Internal Migration and Labor Market Adjustments in the Presence of Nonwage Compensation

Raphael Corbi* Tiago Ferraz† Renata Narita‡

March 16, 2021

Abstract

We investigate the labor market impacts of internal migration in Brazil between 1996 and 2010 using a shift-share approach which combines weather-induced migrant flows from the Semiarid region with the past settlement patterns in each destination based on the migrant's region of origin. Our results indicate that increasing yearly internal migration rate by $1p.p.$ reduces formal employment by $0.30p.p.$, while increases the number of informal jobs by $0.15p.p.$ and self-employment by $0.10p.p.$ The lack of effect on overall employment can be explained by workers reallocating to informality, as well as compensation adjustments through wages and benefits. Consistent with downward wage rigidity in the formal sector, we find no effect on earnings but a negative impact on the share of workers receiving nonwage benefits. Earnings decrease in the informal sector by 1.5% and in self-employment by 1.4%, consistent with classic predictions from perfect competition and the fact that migrants and low skilled natives are close substitutes. In particular, less educated individuals appear to bear most of the welfare losses, both in terms of overall earnings and in offered nonwage benefits.

Keywords: Internal migration, wages, employment, nonwage benefits.
JEL Codes: J2, J3, J61, O15.

*University of São Paulo, Department of Economics, Email: rcorbi@usp.br
†University of São Paulo, Department of Economics, Email: tiago.ferraz@usp.br
‡University of São Paulo, Department of Economics, Email: rnarita@usp.br
For useful comments, we are grateful to Rodrigo Adão, Christian Dustmann, Richard Hornbeck, Peter Hull, Rudi Rocha, and conferences and seminars participants for helpful comments. The authors acknowledge financial support from FAPESP and CAPES. Raphael Corbi thanks the University of Chicago and Renata Narita thanks the Federal Reserve Bank of Chicago for their hospitality where parts of this work were completed.

1 Introduction

Migration, both within and beyond borders, has become an increasingly prominent topic in the international debate. While measuring movements of people is difficult, rough estimates indicate that global internal migration sits around 740 million (UNDP, 2009), approximately three times the estimated number of international migrants (UN DESA, 2017). There is a range of evidence about the positive impacts of migration on human development, through improved access to better economic opportunities, as well as education and health services. Indeed, internal migration can be a powerful tool to fight poverty and an important coping mechanism against climate change, especially for vulnerable populations in rural areas of developing countries (Skoufias et al., 2013; Assunção and Chein, 2016).

A separate but related literature addresses the impacts of migration on the native population at the destination. Classic economics of migration models with perfect competition and substitution between natives and migrants predict full adjustment through changes in wages, and no effect on employment when natives are immobile. In the case of wage rigidity, the supply shock associated with migration is accommodated entirely by job losses. However it is a common but usually hidden assumption to these and other models of perfect or imperfect competition that other labor costs e.g. nonwage benefits are held fixed such that we miss an important margin of labor market adjustment to local labor shocks arising from migration.[1] To the extent that workers value nonwage benefits, changes in this margin also can have important welfare implications.

In this paper we investigate the impacts of internal migration on employment, wage and nonwage compensation of native workers in receiving municipalities in Brazil. In order to address the empirical challenge that migrants tend to move to areas with better labor market opportunities, we build a shift-share instrument strategy in two steps. First, we exploit exogenous rainfall and temperature shocks at the origin to predict the number of individuals leaving each Semiarid's municipality. Then, we use the past settlement patterns to allocate them to destination areas. The resulting predicted inflow of migrants can be used as an instrument for observed migration.[2]

Brazil provides a good setup for our investigation for two main reasons. First, over 3 million people left their hometowns in the Brazilian Semiarid region during our sample period of 1996-2010, mainly due to weather crises. Second, workers are employed in either the unregulated informal labor sector that is supposedly less frictional or in the formal sector where minimum wage is binding and nonwage compensation is frequently offered. Indeed, over 31 million people or 20% of registered workers are

[1] See Clemens (2021) for a recent and fruitful discussion of non-employment margins that can be adjusted in response to minimum wage increases.

[2] In doing so, we build on work by Munshi (2003); Boustan et al. (2010); Kleemans and Magruder (2018), among others.

covered through employer-provided health insurance. After payroll expenses, this is the second highest component of total labor costs (ANS, 2019). Also, 40% of these workers receive food benefits, costing firms about 57% of the minimum wage per worker.[3] Nonwage benefits are also an important part of compensation in developed countries. In the US, employer-provided health insurance and other benefits account for around one-third of compensation costs (Clemens et al., 2018). 74% of firms in Europe paid non-base wage components such as benefits and bonuses in 2013 (Babecký et al., 2019).

Our results show that increasing the share of migrants by one percentage point reduces formal employment among native workers by $0.30p.p.$, while increases the number of informal jobs by $0.15p.p.$ and self-employment by $0.10p.p.$ These results are consistent with a binding minimum wage (above 70% of the median wage) such that migration shocks lead to lower formal employment as formal sector employers cannot adjust wages down, and individuals who lost their formal jobs being absorbed by informal firms or self-employment, which are more competitive labor sectors. We find no effect on overall employment, consistent with workers reallocating to informality as well as adjustments in the wage and nonwage benefits margins.

Indeed we do not find an effect on formal sector earnings but do find a negative impact on the share of formal workers receiving nonwage benefits such as health insurance in the range of $0.2p.p.$ to $0.6p.p.$, food from $0.3p.p.$ to $0.7p.p.$ and transportation subsidy from $0.5p.p.$ to $0.7p.p.$. Despite declines in the provision of non-monetary benefits, which increases labor demand, employment in the formal sector still reduces as we mentioned.[4] This additional adjustment margin of firms also explains why we find null effects on formal sector wages across the entire wage distribution and not only on bottom wages which is explained through minimum wages. For individuals employed in the informal sector or self-employed we show a 1.4% decrease on earnings mostly concentrated on the bottom third of the wage distribution, consistent with classic predictions from perfect competition and less skilled migration.

Heteregoneity analysis shows that all estimated effects are stronger for less educated native workers, which is consistent with the fact that they directly compete with Semi-arid's migrants. When compared to those highly educated, less educated individuals are 30% more likely to exit the formal sector and experience a 40% greater wage reduction. Moreover, as more low education workers earn close or equal to the minimum wage, the negative impact on the most frequently nonwage benefits provided by firms is greater for them. This suggests that welfare declines more for low income workers therefore rising welfare inequality among natives.

Finally, we find that labor force participation reduces by $0.14p.p.$ which may seem at odds with previous results since earnings fall in the informal sector and self-employment,

[3] Arbache and Ferreira (2001) based on various sources estimate the average cost of providing some job benefits in Brazil.

[4] Recent literature, as discussed in Clemens (2021), reports a negative correlation between minimum wage increases and health insurance provision, with the variation of this benefit offsetting about 15% of the cost with minimum wage increases.

and lower benefits are paid in the formal sector. By running separate regressions for head and non-head of the household, we find that almost all the impact on the employment margins comes from head of household workers, while the change in unemployment and inactivity rates are led by the non-head member, consistent with the added worker effect (Lundberg, 1985).[5]

Our work is related to the wide literature that examines the impact of migration flows on labor market outcomes of natives (see Dustmann et al. (2016) for a review). Most studies are concerned with international immigration to high-income countries, with particular attention given to Mexican immigration to the United States (Borjas, 2003) and, more recently, to immigration to Western Europe (Dustmann et al., 2012). Some of these studies find that the wages of natives are harmed by immigration (Borjas and Monras, 2017), while others find only a minor negative effect on native wages (Card, 2001), or even positive (Ottaviano et al., 2013; Foged and Peri, 2016) .[6] A smaller set of works explore environmental shocks to study the causal impact of internal migration on local labor markets in the US (Boustan et al., 2010; Hornbeck, 2012).[7] More closely related to our work is Kleemans and Magruder (2018) who study the impacts of internal migration in a developing country, Indonesia, from a two-sector labor market perspective. They show that internal migration reduces employment in the formal sector and earnings in the informal sector.[8] Our contribution to the economics of migration literature is twofold. First, we show that migration shocks can affect nonwage job attributes. Second, we provide evidence on the effects of internal migration on local labor markets in a large developing country, and show that different adjustment patterns across formal and informal sectors can be explained by mechanisms other than wages and employment.

We also add to the literature that explores how firms use adjustments in non-base wage components when facing adverse economic shocks (Babecký et al., 2019) or as an strategy to offset collective bargaining (Cardoso and Portugal, 2005), particularly when base wages are rigid (Babecký et al., 2012). We show that nonwage benefits are an important margin of adjustment in the case of labor supply shocks due to internal migration.

In terms of empirical strategy our paper relates to many works that uses shift-share instruments to identify the effects of immigration on a wide range of outcomes (Card, 2001, 2009; Ottaviano and Peri, 2012; Ottaviano et al., 2013; Cattaneo et al., 2015; Foged and Peri, 2016). In particular, we take advantage of a recent literature that provides a

[5]The "added worker effect" in a broader sense here refers to an increase in the labor supply of secondary earners (typically wives and children) when the primary earner (husbands) becomes unemployed or lose a formal sector job where benefits, sometimes extended to the family, are provided.

[6]Dustmann et al. (2016) argue that such often contradictory estimates are a result of (i) different empirical specifications (sources of variation), as well as the fact that labor supply elasticity differ across different groups of natives, and immigrants and native do not compete in the labor market within the same education-experience cells.

[7]See also Molloy et al. (2011) for a comprehensive literature review on the determinants of internal migration in the U.S. and Lagakos (2020) on urban-rural internal movements.

[8]This approach relates to the seminal work of Harris and Todaro (1970)

clear framework for distinguishing sufficient conditions for identification according to each source of variation, and how to properly compute standard errors (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2021; Jaeger et al., 2018; Adao et al., 2019).

This paper is organized as follows. In the next section, we first present background information on the Brazilian Semiarid region and labor markets, a simple framework for interpreting our findings, 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 on employment, wages and nonwage wage benefits. Section 5 considers the case in which individuals with lower levels of education may bear most of the welfare losses due to migration. Then we conclude.

2 Background and Data

In this section, we first describe the economic background and weather conditions at the Semiarid region, the functioning of local labor markets in Brazil, and a simple framework 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).[9] A municipality officially belongs to the Semiarid region if at least one of these three criteria holds: (i) annual average precipitation below 800 mm between 1961 and 1990; (ii) aridity index up to 0.5[10]; (iii) risk of drought above 60%[11].

Average historical precipitation in the Semiarid is about 780mm, as opposed to around 1,500 mm for the rest of the country[12], 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.

The Semiarid displays some very bad development indicators: about 80% of the children lived in households below the poverty line and infant mortality reached 31 per 1000 births in 1996, compared to a national average of 25% and 15 per 1000 births, respectively.[13] Municipalities are typically small (median population is around 20,000)

[9]It includes almost all Northeastern states, except for *Maranhão*, plus the northern area of *Minas Gerais*, but it does not includes any state capital.

[10]Measured by Thornthwaite Index, which combines humidity and aridity for a given area, in the same period.

[11]Defined as the share of days under hydric deficit, using the period 1970-1990.

[12]See Figure 11.

[13]See Rocha and Soares (2015).

and their economies are susceptible to suffer from climate shocks. Average human capital level is relatively low: more than 80% of the adult population had less than 8 years of schooling in 1991. Even though the situation improved along the years, this proportion was still above 60% in 2010.

2.2 Labor markets in Brazil

A common feature of labor markets in developing countries is the existence of a two-sector economy, with over 40% of individuals working in the informal sector (those without registration or who do not contribute to social security) including the majority of the self-employed who are also not protected through social security. When firms hire workers under a formal contract they are subject to several legal obligations, like paying minimum wages and complying with safety regulations. Registration also entitles the workers to other benefits such as a wage contract which in Brazil prevents downward adjustment, working up to 44 hours weekly, paid annual leave, paternity or maternity leave, retirement pension, unemployment insurance, and severance payments [e.g. Almeida and Carneiro (2012); Meghir et al. (2015); Narita (2020)]

If firms do not comply with working regulations they may be caught by the labor authorities and have to pay a fine. For example, a firm is fined about one minimum wage for each worker that is found unregistered, or the firm can be fined up to a third of a minimum wage per employee if it does not comply with mandatory contributions to the severance fund (Almeida and Carneiro, 2012).[14]

On the other hand, it is a well-known fact that compliant (formal) firms are those more visible to labor inspectors and thus subject to more inspections whereas informal firms are smaller and thus difficult to get caught. There are also more expected costs for formal firms associated with labor courts in case the worker is fired and decides to file a lawsuit against the firm. Judges decide in favor of workers in nearly 80% of cases (Corbi et al., 2021) All this points to a significant cost of operating in the formal sector, particularly for smaller firms. Imperfect enforcement and costly regulation are associated with high labor informality in the country.

To the extent that formal and informal sectors coexist in the sense that there is a strong overlap of their productivity distributions particularly in lower percentiles (Meghir et al., 2015), both sectors should be affected by the influx of migrants. That is, there are workers who are close substitutes to the migrant workforce thus implying more competition in either sector.

Finally, since internal migrants from the Semiarid are less educated, we anticipate that less educated natives are most negatively impacted by migrants as they compete directly with the incoming migrant workforce.

[14]The minimum wage is above 70% of the median wage in Brazil.

2.3 A simple theory

To interpret our findings, in this section we describe a simple model assuming perfectly competitive labor markets. Migration shifts the aggregate labor supply to the right in the destination region, and both migrants and natives are assumed to be perfect substitutes.

We begin by illustrating a case without institutions (e.g. minimum wages) that affect any labor market adjustment in the regulated sector and holding constant nonwage job attributes that are paid by firms. As Figure 1 shows, labor supply elasticity determines the extent to which migration affects employment vis-à-vis wages. In the extreme case in which supply is inelastic, migration negatively affects wages with no effect on employment of natives and absorbing all migrant workforce. On the other hand, with elastic labor supply, the reduction in wages make jobs less attractive for some native workers such that, at w_1, native employment reduces from L_0 to L_0'. The new equilibrium then determines the employment of migrants, L_1-L_0'.

However, in a real-world scenario, there are downward wage rigidities often imposed by minimum wage laws and collective bargaining agreements. There are also other components of labor costs that firms may adjust given wage constraints (McKenzie, 1980; Clemens, 2021).

