TEXTO PARA DISCUSSÃO

No.702

Internal Migration and Labor Market Adjustments in the Presence of Non-wage Compensation

Raphael Corbi

Tiago Ferraz

Renata Narita

DEPARTAMENTO DE
ECONOMIA
PUC
RIO

Internal Migration and Labor Market Adjustments in the Presence of Non-wage Compensation

Raphael Corbi* Tiago Ferraz† Renata Narita‡

October 1, 2024

Abstract

In this paper, we argue that adjustments in non-wage compensation are empirically relevant and have important implications for understanding the effects of labor supply shocks. We examine the labor market impacts of internal migration in Brazil through a shift-share approach, which combines weather-induced migration with historical settlement patterns at each destination. Our findings indicate that increasing migration inflows lead to a reduction in formal employment while simultaneously increasing informality by a similar magnitude. Unlike previous studies, we observe a significant negative impact on earnings within the formal sector. Additionally, we provide evidence that the proportion of formal workers receiving non-wage benefits declines, underscoring that substantial adjustments take place in the formal sector, even in a context of high informality. We interpret our results within a framework where formal and informal labor inputs are imperfect substitutes and where non-wage benefits generate predictions that align closely with our empirical findings.

Keywords: Internal migration, wages, employment, non-wage benefits.
JEL Codes: J2, J3, J61, O15.

*University of São Paulo, Department of Economics, e-mail: rcorbi@usp.br
†University of São Paulo, Department of Economics, e-mail: tiago.ferraz@usp.br
‡PUC-Rio, Department of Economics, e-mail: rnarita@econ.puc-rio.br

We are grateful to Rodrigo Adão, Luis Alvarez, Bruno Barsanetti, Christian Dustmann, Richard Hornbeck, Peter Hull, Rudi Rocha, Gabriel Ulyssea, and conferences and seminars participants for helpful comments. The authors acknowledge financial support from FAPESP PhD Scholarship 18/12249-7 and CNPq/MCTI/FNDCT research grant 402949/2021-8.

1 Introduction

Migration, both internal and international, significantly influences development, demographics, and economic trends. In response, a vast body of literature has explored its impact on native populations, particularly concerning employment and wages. Borjas (2014), in his extensive work, highlights the costs of immigration for native workers who compete directly with immigrants. In contrast, many scholars argue that migration's effects are more nuanced(Card and Peri, 2016). For example, Card (2009) finds that immigration to the U.S. has only a minimal impact on native wages, while Ottaviano and Peri (2012) identifies minor positive wage effects for certain groups. This balance of perspectives emphasizes the complexity of immigration's economic consequences.

Canonical partial equilibrium models, characterized by perfect competition and substitutability between native and migrant workers, suggest that wage adjustments fully absorb labor market shocks when native workers are immobile. Alternatively, when wages are rigid, these models predict a decline in native employment (for an early example, see Altonji and Card, 1991. To reconcile the seemingly conflicting empirical findings, scholars have expanded the models to include multiple outputs and technological change (Lewis, 2011; Dustmann and Glitz, 2015), and highlighted that different model specifications may capture distinct parameters (Dustmann et al., 2016).

In this paper, we argue that non-wage compensation adjustments, though underexplored in migration literature, are empirically significant and can profoundly influence the analysis of labor supply shocks triggered by migration [1]. Specifically, we examine the effects of internal migration in Brazil on native labor market outcomes in a context characterized by downward wage rigidity, where non-wage benefits play a vital role in compensation, and labor informality is widespread. This setting enables us to explore how firms, when faced with migration-driven labor supply shocks, adjust both earnings and benefits in the formal sector, despite strict wage regulations and pervasive informality.

We motivate our analysis by developing a simple model based on Harris and Todaro (1970) and Fields (1975), where formal and informal employment coexist in destination regions. We extend their framework by incorporating non-wage benefits as a mechanism for adjusting the total compensation of formal sector workers, allowing employers to absorb labor supply shocks more flexibly. Additionally, our model assumes that production combines formal and informal labor along with capital, treating formal and informal workers as substitutes in production rather than maintaining a strict segmentation between these sectors.[2] In the formal sector, firms must comply with

[1] For example, Clemens (2021) suggests that non-wage adjustments may help clarify debates on minimum wage economics.

[2] This assumption is common in the literature, as discussed in Ulyssea (2010, 2018) and Bosch and Esteban-Pretel (2012).

minimum wage laws and frequently offer non-wage benefits (e.g., employer-provided health insurance, food, and transportation subsidies). In contrast, informal workers are not subject to these protections, and their earnings are fully flexible, behaving according to a standard competitive labor market. Our model predicts that, due to competition from incoming migrants, informal sector earnings are likely to decline. Additionally, if the elasticity of substitution between formal and informal labor is high, migration inflows may also suppress the total compensation of formal workers. When formal wages cannot adjust downward because of minimum wage laws, firms may reduce non-wage benefits as a compensatory mechanism to absorb the shock.

We address the econometric concerns associated with the fact that migrants tend to move to areas with better labor market opportunities by taking advantage of a recent body of work that provides a clear framework for distinguishing sufficient conditions for identification and properly computing standard errors (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2021; Jaeger et al., 2018; Adão et al., 2019). In particular, we combine two extensively used identification strategies into a shift-share instrument approach. First, we exploit exogenous rainfall and temperature shocks (or "shift") at the origin to predict the number of individuals leaving a municipality in the Brazilian Semiarid region. Then we leverage the history of the Semiarid as a large source of climate migrants and use the past settlement patterns (or "share") to allocate migration outflows to destination areas (Munshi, 2003; Boustan et al., 2010). The resulting predicted inflow of migrants is an instrument for observed migration. We use this shift-share IV design and annual data at the municipality-year level from 1996-1999 and 2001-2009 to identify the causal effects of immigration on labor market outcomes in the destination areas.

We find that following a weather-induced migration inflow, earnings in both the informal and formal sectors decrease by 0.74% and 0.59%, respectively. Additionally, the share of native workers receiving food benefits declines by 0.22 percentage points, while transport benefits and health insurance decrease by 0.11 percentage points. Although overall employment remains unchanged, formal employment drops by 0.13 percentage points, with a corresponding rise in informality, highlighting a significant substitution between formal and informal labor. Those who lose formal jobs tend to transition into the informal sector, which operates without frictions like minimum wage laws. Contrary to existing literature, we observe negative impacts on formal sector earnings across various samples, except for low-education workers in markets with a high minimum wage bite. Research on internal migration in developing countries typically indicates that wage effects are concentrated in the informal sector or among low-education workers (Kleemans and Magruder, 2018; El-Badaoui et al., 2017). Our findings suggest that notable adjustments occur in the formal sector, even amidst high levels of informality, affecting not only earnings but also the share of workers receiving benefits in most markets.

The heterogeneity analysis reveals that the estimated effects on earnings are more pronounced in municipalities with a higher baseline of informality, aligning with our model predictions. We observe more significant effects on less educated native workers, who directly compete with migrants from Semiarid regions. These lower-educated individuals are more likely to leave the formal sector and suffer greater earnings reductions within the informal sector compared to their higher-educated counterparts. Additionally, since many low-education workers earn close to the minimum wage, the adjustment in benefit shares is anticipated to be more substantial for them. However, it is possible that those low-education workers who lose formal jobs are the ones receiving fewer benefits. This "positive selection" suggests that low-education workers remaining in formal employment are more likely to receive benefits, which mitigates the expected negative impact on benefit shares for this group. Conversely, for high-education workers, we observe negative adjustments in non-wage benefits and earnings. This finding aligns with existing literature that highlights significant cross-effects, where low-skill (high-skill) immigration adversely impacts the earnings of high-skill (low-skill) workers, even in the presence of informality and binding minimum wages (Borjas (2003); Card and Lemieux (2001); Kleemans and Magruder (2018)).

To investigate whether short- and long-term effects of migration differ (Jaeger et al., 2018), we estimate the impact of a contemporaneous migration shock on various outcomes over up to five years. Our findings indicate that the adjustments in earnings and employment tend to stabilize after three to four years.

Brazil provides a suitable environment for our investigation for three reasons. First, more than 3 million people in the Brazilian Semiarid, a historical source of climate migrants, left their hometowns during our sample period of 1996-2009. Second, a within-country analysis minimizes econometric concerns about allocating migrants to particular skill groups (Dustmann et al., 2012). Third, over 40 percent of workers are employed in the less frictional informal labor sector, where firms do not comply with labor market statutes, such as minimum wage and layoff regulations. The rest of the workforce participates in the formal sector where the minimum wage is binding (above 70% of the median wage) and non-wage compensation is frequently offered. In fact, more than 31 million people, or 20% of registered workers, are covered through employer-provided health insurance. After payroll expenses, this is the second highest component of total labor costs (ANS, 2019). In addition, 40% of these workers receive food subsidies, which cost firms about 57% of the minimum wage per worker.[3] To the extent that workers value non-wage benefits, changes in this margin of adjustment can have important welfare implications.

Our work is related to a broad literature that examines the impact of migration flows on native labor market outcomes (see Borjas, 2014 and Dustmann et al., 2016 for a review). Although migration within countries is a larger phenomenon,[4] most studies are concerned with international immigration to high-income countries, with particular attention paid to Mexican immigration to the United States (Borjas, 2003) and,

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

[4] 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).

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; Azoulay et al., 2022).[5] A growing set of studies explores environmental and other economic shocks to study the causal impact of internal migration on labor markets in the US, China, Indonesia, Thailand and Brazil (Boustan et al., 2010; Hornbeck, 2012; Imbert et al., 2022; Kleemans and Magruder, 2018; El-Badaoui et al., 2017; Imbert and Ulyssea, 2024; Albert et al., 2021).[6] Considering the studies on developing countries, Kleemans and Magruder (2018) and El-Badaoui et al. (2017) are closely related to our work showing that immigration leads to higher informality and lower informal wages or earnings in the destination regions, while Imbert and Ulyssea (2024) show that, in the short-run, workers reallocate from the formal to the informal sector, which is consistent with our findings. This agrees with the main predictions of the Harris and Todaro (1970) and Fields (1975) works based on wage rigidity in the formal sector.[7]

In contrast, focusing on long-run changes using decennial data from 2000 to 2010, Imbert and Ulyssea (2024) also find that immigration reduces informality driven by workers reallocating from informal to formal jobs, following a larger number of formal firms and jobs, and no change in labor supply or unemployment. Their results are consistent with a model of firm dynamics and informality that rationalizes their findings; In the long run, labor market frictions constraining formal demand expansion can be alleviated. Relatedly, Albert et al. (2021) argue that frictions in the capital market and spatial labor market frictions indeed constrain the reallocation process of climate migrants from agriculture to manufacturing, limiting potential positive returns of migration even in the long run. Although these two last works focus on Brazil, Imbert et al. (2022) find that migrants quickly find their way into formal manufacturing firms, compatible with a relatively flexible labor market in China.

Our contribution to the economics of migration literature is fourfold. First, we show that firms systematically adjust non-wage benefits in response to labor supply shocks. Second, accounting for such adjustments is key to understanding the effects of migration on natives. Consistent with high substitutability between formal and informal labor inputs, we find offsetting employment responses across formal and informal sectors. Third, we provide evidence on the effects of internal migration on local labor markets in a large developing country and show that these different adjustment

[5]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 differs across different groups of natives, and immigrants and native do not compete in the labor market within the same education-experience cells.

[6]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.

[7]Busso et al. (2021) provide several tests to assess the empirical relevance of the original Harris-Todaro formulation using Brazilian census data and show the importance of adding other aspects such as urban informality and unemployment.

patterns are relevant even in the presence of informality. Fourth, we add to a growing body of evidence that migration is a relevant coping mechanism against climate change, especially for vulnerable populations in rural areas of developing countries (Skoufias et al., 2013; Assunção and Chein, 2016).

Non-wage 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 European firms paid non-base wage components such as benefits and bonuses in 2013 (Babecký et al., 2019). Evidence shows that firms adjust non-wage components when faced with adverse economic shocks (Babecký et al., 2019) or as a strategy to offset collective bargaining (Cardoso and Portugal, 2005), particularly when base wages are rigid (Babecký et al., 2012). We add to this literature by showing that non-wage benefits are a relevant margin of adjustment, and ignoring it would lead us to underestimate the impacts of labor supply shocks due to internal migration.

This paper is organized as follows. The next section presents background information on the Brazilian Semiarid region and local labor markets. Section 3 outlines a simple framework for interpreting our findings. Section 4 describes the data and empirical framework and reports the first-stage estimates that link the observed migration patterns to our predicted migration flows. Next, we present and analyze the main results on employment, earnings, and non-wage wage benefits in Section 5. We also study the sensitivity and heterogeneity of our main estimates. Finally, we interpret our main estimates in light of our simple model and conclude.

2 Background

In this section, we first describe the economic background and weather conditions in the Semiarid region, the functioning of local labor markets in Brazil, and a simple framework to contextualize our analysis. We then discuss the main data sources on labor market outcomes, migration flows, and weather and present descriptive statistics.

