0.116)	(0.131)	(0.114)	(0.216)
90th percentile	0.315***	-0.407***	0.312**	0.752***
	(0.115)	(0.143)	(0.134)	(0.251)
Observations	8,442	8,432	8,434	8,421
Municipalities	2815	2815	2815	2815
Year FE	YES	YES	YES	YES

Notes: Estimated coefficients from a 2SLS regression of the change in the residualized log-wages on migration rate, by decile. Standard errors clustered at municipality level in parentheses. *** significant at the 1% level. ** significant at the 5% level. * significant at the 10% level.

Table 8 – **Impacts of internal migration on earnings inequality**

	(1) P90/P10	(2) P50/P10	(3) Gini
In-migration	1.266***	0.991***	0.248***
	(0.241)	(0.209)	(0.046)
Observations	8,442	8,442	8,442
Municipalities	2815	2815	2815
Year FE	YES	YES	YES

Notes: Estimated coefficients from a 2SLS regression of some measures of earnings inequality on migration rate. In columns (1)-(2) the dependent variable are the change in the ratio of 90th to 10th and 50th to 10th percentiles of residualized log-wages, respectively. In column (3) the dependent variable is the calculated Gini Index of residualized log-wages. Standard errors clustered at municipality level in parentheses. *** significant at the 1% level. ** significant at the 5% level. * significant at the 10% level.

Table 9 – **Heterogeneous effects: Impacts of internal migration, by skill level and age group**

	(1)	(2)	(3)	(4)	(5)
Panel A: Formal Sector	All Formal	Low-Skilled	High-Skilled	Below 30 years	Above 30 years
Employment	-0.170***	-0.202***	0.032*	-0.163***	-0.007
	(0.032)	(0.026)	(0.018)	(0.023)	(0.015)
Wages	0.033	-0.140	0.384***	-0.285**	0.206
	(0.112)	(0.147)	(0.117)	(0.125)	(0.132)

	(1)	(2)	(3)	(4)	(5)
Panel B: Informal Sector	All Informal	Low-Skilled	High-Skilled	Below 30 years	Above 30 years
Employment	-0.064*	-0.008	-0.056***	-0.007	-0.057***
	(0.033)	(0.030)	(0.012)	(0.019)	(0.018)
Wages	-0.233**	-0.064	-0.116	-0.351***	-0.104
	(0.113)	(0.138)	(0.210)	(0.122)	(0.126)
Municipalities	2815	2815	2815	2815	2815
Year FE	YES	YES	YES	YES	YES

Notes: Estimated coefficients from a 2SLS regression of residual proportions of formal and informal employment, by skill level and age group, on migration rate using the predicted in-migration as instrument. Standard errors clustered at municipality level in parentheses. *** significant at the 1% level. ** significant at the 5% level. * significant at the 10% level.

Table 10 – **Robustness Check: Labor market impacts of long distance internal migration**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Employment				Wages			
	Baseline	≥100km	≥200km	≥300km	Baseline	≥100km	≥200km	≥300km
In-migration	-0.298***	-0.231***	-0.267***	-0.307***	-0.146	-0.157	-0.138	-0.139
	(0.057)	(0.054)	(0.069)	(0.078)	(0.143)	(0.118)	(0.144)	(0.161)
First-stage F Statistic	119.7	113.5	71.89	64.32	119.7	113.5	71.89	64.32
Observations	8,442	8,441	8,432	8,417	8,442	8,441	8,432	8,417
Municipalities	2815	2815	2815	2815	2815	2815	2815	2815

Notes: Estimated coefficients from a 2SLS-IV regression. Baseline estimates are the same presented in Table 6. Columns (2)-(4) and (6)-(8) show the estimates of the same specification using different cutoffs for long distance migration. Standard errors clustered at municipality level in parentheses. *** significant at the 1% level. ** significant at the 5% level. * significant at the 10% level.

Table 11 – **Robustness Check: Labor market impacts of internal migration including destination weather shocks**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Employment				Wages			
	Overall	Formal	Informal	Self Employed	Overall	Formal	Informal	Self Employed
In-migration	-0.303***	-0.147***	-0.105**	-0.051	-0.181	-0.054	-0.372**	-0.323
	(0.073)	(0.040)	(0.044)	(0.051)	(0.204)	(0.152)	(0.155)	(0.290)
First-stage F Statistic	102.6	102.6	102.6	102.6	102.6	102.6	102.6	102.6
Observations	7,382	7,382	7,382	7,382	7,382	7,372	7,375	7,362
Municipalities	2479	2479	2479	2479	2479	2479	2479	2479

Notes: Estimated coefficients from a 2SLS-IV regression. Baseline estimates are the same presented in Table 6. Columns (2)-(4) and (6)-(8) show the estimates of the same specification using different cutoffs for long distance migration. Standard errors clustered at municipality level in parentheses. *** significant at the 1% level. ** significant at the 5% level. * significant at the 10% level.