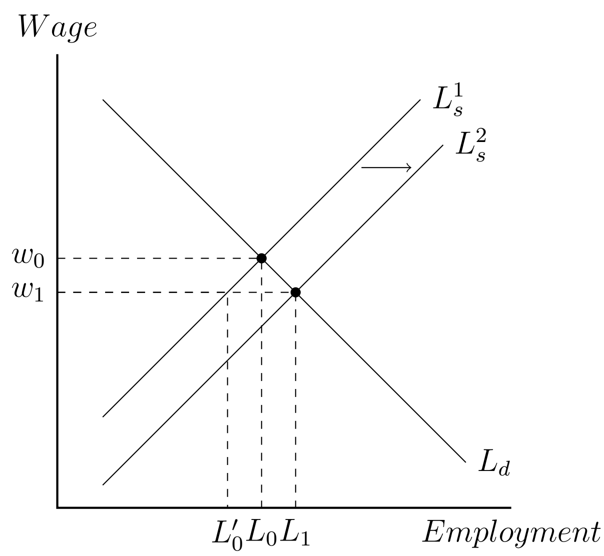
As we introduce wage rigidities, a supply shock due to migration needs to be accommodated by job losses or through cutting labor costs, for example, reducing nonwage benefits, e.g. health insurance. This situation is depicted in Figure 2 starting with an economy where the minimum wage is set at the market-clearing level and a migration shock that causes unemployment of L_1-L_0. Then, as it shows, reductions in nonwage compensation can shift both the supply curve and the demand curve. For firms, lowering nonwage benefits imply a higher labor demand curve because it increases its revenue net of costs. With wages fixed at $\underline{w}$, the new level of employment is L_2. For workers, under the assumption that they value such benefits, labor supply shifts upwards which is consistent with jobs becoming less attractive to workers and with a higher wage to compensate for the loss in benefits. In this case the shift of the supply curve due to adjustments in nonwage benefits may undo the migration supply shock and may even nullify its negative effect on employment. In this special case, demand and supply shifts due to a reduction in amenities bringing the economy to a new equilibrium which pays exactly the minimum wage and employment is at L_2, where there is no unemployment. Migration in this case increases total employment in the receiving region from L_0 to L_2 however some reduction of employment among natives may occur. Importantly, this reduction comes from some workers withdrawing the labor market since they are not willing to work at the lower benefit level.

In sum, in a simple competitive model with no rigidities, which is likely closer to the informal (unregulated) sector case, we expect some negative effects on wages and an increase in total employment. The effects on employment of natives depend on the labor supply elasticity. As we expand this simple model to consider both minimum wage regulations and the possibility of adjustment in nonwage benefits by firms, we find that the model yields ambiguous predictions regarding unemployment. The key aspect that determines this result is the valuation of benefits by workers compared to the cost of providing such benefits by firms.

Empirically, we expect these forces to be true mostly for low-skill workers or those at the bottom-medium but not at the top of the formal wage distribution. Low-skilled workers are the most affected by the migration shock since migrants are generally low-skilled.[15] Their wages are more affected by minimum wage policies and collective bargaining agreements but also benefits such as health insurance and food are disproportionately more generous for low than for high skill workers. Of course, since many firms or many industries combine both high and low skilled workforce there might be spillover effects to the former group due for example to complementarities, but empirically we expect the effects on high skilled workers to be of second order.

Finally, particularly in labor markets in developing countries, if jobs become rationed in the formal sector due to binding minimum wages, then workers can choose to work in the informal sector in which case unemployment may not increase and informal sector wages are expected to go down, consistent with predictions from perfect competition.

Figure 1: The Effects of Migration in a Perfectly Competitive Labor Market

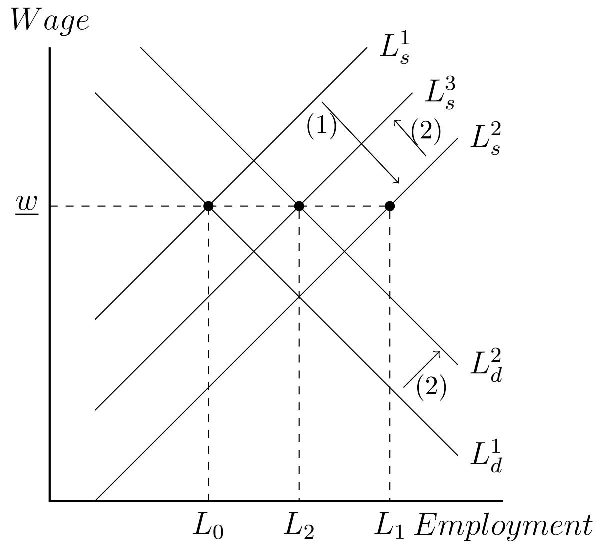


2.4 Migration, labor market and weather data

We collect data from three rounds of the Brazilian Census (1991, 2000 and 2010), provided by the *Instituto Brasileiro de Geografia e Estatística (IBGE)*, to construct a municipality-level panel of migration. For each census-year, they asked individuals who were not born in the municipality (i) how many years they lived in the current municipality; (ii) what is the municipality where they lived before. This allowed us to construct a

[15]This follows the arguments developed by the labor market model in Card and Lemieux (2001) and Borjas (2003).

Figure 2: The Effects of Migration with Binding Minimum Wages and Perfect Adjustment of Nonwage Benefits



NOTE: This figure extends the standard competitive labor market case presented in figure 1 to allow for downward wage constraints e.g. minimum wages ($\underline{w}$) and adjustments in nonwage benefits in response to the labor supply shock due to migration.

measure of yearly migration outflow from each origin municipality in the Semiarid and a measure of inflow to each destination during 1996-2010. 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. Since 1988 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 1991.

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 1991 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 into.

Weather data were retrieved from the Climatic Research Unit at University of East Anglia (Harris et al., 2020). 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 the inverse of the distance to the centroid as weights.[16]

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

[16]This approach is similar to that employed by Rocha and Soares (2015).

$$Rainfall_{ot} = \ln\left(\sum_{\tau \in \{GS\}} r_{o\tau t}\right) - \ln(\bar{r}_o) \tag{1}$$

where $r_{o\tau t}$ is the rainfall in municipality of origin o in month τ of year t, and $\bar{r}_o$ is the municipality's historical average precipitation for the same months. The index τ covers the 6-month growing season (GS).[17] Historical averages are calculated over the period from 1901 to 2010. Temperature shocks are computed in a similar way, using average 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, slightly higher temperatures and are less populated than destination municipalities. On average, 1.0% of Semiarid's population move to larger cities within the country every year.

We gather labor market outcomes from the *Pesquisa Nacional por Amostra de Domicílios (PNAD)* , a survey also conducted by the *IBGE*, covering 851 municipalities in all 27 states. Although the number of municipalities can appear to be small, these are the destination choice of about half of the migrants from the Semiarid region and concentrate around 65% of the employed population in Brazil. The survey is realized every year, except for those when the Census is conducted. Therefore, we have data from 1996 to 1999 and from 2001 until 2009.

We restricted the sample to individuals between 18 and 65 years old, living in the municipality for 10 years or more and we refer to them as natives. To avoid concerns about spatial correlation in the weather shocks we drop from the sample all the observations in municipalities that belong to the Semiarid region. The final sample has 2,153,328 individuals at 684 municipalities.

We use as outcomes the log of earnings and several dummies indicating whether the individual is employed; whether the job is in the formal sector, whether is informal or self-employment; whether she is unemployed and whether is out of the labor force; and some combinations of these dummies with indicators for low/high education level. We also create indicator variables for some forms of nonwage compensation. The survey asks specifically whether the individual received any kind of payment to help cover expenses with food, transport or health insurance. We create a dummy for each type, for those individuals who are employed in the formal sector. Finally, we pooled the 13 years of the survey and took the average of these variables at the municipality-by-year level.

Table 2 provides descriptive statistics for destination municipalities. In our sample, 63% of individuals are employed, 32% have a formal job, 16% have an informal job and 16% are self-employed. Unemployment rate is 13% and 24% of individuals are not in the labor force. The average monthly earning[18] is R$ 637.87, but is substantially higher in the formal sector (R$ 788.18) than in the informal sector (R$ 382.31) and a little higher than that for self-employed natives (R$ 600.83).

[17] We also calculated this measure using the 12 months of the year. Results are virtually the same.

[18] Earnings are measured in R$ (2012).

Among workers employed in the formal sector, 39% receive some help to cover expenses with food, 36% for transport and 20% for health expenditures.

3 Empirical Strategy

In this section, we first describe the empirical framework that allows us to (i) isolate the observed variation in migration induced by exogenous weather shocks, and (ii) the migration flows into destination municipalities determined by past settlements. We then discuss and present supportive evidence on the validity of this shift-share instrument approach based on insights of the recent econometric literature that analyzes its formal structure.

3.1 Main specification and identifying assumptions

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

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

where y_{dt} is a vector of labor outcomes at destination municipality d in year t, m_{dt} is the destination migrant inflow, X_{dt} are municipality-level controls, ψ_t absorb year fixed effects and ϵ_{dt} is the error term. By differencing the outcome variables we can account for time-invariant unobserved characteristics that could be correlated with migration inflows, 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 demand shocks leading to higher wages or job prospects.

We account for this endogeneity problem following a two-step procedure to construct an instrumental variable for the number of migrants entering a destination. First we predict m_{ot}, the migration outflow rate[19] from origin municipality o in year t, using weather shocks in the previous year:

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

where Z is a vector of rainfall and temperature shocks at the origin municipality o in the previous year, ϕ_o and δ_t are municipality and year fixed effects, respectively, and ε_{ot} 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 municipality population reported in the 1991 Census:

$$\widehat{M}_{ot} = \widehat{m}_{ot} \times P_o \tag{4}$$

[19]Defined as the observed number of migrants leaving the municipality divided by the population in the 1991 Census.

In the second step we use the pre-existing network of migrants from the origin o to municipality d in order to distribute them throughout the destination areas by defining our instrument as

$$\widetilde{m_{dt}} = \sum_{o=1}^{O} \frac{s_{od} \times \widehat{M_{ot}}}{P_d} \tag{5}$$

where s_{od} is the share of migrants from origin municipality o who settled in the destination area d in 1991.[20] and P_d is total population at d in 1991.[21] Thus our instrument $\widetilde{m_{dt}}$ can be thought as a combination of exogenous shocks or 'shifts' $\widehat{M_{ot}}$ (weather-driven outflows) and exposure 'shares' ($s_{od} \geq 0$) or past settlement patterns.

The validity of the shift-share instrument approach relies on on some assumptions about the shocks, exposure shares, or both, as discussed by a recent literature which analyzes its formal structure. Goldsmith-Pinkham et al. (2020) demonstrate that a sufficient condition for consistency of the estimator is the strict exogeneity of the shares. Alternatively, Borusyak et al. (2021) show how one can instead use the exogenous variation of shocks for identification by estimating a transformed but equivalent regression - at the origin level in our setup - where shocks are used directly as an instrument.

Based on the these insights, we leverage origin-level weather shocks for identification and define the reduced-form relationship that associates labor market outcomes and the predicted migrant flow at the destination as

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

which is the main specification we use throughout the paper. Regarding hypothesis testing, Adao et al. (2019) show that conventional inference in shift-share regressions are generally invalid because observations which similar exposure shares are likely to have correlated residuals, potentially leading to null hypothesis overrejection. In Appendix B, we take advantage of the equivalence result of Borusyak et al. (2021) to show that these inference issues are not empirically relevant in our setup.[22]

3.2 Weather-induced migration

We use Semiarid municipality-level data to estimate variations of specification 3 and report the results in Table 3. All regressions control for temperature shocks and the log of total population in the previous census; and include time and municipality fixed effects. In columns (2)-(8) we include a flexible trend interacting time dummies with

[20]For every year we use the predetermined share of migrants in the 1991 Census. We also tested another specification with an updated network using the predetermined flow of migrants in the previous Census. There is no much difference in the results, but updating the network may raise concerns about the persistence in migrant flows as discussed by Jaeger et al. (2018).

[21]In appendix A we further discuss our shift-share instrument in more detail.

[22]By estimating shift-share coefficients with a shock-level IV regression, Borusyak et al. (2021) also show that one obtains standard errors that are asymptotically valid regardless of the correlation structure of the error term.

1991 characteristics (age and the shares of high school and college educated individuals). Columns (3)-(6) include up to three lags, contemporaneous and one lead of rainfall and temperature shocks. Standard errors are clustered at the grid level to account for the fact that municipalities in the same grid will have similar shocks[23].

As expected, rainfall shocks in the previous year are negatively correlated with migration outflows indicating that Semiarid's inhabitants leave the region during drought periods. Coefficient estimates are remarkably stable across specifications and adding more lags do not change the baseline 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 $rainfall_{t+1}$ reported in column (6) is small in magnitude and not statistically significant, while the coefficient for $rainfall_{t-1}$ remains almost unchanged. Our estimates indicate that a municipality where annual rainfall is 10% below historical average will experience an increase of $1.0p.p.$ in migration outflow rate.

In the second step, we distribute the predicted migration outflows shock using the pre-existing network of migrants from origin municipality o to destination d. A *sine qua non* requirement implicit in our empirical framework is that both predicted migration inflow and outflow rates, $\widetilde{m_{ot}}$ and $\widetilde{m_{dt}}$ respectively, should be strongly correlated with their observed counterparts. Figure 3 illustrates that our prediction provides a reasonable approximation of the observed migration. Panel (a) shows the relationship between the predicted and observed number of migrants leaving the Semiarid region and entering all non-Semiarid municipalities, accumulated along the period 1996-2010. Panel (b) shows the predicted and observed numbers of incoming Semiarid migrants for destination municipalities with a maximum native population of 200,000 people.

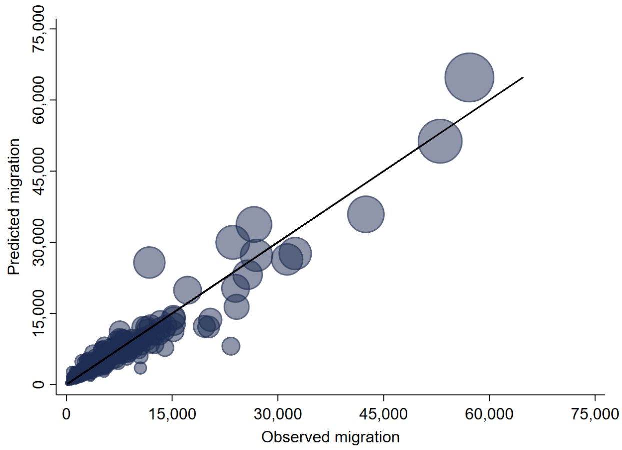
In Appendix A we describe our data source for weather shocks, discuss alternative measures of weather, and present further details about how we constructed our instrument including predicted and past settlement patterns.

Overall, this analysis shows that our strategy provides a strong first-stage as predicted migration rates, $\widetilde{m_{dt}}$, are strongly correlated with observed migration. For simplicity, as our main results we reports estimates of the effects of predicted migration on labor market outcomes in the next section. In section 4.1 we report the corresponding IV estimates (instrumenting observed migration by predicted migration) for our main findings. In particular, Table 9 reveals that our first-stage point estimates are close to a one-to-one relationship (0.92) - making the magnitude of reduced-form and IV estimates almost identical - and have an F-stat of 220.[24]

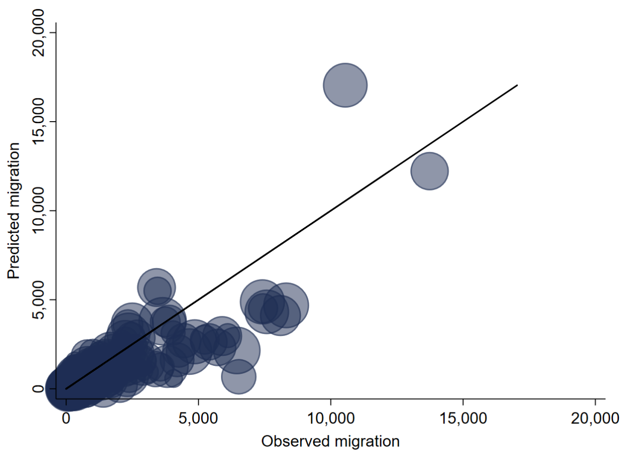
[23]Similar but no equal, because the shocks are computed by taking the average of the grid's four nodes, weighted by the inverse of the distance from each node to the municipality centroid.