2.1 Brazilian Semiarid

The Brazilian Semiarid encompasses 960 municipalities spread over nine states, covering an area of around 976,000km^2.[8] According to the official definition by the Ministry of National Integration, a municipality qualifies as Semiarid 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[9]; (iii) risk of drought above 60%[10]. The average historical precipitation in the Semiarid is about 780mm, as opposed to around 1,600

[8]That is roughly the same as the territory of Germany and France combined. The semiarid comprises 11 percent of the Brazilian territory and includes parts of almost all Northeastern states (except for *Maranhão*) plus the northern area of *Minas Gerais*, but it does not cover any state capital.

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

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

mm for the rest of the country, while the average temperature is around 25°C. The rainy season occurs between November and April, with the highest precipitation levels after February, when the sowing season usually starts.

Municipalities are relatively small, with a median population of around 20,000, and have economies mainly based on agriculture and cattle ranching in small subsistence properties. Local economic activity is particularly susceptible to weather shocks (Wang et al., 2004), with some studies showing a loss of up to 80% of agricultural production in long drought (Kahn and Campus, 1992). About 80% of the children lived below the poverty line, and infant mortality reached 31 per 1,000 births in 1996, compared to a national average of 25% and 15 per 1,000 births, respectively (Rocha and Soares, 2015). More than 80% of the adult population had less than eight years of schooling in 1991.

Such poor socioeconomic indicators associated with periods of extreme drought have historically driven large outflows of migrants - or so-called *retirantes* - from the Semiarid to other areas of the country (Barbieri et al., 2010). During the 1960s and 1970s, net migration out of Northeastern states (where most of the Semiarid is located) was 2.2 and 3.0 million individuals (Carvalho and Garcia, 2002), which correspond to net migration rates of -7.6 and -8.7%, respectively. Between 1996 and 2010, around 3.0 million people left the Semiarid alone, searching for better conditions elsewhere in the country. Appendix Figure C1 shows that these migrants tend to be historically concentrated in some states. In the last four decades, São Paulo alone harbored almost 40 percent of the people arriving from the Semiarid. However, in relative terms, the arrival of migrants represented an increase in the population above 2% for the top 10 receiving states.

2.2 Labor Markets in Brazil

A common feature of labor markets in developing countries is the existence of a two-sector economy where the informal sector accounts for one to two-thirds of the GDP (see Perry et al. (2007) and Ulyssea (2020) for a review). In Brazil, more than 40% of the individuals work in the informal sector (those without registration or who do not contribute to social security), including most of the self-employed who are not protected through social security. When firms hire workers under a formal contract, they are subject to several legal obligations, such as paying minimum wages and complying with safety regulations. The registration also entitles workers to other benefits, such as a wage contract, which in Brazil prevents downward adjustment, working up to 44 hours a week, paid annual leave, paternity or maternity leave, retirement pension, unemployment insurance and severance payments (e.g. Gonzaga, 2003; 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 who is found to be 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).[11] 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, while informal firms are smaller and thus difficult to get caught (Cardoso and Lage, 2006). There are also other expected costs for formal firms associated with labor courts if the worker is fired and decides to file a lawsuit against the firm. The judges decide in favor of the workers in almost 80% of the cases (Corbi et al., 2022). This indicates a significant operating cost in the formal sector, particularly for smaller firms. Imperfect enforcement and costly regulation are associated with high informality in the country.

Finally, as there is a substantial overlap between the productivity distributions of formal and informal sectors (Meghir et al., 2015), even for lower percentiles of the overall distribution, both sectors should be affected by the influx of migrants. In other words, both sectors have workers who are close substitutes to the migrant workforce and thus will experience competition.

Non-wage compensation. Our empirical analysis focuses on three main fringe benefits we observe in the data: private health insurance, food, and transport subsidies. In Brazil, benefits became popular in the 1980s, as providing food subsidies and employer-provided health insurance became more frequent among private sector firms (Arbache, 1995). Data from PNAD surveys for 1996-2009 indicate that 12% of working-age native population receive food benefits, 11% receive transport benefits, and 7% get private health insurance through their employers.[12] Arbache and Ferreira (2001) estimate that benefits like food subsidies for instance cost around 57% of one minimum wage (around 16% of average total compensation). Similarly, Brazilian Federal Health Agency data (ANS, 2018) show that employer-provided health insurance cost on average R$582 in 2018, which is 17% of total compensation in that same year. These numbers imply that depending on how firms opt to mix benefits in the workers' package, these expenses may add up to above 30% of the total payroll cost. In the US, benefits, including employer-provided health insurance, account for around one-third of compensation costs (Clemens, 2021).

At least two reasons can explain the use of fringe benefits in workers' compensation. First, these benefits in Brazil are not subject to payroll taxes, reducing total labor costs. Second, labor legislation is generally more flexible regarding the provision of benefits such that it is easier to adjust benefits than wages (Arbache, 1995). Even though regulations for fringe benefits provision are considered less rigid than for wages, collective bargaining agreements (CBA) sometimes include clauses about these benefits. In particular, the third most common clause type among extended firm-level CBA includes wage supplements such as food subsidy (Lagos, 2020) Also, around 10% of all formal sector firms are under CBA with a clause on health plan/insurance (Marinho, 2020).

But how important should adjustments be in the non-wage compensation margin

[11] The minimum wage is above 70% of Brazil's median wage.

[12] Among formal native workers, these fractions are 39%, 36% and 21%, respectively.

in Brazil compared to other settings? Non-wage benefits should be more prominent in countries with high minimum wages. The minimum wage as a fraction of the median wage ranges is lower in the United States (32%), Japan (44%), and Mexico (46%) than in Brazil (75%) and many European countries, for example, the UK (55%) and France (61%).[13]

Although transport subsidies have been a mandated benefit in Brazil since 1985, we treat this as a benefit that firms can adjust to. This is likely the case since we observe that only 36% of formal sector workers report they receive this benefit. Firms may not fully comply with all aspects of labor regulations. Also, as transport benefits are non-wage compensation, firms do not incur payroll taxes. In addition, firms may deduct the cost with the offered subsidy from the base for income taxation and their operational cost, lowering net revenues, which are the basis for other corporate and payroll taxes.[14] This implies that firms have incentives to offer transport benefits and a further incentive to adjust it at the intensive margin by providing better transportation or increasing the benefit in cash.

3 A Simple Theory

In this section, we describe a simple model assuming perfectly competitive labor markets to guide our analysis. We assume migrants and natives as perfect substitutes and investigate the consequences of a migration shock that shifts the aggregate labor supply to the right. Then, we introduce intersectoral linkages where formal and informal workers are inputs that can be combined in a production function.

We begin with a standard competitive market model, in which the wage is determined, such as when labor supply equals labor demand and migration negatively affects wages. In the extreme case of an inelastic supply, migration does not affect the employment of natives, and the entire migrant workforce is absorbed. On the other hand, with an upward-sloping labor supply, the wage decline makes jobs less attractive for some native workers, and native employment falls.

Of course, this benchmark vastly simplifies how the labor market works. Downward wage rigidities are often present in reality due to minimum wage laws and collective bargaining agreements. In this case, migration shocks can be accommodated by job losses or lower labor costs, for example, reducing non-wage benefits (McKenzie, 1980; Clemens, 2021). In developing countries' labor markets, excess workers can also be accommodated in the competitive informal sector.

This traditional view of informality considers that the formal and informal sectors are segmented, which masks important intersectoral linkages, in particular, on the production side (see Ulyssea (2010, 2018) and Bosch and Esteban-Pretel (2012)). To address such context, we develop an extension of a model with informality in which

[13] Brazil (Source: PNAD Continua 2015). Data for OECD countries in 2019 (Source: OECD.Stat, https://stats.oecd.org/Index.aspx?DataSetCode=MIN2AVE

[14] The income tax due cannot be reduced by more than 10%.

the regulated sector is subject to minimum wage laws. This is guided by previous work by Kleemans and Magruder (2018) that builds on Harris and Todaro (1970) among others[15], and we modify their model in two ways.[16] First, we add non-wage benefits as a source of adjustment of total compensation for formal workers in the presence of binding minimum wages. Second, we assume that production combines formal and informal labor besides capital. We follow the literature and allow formal and informal inputs to substitute in production.[17] Firms comply with minimum wages by paying formal workers at least $\underline{W}$, where this is binding and can also offer non-wage benefits frequently observed in the data (e.g., employer-provided health insurance and food). On the other hand, informal workers are not entitled to minimum wages or non-wage benefits, and their wages are completely flexible as in a standard competitive setting.

Production is given by

$$Y = K^{1-\alpha} L^{\alpha}, \ 0 < \alpha < 1 \tag{1}$$

where labor combines two inputs, formal and informal, that can be substituted with elasticity $1/(1-\nu)$ as described below,

$$L = (\theta_f L_f^{\nu} + \theta_i L_i^{\nu})^{1/\nu} \tag{2}$$

with and θ_f and θ_i denoting the share parameters of formal and informal labor inputs, respectively, and $0 < \nu < 1$ reflecting heterogeneities across formal and informal labor.

Wages (or total compensation) are determined by marginal products for workers of each type. For formal sector workers, if we assume that firms can freely adjust non-wage compensation but can only reduce wages up to the binding minimum, then total compensation equates to their marginal product, such that

$$W_f = \underline{W} + B = \frac{\alpha \theta_f K^{1-\alpha}}{L_f^{1-\nu}} (\theta_f L_f^{\nu} + \theta_i L_i^{\nu})^{(\alpha-\nu)/\nu} \tag{3}$$

where B is the cost of non-wage benefits to employers.

For informal workers, wages are freely determined. This implies that wages are

$$W_i = \frac{\alpha \theta_i K^{1-\alpha}}{L_i^{1-\nu}} (\theta_f L_f^{\nu} + \theta_i L_i^{\nu})^{(\alpha-\nu)/\nu} \tag{4}$$

Now, consider that the immigration labor supply shock affects informal workers, L_i.[18] This is consistent migrants and informal workers being more similar to low-

[15]e.g. Fields (1975) and Mazumdar (1976). In the same spirit, Busso et al. (2021) considers two urban sectors - formal and informal - and unemployment.

[16]Kleemans and Magruder (2018) also follow Card and Lemieux (2001) and Borjas (2003)'s model to address the differences in the impacts of immigration by skill, which in their context (Indonesia) is crucial given that the poor natives were less educated than the migrant workforce.

[17]In Brazil and Mexico, respectively, 40% and 44% of informal labor are employed in formal firms (Ulyssea, 2018).

[18]This can be thought as a supply shock in the number of low-education workers. Given that the informal sector is more intensive in low-skilled labor, it should be the one affected directly.

education natives. Our model leads to the following predictions derived in Appendix A:

(i) $\frac{\partial W_i}{\partial L_i} < 0$

(ii) $\alpha - \nu < 0 \Rightarrow \frac{\partial B}{\partial L_i} < 0$

These statements show a basic intuition for the effects of internal migration of workers that are similar to native workers in the informal sector. First, informal wages are expected to decline due to competition with similar incoming migrants. Second, the effect on non-wage compensation in the formal sector depends on the degree of substitution between formal and informal inputs. If the elasticity of substitution is higher than the labor share then an inflow of informal workers lowers the compensation of formal workers. Additionally, in cases where wages cannot fully adjust due to binding minimum wage laws, non-wage benefits are anticipated to decrease as well.

The model also accounts for variations in outcomes across submarkets; for instance, in markets characterized by higher baseline informality (high θ_i), the negative impacts on wages or total compensation are shown to intensify as the share of informality rises (see Appendix A).

4 Data and Empirical Strategy

In this section, we begin by listing the main sources of data used in our analysis and showing some descriptive statistics. Then, we describe the empirical framework and report first-stage estimates that link observed migration patterns to our predicted migration flows.

Migration We draw data from three waves of the Brazilian Census (1991, 2000, and 2010), provided by the *Instituto Brasileiro de Geografia e Estatística (IBGE)*, to construct two of the main variables used in our study.[19] First, we leverage Census answers about the municipality of origin and year of migration to construct a measure of yearly migration outflow from each municipality in the Semiarid and a measure of inflow to each destination (all but Semiarid) during the 1996-2010 period. Second, we use the 1991 Census to build a "past settlement" measure by associating the share of migrants from each Semiarid municipality who reside in each destination. In Appendix C, we provide more details on how we structure our yearly migration dataset.

Weather shocks Weather data were retrieved from the Climatic Research Unit at the University of East Anglia (Harris et al., 2020). The CRU Time Series provides worldwide

[19] As several municipalities were split into new ones during the 1990s, we aggregate our data using the original municipal boundaries as they were in 1991 (so-called "minimum comparable areas" or MCA) to avoid potential miscoding regarding migration status or municipality of origin. We use municipality and MCA as synonyms throughout the paper.

monthly gridded data of precipitation and temperature at the $0.5^{\circ} \times 0.5^{\circ}$ 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.[20] We define rainfall and temperature shocks as deviations from the historical average.[21]

Labor outcomes Our primary source of data for the outcome variables is the *Pesquisa Nacional por Amostra de Domicílios - PNAD*, a national household survey also collected by the *IBGE*, the bureau responsible for the Census. The *PNAD* survey is conducted yearly, except during Census years. Thus, our data spans from 1996 to 1999 and from 2001 until 2009. Interviews take place in 808 municipalities in all 27 Brazilian states and cover several dimensions such as education, labor, income, fertility, and household infrastructure. On average, around 300,000 people are interviewed in each round.