[24]A sufficiently high F-stat avoids weak instrument concerns, especially in the light of the recent discussion in Lee et al. (2020) who show that a 5 percent test requires a F statistic of 104.7, significantly higher than the broadly accepted threshold of 10.

Figure 3: Observed vs predicted migration



(a) Migration outflow



(b) Migration inflow

Notes: This figure presents the relationship between the predicted and observed migration flows among Brazilian municipalities, accumulated between 1996 and 2010. Panel (a) shows the number of migrants leaving the Semiarid region to non-Semiarid municipalities. Panel (b) shows the number of migrants from the Semiarid entering each destination (destinations smaller than 200,000 natives in 1991). Circle size represents the municipality's total population in 1991. *Data source*: Census microdata (IBGE).

4 Migration Flows and Labor Market Effects

Now we turn our attention to labor markets at the destination and investigate how internal migration affects earnings, employment and nonwage benefits of native workers.

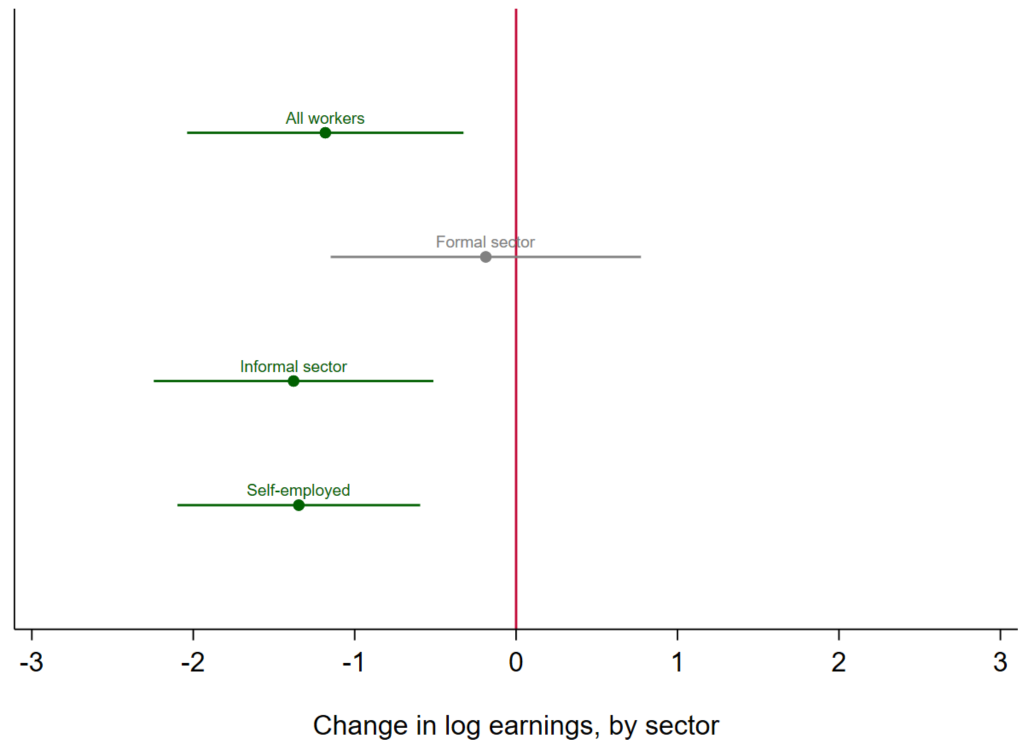
Effects on earnings. To test the implications of the simple model presented in Section 2.3, we first evaluate whether wages adjust at all in response to the migration flows. Figure 4 summarizes our main findings, while in Table 4 we present the estimates from several specifications for the reduced form equation (6). Column (1) displays a flexible specification, without any control. In column (2) we control for the log of the native population in the previous year and for changes in a set of municipality-level demographics (the share of individuals with high school and college education; the share of female, white and black individuals; and a cubic polynomial of age). Column (3) also controls for these changes in demographics and includes time dummies. Finally, there are two important changes occurring during our period that could be a concern for our identification: population aging and an increase in education levels. To account for this, in Column (4) we include a linear trend of age and the shares of high school and college educated individuals. All regressions are weighted by the 1991 native population to account for the fact that heterogeneity in municipality size could lead to differential effects.

Panel A reveals a strong negative effect of the predicted inflow of migrants on average log earnings for native workers. The inclusion of controls diminishes a little the magnitude of our coefficient of interest, but that does not change substantially our main conclusion. One percentage point increase in the predicted number of migrants reduces earnings by 1.3%. In Panel B we restricted our analysis to native workers holding a formal job, while in Panels C and D we focus on those in the informal sector and self-employed, respectively. There is no significant effect on the earnings of formal workers, but for those individuals in the informal sector and self-employed earnings are reduced by 1.5% and 1.4%, respectively.

Any downward wage restrictions such as minimum wages or collective bargaining agreements may prevent wages from falling in the formal sector. This explains the negative however small and insignificant effect of migration on formal sector wages. However, in the informal sector, the large negative impact on wages is consistent with absence of downward wage rigidity in this sector such that the classic predictions from perfect competition prevail. Regarding self-employment, if the demand in the product market is fixed, there is more competition faced by native self-employed workers therefore the supply shock due to migration causes earnings to fall.

Effects on employment. Our wage results imply that it is more expensive to hire workers in the formal than in the informal sector. If the informal sector and self-employment absorbs all workforce who left the formal sector, we should not expect impacts on total

Figure 4: **Effects of predicted migration on earnings**



Notes: This figure shows the relationship between changes in log earnings, by sector, and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Controls include time dummies, the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Standard errors are clustered at the municipality level. Green markers are statistically significant at the 5% level.

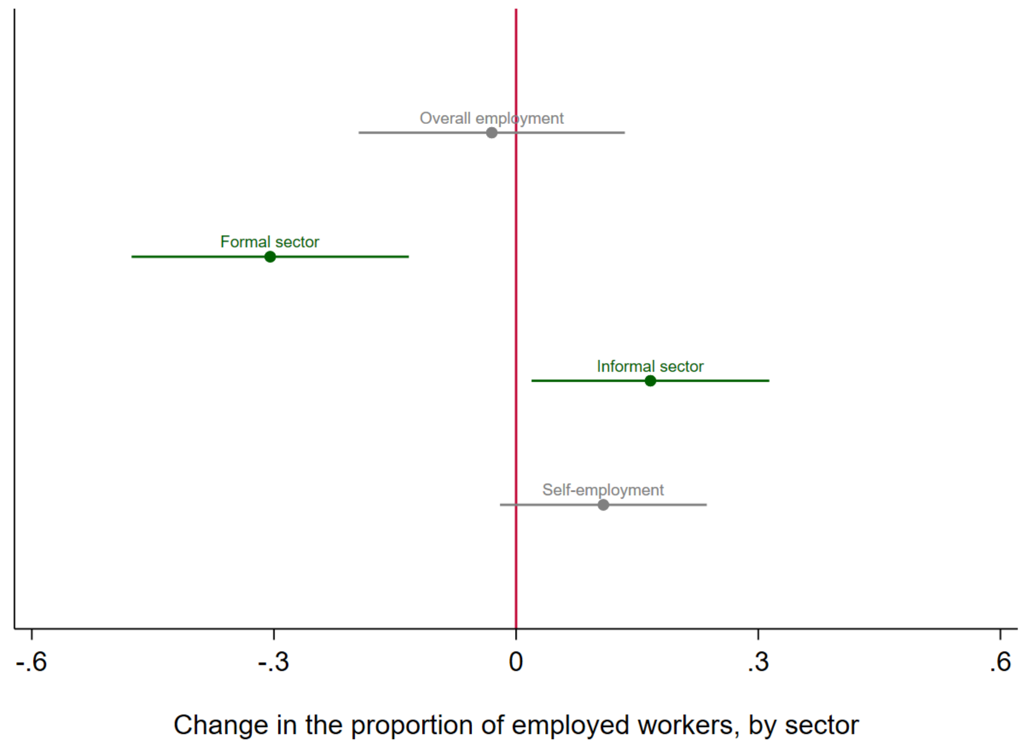
employment. Figure 5 summarizes our main employment results, while in Table 5 we present the point estimates.

As we see, there is a small negative but insignificant effect on overall employment, however predicted migration does affect the composition of workers across sectors. Our estimates in Panels B to D imply that one percentage point increase in the predicted inflow reduces the share of formal employment by $0.3p.p.$, and increases the share of informal workers and those in self-employment by $0.15p.p.$ and $0.10p.p.$, respectively. Most of these estimates are precise and stable across specifications.

The results by sector are consistent with the two versions of the model we presented in Section 2.3. When the labor market in a receiving municipality is hit by a supply shock, native workers face more competition. In a simple competitive labor market setting with elastic labor supply, employment of natives is expected to reduce. When we introduce downward wage constraints in the model, in absence of any other margin of adjustment, then employment may further reduce. This explains the negative impact on formal employment.

On the other hand, the share of native workers in the informal sector and self-

Figure 5: **Effects of predicted migration on employment**



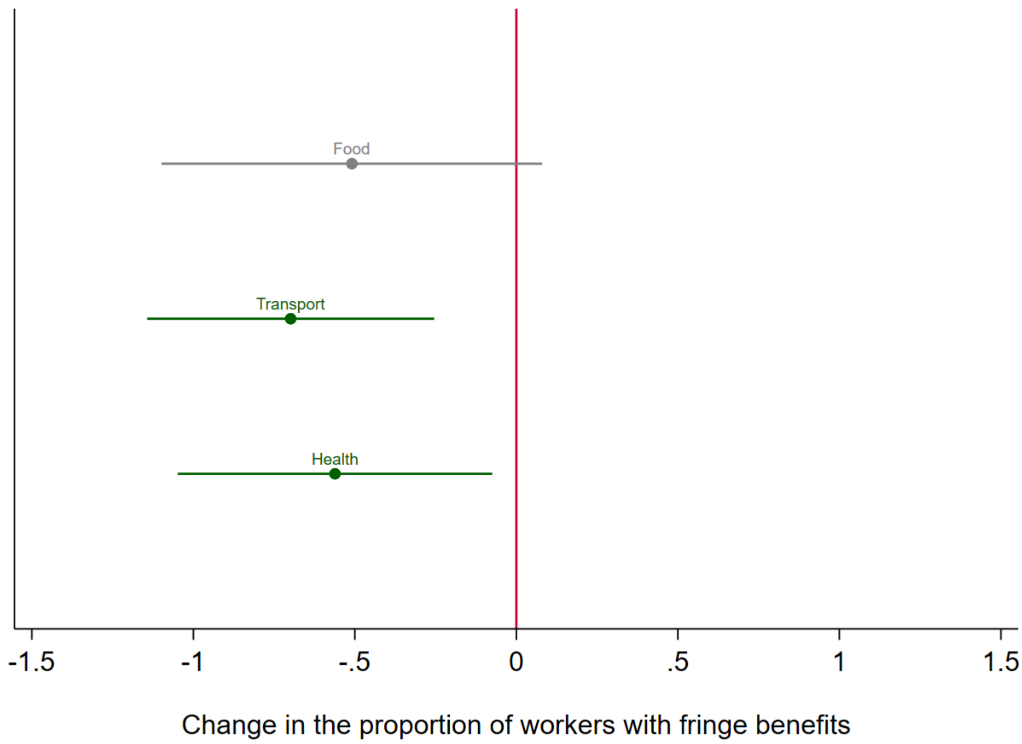
Notes: This figure shows the relationship between changes in employment rate, by sector, and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Controls include time dummies, the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Standard errors are clustered at the municipality level. Green markers are statistically significant at the 5% level.

employment rise consistently with the existence of a dual labor market, with an unregulated "sector", where the minimum wage and other labor laws do not apply.

Effects on nonwage compensation. Our paper exploits an additional important margin of adjustment due to migration shocks. Because firms operating in the formal sector cannot reduce wages below the legal minimum they may adjust to labor supply shocks by reducing their fringe benefits offerings. We focus on individuals who are currently holding a formal job because these benefits are typically offered by firms operating in this sector. Figure 6 presents the main results for this mechanism. Corresponding estimates are provided in Table 6. Although in this case the results are more sensitive to the specification, all point estimates are negative and suggest an important effect. We find that a one percentage point increase in the predicted number of migrants reduces the share of workers receiving health insurance in the range of $0.2p.p.$ to $0.6p.p.$, transportation from $0.5p.p.$ to $0.7p.p.$, and food subsidy from $0.3p.p.$ to $0.7p.p.$.

In summary, the model we presented in Section 2.3 generates predictions broadly consistent with our findings. In the formal sector, we observe no change in earnings, a decline in nonwage job benefits, and a decrease in employment of natives, which occurs due to wage rigidity and an imperfect adjustment of the benefit margin. In the informal sector and self-employment, where the minimum wage does not apply, we find a decrease in earnings. We also find an increase instead of a decrease in employment, consistent with workers reallocating from the formal to the informal sector or self-employment.

Figure 6: **Effects of predicted migration on nonwage compensation**



Notes: This table shows the relationship between changes in the proportions of working-age native population, employed in the formal sector, who receive some assistance from the employer to pay for food, transport or health expenses and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Controls include time dummies, the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Standard errors are clustered at the municipality level. Green markers are statistically significant at the 5% level.

Differential effects along the earnings distribution. We also investigate the existence of differential effects according to the native worker's position in the earnings distribution and present the estimates in Figure 7. We estimate the reduced form on each decile of earnings, by sector. For those native workers employed in the formal sector, there is no statistically significant effect along the entire distribution. Benefits' adjustment can help explaining why we find null effects on formal sector wages across the entire wage distribution and not only on bottom wages which is explained through minimum wages.

For informal and self-employed workers, the impact is substantially stronger for those at the bottom third of both distributions, consistent with classic predictions from perfect competitition and greater substitutability between migrants and less skilled natives in these sectors. To a smaller extent, migration also affects higher earnings deciles of informal sector workers and self-employed. The negative impact of migration in this case is attenuated due to some formal sector workers moving into informality or self-employment. Since formal sector workers are on average more productive, this increases earnings in higher percentiles when they move to other sectors.

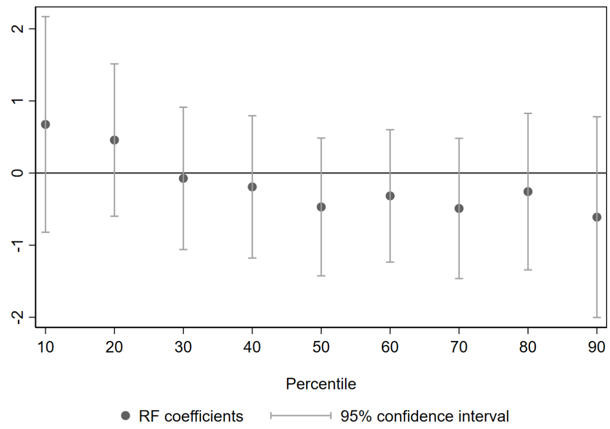
Effects on unemployment and labor force participation. Finally, to draw a more complete picture we also look at the impacts on unemployment and labor force participation. In Panel A from Table 7 the dependent variable is the change in the municipality-level proportion of unemployed native workers. We observe that migration inflows lead to an increase in the unemployment rate ($0.16 p.p.$). In Panel B, the outcome is the change in the proportion of individuals out of the labor force and the point estimate is almost identical, even though with opposite sign. But, in this case is hard to tell ex-ante what should be the most likely effect. On one hand, increased competition in the labor market could discourage native individuals to work if wages or benefits fall down. On the other hand, if the primary earner in the household looses his/her job because of the increased competition, then it is possible that other members of the household would enter the market, a phenomenon known as the added worker effect (Lundberg, 1985). We test this second mechanism running the same regressions separately for individuals identified as head or non-head of the household. As we can see in Table 8, almost all the impact on the employment margins comes from those native workers who are head of the household, while the change in unemployment and inactivity rates are led by the non-head members. This confirms our intuition that the second channel prevails. Also, the symmetry between the effects on unemployment and inactivity suggests that once secondary earners enter the market, it takes time for them to actually find a job.

4.1 IV Estimates

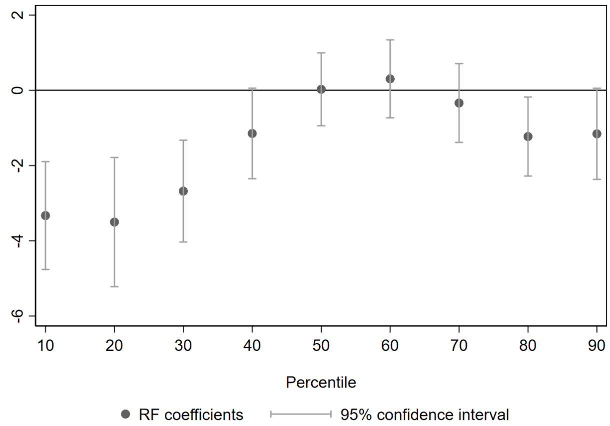
In this section we present the two-stage least squares (2SLS) estimates of the parameters of interest using the predicted migration rate as an instrument for the observed inflow. In Table 9 we report OLS estimates in columns (1), (3), (5) and (7) and IV estimates in columns (2), (4), (6) and (8), as well as the associated first-stage coefficients and F-statistics. All regressions are weighted by the total native population in 1991 and include time and state dummies, as well as demographic controls. Standard errors were calculated via bootstrapping to account for the fact that our instrumental variable was created with an estimated shock.

As illustrated by Figure 3 and briefly discussed in the end of Section 3.2, Table 9 shows that our first-stage point estimates are close to a one-to-one relationship (0.92) - making the magnitude of IV and reduced-form estimates (reported in Table 5) almost

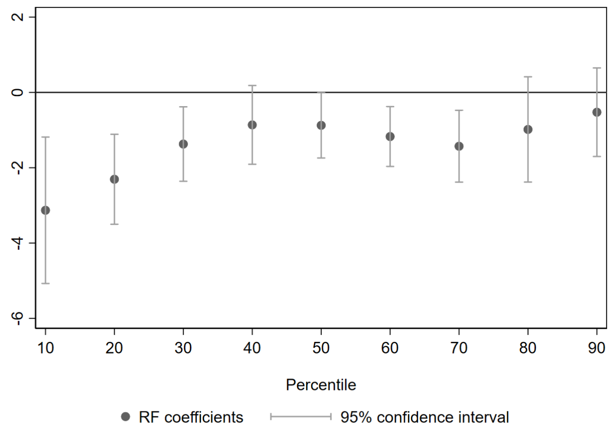
Figure 7: Effects of predicted migration along the earnings distribution



(a) Formal sector



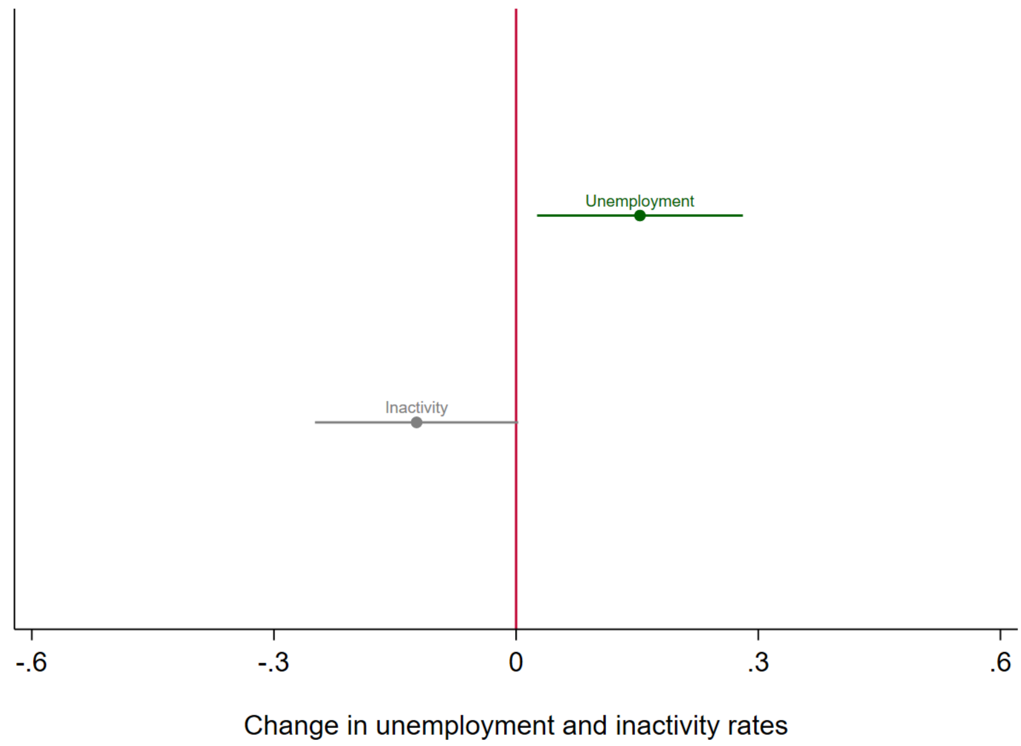
(b) Informal sector



(c) Self-employment

Notes: This figure presents the impacts of predicted migration inflow along the earnings distribution, by sector. Panel (a) uses only formal workers, while in Panels (b) and (c) we restrict the sample to individuals holding an informal job and self-employed, respectively. Each point is the coefficient from a regression of the change in the average of log earnings, in each decile, on the predicted number of migrants from the Semiarid, measured as a fraction of the native working-age population in 1991. Controls include time dummies, the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. The capped lines show the 95% confidence interval. Standard errors are clustered at the municipality level.

Figure 8: **Effects of predicted migration on unemployment and participation**



Notes: This figure shows the relationship between changes in unemployment and inactivity rates and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Controls include time dummies, the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Standard errors are clustered at the municipality level. Green markers are statistically significant at the 5% level.

identical. A F-stat of 220 avoids weak instrument concerns, especially in the light of the recent discussion in Lee et al. (2020).

At a first glance, it is not clear what to expect from the direction of the OLS bias in our setup. On the one hand, one could expect that migrants moving into booming municipalities would lead OLS estimates to be biased upwards. On the other hand, attenuation bias could dominate if classical measurement error is relevant in observed migration. Indeed, coefficients in Tables 9 show that OLS estimates are biased toward zero across most specifications across formal, informal and self-employment workers. The same is true for the estimates on labor force participation reported in Table 10, but not for unemployment.

4.2 Sensitivity checks

Now we summarize a series of robustness checks we have performed to assess the validity of our main findings.

The first issue we address is whether a shift in local labor demand may be confounding our identification. If that was the case, then we should expect that migrants from other regions outside the Semiarid would be attracted for the same destinations. In other words, we should observe a positive correlation between migrant inflows from the Semiarid and that from other regions. In Table 11 we show the coefficients from a regression of the migration inflow rate of migrants from other regions on our instrument. Column (1) includes time and municipality fixed effects. In Column (2) we add the same set of controls from our main results and Column (3) also includes the same linear trends as before. There is no statistically significant impacts in any specification.

The second issue is that our strategy relies on the assumption that rainfall at origin municipalities affects destination labor markets only through internal migration. One possible violation of this assumption would be if a negative income shock at the origin, due to low rainfall levels, had reduced trade flows with some of the destination areas, for instance. In this case, one should expect higher effects in those industries more exposed to trade shocks, like agricultural or manufactured goods. In Table 12 we report the coefficients from a regression of the predicted inflow rate on changes in log earnings by industry. Panel A and B show the effects on native workers employed in the agricultural and manufacturing sector, respectively. In both cases, although the coefficients display a negative sign, they are not significant. In contrast, in Panel C we show that the effect on those employed in services have larger magnitude and are statistically significant across all colums.

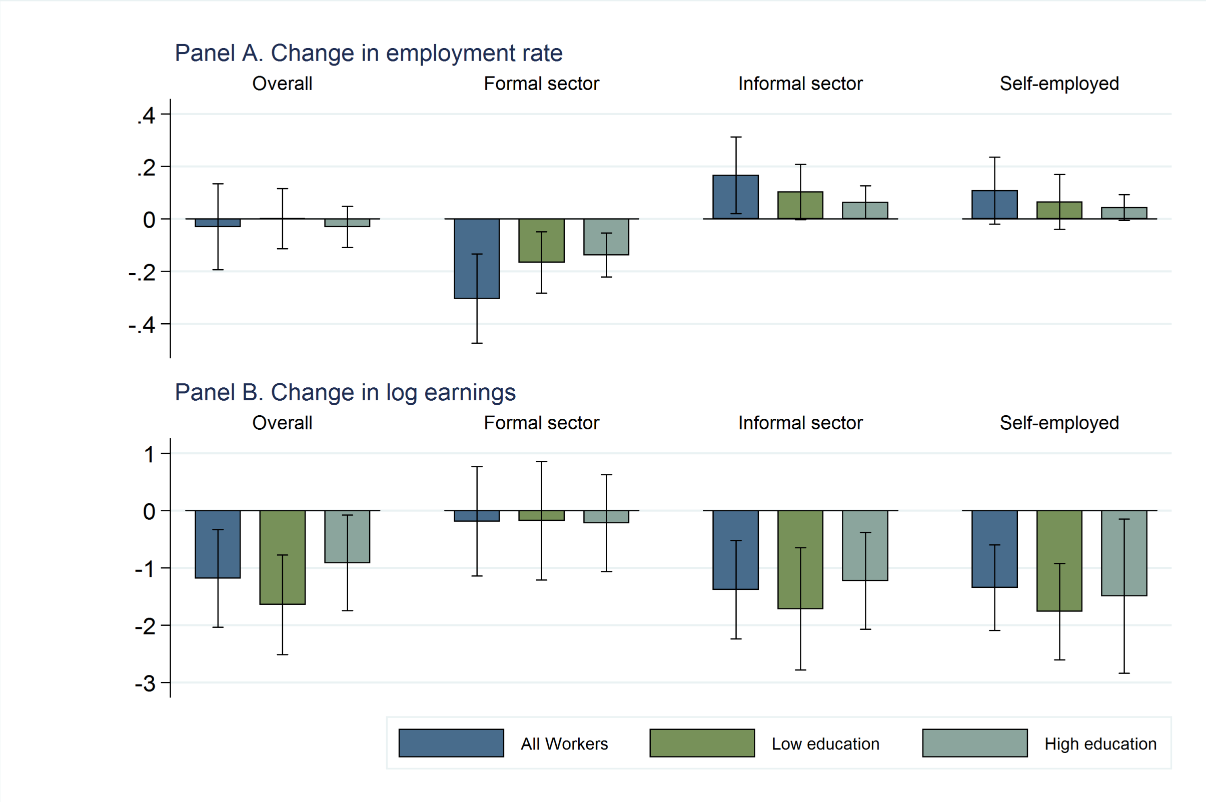
Finally, in Appendix B we address the issue raised by Adao et al. (2019) who argue that the typical shift-share design standard errors may be underestimated because the regression residuals are by design correlated across units with similar shares. By reestimating our regressions at the origin/shock level as suggested by Borusyak et al. (2021), we show that the shock-level standard errors are nearly the same as in our main tables, suggesting that these inference issues are not empirically relevant in our setup.

5 Differential Effects according to Education

The results reported so far represent estimates of the impact of an exogenous increase in the inflow of migrants from the Semiarid on natives, on average. In this section, we consider the case in which individuals with different levels of education may experience different effects. In particular, we expect that native workers whose skill level are similar to migrants, i.e. those with low education, to be close substitutes. Thus we reestimate the effect of migration on local labor market outcomes of natives with low and high education, separately. We define as low educated those with up to 7 years of schooling, which is equivalent to an incomplete elementary education. In our sample, 59% of natives are less educated.

Figure 9 illustrates the estimates by education. Panel A shows the effect of predicted migration on the changes in employment rates, by sector and education group. Low-educated native individuals are more likely to exit the formal sector and to become informal sector workers compared to those who have higher education. In Panel B

Figure 9: **Effects of migration on employment and earnings, by education level**

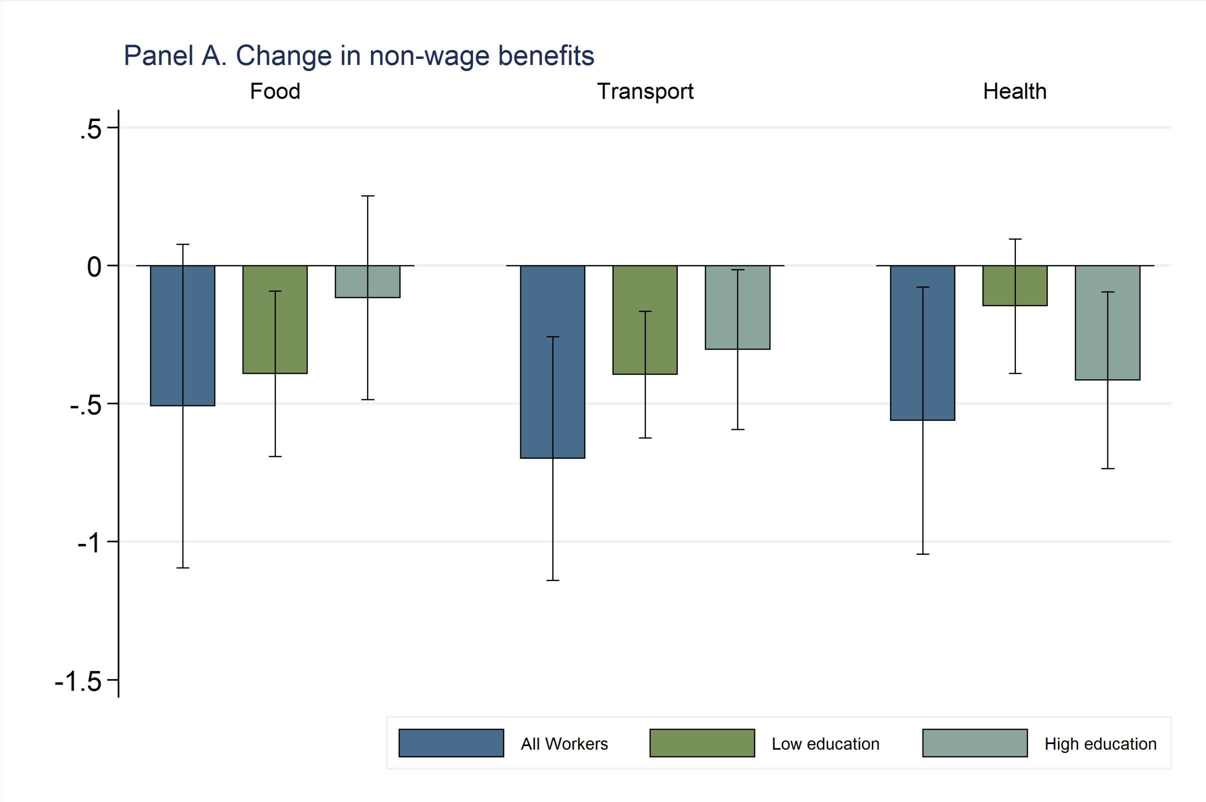


Notes: This figure shows the relationship between changes in labor market outcomes, by education level, and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. In Panel A, the dependent variables are the changes in employment rates while in Panel B we present estimates for changes in log earnings, for each sector. Each bar represents the reduced form coefficient by education level (low education = up to 7 years of schooling). All regressions are weighted by total native population in 1991, include time dummies and control for the log of native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. The capped lines show the 95% confidence intervals.

we analyze the differential effects on log earnings. In the formal sector, there is no significant impact on native workers across education levels. This is again consistent wage rigidity due to minimum wages or contractual wages preventing downward adjustments in the formal sector. On the other hand, native workers with low education have a relatively higher loss in informal and self-employment earnings, consistent with the conjecture that they compete more directly with (low-skilled) migrants.