Two key features of the PNAD warrant further examination concerning the suitability of its data for our analysis. First, it is important to note that less than 20% of all Brazilian municipalities are represented in the survey. Nevertheless, these municipalities account for nearly 80% of the destinations chosen by migrants from the Semiarid region and encompass 65% of the employed population in Brazil. This concentration indicates that, despite the limited geographic scope, the survey captures a significant portion of migration trends and labor market dynamics. Second, the PNAD is not structured to be a representative survey at the municipality level. To validate its use as our primary source for labor outcome measures, we present simple descriptive statistics and empirical distributions of the main variables from both the PNAD and the Census in Appendix B. Appendix Table B1 provides these sociodemographic statistics. Furthermore, in Figure B1 and B2, we compare the municipality-level employment rates and log earnings in both formal and informal sectors from the 2009 PNAD with corresponding data from the 2010 Census for the same municipalities, demonstrating significant overlap in both formal and informal labor markets. Unfortunately, the Census lacks information on non-wage benefits. Additionally, Figures B3 and B4 present the correlation between employment rates and earnings in both sectors, utilizing PNAD data from 1996, 2001, and 2009, which align more closely with Census samples from 1991, 2000, and 2010. We argue that the outcomes from these datasets exhibit similar patterns.

Our main outcomes comprise earnings and employment. We explore whether the worker is an employee in the registered formal sector, in the informal sector, or self-employed. We also use information on non-wage compensation. The survey asks specifically whether the individual received any kind of payment or help to cover food or transport expenses and whether the employer provides health insurance. Unfortunately, data on the intensive margin of non-wage compensation are not available.

We restricted our attention to individuals between 18 and 65 years of age, who live

[20]This approach is similar to the one used by Rocha and Soares (2015).

[21]See Appendix G for a detailed description and discussion on this measure.

in the municipality for 10 years or more, and we refer to them as *natives*. We consider as destinations all PNAD municipalities that are not in the Semiarid region to minimize concerns about spatial correlation in weather shocks. We pool the 13 years of individual survey data and take averages at the municipality-year level. As our main regressions are first differences, the final destination sample has 684 unique municipalities and 8,190 municipality-year observations, averaging 2,152,950 individuals.

Table 1 describes municipality-level data for the origin (Panel A) and the destination (Panel B) municipalities. Semiarid areas show lower levels of rainfall and slightly higher temperatures and are less populated than destination municipalities. On average, 1.0 p.p. of Semiarid's population leaves every year, resulting in an average increase of 0.30 p.p. of the labor force in the destination.

Table 2 provides descriptive statistics for native individuals in the destination municipalities. In our sample, 63% of individuals are employed - with 31% having a formal job, the same proportion of informal workers. The unemployment rate is 13%, and 24% of individuals are not in the labor force. The average monthly earning is R$ 637.89, with the formal sector having a substantially higher average (R$ 788.22) than the informal sector (R$ 491;28).[22] Among the native individuals, 12% receive financial help to cover expenses with food, 11% for transport, and 7% for health expenditures.[23]

Finally, Table 3 compares migrants to low- and high-education natives. We consider low-education individuals with up to seven years of schooling, roughly equivalent to an incomplete elementary level. Migrants earn slightly more than less-educated natives. They also have a similar likelihood of working part-time and are slightly more likely to be employed in the formal sector when compared to low-education natives. On the other hand, high-education natives are more likely to work in the formal sector and have considerably higher pay. Appendix Table C1 shows that top occupations for migrants (e.g., typically bricklayer for men, domestic worker for women) also represent a large share of the occupations among low-education natives, but not for the high-educated. Also, the same five industries that concentrate almost 70% of working migrants also employ a similar share of low-education workers (see Appendix Table C2). Overall, this characterization is consistent with greater substitutability between migrants and less educated natives in the labor market.

4.1 Empirical Strategy

Here, 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. Next, we discuss and present supportive evidence on the validity of this shift-share instrument approach based on insights from the recent econometric literature that analyzes its formal structure.

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

[23]Less than 1% of informal and self-employed workers receive any form of non-wage compensation.

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

$$\Delta y_{dt} = \alpha + \beta m_{dt} + \gamma X_d + \psi_t + \epsilon_{dt} \tag{5}$$

where y_{dt} is a vector of labor outcomes at destination municipality d in year t, m_{dt} is the destination migrant inflow from the Semiarid region, X_d are destination-level controls, ψ_t absorb time fixed effects and ϵ_{dt} is the error term. The main challenge to identifying β is that the observed migration, m_{dt}, is the equilibrium between the demand and supply of migrants. Another issue is that the error term, ϵ_{dt}, may include unobserved characteristics that could be correlated with migration inflows. In particular, migrants could choose a specific destination municipality due to demand shocks leading to higher wages or job prospects. By differencing the outcome variables, we can account for time-invariant unobserved characteristics that could be correlated with migrant inflows but not the time-varying confounders, which would potentially bias OLS estimates.

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, defined as the observed number of migrants leaving the municipality divided by the population in the 1991 Census, from origin municipality o in year t, using weather shocks in the previous year:

$$m_{ot} = \alpha + \beta' Z_{ot-1} + \phi_o + \delta_t + \varepsilon_{ot} \tag{6}$$

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:

$$\widehat{M}_{ot} = \widehat{m}_{ot} \times P_o \tag{7}$$

In the second step, we use the past settlements of migrants from the origin o to municipality d to distribute them throughout the destination areas. More formally, we define this share of migrants from origin o settling into destination d in 1991. We fix our past settlement measure in 1991 across the period of our sample to avoid concerns about the persistence in migrant flows as discussed by Jaeger et al. (2018).

$$s_{od} = \frac{M_{od}}{\sum_d M_{od}} \tag{8}$$

allowing us to define our shift-share instrumental variable (SSIV) as

$$\widehat{m}_{dt} = \sum_{o=1}^{O} \frac{s_{od} \times \widehat{M}_{ot}}{N_d} \tag{9}$$

where N_d is the total native population at d in 1991. Thus, our instrument $\widehat{m}_{dt}$ can be thought of as a combination of exogenous shocks or 'shifts' $\widehat{M}_{ot}$ (weather-driven outflows) with 'shares' ($s_{od} \geq 0$) or 'ethnic enclaves' as in Card (2001).[24]

The validity of the shift-share instrument approach relies on assumptions about the shocks, exposure shares, or both, as discussed by 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 these insights, we leverage origin-level weather shocks[25] for identification and define the reduced-form relationship that associates labor market outcomes and the predicted migrant flow at the destination as

$$\Delta y_{dt} = \alpha + \beta \widehat{m}_{dt} + \gamma X_d + \psi_t + \epsilon_{dt} \tag{10}$$

We follow Borusyak et al. (2021) and calculate an origin-level weighted average version of equation 10, that uses the exposure shares s_{od} as weights, and results in the transformed reduced form relationship

$$\bar{y}_{ot} = \alpha + \beta \widehat{M}_{ot} + \bar{\varepsilon}_{ot} \tag{10'}$$

In Appendix H, we provide a detailed derivation of the transformation performed. As discussed in Borusyak et al. (2021), the consistency of our shift-share approach is based on two conditions:

Assumption 1(*Quasi-random shock assignment*): $\mathbb{E}[Z_{\tilde{o}}|\bar{e}, s] = \mu$ for all $\tilde{o}$.

Assumption 2 (*Many uncorrelated shocks*): $\mathbb{E}[\sum_o s_o^2] \to 0$ and $Cov[Z_{\tilde{o}}, Z_{\tilde{o}'}|\bar{e}, s] = 0$ for all $\tilde{o}, \tilde{o}'$.

where $\tilde{o} = (o, t)$, $\bar{e} = \{\bar{e}_{\tilde{o}}\}_{\tilde{o}}$, $s_o = \sum_d s_{od}$ and $s = \{s_o\}_o$. As in Borusyak et al. (2021), $\bar{e}_{ot}$ correspond to the error term from equation 5 computed at the level of shocks (e.g., municipality of origin). Assumption 1 guarantees that our shift-share IV is valid when weather shocks are as-good-as-randomly assigned, which comes from standard natural shocks arguments. Given identification, Assumption 2 gives us consistency when the number of observed shocks is large and when shocks are mutually uncorrelated given the unobservables and s_o. In the Appendix Table H1 we show that the effective sample size[26] is sufficiently large, reassuring us that exposure concentration is not a relevant issue in our setting. Also, in Appendix Tables I2-I4, we present evidence that the

[24] In appendix G we discuss this approach in more detail.

[25] Figure 1 illustrates the variation in weather shocks that are used for identification.

[26] According to Borusyak et al. (2021), the effective sample size is the inverse of the Herfindahl index of concentration of migrants: $H = \frac{1}{\sum_o s_o^2}$.

shocks we are using can be treated as uncorrelated, which supports the validity of the Assumption 2 in our setting.

One additional advantage of using the origin-level shocks concerns hypothesis testing. Adão et al. (2019) show that conventional inference in shift-share regressions is generally invalid because observations with similar exposure shares are likely to have correlated residuals, potentially leading to null hypothesis overrejection. Borusyak et al. (2021) show that by using the shock-level relationship instead of the destination-level, one can obtain standard errors that converge to those obtained by the Adão et al. (2019)'s correction procedure.

4.2 Weather-induced Migration

We begin the exploration of our first-stage results by estimating variations of specification 6 and report the estimates in Table 4. All regressions control for temperature shocks and the log of the 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 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. For brevity, we omit (mostly insignificant) coefficients associated with temperature shocks in Table 4. Standard errors are clustered at the grid level to account for the fact that municipalities in the same grid will have similar shocks.[27]

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 does not change the baseline results. More important to our identification, we include rainfall and temperature shocks one year forward as controls 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 the historical average will experience an increase of $1p.p.$ in migration outflow rate.

Next, we distribute the predicted migration outflow shock using past settlement patterns of migrants from the origin municipality o to destination d. A *sine qua non* requirement implicit in our empirical framework is that both predicted migration outflow and inflow rates, $\widehat{m}_{ot}$ and $\widehat{m}_{dt}$ respectively, should be strongly correlated with their observed counterparts. Figure 2 illustrates that our predictions provide a strong fit for the observed migration. Panel (a) shows the relationship between the predicted and observed number of migrants leaving the Semiarid region and entering non-Semiarid

[27]Similar, but not identical, as 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. Therefore, two municipalities inside the same grid have different shocks because the distance to the centroid is not the same.

municipalities accumulated over the period 1996-2010. Panel (b) shows the predicted and observed numbers of incoming Semiarid migrants for destination municipalities.

In Appendix G, we describe in more detail our data source for weather shocks, discuss alternative weather measures, 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, $\widehat{m}_{dt}$, strongly correlate with observed migration. Appendix Table H1 reveals that our first-stage point estimates are close to a one-to-one relationship (0.93) - making the magnitude of reduced-form and IV estimates almost identical - and have an F-stat of 2,172.[28]

5 Labor Market Effects of Migration Inflows

Now, we turn our attention to labor markets at the destination and investigate how migration inflows affect the earnings and employment of native workers. Next, we explore how labor markets adjust to migration shocks regarding non-wage compensation.

First we investigate how native workers' earnings adjust to exogenous migration inflows. Table 5 reports the OLS and SSIV estimates for employment (columns 1-3) and earnings (columns 4-6). In all regressions, we control for a vector of destination-level characteristics measured in 1991 (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) . The regressions were weighted by the working-age native population in 1991. Standard errors are clustered at the origin municipality level.

Panel A shows that the OLS results on employment and earnings are not statistically significant. Comparing the OLS with the SSIV results in Panel B, we find strong evidence of "moving to opportunity bias" in the earnings and formal employment OLS estimates.[29] The results in Panel B reveal a negative effect of the inflow of Semiarid migrants on log earnings for native workers. A one percentage point increase in the number of migrants reduces earnings by 0.87%. When we look at native workers holding a formal job or those in the informal sector, including self-employed workers, we find that increasing the migrant share by 1 percentage point reduces earnings of formal workers by 0.59% and of informal workers by 0.74%. The negative effect on informal earnings is greater, albeit only slightly. That is consistent with the absence of downward

[28] 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 an F statistic of 104.7, significantly higher than the broadly accepted threshold of 10.

[29] Often referred to as the 'moving to opportunity bias', regressing labor market outcomes on immigrant stocks may be confounded by the tendency of migrants to be attracted to areas with better labor market opportunities (Kleemans and Magruder, 2018). OLS estimates of migration effects are, therefore, likely to be biased in the positive direction.

wage rigidity in this sector, such that the classic predictions from perfect competition prevail.

The magnitude of our estimates for the formal sector is larger than that found by Imbert et al. (2022) (a reduction of 0.15% including pensions and housing benefits). An important difference between our approaches is that their estimates come from changes over six years, while we use year-on-year changes. Similarly, our estimates for the formal sector are larger than those from Kleemans and Magruder (2018) who find no effects, even though in Indonesia migrants were more educated than natives. On the other hand, they find larger negative effects on wages of informal workers. In their paper, they estimate the effects on the log income per hour - so they may also capture the adjustment on hours worked - while we focus on log monthly earnings.