In terms of adjustments on the nonwage benefits margin, it is less clear why they should differ by worker skills. In principle, working in the same firm implies that workers of different skills are offered a common benefits package. However, if there is a positive matching in the labor market with low (high) education workers selecting into less (more) productive and small (large) firms, then we should expect the less educated workers to be the most affected as the minimum wages bind more tightly in the firms where they work. In Figure 10 we show that the negative impact on

Figure 10: **Effects of predicted migration on nonwage benefits, by education level**



Notes: This figure shows the relationship between changes in nonwage benefits, by education level, and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. The dependent variables are the changes in the proportions of native workers in the formal sector who received some help to cover expenses with food, transport or health insurance. Each bar represents the reduced form coefficient by education level (low education = up to 7 years of schooling). All regressions are weighted by total native population in 1991, include time dummies and control for the log of native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. The capped lines show the 95% confidence intervals.

food and transportation benefits are indeed stronger and relatively more precise for low education workers. In contrast, high education workers have a clear reduction in employer-provided health insurance which is consistent again with some selection of these workers in large firms which tend to offer health insurance and where there is a mix of high and some low education workforce. A possible explanation is that the inflow of migrants competing with native low education workers in large firms pressures wages down. However under minimum wage restrictions, the adjustment occurs through lowering health insurance.

Changes in the benefits can have important welfare implications. We found that migration lowers the provision of food and transportation benefits to less educated individuals. On the other hand, we show that health insurance is not significantly changed for low education workers on average, while it is less offered for the high education workers. Considering that food and transportation are the two most offered benefits in the data (as shown in Table 2) and to the extent that workers value these benefits, their reduction together with a stronger negative impact on earnings for the low education workers suggest that the welfare of the less educated workers declines more than for high education workers.

6 Conclusion

In this paper we investigate the labor market impacts of weather-induced internal migration in Brazil. We use a shift-share instrument approach combining variation in the number of people leaving their hometowns, driven by weather shocks, with past settlement patterns to exploit exogenous variation in the number of migrants entering each destination municipality.

We find that internal migration reduces formal employment and raises the share of informal sector workers and self-employed. We also find a reduction in earnings for natives in the informal sector or self-employed, but no effect on formal workers. For them, adjustment occurs through reductions in nonwage benefits such as health insurance, food and transportation. As a result, there is no change in overall employment. This is consistent with workers reallocating to the informal sector or self-employment as well as adjustments in wage and nonwage benefit margins. Most effects are stronger for low educated individuals, which are more likely to be substituted by migrants escaping droughts. In particular, evidence that internal migration reduces informal sector earnings and nonwage job benefits more strongly for low education workers suggest that welfare inequality among natives rises.

References

Adao, R., M. Kolesár, and E. Morales (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics 134*(4), 1949–2010.

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

ANS (2019). Dados e indicadores do setor. Agência Nacional de Saúde Suplementar, http://www.ans.gov.br/perfil-do-setor/dados-e-indicadores-do-setor. Last accessed on 28/10/2020.

Arbache, J. S. and D. M. Ferreira (2001). Salários indiretos e previdência social no brasil. *Estudos Empresariais 6*(1), 37–44.

Assunção, J. and F. Chein (2016). Climate change and agricultural productivity in brazil: future perspectives. *Environment and Development Economics 21*(5), 581–602.

Babecký, J., C. Berson, L. Fadejeva, A. Lamo, P. Marotzke, F. Martins, and P. Strzelecki (2019). Non-base wage components as a source of wage adaptability to shocks: evidence from european firms, 2010–2013. *IZA Journal of Labor Policy 8*(1), 1.

Babecký, J., P. Du Caju, T. Kosma, M. Lawless, J. Messina, and T. Rõõm (2012). How do european firms adjust their labour costs when nominal wages are rigid? *Labour Economics 19*(5), 792–801. Special Section on: Price, Wage and Employment Adjustments in 2007-2008 and Some Inferences for the Current European Crisis.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics 118*(4), 1335–1374.

Borjas, G. J. and J. Monras (2017). The labour market consequences of refugee supply shocks. *Economic Policy 32*, 361—-413.

Borusyak, K., P. Hull, and X. Jaravel (2021). Quasi-experimental shift-share research designs. *The Review of Economic Studies*, Forthcoming.

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

Carabetta, J., R. H. M. Pereira, and C. N. Goncalves (2020). geobr: Loads shapefiles of official spatial data sets of brazil. python package version 0.1.5. *The Python Package Index (PyPI).*

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

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

Cardoso, A. and P. Portugal (2005). Contractual wages and the wage cushion under different bargaining settings. *Journal of Labor Economics 23*(4), 875–902.

Cattaneo, C., C. V. Fiorio, and G. Peri (2015). What happens to the careers of european workers when immigrants "take their jobs"? *Journal of Human Resources 50*(3), 655–693.

Cavalcanti, F. (2018, March). Voters sometimes provide the wrong incentives. The lesson of the Brazilian drought industry. MPRA Paper 88317, University Library of Munich, Germany.

Clemens, J. (2021, February). How do firms respond to minimum wage increases? understanding the relevance of non-employment margins. *Journal of Economic Perspectives 35*(1), 51–72.

Clemens, J., L. B. Kahn, and J. Meer (2018, May). The minimum wage, fringe benefits, and worker welfare. Working Paper 24635, National Bureau of Economic Research.

Corbi, R., R. Ferreira, R. Narita, and D. Souza (2021). Labor courts, job search and employment: Evidence from a labor reform in brazil.

Correia, S. (2017). Linear models with high-dimensional fixed effects: An efficient and feasible estimator. Working paper. http://scorreia.com/research/hdfe.pdf.

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

Dustmann, C., U. Schönberg, and J. Stuhler (2016). The impact of immigration: Why do studies reach such different results? *The Journal of Economic Perspectives 30*(4), 31–56.

Foged, M. and G. Peri (2016). Immigrants' effect on native workers: New analysis on longitudinal data. *American Economic Journal: Applied Economics 8*(2), 1–34.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020, August). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624.

Harris, I., T. J. Osborn, P. Jones, and D. Lister (2020). Version 4 of the cru ts monthly high-resolution gridded multivariate climate dataset. *Scientific data 7*(1), 1–18.

Harris, J. R. and M. P. Todaro (1970). Migration, unemployment and development: A two-sector analysis. *American Economic Review 60*(1), 126–142.

Hornbeck, R. (2012). The enduring impact of the american dust bowl: Short-and long-run adjustments to environmental catastrophe. *American Economic Review 102*(4), 1477–1507.

Jaeger, D. A., J. Ruist, and J. Stuhler (2018, February). Shift-Share Instruments and the Impact of Immigration. Working Paper 24285, National Bureau of Economic Research.

Kleemans, M. and J. Magruder (2018). Labour market responses to immigration: Evidence from internal migration driven by weather shocks. *The Economic Journal 128*(613), 2032–2065.

Lagakos, D. (2020, August). Urban-rural gaps in the developing world: Does internal migration offer opportunities? *Journal of Economic Perspectives 34*(3), 174–92.

Lee, D. S., J. McCrary, M. J. Moreira, and J. Porter (2020). Valid t-ratio inference for iv. Papers, arXiv.org.

Lundberg, S. (1985). The added worker effect. *Journal of Labor Economics 3*(1), 11–37.

McKenzie, R. (1980). The labor market effects of minimum wage laws: A new perspective. *Journal of Labor Research 1*, 255–264.

Meghir, C., R. Narita, and J.-M. Robin (2015, April). Wages and informality in developing countries. *American Economic Review 105*(4), 1509–46.

Molloy, R., C. L. Smith, and A. Wozniak (2011, September). Internal migration in the united states. *Journal of Economic Perspectives 25*(3), 173–96.

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

Narita, R. (2020). Self-employment in developing countries: A search-equilibrium approach. *Review of Economic Dynamics 35*, 1–34.

Ottaviano, G. I. and G. Peri (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association 10*(1), 152–197.

Ottaviano, G. I., G. Peri, and G. C. Wright (2013). Immigration, offshoring, and american jobs. *American Economic Review 103*(5), 1925–59.

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

Skoufias, E., K. Vinha, and H. Conroy (2013). The impacts of climate variability on household welfare in rural mexico. *Population and Environment 34*(3), 370–399.

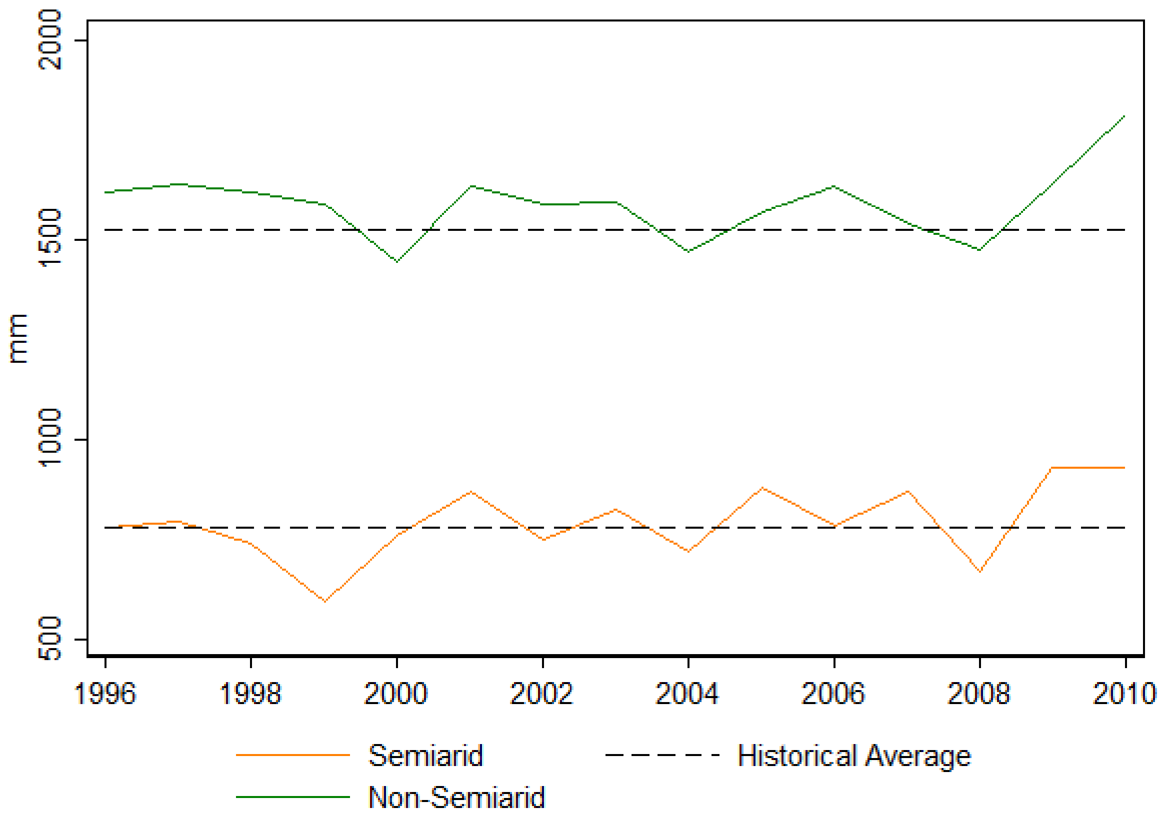
UN DESA (2017). International migration report 2017. United Nations Department of Economic and Social Affairs, Population Division.

UNDP (2009). Overcoming barriers: Human mobility and development. United Nations Development Program, Human Development Report 2009. http://hdr.undp.org/en/content/human-development-report-2009. Last accessed on 03/09/2021.

Xavier, A. C., C. W. King, and B. R. Scanlon (2016). Daily gridded meteorological variables in brazil (1980–2013). *International Journal of Climatology 36*(6), 2644–2659.

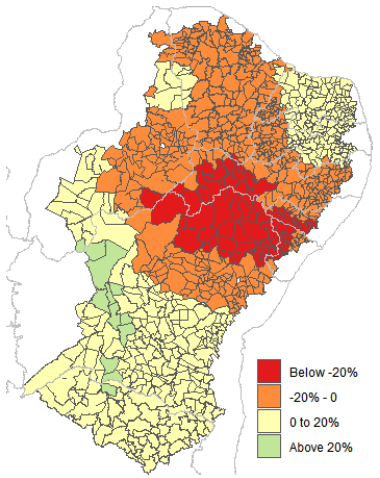
Figures and tables

Figure 11: **Precipitation level: Semiarid vs Non-Semiarid**

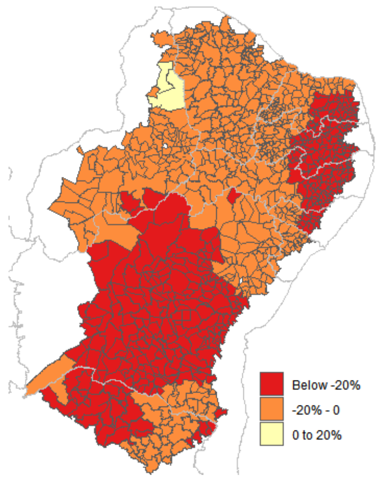


Notes: This figure compares the average precipitation level for the Semiarid region and the rest of the country, from 1996 to 2010. *Data source*: CRU Time Series v4 (Harris et al., 2020).

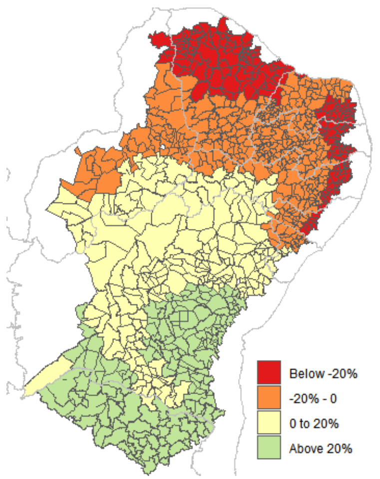
Figure 12: Precipitation levels in the Semiarid region for selected years



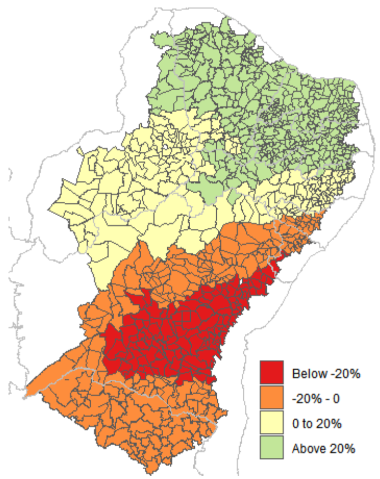
(a) 1997



(b) 2001



(c) 2005



(d) 2009

Notes: This figure presents the distribution of rainfall across the Semiarid region municipalities for selected years. Rainfall is measured as the log-deviations from historical averages. *Data source*: CRU Time Series v4 (Harris et al., 2020).

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

Panel A: Semiarid region	Mean	Std. Dev.	Min	Max	Obs
Annual rainfall	791.77	213.94	266.31	2,052.87	14,400
Historical rainfall	779.62	163.06	355.00	1,431.38	14,400
Annual temperature	25.55	1.39	21.42	28.93	14,400
Historical temperature	25.20	1.48	21.22	27.93	14,400
Migration outflow	214.16	323.66	0.00	5,773.00	14,400
Predicted migration outflow	211.66	284.16	12.66	4,551.83	14,400
Migration outflow rate (p.p.)	1.05	0.62	0.00	7.22	14,400
Predicted migration outflow rate (p.p.)	1.05	0.03	0.99	1.14	14,400
Population	21,260.91	30,092.29	1,265.00	480,949.00	14,400
Panel B: Non-Semiarid region	**Mean**	**Std. Dev.**	**Min**	**Max**	**Obs**
Migration inflow	146.69	896.95	0.00	25,423.00	8,190
Predicted migration inflow	135.47	929.32	0.00	21,543.50	8,190
Migration inflow rate (p.p.)	0.27	0.90	0.00	24.30	8,190
Predicted migration inflow rate (p.p.)	0.20	0.41	0.00	4.77	8,190
Native population	56,382.74	248,912.77	312.42	5,113,798.35	8,190

Notes: Rainfall is measured in mm. Temperature is measured in degrees Celsius. Population observed in the previous Census.

Table 2: **Summary statistics: Native individuals in destination municipalities**

	Mean	Std. Dev.	Min	Max	Obs
	Individual Characteristics				
	Mean	Std. Dev.	Min	Max	Obs
Female	51.08	3.64	0	72.72	8,190
Black	6.23	5.98	0	53.85	8,190
Age	37.45	1.96	30.15	55	8,190
Schooling	6.58	1.78	0	13.52	8,190
Low educated	65.32	15.74	4.71	100	8,190
	Employment				
	Mean	Std. Dev.	Min	Max	Obs
Overall	62.72	7.95	10	100	8,190
Formal sector	31.34	11.85	0	100	8,190
Informal sector	15.84	6.49	0	64.7	8,190
Self-employed	15.54	6.50	0	71.93	8,190
Unemployed	13.05	7.73	0	80	8,190
OLF	24.23	7.08	0	58.14	8,190
	Earnings				
	Mean	Std. Dev.	Min	Max	Obs
Overall	637.87	349.00	60.88	3,582.08	8,190
Formal sector	788.18	439.50	58.67	15,167.10	8,174
Informal sector	382.31	238.34	20	6,000.00	8,172
Self-employed	600.83	385.11	20	6,384.64	8,155
	Nonwage benefits				
Food	38.89	21.06	0	100	8,165
Transport	36.39	25.40	0	100	8,165
Health	20.86	16.41	0	100	8,165

Notes: Each observation is a municipality-year cell. Low-educated individuals are those with incomplete elementary schooling. Earnings are measured in R$ of 2012. Nonwage benefits are calculated only for native workers employed in the formal sector.

Table 3: **Migration outflows induced by weather shocks**

	(1)	(2)	(3)	(4)	(5)	(6)
Rainfall_{t-1}	-0.099***	-0.086***	-0.087***	-0.086***	-0.088***	-0.094***
	(0.028)	(0.030)	(0.030)	(0.030)	(0.030)	(0.031)
Rainfall_{t-2}			0.042	0.053		
			(0.032)	(0.033)		
Rainfall_{t-3}				0.056*		
				(0.031)		
Rainfall_{t}					-0.041	
					(0.032)	
Rainfall_{t+1}						-0.018
						(0.041)
Observations	14,400	14,400	14,400	14,400	14,400	14,400
Municipalities	960	960	960	960	960	960
R-Squared	0.461	0.472	0.472	0.473	0.473	0.472
F Stat	8.372	5.987	3.134	2.291	4.362	3.049
Time dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Temperature shocks	✓	✓	✓	✓	✓	✓
Covariates trends		✓	✓	✓	✓	✓

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 1991 Census. Rainfall is measured as log-deviation from historical average. All specifications include controls for temperature shocks, municipality and year fixed effects. Columns (2)-(6) also control for the log of population in the previous census and include interactions between time dummies and 1991 municipality-level characteristics (age and the share of high school and college educated individuals). Standard errors are clustered at the grid level. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table 4: **Effects of migration on earnings**

	(1)	(2)	(3)	(4)
	A. Change in log earnings			
Predicted inflow	-1.655***	-1.323***	-1.182***	-1.265***
	(0.576)	(0.438)	(0.436)	(0.411)
Observations	8,190	8,190	8,190	8,190
Municipalities	684	684	684	684
	B. Change in log earnings, formal sector			
Predicted inflow	-0.848	-0.300	-0.188	-0.462
	(0.674)	(0.488)	(0.489)	(0.444)
Observations	8,162	8,162	8,162	8,162
Municipalities	684	684	684	684
	C. Change in log earnings, informal sector			
Predicted inflow	-1.929***	-1.517***	-1.379***	-1.490***
	(0.607)	(0.439)	(0.441)	(0.390)
Observations	8,156	8,156	8,156	8,156
Municipalities	684	684	684	684
	D. Change in log earnings, self-employed			
Predicted inflow	-1.557***	-1.516***	-1.346***	-1.343***
	(0.450)	(0.391)	(0.383)	(0.405)
Observations	8,133	8,133	8,133	8,133
Municipalities	683	683	683	683
Demographics		✓	✓	✓
Time dummies			✓	✓
Covariates trends				✓

Notes: This table shows the relationship between changes in log earnings and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Each observation is a municipality-year cell. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Column (3) also include time dummies. In column (4) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). All regressions are weighted by the 1991 native population. Standard errors clustered at the municipality level in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table 5: **Effects of migration on employment**

	(1)	(2)	(3)	(4)
	A. Change in employment rate			
Predicted inflow	-0.014	-0.033	-0.030	-0.017
	(0.091)	(0.084)	(0.084)	(0.087)
Observations	8,190	8,190	8,190	8,190
Municipalities	684	684	684	684
	B. Change in formal employment rate			
Predicted inflow	-0.316***	-0.323***	-0.305***	-0.289***
	(0.104)	(0.088)	(0.087)	(0.096)
Observations	8,190	8,190	8,190	8,190
Municipalities	684	684	684	684
	C. Change in informal employment rate			
Predicted inflow	0.160**	0.181**	0.166**	0.153**
	(0.079)	(0.075)	(0.075)	(0.076)
Observations	8,190	8,190	8,190	8,190
Municipalities	684	684	684	684
	D. Change in self-employment rate			
Predicted inflow	0.143**	0.109*	0.108*	0.119
	(0.064)	(0.066)	(0.065)	(0.077)
Observations	8,190	8,190	8,190	8,190
Municipalities	684	684	684	684
Demographics		✓	✓	✓
Time dummies			✓	✓
Covariates trends				✓

Notes: This table shows the relationship between changes in employment rate and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Each observation is a municipality-year cell. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Column (3) also include time dummies. In column (4) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). All regressions are weighted by the 1991 native population. Standard errors clustered at the municipality level in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table 6: **Effects of migration on nonwage benefits**

	(1)	(2)	(3)	(4)
	A. Food			
Predicted inflow	-0.342	-0.472	-0.509*	-0.701**
	(0.313)	(0.299)	(0.300)	(0.301)
Observations	8,147	8,147	8,147	8,147
Municipalities	684	684	684	684
	B. Transportation			
Predicted inflow	-0.540*	-0.655***	-0.698***	-0.542*
	(0.306)	(0.228)	(0.226)	(0.278)
Observations	8,147	8,147	8,147	8,147
Municipalities	684	684	684	684
	C. Health			
Predicted inflow	-0.443**	-0.546**	-0.562**	-0.226
	(0.212)	(0.239)	(0.248)	(0.218)
Observations	8,147	8,147	8,147	8,147
Municipalities	684	684	684	684
Demographics		✓	✓	✓
Time dummies			✓	✓
Covariates trends				✓

Notes: This table shows the relationship between changes in the proportions of working-age native population, employed in the formal sector, who receive nonwage benefits and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Column (3) also include time dummies. In column (4) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). All regressions are weighted by the 1991 native population. Standard errors clustered at the municipality level in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table 7: **Effects of migration on unemployment and participation**

	(1)	(2)	(3)	(4)
	A. Change in unemployment rate			
Predicted inflow	0.177**	0.161**	0.153**	0.156**
	(0.084)	(0.063)	(0.065)	(0.064)
Observations	8,190	8,190	8,190	8,190
Municipalities	684	684	684	684
	B. Change in inactivity rate			
Predicted inflow	-0.163**	-0.128**	-0.123*	-0.139**
	(0.079)	(0.064)	(0.064)	(0.067)
Observations	8,190	8,190	8,190	8,190
Municipalities	684	684	684	684
Demographics		✓	✓	✓
Time dummies			✓	✓
Covariates trends				✓

Notes: This table shows the relationship between changes in unemployment and inactivity rates and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Each observation is a municipality-year cell. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Column (3) also include time dummies. In column (4) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). All regressions are weighted by the 1991 native population. Standard errors clustered at the municipality level in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table 8: **Effects of migration on labor market outcomes, by status in the household**

	(1) Employment	(2) Formal	(3) Informal	(4) Self-Employment	(5) Unemployment	(6) Inactivity
	A. Head					
Predicted inflow	-0.067	-0.220***	0.088**	0.065	0.009	0.021
	(0.054)	(0.067)	(0.043)	(0.041)	(0.032)	(0.035)
Observations	8,190	8,190	8,190	8,190	8,190	8,190
Municipalities	684	684	684	684	684	684
	B. Non-head					
Predicted inflow	0.036	-0.085*	0.079**	0.043	0.144***	-0.144**
	(0.064)	(0.050)	(0.040)	(0.033)	(0.043)	(0.059)
Observations	8,190	8,190	8,190	8,190	8,190	8,190
Municipalities	684	684	684	684	684	684
Demographics	✓	✓	✓	✓	✓	✓
Time dummies	✓	✓	✓	✓	✓	✓

Notes: This table shows the relationship between changes in employment (by sector), unemployment and inactivity rates and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Each observation is a municipality-year cell. In Panel A we use only individuals identified as the head of the household while in Panel B only those identified as non-head are used. All regressions control for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; a cubic polynomial of age, include time dummies and are weighted by the 1991 native population. Standard errors clustered at the municipality level in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table 9: **Impacts of internal migration on labor markets: IV estimates**

	Overall		Formal		Informal		Self-employment	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. Change in employment rates							
Migrant inflow	-0.114	-0.033	-0.218***	-0.330***	0.087	0.180**	0.017	0.117*
	(0.078)	(0.095)	(0.075)	(0.093)	(0.063)	(0.086)	(0.056)	(0.071)
First-stage coefficient		.92		.92		.92		.92
First-stage F-statistic		220.6		220.6		220.6		220.6
Observations	8,190	8,190	8,190	8,190	8,190	8,190	8,190	8,190
Municipalities	684	684	684	684	684	684	684	684
	B. Change in log earnings							
Migrant inflow	-0.695**	-1.281***	-0.265	-0.203	-1.147***	-1.494***	-0.693	-1.459***
	(0.295)	(0.465)	(0.309)	(0.527)	(0.392)	(0.485)	(0.465)	(0.441)
First-stage coefficient		.92		.92		.92		.92
First-stage F-statistic		220.6		220.2		220.8		220.4
Observations	8,190	8,190	8,162	8,162	8,156	8,156	8,133	8,133
Municipalities	684	684	684	684	684	684	683	683
Time dummies	✓	✓	✓	✓	✓	✓	✓	✓
Demographics	✓	✓	✓	✓	✓	✓	✓	✓

Notes: This table presents OLS and IV estimates of the relationship between the number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991, and labor market outcomes. Columns (1), (3), (5) and (7) show OLS estimates and columns (2),(4),(6) and (8) present the IV coefficients. In Panel A the dependent variables are the changes in employment rates while in Panel B are the log earnings, for each sector. All specifications are weighted by the total native population in 1991, include time dummies and control for the log of total native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Bootstrapped standard errors in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table 10: **Effects of migration on unemployment and participation: IV estimates**

	Unemployment		Inactivity	
	OLS	IV/2SLS	OLS	IV/2SLS
	(1)	(2)	(3)	(4)
Migrant inflow	0.173**	0.166**	-0.059	-0.133*
	(0.072)	(0.072)	(0.067)	(0.075)
First-stage coefficient		.92		.92
First-stage F-statistic		220.6		220.6
Observations	8,190	8,190	8,190	8,190
Municipalities	684	684	684	684
Time dummies	✓	✓	✓	✓
Demographics	✓	✓	✓	✓

Notes: This table presents OLS and IV estimates of the relationship between changes in unemployment and inactivity rates and the number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Columns (1) and (3) show OLS estimates and columns (2) and (4) present the IV coefficients. All specifications are weighted by the total native population in 1991, include time dummies and control for the log of total native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Bootstrapped standard errors in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table 11: **Correlation between predicted migration from the Semiarid and other regions**

	(1)	(2)	(3)
	Migrants from other regions		
Predicted inflow	0.217	0.006	0.198
	(2.854)	(2.823)	(2.846)
Observations	8190	8190	8190
Municipalities	684	684	684
Time dummies	✓	✓	✓
Municipality dummies	✓	✓	✓
Demographics		✓	✓
Covariates Trends			✓

Notes: This table shows the relationship between the predicted number of migrants from the Semiarid region in each destination municipality and the inflow of migrants coming from other regions, both measured as a fraction of the native working-age population in 1991. All specifications are weighted by total native population in 1991 and include municipality and year fixed effects. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. In column (3) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). Standard errors clustered at the municipality level in parenthesis.*** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table 12: **Migration and changes in log earnings, by industry**

	(1)	(2)	(3)	(4)
	A. Agriculture			
Predicted inflow	-2.476	-0.709	-0.693	-0.391
	(2.222)	(1.863)	(1.821)	(1.866)
Observations	6,251	6,251	6,251	6,251
Municipalities	661	661	661	661
	B. Manufacturing			
Predicted inflow	-1.003*	-0.866	-0.667	-0.467
	(0.584)	(0.578)	(0.578)	(0.579)
Observations	7,902	7,902	7,902	7,902
Municipalities	683	683	683	683
	C. Services			
Predicted inflow	-1.319**	-0.971**	-0.830**	-1.042***
	(0.570)	(0.386)	(0.386)	(0.357)
Observations	8,175	8,175	8,175	8,175
Municipalities	684	684	684	684
Demographics		✓	✓	✓
Time dummies			✓	✓
Covariates trends				✓

*Notes:*This table shows the relationship between changes in log earnings by industry and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Each observation is a municipality-year cell. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Column (3) also include time dummies. In column (4) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). All regressions are weighted by the 1991 native population. Standard errors clustered at the municipality level in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Appendix A Shift-share instrument

In this section we present the details about how we constructed our instrument using two sources of exogenous variation.