We also investigate the differential effects according to the native worker's position in the earnings distribution. We calculate the municipality-level thresholds of each decile of the earnings distribution, separately for each sector, and run a regression on the changes of each threshold. Figure 3 reports our results. For those workers employed in the formal sector, we find smaller impacts at the bottom of the distribution. This is consistent with wage rigidity in the formal sector limiting the negative impacts for low-paid workers. The entry of low-skilled migrants competing for less paid jobs in this sector can be enough also to reverberate and have amplified effects all the way up the earnings distribution, if low and high-skill workers are substitutes. For informal workers, the impact is substantially stronger for those at the bottom third of the distribution, consistent with classic predictions from perfect competition and greater substitutability between migrants and less skilled natives in this sector.

Our results for employment in Panel B reveal no effect on overall employment, however we find offsetting employment responses across formal and informal sectors. A percentage point increase in the inflow of migrants reduces the share of formal employment by 0.13p.p. and increases the share of informal by almost the same amount (0.11p.p.) Kleemans and Magruder (2018) find a larger negative effect on formal employment (-0.33p.p.) and no significant impact on the informal sector. In contrast, we find evidence consistent with the reallocation of workers between sectors. This result is also consistent with important cross-effects we find on formal sector that is predicted by our model in the case of high substitutability between formal and informal inputs.

To draw a more complete picture, we also estimate the impacts on unemployment and labor force participation reported in Table D1. An increase of 1 percentage point in the migration share leads to an increase of 0.09p.p. in the unemployment rate and a decrease of 0.08p.p. in the proportion of out-of-labor-force individuals. One reason behind these estimates could be that, if the primary earner in the household loses 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 confirm this mechanism by running the same regressions separately for individuals identified as head or non-head of the household. According to Table D2, almost all employment effects come from the head of households, while the changes in unemployment and inactivity rates stem from non-head members. This change in composition of workers can be at least part of the explanation for the drop in earnings we observe at bottom of the formal earnings distribution in Figure 3.

5.1 Non-wage Compensation.

We now explore an additional margin of adjustment due to migration shocks. As firms operating in the formal sector cannot reduce wages below the legal minimum, they may adjust to labor supply shocks by reducing fringe benefits, as discussed in Section 2.2. We construct the share of working-age native workers who receive each benefit. Table 6 reports the estimates. A one percentage point increase in the predicted number of migrants reduces the share of workers receiving food subsidy by $0.22p.p.$, and transport and health insurance by $0.11p.p.$, consistent with our model prediction that a labor supply shock should reduce formal total compensation in the case of high elasticity of substitution between formal and informal inputs.

Next, we complement these estimates by focusing on firms' behavior as health insurance providers to their employees. Around 20% of workers get private health insurance through their employers. In 2018, the average employer-provided health insurance benefit cost, on average, R$582, or 17% of total compensation in that same year (ANS, 2018).[30] Instead of relying on survey data, here we turn to firm-level administrative data on health insurance contracts obtained from *Agência Nacional de Saúde Suplementar (ANS)*, the Brazilian regulatory agency responsible for overseeing the private health industry. They provide information about every employer-sponsored contract signed as far back as 1940. We have data on the date when the contract was signed and the firm unique identifier, which we can use to merge with *RAIS*, an employer-employee matched dataset obtained from the Ministry of Labor that provides firm-level data on the near universe of formal employment contracts. We define an indicator variable $y_{idt} = \mathbb{1}(t \geq t^s)$ for each firm i in the destination municipality d at year t, with t^s being the year when the health insurance is hired. Then we estimate how migration inflow rates at destination municipality d affect changes in y_{idt}, that is, the variation in the likelihood that firm i provides health insurance to its employees. Formally, the estimated equation is

$$\Delta y_{idt} = \alpha + \beta \widehat{m}_{dt} + \phi_i + \delta_t + \mu_{idt} \tag{11}$$

where $\widehat{m}_{dt}$ is the predicted incoming migration rate; ϕ_i and δ_t are firm and year dummies. All the regressions are weighted by the number of employees in the firm in 1996, the first year in our sample.

In column 1 of Table D3, we find that firms operating in a municipality that receives more incoming migrants are, on average, less likely to provide health insurance to employees. An increase of 1 percentage point in the migration rate (equal to one standard

[30] See section 2.2 for more details.

deviation as reported in Table 1) implies a 1.5p.p. decrease in the share of firms that provide health insurance, roughly the average of y_{idt}. In Columns 2-5, we restrict the sample to different bins of firm size. The effect is close to zero and insignificant for firms below 100 employees but negative and of greater magnitude for larger firms. Firms above 100 employees are at least six times more likely to provide health insurance as compensation.

All such evidence is consistent with formal firms systematically adjusting non-wage benefits and earnings in response to labor supply shocks, reinforcing that important adjustments occur in formal contracts even in the presence of informality.

5.2 Exploring the Mechanism

As discussed earlier in this section, earnings fall in both sectors. However, the magnitude of this fall might depend on the extent of the labor market rigidity due to minimum wages. The adjustment should occur mostly via formal employment and informal earnings in places with a more binding minimum wage. In regions where the minimum wage bite is lower, it works primarily through total compensation, while the impacts on employment would be limited. Finally, non-wage compensation may help introduce some wage flexibility, which may, in principle, reduce job losses.

To provide a formal test for this mechanism, we calculate the baseline Kaitz index (minimum-to-median wage rate) in 1996 and divide the sample into two groups. Municipalities with high (low) minimum wage bite are defined as those above (below) the median of the Kaitz index. Then, we run the same SSIV specification in Panel B from Tables 5 and 6 separately for each group. Tables F1 and F2 present our findings.

In regions where the minimum wage is more binding (as shown in columns 1-3), the informal sector experiences a significantly stronger negative impact on earnings. A 1 percentage point increase in the migration rate leads to a 1.6% decrease in informal sector earnings. On the employment side, the increase in informal sector jobs mirrors the decline in formal employment, suggesting that native workers who lost their formal jobs transition into the informal sector. Additionally, there is a slight decrease in the share of workers receiving food benefits, with no changes in other non-wage benefit categories. We attribute this lack of adjustment in non-wage compensation to selection effects, primarily driven by a large reduction in the formal employment share. Our findings indicate that formal workers who lost their jobs are less likely to receive non-wage benefits and are more likely to move into the informal sector, where such benefits are not available. Moreover, positive selection among workers who remain in the formal sector helps mitigate the negative effects predicted by our model.

When the minimum wage is less binding (as seen in columns 4-6), there is minimal impact on informal employment and earnings. In line with the mechanism previously discussed, the adjustment in the formal sector primarily occurs through changes in total compensation, including both wages and non-wage benefits. A 1 percentage point increase in the in-migration rate results in a 0.72% decrease in formal sector earnings. Additionally, the share of workers receiving non-wage benefits declines, with reductions of 0.24 percentage points for food subsidies, 0.13 percentage points for transport subsidies, and 0.17 percentage points for health insurance coverage.

In principle, the varying impact of how binding the minimum wage was at baseline should primarily affect low-skill workers. To test this, we ran regressions by minimum wage bite for workers with low and high education levels separately. As shown in Table F3, we found no significant impact on formal earnings for low-education workers in markets where the minimum wage bite was high, suggesting that the minimum wage only binds for this group. Conversely, in low minimum wage bite markets, formal earnings for low-education workers dropped by 1.3%, as expected. For high-education workers, most earnings effects were negative, regardless of how binding the minimum wage was at baseline. Although this group should theoretically be unaffected by minimum wages, we still observed substantial negative earnings effects, particularly in high minimum wage markets. This pattern is consistent with high substitutability between high- and low-education workers, a concept that will be discussed further in the next section.

Our findings indicate that when firms in the formal sector are unable to reduce wages, they may dismiss formal workers. These workers often move to the informal sector, where average earnings are lower. Additionally, firms tend to replace some high-skill native workers with low-skill informal labor. A 1 p.p. increase in the migrant share leads to a 0.18 p.p. rise in total employment among low-education workers, while employment for high-education workers declines by a similar rate.

In terms of non-wage benefits, Table F4 indicates that non-wage benefits do not adjust for low-education workers in markets where the minimum wage is highly binding. As we previously discussed, this is likely a result of selection, driven by a decrease in the formal employment share. Our findings suggest that this selection effect is more pronounced among low-education workers. In particular, less educated workers who exit formal employment in these high minimum wage bite markets are less likely to receive non-wage benefits.

5.3 Sensitivity and Heterogeneity Analysis.

This section provides an overview of the robustness checks conducted to validate our key findings. Following this, we explore the heterogeneity of our primary estimated effects, focusing on the level of informality in the labor market and the differences by workers' education levels.

The first issue we need to address is the possibility that labor market outcomes in destinations receiving incoming migrants might have evolved differently than they would have in the absence of this migration shock. Following the advice of Borusyak et al. (2021), we investigate the common trend assumption by regressing lagged outcomes in differences on our SSIV while using the same set of covariates as in the main specifications. The results, presented in Table E1, reveal that several of these pre-trends are correlated with the instrument. To reinforce the validity of our identification strategy, we reestimate our main specifications, adjusting for these pre-trends. As shown in Table E2 and E3, the results remain robust, providing additional confidence in our findings.

The second potential threat to our identification strategy concerns the possibility of a shift in local labor demand that could confound our results. If this were occurring, migrants from regions outside the Semiarid would likely be attracted to the same destinations as those from the Semiarid. In this scenario, we would expect a positive correlation between the inflow of migrants from the Semiarid and the inflow of migrants from other regions. To demonstrate that this is not the case, we conduct a regression analysis, where we regress the inflow migration rate from other areas on the predicted inflow rate of migrants from the Semiarid and show the estimates in Table E4. Column (1) includes time and municipality fixed effects, while in Column (2), we add the same set of controls from our main results. Point estimates are close to zero and not statistically significant in any specification.

The third threat to our identification strategy arises from the assumption that rainfall shocks in the origin municipalities impact destination labor markets solely through internal migration. A potential violation of this assumption could occur if a negative income shock in the origin, caused by lower rainfall, also reduced trade flows to certain destination areas. In such a scenario, industries more exposed to trade shocks, such as agriculture or manufacturing, might be more severely impacted. To investigate this, we present in Table E5 the coefficients from regressions of changes in log earnings based on the predicted migrant inflow rate, disaggregated by industry. Our results indicate that there are no statistically significant effects on the earnings of workers in the agricultural or manufacturing sectors. The observed impacts are concentrated among native workers employed in the service sector, who are less likely to be affected by negative shocks in the origin municipalities.

Finally, we explore the sensitivity of our results according to the degree of aggregation of regions of origin. In Appendix H, we argue that the consistency of our shift-share instrument needs the origin-level shocks to be mutually uncorrelated. As rainfall shocks are likely correlated across smaller geographical units, in Appendix I we investigate this issue by re-constructing our instrument according to larger catchment areas of origin of a migrant - such as a microregion or a mesoregion - instead of a municipality.[31] First, we document that spatial correlation among shocks decreases dramatically as we consider larger areas. Second, Tables I2-I4 show that our results associating migration and rainfall, earnings, employment, and non-wage benefits remain virtually unchanged, indicating that spatial correlation among rainfall shocks in origin municipalities are irrelevant to our results.

[31] IBGE (1990) defines microregions as "groups of economically integrated municipalities sharing borders and structure of production". Mesoregions are collections of microregions of which not all municipalities share borders. The Semiarid has 960 municipalities, 137 micro, and 35 mesoregions.

Heterogeneity by informality levels. In the model presented in Section 3, the parameter θ_i represents the exogenous informal input share, which serves as a proxy for informality at the market level. We demonstrate that the negative impacts on earnings and total compensation in both sectors increase with θ_i, and our empirical results support this prediction. Table F5, Panel A, shows stronger earnings effects in municipalities with higher baseline informality. Panel B reveals a larger adjustment in formal employment in high-informality municipalities. The fact that food and health benefits do not experience greater impacts in these areas can be explained by selection (see Table F6). Specifically, if formal job losses predominantly affect workers less likely to receive benefits, positive selection could mitigate the negative impact on non-wage benefits, especially in high-informality labor markets.

Heterogeneity by education. Next, we assess whether individuals with different levels of education may experience differential impacts. In our study context, we expect low-education native workers to be close substitutes for migrants. We define low-education as those with up to 7 years of schooling, equivalent to an incomplete elementary education. In our sample, 56% of natives are low-educated. We then reestimate the effect of migration on local labor market outcomes of natives with low and high education separately.

Table F7 presents the employment and earnings effects segmented by education level. In columns 1-3, the employment results suggest that less educated native individuals are more likely to exit the formal sector and move into informal sector employment, compared to those with higher education levels. Columns 4-6 examine the differential effects on log earnings. In the formal sector, earnings among native workers decline uniformly across education levels. However, the earnings loss in the informal sector is notably larger for low-education native workers. This aligns with the hypothesis that less educated natives face greater competition with migrants in the informal labor market, amplifying their earnings losses.

Regarding adjustments in non-wage benefits, Table F8 indicates that migration inflows negatively affect nearly all types of benefits across both education groups. Interestingly, the impact is more pronounced among high-education workers, despite the absence of direct competition from immigrants in this group. This phenomenon may be explained by the assumption of positive selection among low-education workers, who tend to exit the formal sector and are less likely to receive benefits. As a result, the negative impacts on non-wage benefits for low-education workers may be overestimated, reflecting less or no adjustment in their benefits.

Although we do not explicitly model skills, we can consider the heterogeneity of skills within both the formal and informal sectors, using education as a proxy, as suggested by extensive literature (e.g. Borjas (2003); Card and Lemieux (2001); Kleemans and Magruder (2018)). Assuming a high elasticity of substitution between low-education and high-education workers within each sector, we expect that the effects on earnings or total compensation for high-education workers should be negative. This assumption aligns with our findings.

5.4 Dynamic Effects

Here, we focus on the dynamics of migration's impact on local labor markets in Brazil. The short-run and long-run effects may differ as markets adjust to current shocks. Jaeger et al. (2018) report that the short-run local impacts of migration inflows in the US during the 1970s were more negative than previously documented, indicating a potentially significant initial impact on native workers. However, they also demonstrate that much of this decline is reversed in subsequent periods.

In our SSIV setup, accounting for dynamic effects proves challenging due to the serial correlation of migration shocks by design. This characteristic complicates the disentangling of contemporaneous effects from those shocks that occurred in the past but have persisted over time. To better understand how the adjustment path evolves, we estimate the following equation:

$$y_{dt+k} - y_{dt-1} = \alpha + \theta_t m_{dt} + \gamma X_d + \varepsilon_{dt} \tag{12}$$

The coefficient θ_t represents the cumulative effect of an incoming migration shock at year t on the changes in outcomes for destination municipality d propagating up to the year $t+k$, where $k \in [0;5]$. Equation 12 is estimated separately for each k using a sample restricted to the period 1996-1999 and 2001-2004.[32]

This strategy builds on the work of Autor et al. (2014), who examined the impact of exposure to trade shocks on the cumulative outcomes of U.S. workers on a rolling annual basis from 1991 to 2007. It also relates to the study by Dix-Carneiro and Kovak (2017), which analyzed the effects of trade liberalization on Brazilian labor markets between 1990 and 1995 by estimating the empirical impulse response function for cumulative impacts post-liberalization. While Autor et al. (2014) focused on the change in exposure to trade shocks throughout the entire period as the primary explanatory variable, Dix-Carneiro and Kovak (2017) assessed the effects of a specific trade policy on cumulative outcomes up to 15 years after its full implementation in 1995. Our approach is situated between these two methodologies, as we estimate the cumulative effects of incoming migration shocks on outcomes up to k years forward.However, unlike the previous studies, we allow these shocks to vary annually, producing estimates that reflect the average adjustment path over time.

We show in Figure 4 that earnings experience a downward adjustment during the first three years following the shock, ultimately stabilizing in both formal and informal sectors. Notably, the decline is significantly more pronounced among informal sector workers, which aligns with the existence of wage rigidities in the formal sector. Regarding employment, Figure 5 shows that the adjustment persists for up to four years

[32] PNAD data are not available for years when the Census is collected (2000 and 2010), so the final year in our sample is 2009.

following the shock before stabilizing. The effects appear nearly symmetric, indicating a gradual reallocation of workers from the formal sector to the informal sector over time. Finally, Figure 6 indicates that adjustments in non-wage benefits persist even after the impact on earnings in the formal sector diminishes.

6 Discussion and Concluding Remarks

In this paper, we examine the labor market effects of weather-induced internal migration in Brazil. Our methodology employs a shift-share instrument approach, which combines the variation in the number of individuals departing their hometowns—prompted by weather shocks—with historical settlement patterns. This allows us to exploit exogenous variation in the number of migrants entering each destination municipality.

Overall, our findings indicate that an exogenous supply shock of low-skill workers leads to a decrease in earnings within the unregulated informal sector. We also observe declines in the formal sector, with near-zero earnings estimates for low-education workers in regions with a high minimum wage bite. In these areas, reductions in formal employment are readily absorbed by informal jobs.

However, our estimates regarding non-wage benefits—such as food, transportation, and health insurance—are viewed as conservative due to potential selection effects, particularly in high minimum wage bite regions and informal markets involving low-education workers. In these contexts, the composition of formal employment shifts, and if we assume positive selection (i.e., workers less likely to receive benefits are more prone to exit formal jobs), the negative impact observed on average across most markets is likely mitigated. In the absence of selection effects, incorporating non-wage adjustments suggests even more significant reductions in formal compensation compared to analyses that overlook these adjustments.

We interpret our results within a framework where formal and informal labor inputs are imperfect substitutes and where non-wage benefits generate predictions that align closely with our empirical findings.

Our findings highlight that non-wage benefits serve as a significant margin of adjustment for firms, particularly in markets where fringe benefits are prevalent, such as in many European and Latin American countries. These benefits enable employers to absorb part of economic shocks, thereby mitigating the impact on employment levels.

References

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

Albert, C., P. Bustos, and J. Ponticelli (2021). The effects of climate change on labor and capital reallocation. Technical report, National Bureau of Economic Research.

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

Altonji, J. G. and D. Card (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market*, pp. 201–234. University of Chicago Press.

ANS (2018). Painel de precificação: Planos de saúde. Agência Nacional de Saúde Suplementar, http://www.ans.gov.br/materiais-publicados/periodicos.

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. (1995). Uma introdução aos salários indiretos no brasil. Anais, Encontro Nacional de Estudos do Trabalho, São Paulo. p. 1297-1322.

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.

Autor, D. H., D. Dorn, G. H. Hanson, and J. Song (2014). Trade adjustment: Worker-level evidence. *The Quarterly Journal of Economics 129*(4), 1799–1860.

Azoulay, P., B. Jones, J. D. Kim, and J. Miranda (2022). Immigration and entrepreneurship in the united states. *American Economic Review: Insights 4*(1), 71–88.

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.

Barbieri, A. F., E. Domingues, B. L. Queiroz, R. M. Ruiz, J. I. Rigotti, J. A. M. Carvalho, and M. F. Resende (2010). Climate change and population migration in brazil's northeast: scenarios for 2025–205. *Population and Environment 31*, 344–370.

Beenstock, M., D. Felsenstein, et al. (2019). *The econometric analysis of non-stationary spatial panel data*. Springer.

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. (2014). *Immigration Economics*. Cambridge: Harvard University Press.

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.

Bosch, M. and J. Esteban-Pretel (2012). Job creation and job destruction in the presence of informal markets. *Journal of Development Economics 98*(2), 270–286.

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.

Busso, M., J. P. Chauvin, and N. Herrera L (2021). Rural-urban migration at high urbanization levels. *Regional Science and Urban Economics 91*, 103658.

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.

Card, D. and T. Lemieux (2001). Can falling supply explain the rising return to college for younger men? a cohort-based analysis. *The Quarterly Journal of Economics 116*(2), 705–746.

Card, D. and G. Peri (2016). Immigration Economics by George J. Borjas: A Review Essay. *Journal of Economic Literature 54*(4), 1333–49.

Cardoso, A. and T. Lage (2006). Labor inspection in Brazil. *Dados 2*, Rio de Janeiro.

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

Carvalho, J. A. and R. A. Garcia (2002). Estimativas decenais e quinquenais de saldos migratorios e taxas líquidas de migração do brasil. Working paper, Cedeplar/UFMG.

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 (2022). Labor courts, job search and employment: Evidence from a labor reform in brazil. Working paper, University of Sao Paulo.

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

Dix-Carneiro, R. and B. K. Kovak (2017). Trade liberalization and regional dynamics. *American Economic Review 107*(10), 2908–2946.

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. and A. Glitz (2015). How do industries and firms respond to changes in local labor supply? *Journal of Labor Economics 33*(3), 711–50.

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.

El-Badaoui, E., E. Strobl, and F. Walsh (2017). Impact of internal migration on labor market outcomes of native males in thailand. *Economic Development and Cultural Change 66*(1), pp. 147–177.

Fields, G. (1975). Rural-urban migration, urban unemployment and underemployment, and job-search activity in ldcs. *Journal of Development Economics 2*(2), 165–187.

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.

Gonzaga, G. (2003). Labor turnover and labor legislation in Brazil. *Economía (Journal of the Latin American and Caribbean Economic Association) 4*(1), 165–222.

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.

IBGE (1990). Divisão do brasil em mesorregiões e microrregiões geográficas. Instituto Brasileiro de Geografia e Estatística, https://www.ibge.gov.br/geociencias/organizacao-do-territorio/divisao-regional/15778-divisoes-regionais-do-brasil.html?edicao=16163&t=downloads. Last accessed on 13/08/2021.

Imbert, C., M. Seror, Y. Zhang, and Y. Zylberberg (2022). Migrants and firms: Evidence from china. *American Economic Review 112*(6), 1885–1914.

Imbert, C. and G. Ulyssea (2024). Rural migrants and urban informality: Evidence from brazil. Working paper, University College London.

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.

Kahn, A. S. and R. T. Campus (1992). Effects of drought on agricultural sector of northeast brazil. Working paper.

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.

Lagos, L. (2020). Labor market institutions and the composition of firm compensation: Evidence from brazilian collective bargaining. Working paper, Columbia University.

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

Lewis, E. (2011). Immigration, skill mix, and capital skill complementarity. *Quarterly Journal of Economics 126*(2), 1029–69.

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

Marinho, C. R. V. (2020). Essays on health economics. PhD Dissertation, Insper.

Mazumdar, D. (1976). The urban informal sector. *World Development 4*(8), 655–679.

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.

Perry, G., W. Maloney, O. Arias, P. Fajnzylber, A. Mason, and J. Saavedra-Chanduvi (2007). *Informality: Exit or Exclusion*. Washington, DC: World Bank.

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.

Ulyssea, G. (2010). Regulation of entry, labor market institutions and the informal sector. *Journal of Development Economics 91*, 87–99.

Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from brazil. *American Economic Review 108*(8), 2015–47.

Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics 12*(1), 525–546.

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

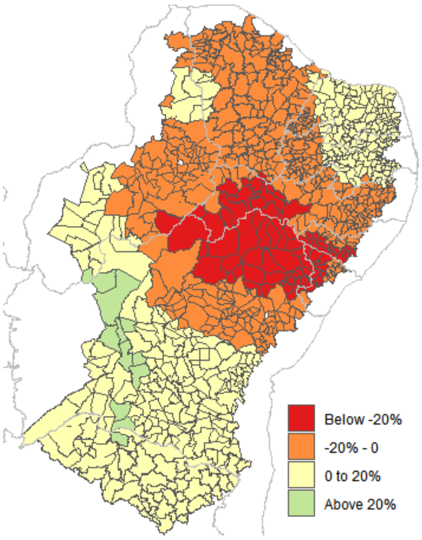
UNDP (2009). Overcoming barriers: Human mobility and development. United Nations Development Program, Human Development Report 2009. http://hdr.undp.org/en/content/human-development-report-2009. Last accessed on 03/09/2021.

Vicente-Serrano, S. M., S. Beguería, and J. I. López-Moreno (2010). A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of Climate 23*(7), 1696–1718.

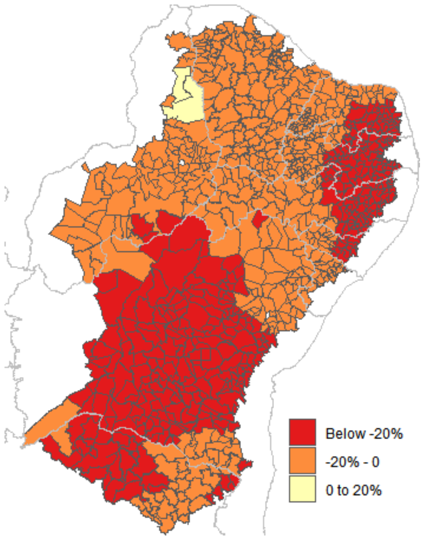
Wang, X., A. S. Auler, R. L. Edwards, H. Cheng, P. S. Cristalli, P. L. Smart, D. A. Richards, and C.-C. Shen (2004). Wet periods in northeastern brazil over the past 210 kyr linked to distant climate anomalies. *Nature 432*, 740–743.