A.1 Weather data

Our main data source for weather shocks comes from the CRUTS v4, a gridded dataset produced by the Climatic Research Unit at the University of East Anglia (Harris et al., 2020). This dataset provides information on monthly precipitation and temperature covering the whole globe[25], from 1901 to 2018. The grid resolutions is 0.25° × 0.25° (around 56km^2) and is created by interpolation from ground-based weather stations around the world.

We use the R package 'geobr' (Carabetta et al., 2020) to download the shapefile of Brazilian municipalities and georreference the coordinates from each municipality's centroid and keep only municipalities that belong to the Semiarid region. Then, for each municipality, we find the grid's four points which are closest to it's centroid and calculate the average level of precipitation and temperature from this points, weighted by the inverse distance to the centroid.

This procedure results in a dataset of monthly averages of precipitation and temperature for each municipality in the Semiarid, from 1901 to 2018, which we aggregate in yearly measures. Precipitation is defined as the sum of monthly levels and temperature as the average. For each municipality we calculate the historical mean from both variables and take log of the levels and long term averages.

Finally, our weather shock variables are defined as

$$Rainfall_{ot} = \ln\left(\sum_{\tau\in\{GS\}} r_{o\tau t}\right) - \ln(\bar{r}_o) \tag{A1}$$

where $r_{o\tau t}$ is the rainfall in municipality of origin o in month τ of year t, and $\bar{r}_o$ is the municipality's historical average precipitation for the same months. The index τ covers the 6-month growing season (GS). Temperature is calculated in a similar way, but using the average instead of summation to create yearly data. Our main specification focus on the growing season, but results are very similar when we use the full year.[26]

[25]Except for the Antartica.

[26]The growing season in the Semiarid region occurs from November to April.

Table A1: **Migration outflows induced by weather shocks (12 months)**

	(1)	(2)	(3)	(4)	(5)	(6)
Rainfall$_{t-1}$	-0.126***	-0.120***	-0.118***	-0.128***	-0.121***	-0.124***
	(0.033)	(0.035)	(0.036)	(0.036)	(0.035)	(0.035)
Rainfall$_{t}$			-0.032	-0.055	-0.029	
			(0.041)	(0.042)	(0.040)	
Rainfall$_{t-2}$			0.053	0.069*		
			(0.040)	(0.041)		
Rainfall$_{t-3}$				0.006		
				(0.035)		
Rainfall$_{t+1}$						-0.040
						(0.042)
Observations	14,400	14,400	14,400	14,400	14,400	14,400
Municipalities	960	960	960	960	960	960
R-Squared	0.461	0.469	0.469	0.470	0.469	0.469
F Stat	8.033	7.232	3.626	3.491	4.398	4.687
Time dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Temperature shocks	✓	✓	✓	✓	✓	✓
Covariates trends		✓	✓	✓	✓	✓

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 1991 Census. Rainfall is measured as log-deviation from historical average. All specifications include controls for temperature shocks, municipality and year fixed effects. Columns (2)-(6) also control for the log of population in the previous census and include interactions between time dummies and 1991 municipality-level characteristics (age and the share of high school and college educated individuals). Standard errors are clustered at the grid level. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

A.2 Alternative measures of weather

One possible concern about our measure of weather is that we focus on rainfall levels, controlling for temperature variation, to predict the flow of migrants leaving the Semiarid region. This may be problematic because we cannot account for the presence of groundwater or any other factors that influence water balance. To circumvent this issue we gather new data from Xavier et al. (2016), who provides a gridded dataset with daily averages of precipitation and potential evaporation, from 1980 to 2013, based on ground data from weather stations interpolated to create high-resolution grids (0.25° $\times$ 0.25°) across the Brazilian territory. They calculate potential evaporation using maximum and minimum temperatures, solar radiation, relative humidity and wind speed. We aggregate the daily precipitation and evaporation data into monthly measures and follow Cavalcanti (2018) to construct a measure of drought severity, the aridity index, as follows:

$$AI_{mt} = \frac{\sum_{\tau \in \{GS\}} PE_{m\tau t}}{\sum_{\tau \in \{GS\}} Pr_{m\tau t}} \tag{A2}$$

where $PE_{m\tau t}$ is the potential evaporation in the municipality m, at the month τ of the growing season(GS) in year t. Then we standardize this measure to simplify interpretation and calculate de aridity index z-score as

$$Z^{AI}_{mt} = \frac{\left(AI_{mt} - \overline{AI}\right)}{AI_{sd}} \tag{A3}$$

We show in Table A2 that this alternative measure is also strongly correlated with the migration outflow rate. Including up to three lags and one lead does not affect the main coefficient, neither does the inclusion of controls. In Panel B we regress outflow rate on a categorical variable indicating the quartile of the Aridity Index z-score. Our estimates show that extreme events of drought increase migration even further.

More importantly, in Figure A1 we compare the predicted number of migrants using this alternative measure to the prediction we had before and show they are virtually identical, which means that our instrument is robust to measurement errors in water availability.

A.3 Predicting migration outflows

We use the weather shocks to predict the number of migrants leaving the Semiarid municipalities and settling in other places across the country. We estimate the following two-way fixed effects regression:

$$m_{ot} = \alpha + \beta' Z_{ot-1} + \phi_o + \delta_t + \varepsilon_{ot} \tag{A4}$$

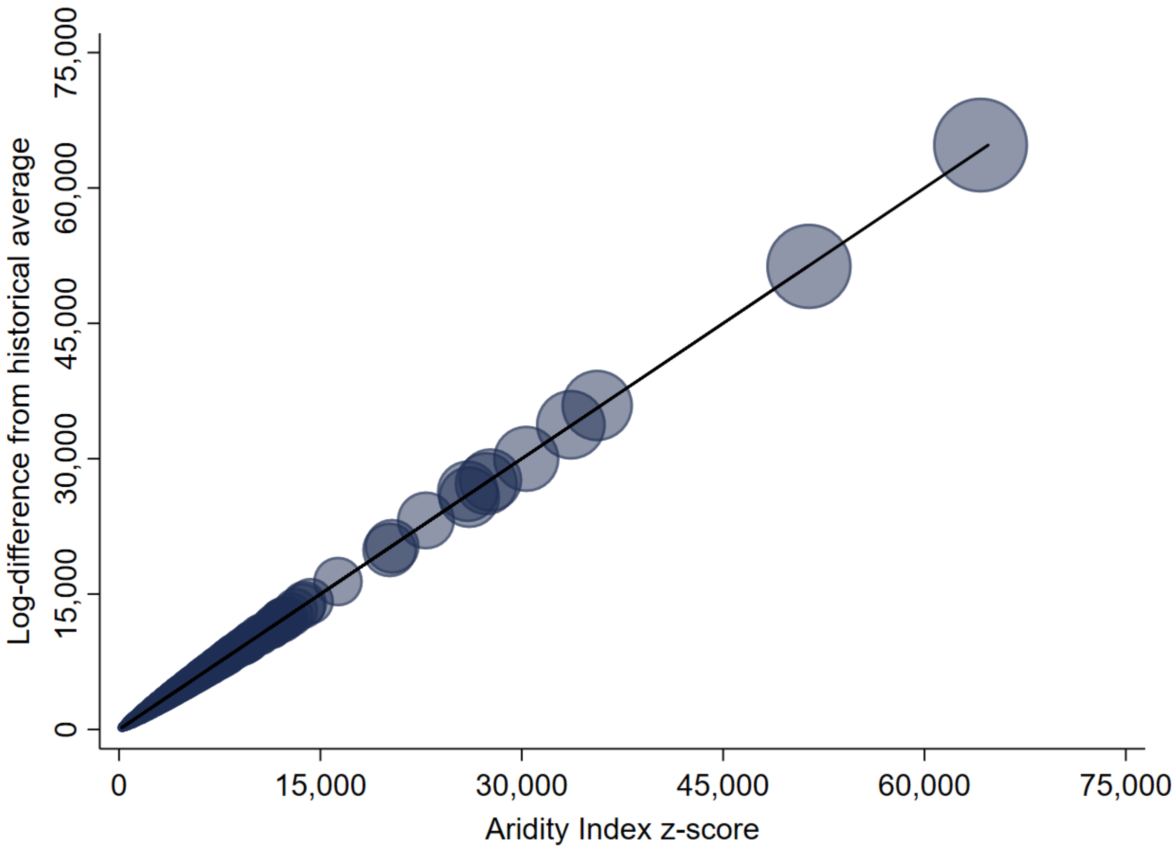
where m_{ot} is the migration rate in origin municipality o and year t, Z is a vector of rainfall and temperature shocks in the previous year, ϕ_o and δ_t are municipality and year fixed effects, respectively, and ε_{ot} is a random error term. Then we project the migration

Table A2: **Migration outflows induced by weather shocks: Aridity Index**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Panel A: *Continuous Z-score*				
Aridity Index$_t$	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	
Aridity Index$_{t-1}$		0.003***	0.003***	0.003***			
		(0.001)	(0.001)	(0.001)			
Aridity Index$_{t-2}$			0.001	0.001			
			(0.001)	(0.001)			
Aridity Index$_{t-3}$				0.000			
				(0.001)			
Aridity Index$_{t+1}$					0.001		
					(0.001)		
Constant	1.020***	0.995***	0.987***	0.985***	1.015***	-2.022**	
	(0.007)	(0.012)	(0.015)	(0.018)	(0.011)	(0.851)	
			Panel B: *Drought severity*				
Second quartile							0.028**
							(0.014)
Third quartile							0.010
							(0.016)
Fourth quartile							0.076***
							(0.019)
Constant							1.024***
							(0.011)
Observations	14,400	14,400	14,400	14,400	14,400	14,400	14,400
Municipalities	960	960	960	960	960	960	960
R-Squared	0.461	0.462	0.462	0.462	0.461	0.470	0.462
Time dummies	✓	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓	✓
Demographics						✓	

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 1991 Census. Aridity Index is measured as the municipality z-score of the ratio between evaporation and precipitation accumulated from November to April. All specifications include municipality and year fixed effects. Column (6) also control for the log of population in the previous census and include interactions between time dummies and 1991 municipality-level characteristics (age and the share of high school and college educated individuals). Drought severity measures are the quartiles of the Aridity Index z-score. Standard errors are clustered at the municipality level. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Figure A1: **Predicted migration outflows using alternative measures of weather**



Notes: Y-axis shows the number of migrants predicted by rainfall and temperature log-differences from historical average. X-axis presents the number of migrants predicted by the aridity index z-score. Circle size represents the municipality's total population in 1991. *Data sources*: CRU TS v4 (Harris et al., 2020) and Xavier et al. (2016).

rate and multiply by the 1991 population to obtain the predicted number of migrants who left the Semiarid region because of weather shocks.

$$\widehat{M}_{ot} = \widehat{m}_{ot} \times P_o \tag{A5}$$

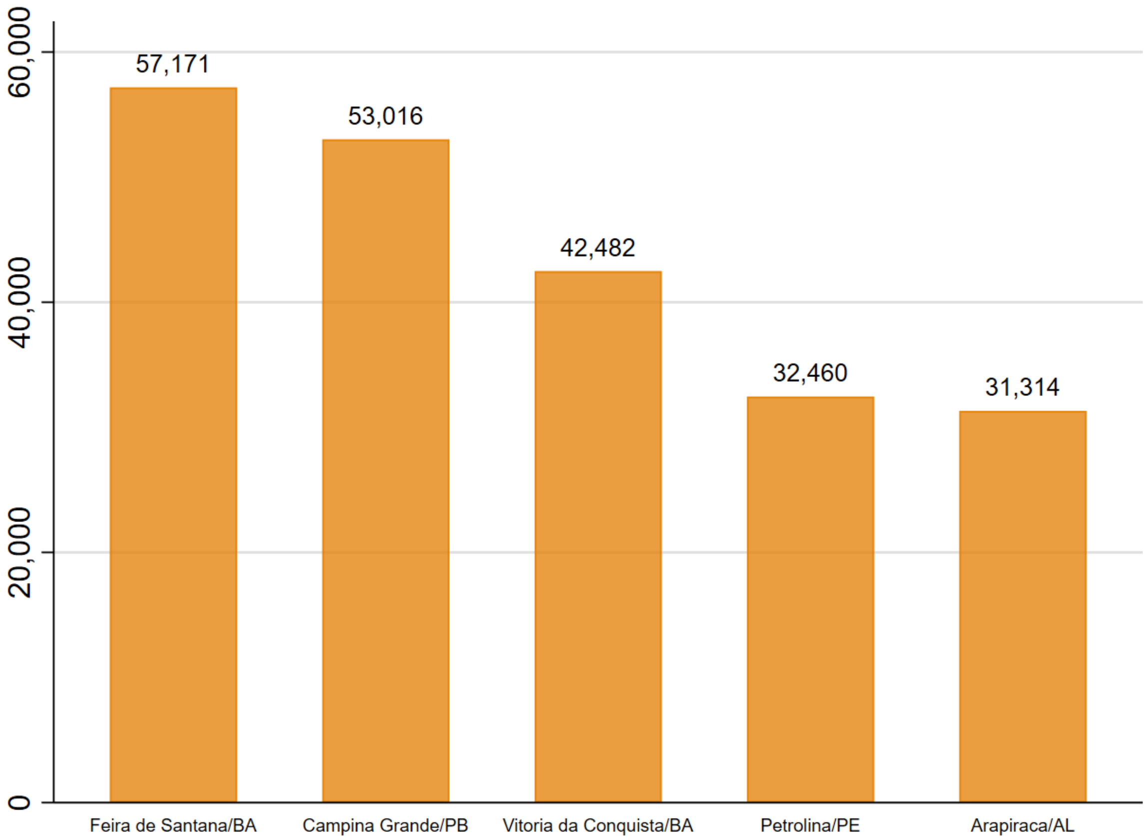
As shown in Figure 3a this procedure results in a good approximation of the observed migration.

From 1996 to 2010 over 1.5 million people left the Semiarid region to settle elsewhere. Figure A2 indicates the main origin municipalities, responsible for around 7% of the out-migrants in our sample.