Figures and tables

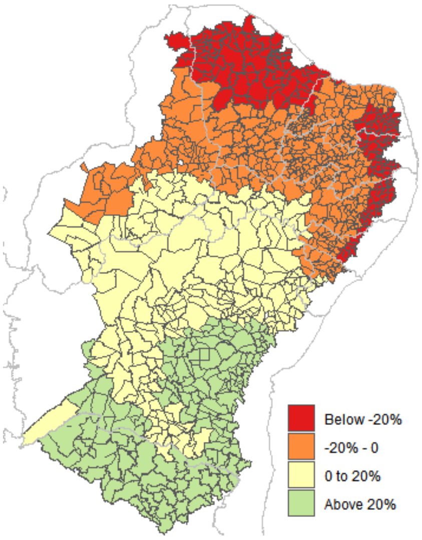
Figure 1: Precipitation levels in the Semiarid region for selected years



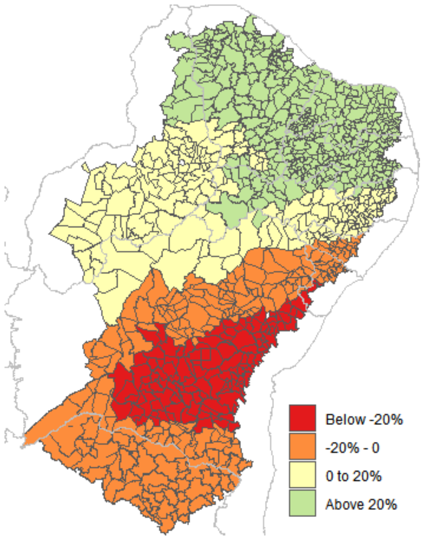
(a) 1997



(b) 2001



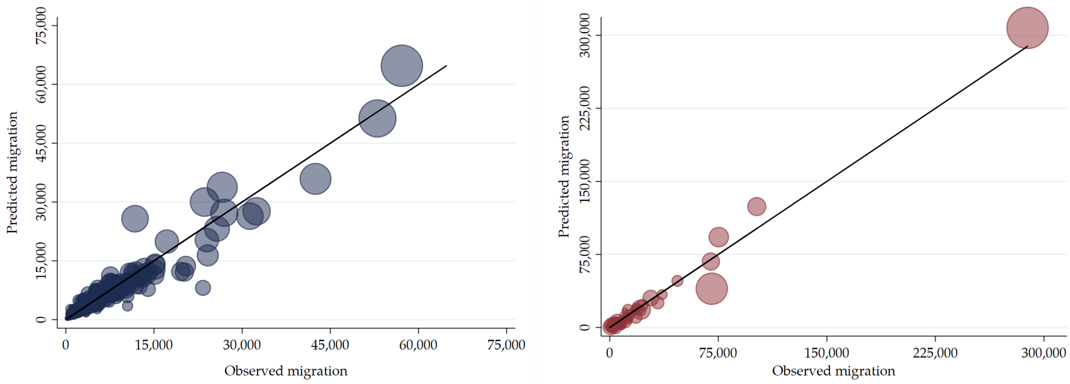
(c) 2005



(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).

Figure 2: Observed vs predicted migration

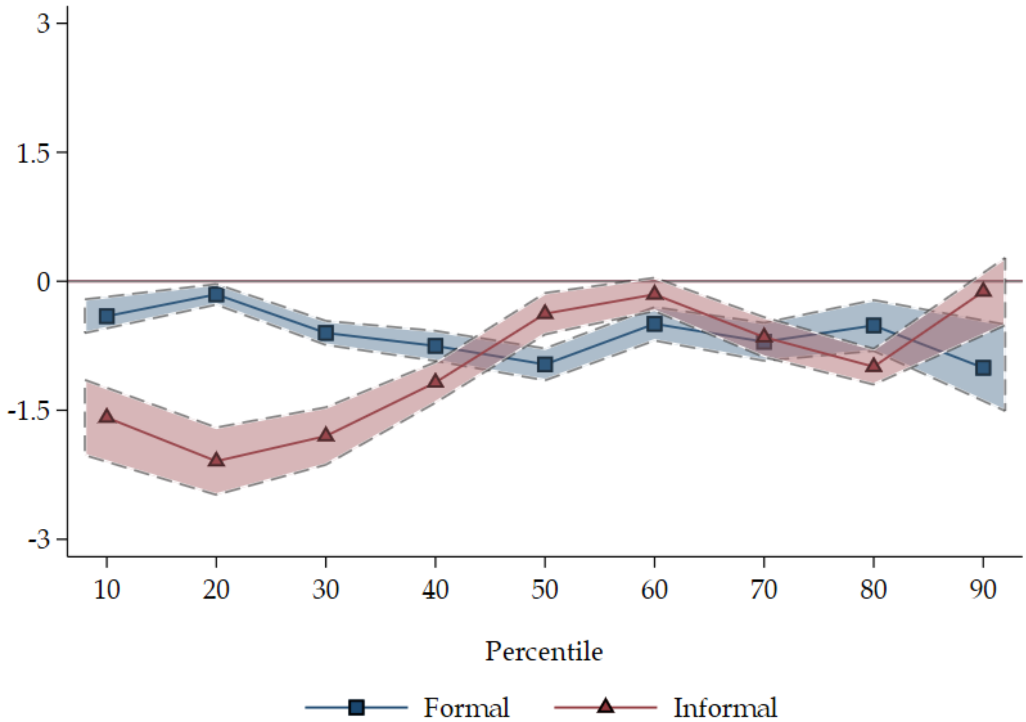


(a) Migration outflow

(b) Migration inflow

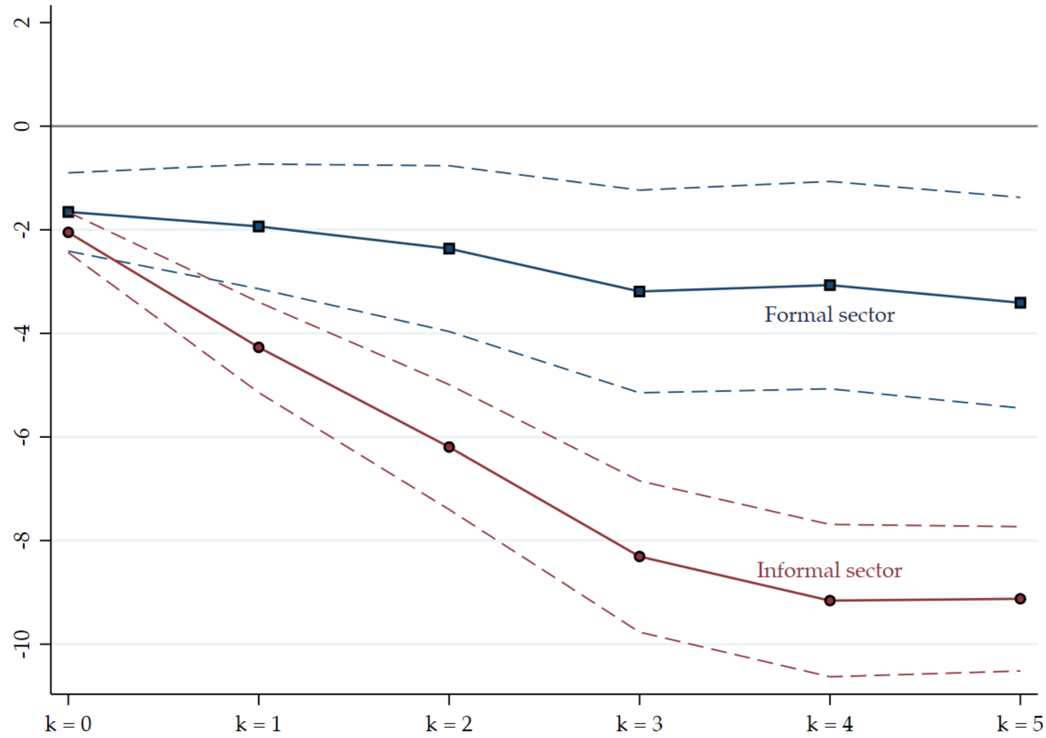
Notes: This figure presents the relationship between the accumulated predicted and observed migration flows across Brazilian municipalities from 1996 to 2010. Panel (a) shows the number of migrants leaving the Semiarid region to non-Semiarid municipalities. Panel (b) shows the number of incoming Semiarid migrants for destination municipalities. The circle size represents the municipality's total population in 1991. *Data source*: Census microdata (IBGE).

Figure 3: Effects of predicted migration along the earnings distribution



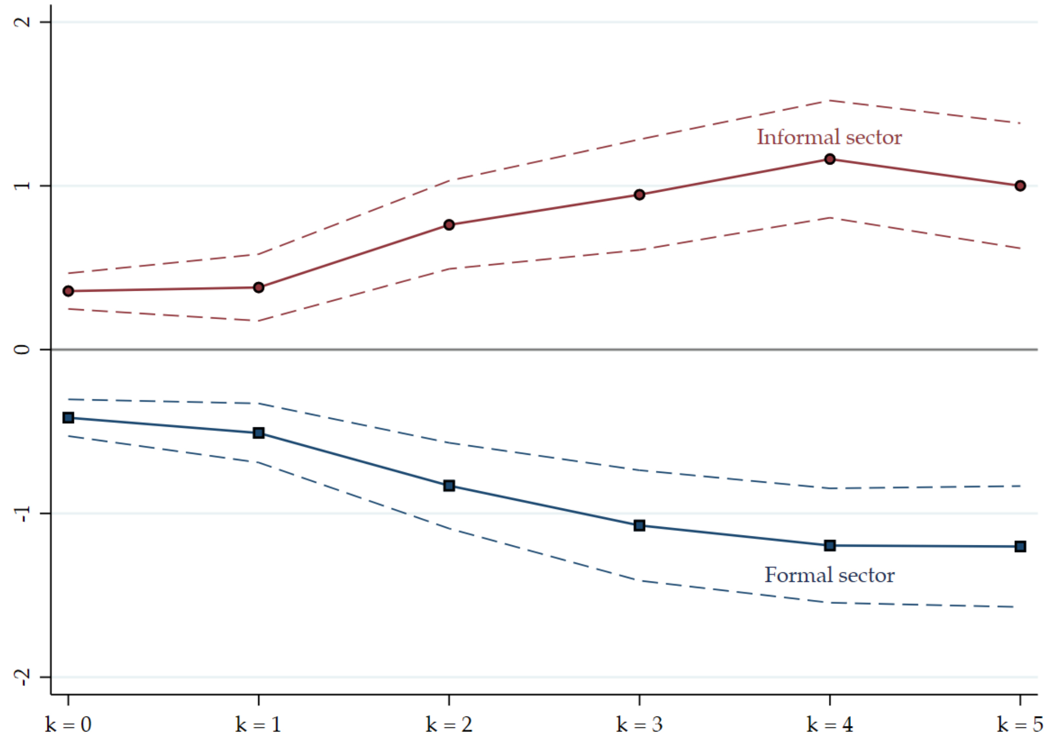
Notes: This figure plots origin-level SSIV coefficients of the incoming migration rate in the changes of the cutoffs of the earnings distribution, by sector. The informal sector also includes self-employed workers. Controls include time dummies and destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water).

Figure 4: **Dynamic effects of internal migration on earnings**



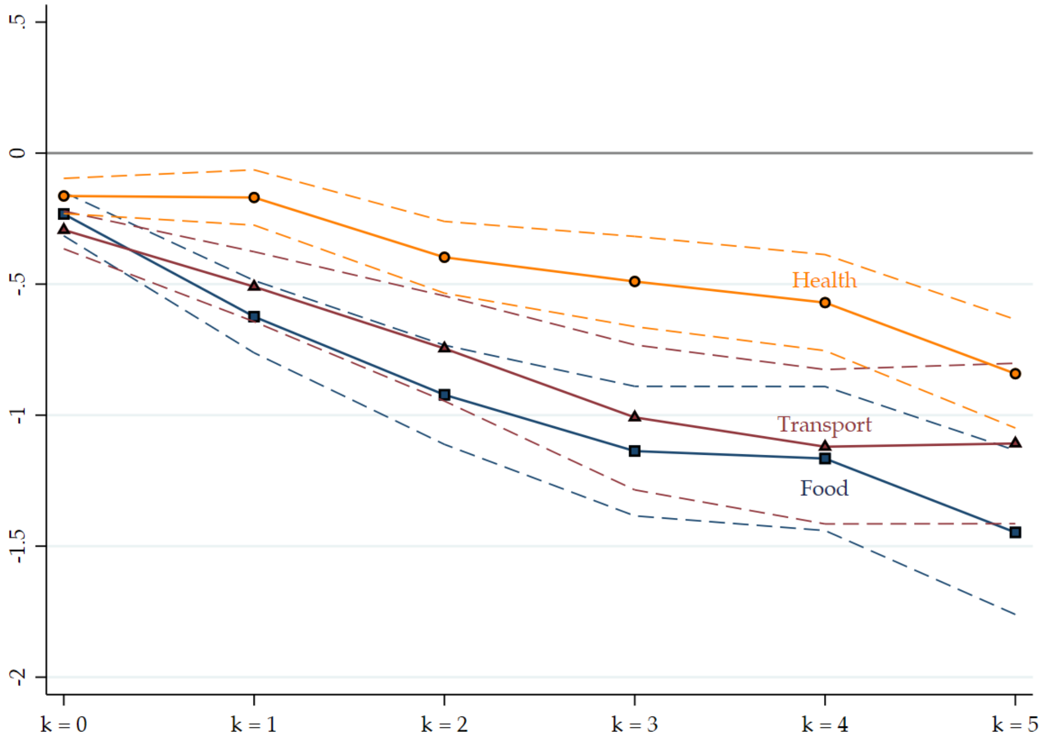
Notes: This figure plots origin-level SSIV coefficients on the change in log earnings, by sector, up to five years after the incoming migration shock. Each point corresponds to a specific value of θ_t from equation 12. The informal sector includes self-employed workers. Controls include time dummies and destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water). Dashed lines represent the 95% confidence interval.

Figure 5: **Dynamic effects of internal migration on employment**



Notes: This figure plots origin-level SSIV coefficients on the change in the employment rate, by sector, up to five years after the incoming migration shock. Each point corresponds to a specific value of θ_t from equation 12. The informal sector includes self-employed workers. Controls include time dummies and destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water). Dashed lines represent the 95% confidence interval.

Figure 6: **Dynamic effects of internal migration on non-wage compensation**



Notes: This figure plots origin-level SSIV coefficients on the change in the proportion of working-age natives who receive health insurance, food, or transport subsidies up to five years after the incoming migration shock. Each point corresponds to a specific value of θ_t from equation 12. Controls include time dummies and destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water). Dashed lines represent the 95% confidence interval.

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

Panel A: Origin (Semiarid)	Mean	Std. Dev.	Min	Max	Obs
Annual Rainfall	782.33	248.71	165.49	1,953.17	14,400
Rainfall shock	-0.02	0.19	-0.73	0.48	14,400
Annual Temperature	25.54	1.39	21.42	28.93	14,400
Temperature shock	0.01	0.01	-0.01	0.05	14,400
Out-migration	214.16	323.66	0.00	5,773	14,400
Out-migration rate (p.p.)	1.05	0.62	0.00	7.22	14,400
Population	21,377	30,386	1,265	480,949	14,400
Panel B: Destination (Non-Semiarid)	**Mean**	**Std. Dev.**	**Min**	**Max**	**Obs**
Annual Rainfall	1,610.44	401.69	660.63	3,618.55	8,190
Rainfall shock	0.04	0.16	-0.77	0.65	8,190
Annual Temperature	23.15	2.82	15.82	28.77	8,190
Temperature shock	0.03	0.02	-0.03	0.08	8,190
In-migration	146.69	896.95	0.00	25,423	8,190
In-migration rate (p.p.)	0.30	1.00	0.00	27.95	8,190
Native population	51,963	231,296	290	4,771,965	8,190

Notes: Rainfall is measured in mm. Temperature is measured in degrees Celsius. Migration outflow (inflow) rate is the share of migrants over the local (native) population.

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

	Mean	Std. Dev.	Min	Max	Obs
Individual Characteristics					
Female	51.08	3.65	0.00	72.72	8,190
Black	6.23	5.98	0.00	53.85	8,190
White	52.81	25.46	0.00	100.00	8,190
Age	37.45	1.96	30.15	55.00	8,190
Years of schooling	6.58	1.78	0.00	13.52	8,190
Less than elementary	55.49	17.44	3.19	100.00	8,190
Employment					
Any employment	62.72	7.95	10.00	100.00	8,190
Formal	31.34	11.85	0.00	100.00	8,190
Informal	31.38	9.05	0.00	81.80	8,190
Unemployed	13.05	7.73	0.00	80.00	8,190
Out of labor force	24.23	7.08	0.00	58.14	8,190
Earnings					
Overall	637.89	348.99	60.88	3,582.08	8,190
Formal	788.22	439.49	58.67	15,167.10	8,174
Informal	491.28	284.28	20.00	4,941.10	8,184
Non-wage benefits					
Food	12.11	8.80	0.00	58.34	8,190
Transport	11.50	9.49	0.00	53.67	8,190
Health	6.63	6.32	0.00	100.00	8,190

Notes: Each observation is a destination municipality-year cell. Data come from PNAD in 1997-1999 and 2001-2009. Earnings are measured in R$ of 2012. The informal sector also includes self-employed workers. Non-wage benefits are percentages from the working-age native population.

Table 3: **Comparative characteristics: natives vs migrants, by education**

	Migrants	Low-ed. natives	High-ed. natives
Individual Characteristics			
Age	32.63	39.30	31.24
	(6.88)	(3.44)	(2.48)
Female	45.96	47.36	53.50
	(25.01)	(2.77)	(5.82)
Non-white	59.40	47.29	36.78
	(31.27)	(26.82)	(26.96)
Married	66.92	70.03	54.56
	(28.01)	(6.02)	(8.01)
Number of children	2.80	3.23	1.83
	(1.55)	(0.51)	(0.51)
Labor market conditions			
Share of employed	62.13	56.71	69.18
	(26.91)	(9.73)	(9.31)
Share of formal	29.01	18.68	35.83
	(28.32)	(11.04)	(12.89)
Earnings	904.02	745.32	1247.20
	(1068.43)	(290.59)	(513.77)
Work less than 40h/week	9.13	9.72	16.89
	(16.72)	(4.61)	(7.47)

Notes: This table shows comparisons of migrants from the Semiarid region and native individuals by education level, using data from the Census (1991, 2000, and 2010) only for the municipalities also covered by *PNAD*. Each observation is an individual-year cell. Low-educated individuals are those with incomplete elementary schooling. Earnings are measured in R$ of 2012. The share of formal employment, earnings, and worked hours are calculated only for employed individuals.

Table 4: **Migration outflows induced by weather shocks**

	(1)	(2)	(3)	(4)	(5)	(6)
Rainfall$_{t-1}$	-0.099***	-0.092***	-0.093***	-0.092***	-0.093***	-0.096***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.030)
Rainfall$_{t-2}$			0.008	0.022		
			(0.030)	(0.031)		
Rainfall$_{t-3}$				0.059**		
				(0.028)		
Rainfall$_{t}$					-0.047	
					(0.031)	
Rainfall$_{t+1}$						-0.059
						(0.036)
Observations	14,400	14,400	14,400	14,400	14,400	14,400
Municipalities	960	960	960	960	960	960
R-Squared	0.461	0.465	0.465	0.466	0.465	0.466
Time dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Temperature shocks	✓	✓	✓	✓	✓	✓
Demographics		✓	✓	✓	✓	✓

Notes: Each observation is an origin municipality-year cell. The 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 the log-deviation from the historical average (for the 6 months in the crop growing season). All specifications include controls for temperature shocks, municipality, and year fixed effects. Columns (2)-(6) also control for municipality-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) interacted with time dummies. Standard errors are clustered at the grid level. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

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

	Employment			Earnings		
	Overall	Formal	Informal	Overall	Formal	Informal
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
In-migration rate	-0.084	-0.032	-0.052	0.146	0.296	-0.231
	(0.084)	(0.089)	(0.075)	(0.405)	(0.438)	(0.424)
Observations	8,190	8,190	8,190	8,190	8,162	8,179
Municipalities	684	684	684	684	684	684
Panel B: SSIV						
In-migration rate	-0.017	-0.125***	0.108***	-0.867***	-0.591***	-0.744***
	(0.034)	(0.037)	(0.034)	(0.197)	(0.198)	(0.123)
Observations	11,460	11,460	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955	955	955

Notes: This table shows destination-level OLS estimates and origin-level SSIV coefficients by sector. Each observation is a municipality-year cell. The informal sector also includes self-employed workers. In Columns (1)-(3), the dependent variable is the change in the employment rate, while in Columns (4)-(6), it is the change in log earnings. All regressions control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) and are weighted by the working-age native population in 1991. Standard errors clustered at the municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table 6: **Effects of migration on non-wage benefits**

	Food	Transport	Health
	(1)	(2)	(3)
<u>Panel A: OLS</u>			
In-migration rate	-0.063	0.018	-0.049
	(0.088)	(0.077)	(0.076)
Observations	8,190	8,190	8,190
Municipalities	684	684	684
<u>Panel B: SSIV</u>			
In-migration rate	-0.222***	-0.106***	-0.109***
	(0.034)	(0.024)	(0.024)
Observations	11,460	11,460	11,460
Municipalities	955	955	955

Notes: This table shows destination-level OLS estimates and origin-level SSIV coefficients of changes in the proportion of working-age native workers receiving non-wage benefits. Each observation is a municipality-year cell. All regressions control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) and are weighted by the working-age native population in 1991. Standard errors clustered at the municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Internal Migration and Labor Market Adjustments in the Presence of Non-wage Compensation

Raphael Corbi, Tiago Ferraz, and Renata Narita

ONLINE APPENDIX

Appendix A Proofs

In this section we provide the proofs of predictions for the effects of migration inflows from the Semiarid regions on the informal and formal sectors at the destination, as discussed in Section 3.

A.1 Prediction (i)

$$\frac{\partial W_i}{\partial L_i} < 0.$$

By differentiating (4) with respect to L_i, we obtain:

$$\frac{\partial W_i}{\partial L_i} = \alpha\theta_i K^{1-\alpha}(\theta_f L_f^\nu + \theta_i L_i^\nu)^{\frac{\alpha-2\nu}{\nu}} L_i^{\nu-2}\left[(\alpha-\nu)\theta_i L_i^\nu - (1-\nu)(\theta_f L_f^\nu + \theta_i L_i^\nu)\right] < 0$$

$$(\alpha-\nu)\theta_i L_i^\nu < (1-\nu)(\theta_f L_f^\nu + \theta_i L_i^\nu)$$

which is always true since $0 < \alpha < 1$, and $\theta_i L_i^\nu < \theta_f L_f^\nu + \theta_i L_i^\nu$.

A.2 Prediction (ii)

$$\alpha - \nu < 0 \Rightarrow \frac{\partial B}{\partial L_i} < 0.$$

Now, differentiating (3) with respect to L_i,

$$\frac{\partial W_f}{\partial L_i} = \frac{\alpha(\alpha-\nu)\theta_f\theta_i K^{1-\alpha}(\theta_f L_f^\nu + \theta_i L_i^\nu)^{\frac{\alpha-2\nu}{\nu}}}{(L_f L_i)^{1-\nu}}$$

which is necessarily negative if $\alpha - \nu < 0$. In the presence of binding minimum wages, $\frac{\partial B}{\partial L_i} < 0$.

Moreover, the adverse effects on earnings or total compensation are increasing in the informality share, θ_i, as the elasticities show:

$$\frac{\partial W_i}{\partial L_i}\frac{L_i}{W_i} = \frac{(\alpha-\nu)\theta_i L_i^\nu}{(\theta_f L_f^\nu + \theta_i L_i^\nu)} - (1-\nu)$$

$$\frac{\partial W_f}{\partial L_i}\frac{L_i}{W_f} = \frac{(\alpha-\nu)\theta_i L_i^\nu}{(\theta_f L_f^\nu + \theta_i L_i^\nu)}$$

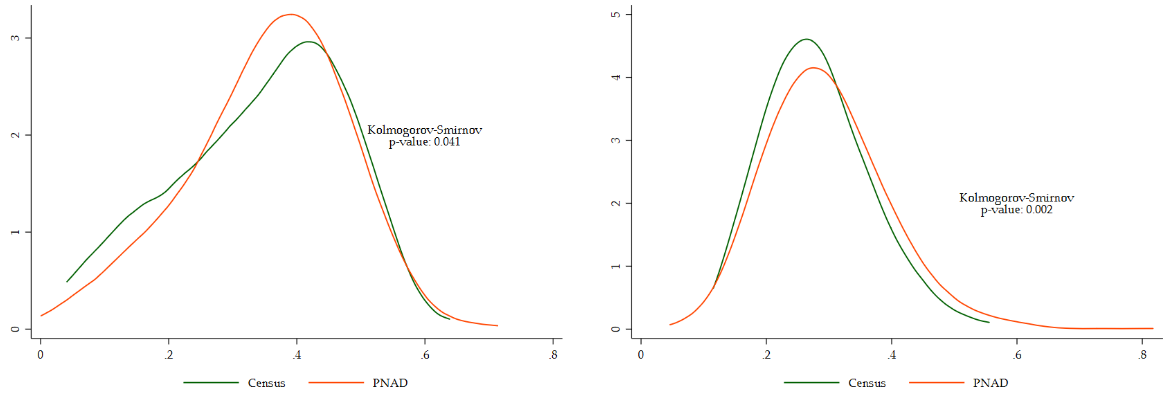
Appendix B Comparing PNAD vs Census data

Table B1: **Summary statistics: Native individuals in destination municipalities (PNAD vs Census)**

	Census	PNAD
Age	36.76	37.47
	(1.64)	(2.06)
Female	0.50	0.51
	(0.02)	(0.04)
Non-white	47.25	46.56
	(25.67)	(25.88)
Married	64.23	78.18
	(4.93)	(7.61)
Number of children	2.84	1.18
	(0.56)	(0.29)
Less than elementary	57.69	58.01
	(17.44)	(18.46)

Notes: Each observation is a destination municipality-year cell. PNAD data from 1996, 2001, and 2009 and Census from 1991, 2000, and 2010, only for the municipalities also covered by *PNAD*. Standard deviations are in parentheses.

Figure B1: Distribution of employment rates, PNAD vs. Census

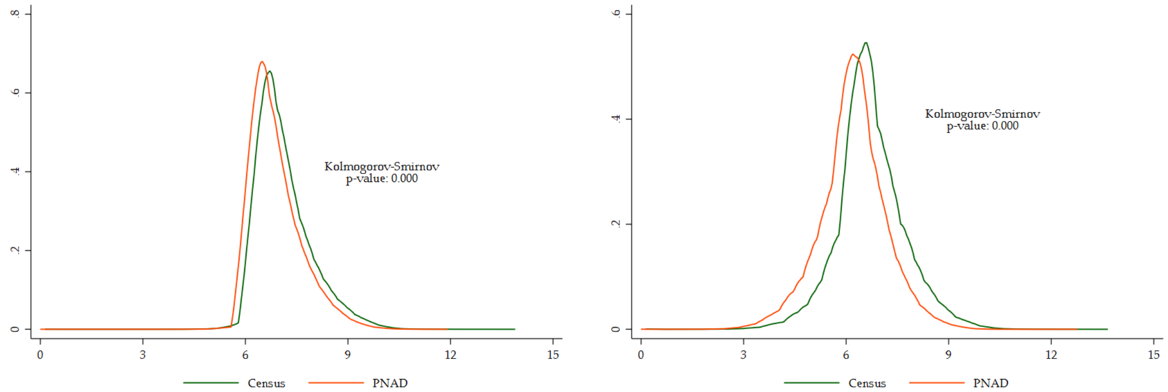


(a) Formal sector

(b) Informal sector

Notes: Municipality-level employment rate in the formal and informal sector from PNAD (2009) and Census (2010).