A.4 Past settlement patterns (1991)

We use the 1991 Census' microdata to construct the predetermined network of migrants from the Semiarid region settling in the destination municipalities. Individuals were asked how many years they lived in the current municipality and what was the

Figure A2: **Top 5 sending municipalities in the Semiarid region**



Notes: This figure shows the top five sending municipalities. *Data source*: Census microdata.

municipality where they lived before[27]. With this information we construct a yearly panel of migration for every pair of origin-destination municipalities in Brazil.

For each destination d we define the share of Semiarid's migrants from origin o as

$$s_{od} = \frac{m_{od}}{\sum\limits_{o} m_{od}} \tag{A6}$$

By construction, we have $\sum\limits_{o} s_{od} = 1$, meaning that every individual predicted to migrate will be allocated in some destination. The final step is to use these shares to allocate the predicted number of migrants from the Semiarid region into the destination municipalities

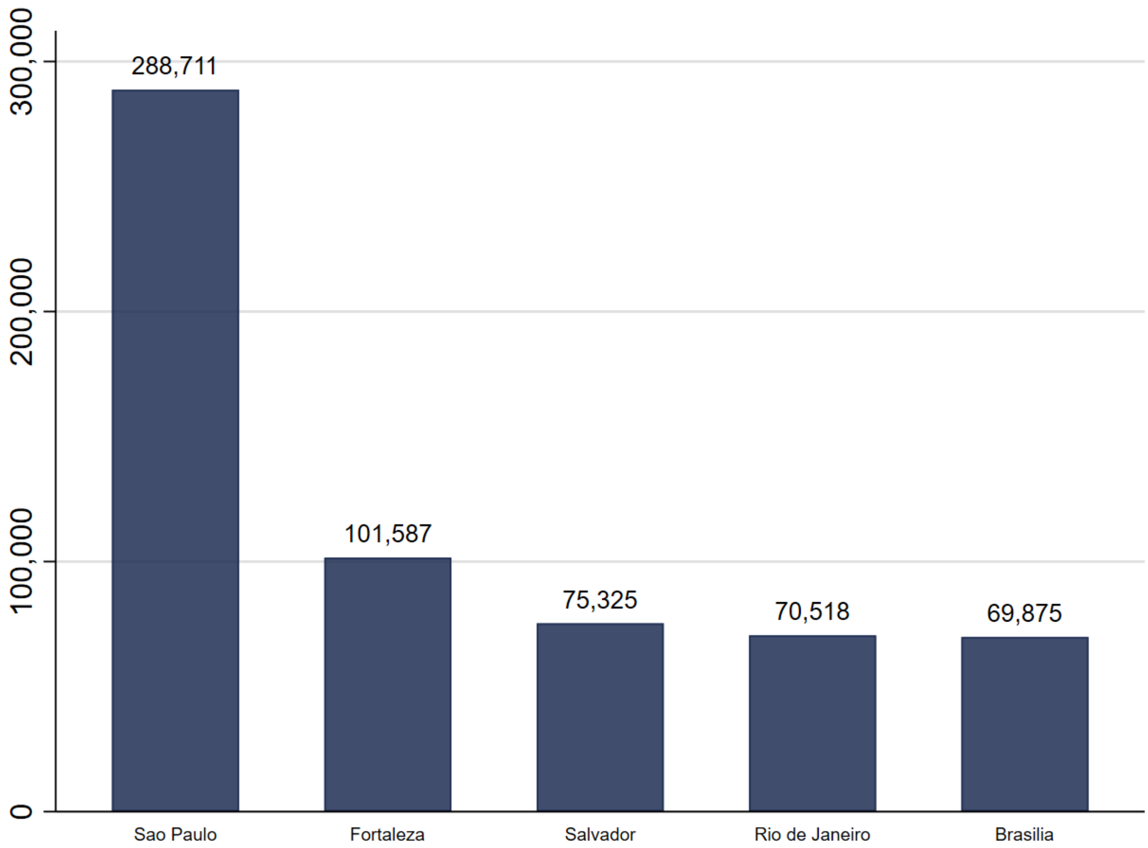
$$\widetilde{m_{dt}} = \sum_{o=1}^{O} \frac{s_{od} \times \widehat{M_{ot}}}{P_d} \tag{A7}$$

where $\widehat{M_{ot}}$ is obtained from equation A5 and P_d is the 1991 population in the destination municipality, which we use as normalization.

[27]Microdata only records the previous municipality for those individuals who moved during the decade covered by the 1991 Census. Therefore, we can only track individuals as far as 1982.

Figure **??** already shown that we obtained a good prediction of migration inflows. We emphasize that destination areas are very concentrated. Figure A3 depicts the top five settling municipalities for Semiarid migrants during our period, which concentrate around 40% of the total migration in our sample. All of them are state capitals, but only two (Fortaleza and Salvador) are in the Northeast region.

Figure A3: **Top 5 receiving municipalities**



Notes: This figure shows the top five receiving municipalities. *Data source*: Census microdata.

Appendix B Inference correction

In this appendix we address the issue raised by Adao et al. (2019) who demonstrated that in the typical shift-share design standard errors may be underestimated because the regression residuals are correlated across units with similar shares.

B.1 Shock-level IV regressions

In this section we apply the method developed by Borusyak et al. (2021) to correct the standard errors in our estimates and show that our main results remain valid. Applying their procedure to our setting, we estimate the following IV regression

$$\bar{y}_{ot} = \alpha + \beta \bar{m}_{ot} + \bar{\varepsilon}_{ot} \tag{B1}$$

where $\bar{v}_{ot} = \frac{\sum_d s_{od} v_{dt}}{\sum_d s_{od}}$ is a share-weighted average of the outcomes y or the instrument m at the destination level and the regression is weighted by the average exposure to the shocks $s_o = \sum_d s_{od}$. Put in a more simplistic way, in the traditional shift-share design we take a shock at the origin level and calculate a share-weighted average to use as instrumental variable at the destination. What Borusyak et al. (2021) propose is to take the destination-level outcomes and endogenous variables, calculate its weighted averages for each origin municipality and use the shocks as instruments to estimate the parameter β. They demonstrate that even if the shares are not exogenous their approach renders a consistent estimator. More importantly for this section, they also show that this procedure allows to obtain standard errors which are asymptotically valid in the framework of Adao et al. (2019).

We calculate share-weighted averages of our outcomes and instrument and estimate all the same specifications at the origin-level and present the results in Tables B1-B8.

Table B1: **Inference a la Borusyak et al. (2021): Reduced form estimates on employment**

	(1)	(2)	(3)	(4)
	A. Change in employment rate			
Predicted inflow	-0.014	-0.033	-0.030	-0.017
	(0.062)	(0.064)	(0.046)	(0.058)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	B. Change in formal employment rate			
Predicted inflow	-0.316***	-0.323***	-0.305***	-0.289***
	(0.064)	(0.060)	(0.041)	(0.052)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	C. Change in informal employment rate			
Predicted inflow	0.160**	0.181**	0.166**	0.153**
	(0.079)	(0.075)	(0.075)	(0.076)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	D. Change in self-employment rate			
Predicted inflow	0.143***	0.109***	0.108***	0.119***
	(0.037)	(0.037)	(0.035)	(0.044)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
Demographics		✓	✓	✓
Time dummies			✓	✓
Covariates trends				✓

Notes: This table shows the relationship between changes in employment rate and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Each observation is a municipality-year cell. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Column (3) also include time dummies. In column (4) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). All regressions are weighted by the 1991 native population. Robust standard errors calculated by the procedure in Borusyak et al. (2021) in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table B2: **Inference a la Borusyak et al. (2021): Reduced form estimates on earnings**

	(1)	(2)	(3)	(4)
	A. Change in log earnings			
Predicted inflow	-1.655***	-1.323***	-1.182***	-1.265***
	(0.242)	(0.224)	(0.150)	(0.190)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	B. Change in log earnings, formal sector			
Predicted inflow	-0.848***	-0.300*	-0.188	-0.462***
	(0.193)	(0.178)	(0.125)	(0.157)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	C. Change in log earnings, informal sector			
Predicted inflow	-1.929***	-1.516***	-1.378***	-1.489***
	(0.241)	(0.243)	(0.188)	(0.237)
Observations	8,156	8,156	8,156	8,156
Municipalities	684	684	684	684
	D. Change in log earnings, self-employed			
Predicted inflow	-1.557***	-1.516***	-1.345***	-1.342***
	(0.416)	(0.425)	(0.326)	(0.412)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
Demographics		✓	✓	✓
Time dummies			✓	✓
Covariates trends				✓

Notes: This table shows the relationship between changes in log earnings and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Each observation is a municipality-year cell. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Column (3) also include time dummies. In column (4) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). All regressions are weighted by the 1991 native population. Robust standard errors calculated by the procedure in Borusyak et al. (2021) in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table B3: **Inference a la Borusyak et al. (2021): Reduced form estimates on unemployment and participation**

	(1)	(2)	(3)	(4)
	A. Change in unemployment rate			
Predicted inflow	0.177***	0.161***	0.153***	0.156***
	(0.055)	(0.057)	(0.040)	(0.051)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	B. Change in inactivity rate			
Predicted inflow	-0.163***	-0.128**	-0.123***	-0.139**
	(0.053)	(0.053)	(0.046)	(0.058)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
Demographics		✓	✓	✓
Time dummies			✓	✓
Covariates trends				✓

Notes: This table shows the relationship between changes in unemployment and inactivity rates and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Each observation is a municipality-year cell. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Column (3) also include time dummies. In column (4) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). All regressions are weighted by the 1991 native population. Robust standard errors calculated by the procedure in Borusyak et al. (2021) in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table B4: **Inference a la Borusyak et al. (2021): Reduced form estimates on nonwage benefits**

	(1)	(2)	(3)	(4)
	A. Food			
Predicted inflow	-0.342**	-0.472***	-0.508***	-0.700***
	(0.141)	(0.144)	(0.107)	(0.134)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	B. Transportation			
Predicted inflow	-0.540***	-0.655***	-0.698***	-0.541***
	(0.137)	(0.135)	(0.099)	(0.124)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	C. Health			
Predicted inflow	-0.443**	-0.546***	-0.561***	-0.226
	(0.176)	(0.184)	(0.125)	(0.157)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
Demographics		✓	✓	✓
Time dummies			✓	✓
Covariates trends				✓

Notes: This table shows the relationship between changes in the proportions of working-age native population, employed in the formal sector, who receive nonwage benefits and the predicted number of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Column (3) also include time dummies. In column (4) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). All regressions are weighted by the 1991 native population. Robust standard errors calculated by the procedure in Borusyak et al. (2021) in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table B5: **Inference a la Borusyak et al. (2021): IV/2SLS estimates on employment**

	(1)	(2)	(3)	(4)
	A. Change in employment rate			
Migrant inflow	-0.015	-0.035	-0.033	-0.018
	(0.068)	(0.069)	(0.050)	(0.063)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	B. Change in formal employment rate			
Migrant inflow	-0.347***	-0.350***	-0.330***	-0.310***
	(0.071)	(0.065)	(0.045)	(0.056)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	C. Change in informal employment rate			
Migrant inflow	0.175***	0.196***	0.180***	0.164***
	(0.048)	(0.049)	(0.035)	(0.044)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	D. Change in self-employment rate			
Migrant inflow	0.156***	0.119***	0.117***	0.128***
	(0.041)	(0.041)	(0.038)	(0.047)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
Demographics		✓	✓	✓
Time dummies			✓	✓
Covariates trends				✓

Notes: This table shows the relationship between changes in employment rate and the inflow of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Each observation is a municipality-year cell. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Column (3) also include time dummies. In column (4) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). All regressions are weighted by the 1991 native population. Robust standard errors calculated by the procedure in Borusyak et al. (2021) in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table B6: **Inference a la Borusyak et al. (2021): IV/2SLS estimates on earnings**

	(1)	(2)	(3)	(4)
	A. Change in log earnings			
Migrant inflow	-1.815***	-1.433***	-1.281***	-1.359***
	(0.268)	(0.244)	(0.164)	(0.205)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	B. Change in log earnings, formal sector			
Migrant inflow	-0.930***	-0.325*	-0.203	-0.496***
	(0.210)	(0.192)	(0.135)	(0.168)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	C. Change in log earnings, informal sector			
Migrant inflow	-2.116***	-1.642***	-1.494***	-1.599***
	(0.265)	(0.263)	(0.204)	(0.255)
Observations	8,156	8,156	8,156	8,156
Municipalities	684	684	684	684
	D. Change in log earnings, self-employed			
Migrant inflow	-1.708***	-1.642***	-1.458***	-1.441***
	(0.458)	(0.463)	(0.354)	(0.443)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
Demographics		✓	✓	✓
Time dummies			✓	✓
Covariates trends				✓

Notes: This table shows the relationship between changes in log earnings and the inflow of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Each observation is a municipality-year cell. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Column (3) also include time dummies. In column (4) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). All regressions are weighted by the 1991 native population. Robust standard errors calculated by the procedure in Borusyak et al. (2021) in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table B7: **Inference a la Borusyak et al. (2021): IV/2SLS estimates on unemployment and participation**

	(1)	(2)	(3)	(4)
	A. Change in unemployment rate			
Migrant inflow	0.194***	0.174***	0.166***	0.168***
	(0.060)	(0.061)	(0.044)	(0.055)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	B. Change in inactivity rate			
Migrant inflow	-0.179***	-0.139**	-0.133***	-0.150**
	(0.058)	(0.058)	(0.050)	(0.062)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
Demographics		✓	✓	✓
Time dummies			✓	✓
Covariates trends				✓

Notes: This table shows the relationship between changes in unemployment and inactivity rates and the inflow of migrants from the Semiarid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Each observation is a municipality-year cell. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Column (3) also include time dummies. In column (4) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). All regressions are weighted by the 1991 native population. Robust standard errors calculated by the procedure in Borusyak et al. (2021) in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table B8: **Inference a la Borusyak et al. (2021): IV/2SLS estimates on nonwage benefits**

	(1)	(2)	(3)	(4)
	A. Food			
Migrant inflow	-0.375**	-0.511***	-0.551***	-0.751***
	(0.155)	(0.158)	(0.116)	(0.145)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	B. Transport			
Migrant inflow	-0.592***	-0.709***	-0.756***	-0.581***
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
	C. Health			
Migrant inflow	-0.486**	-0.591***	-0.608***	-0.242
	(0.194)	(0.200)	(0.136)	(0.169)
Observations	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955
Demographics		✓	✓	✓
Time dummies			✓	✓
Covariates trends				✓

Notes: This table shows the relationship between changes in the proportions of working-age native population, employed in the formal sector, who receive nonwage benefits and the inflow of migrants from the Semi-arid region in each destination municipality, measured as a fraction of the native working-age population in 1991. Column (2) controls for the log of the native population in the previous year and changes in the share of individuals with high school and college education; the shares of female, white and black individuals; and a cubic polynomial of age. Column (3) also include time dummies. In column (4) we add a linear trend of 1991 municipality-level characteristics (age and the shares of high school and college educated individuals). All regressions are weighted by the 1991 native population. Robust standard errors calculated by the procedure in Borusyak et al. (2021) in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.