Figure B2: Distribution of log earnings, PNAD vs. Census

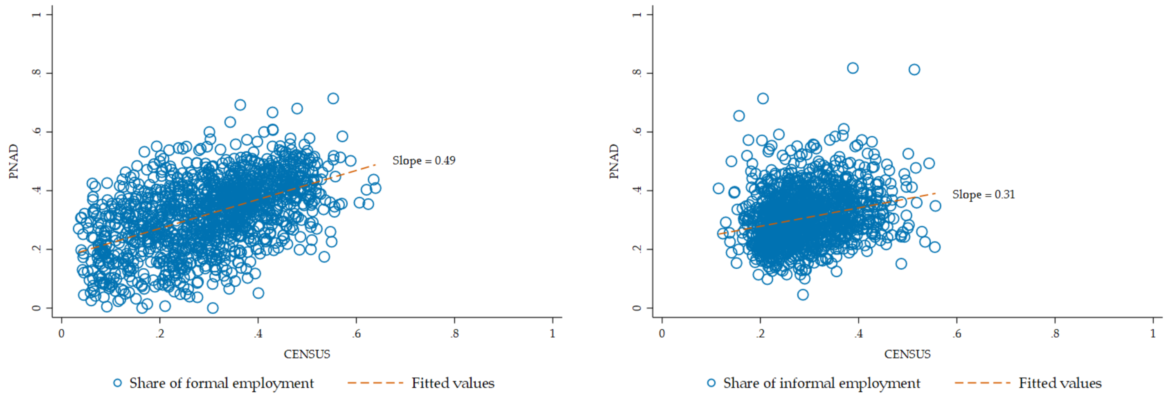


(a) Formal sector

(b) Informal sector

Notes: Individual-level log earnings in the formal and informal sector from PNAD (2009) and Census (2010).

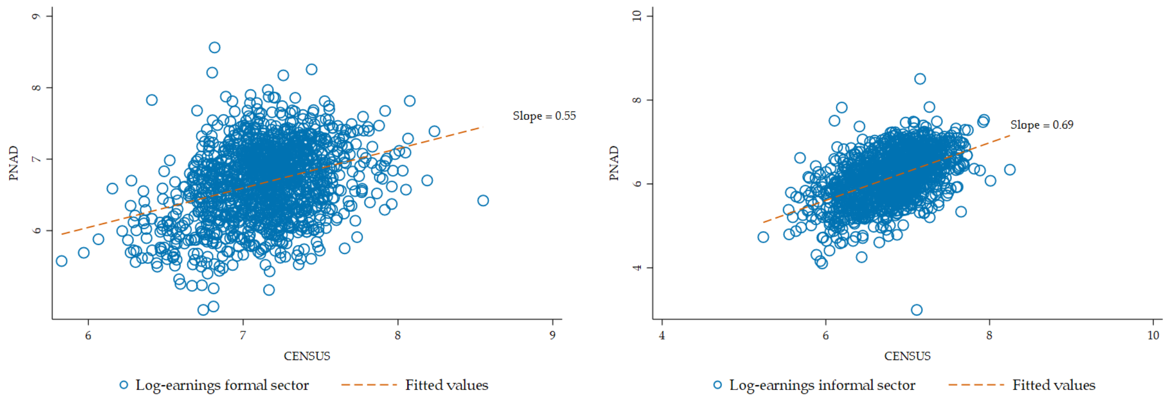
Figure B3: Relationship between PNAD and Census employment rates



(a) Formal sector (b) Informal sector

Notes: Relationship between municipality-level employment rate in the formal and informal sector from PNAD (1996, 2001, and 2009) and Census (1991, 2000, and 2010).

Figure B4: Relationship between PNAD and Census log earnings



(a) Formal sector (b) Informal sector

Notes: Relationship between municipality-level log earnings in the formal and informal sector from PNAD (1996, 2001, and 2009) and Census (1991, 2000, and 2010).

Appendix C Migrant flows from the Semiarid region

In this section, we discuss our measure of migration between cities in more detail and how we structure a yearly panel dataset from the 2000 and 2010 censuses.

C.1 Migration from the Semiarid region

In every round of the Census, two questions allow us to track the migrants and establish their municipalities of origin and destination, as well as the year when they moved.

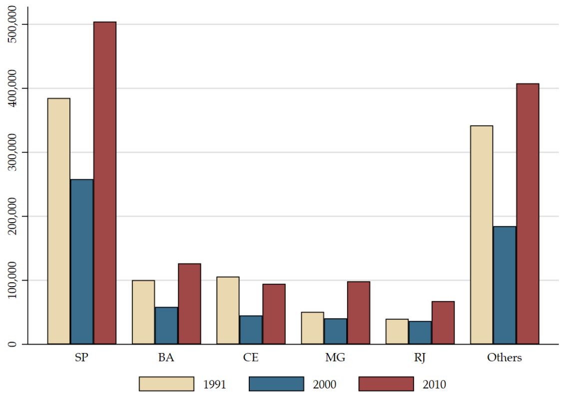
First, in the 2010 Census, respondents were asked how many years they had lived in the current municipality (from one up to ten). This variable allows us to calculate the year the individual migrated. We consider a migrant an individual who moved to the current municipality in the previous ten years. In the 2000 Census, interviewees were asked about the municipality where they were living five years ago instead of the last place where they lived, so we can only identify migrants who came as far as 1996. This is not a major concern in our analysis as 1996 is the first year for which *PNAD* data - the source from which we draw labor market outcomes information - is available.

Second, they were asked what municipality they lived in before. Thus, if an individual has migrated from an origin municipality in the Semiarid region, she will be counted as a Semiarid migrant. A limitation is that we can only track one origin location for each person, probably the last municipality where she lived.

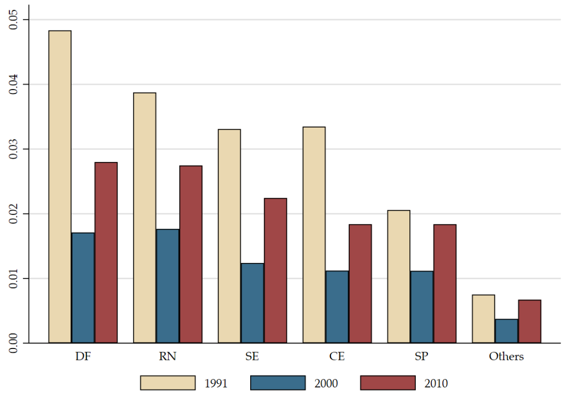
The Semiarid region has always been an important source of migrants for the rest of the country. Figure C1 shows that these migrants tend to be historically concentrated in some states. In the last four decades, São Paulo alone harbored over 30 percent of the people arriving from the Semiarid. However, in relative terms, incoming migrants represented a population increase of above 2% for the top 10 receiving states.

Table 3 compares migrants to low and high-education natives. Migrants earn slightly less than less-educated natives and also have a similar likelihood of working part-time, but they are more likely to be in the formal sector when compared to low-education natives. On the other hand, high-education natives are more likely to work in the formal sector and have considerably higher pay than the migrants from the Semiarid region. Table C1 shows that top occupations for migrants (e.g., typically bricklayer for men, domestic worker for women) are also top occupations for low-education natives, but not for the skilled. Also, the same five industries that concentrate over 80% of working migrants also employ a similar share of low-education workers (see Table C2).

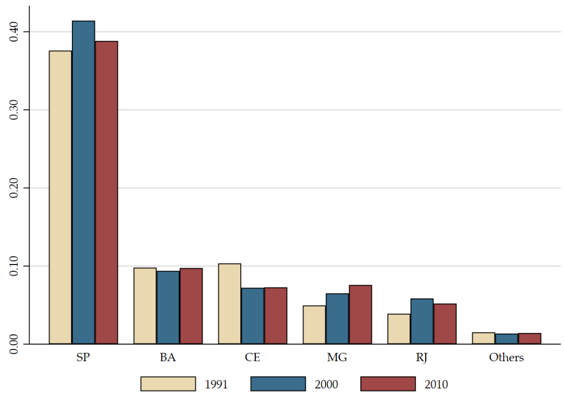
Figure C1: Top destinations for migrants from the Semiarid region



(a) Absolute number of Semiarid's migrants



(b) Semiarid's migrants as a fraction of total population



(c) Semiarid's migrants as a share of total migration

Notes: This figure presents the main destination states chosen by migrants from the Semiarid region. Panel (a) shows the absolute number of migrants leaving the Semiarid region for non-Semiarid areas. Panel (b) presents the same inflow measured as a fraction of the total population in the state, while in Panel (c), that number is measured as a share of the total number of migrants in each state. In each panel, states are ranked by the respective average across years. *Data source*: Census microdata (IBGE).

Table C1: **Main occupations for employed people: Migrants vs Natives**

	Position	Occupation	Share of employment	Cumulative
Migrants	1	Domestic worker	12.5	12.5
	2	Bricklayer	11.0	23.5
	3	Retail worker	10.4	33.9
	4	Rural worker	8.0	41.9
	5	Low-level office	4.2	46.1
	6	Military	4.0	50.1
	7	Non-specified occupations	2.5	52.6
	8	Manager	2.3	54.9
	9	Tailor	2.2	57.1
	10	Janitor	2.1	59.2
Low-ed. Natives	1	Rural worker	22.3	22.3
	2	Bricklayer	10.6	33.0
	3	Domestic worker	10.1	43.1
	4	Retail worker	7.8	50.9
	5	Tailor	2.9	53.7
	6	Military	2.7	56.4
	7	Manager	1.9	58.3
	8	Driver	1.8	60.1
	9	Janitor	1.7	61.8
	10	Low-level office	1.7	63.5
High-ed. Natives	1	Retail worker	10.5	10.5
	2	Low-level office	10.3	20.8
	3	Manager	7.2	28.0
	4	Teacher	6.2	34.2
	5	Military	5.6	39.8
	6	Bricklayer	3.2	43.0
	7	Rural worker	3.2	46.2
	8	Domestic worker	3.0	49.2
	9	Tailor	1.8	51.0
	10	Non-specified occupations	0.7	51.7

Notes: This table presents the top ten occupations for workers in the destination municipalities, using data from the Census (1991, 2000, and 2010).

Table C2: **Main industries for employed people: Migrants vs Natives**

	Position	Occupation	Share of employment	Cumulative
Migrants	1	Other Services	20.2	20.2
	2	Manufacturing	14.4	34.6
	3	Retail	12.8	47.4
	4	Transport/Communication	12.8	60.2
	5	Agriculture/Mining	8.9	69.1
	6	Hospitality	6.4	75.5
	7	Construction	6.3	81.8
	8	Health/Education	3.1	84.9
	9	Public Sector	0.9	85.8
Low-ed. Natives	1	Agriculture/Mining	24.0	24.0
	2	Other Services	16.0	40.0
	3	Manufacturing	13.0	53.0
	4	Retail	11.1	64.1
	5	Transport/Communication	11.1	75.1
	6	Construction	4.8	79.9
	7	Hospitality	4.2	84.1
	8	Health/Education	2.1	86.2
	9	Public Sector	1.0	87.2
High-ed. Natives	1	Other Services	32.0	32.0
	2	Transport/Communication	17.4	49.4
	3	Manufacturing	12.4	61.8
	4	Health/Education	8.3	70.1
	5	Retail	6.5	76.6
	6	Agriculture/Mining	4.0	80.6
	7	Hospitality	3.6	84.2
	8	Construction	2.9	87.0
	9	Public Sector	1.6	88.6

Notes: This table presents the top ten industries for workers in the destination municipalities, using data from the Census (1991, 2000, and 2010).

Appendix D Additional results

Table D1: **Effects of migration on unemployment and participation**

	OLS		SSIV	
	Unemployed	Inactive	Unemployed	Inactive
	(1)	(2)	(3)	(4)
In-migration rate	0.120	-0.036	0.094***	-0.076***
	(0.077)	(0.067)	(0.020)	(0.029)
Observations	8,190	8,190	11,460	11,460
Municipalities	684	684	955	955

Notes: This table shows destination-level OLS and origin-level SSIV coefficients on changes in unemployment and inactivity rates. Each observation is an origin municipality-year cell. Columns (1)-(2) show OLS estimates, while Columns (3)-(4) present the SSIV coefficients. All regressions are weighted by the working-age native population in 1991, include time dummies and control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water). Standard errors clustered at the respective municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table D2: **Effects of migration on labor market outcomes, by status in the household**

	Employment	Formal	Informal	Unemployment	Inactivity
	(1)	(2)	(3)	(4)	(5)
Panel A: Head					
In-migration rate	-0.028*	-0.112***	0.085***	0.018*	0.032**
	(0.015)	(0.021)	(0.019)	(0.011)	(0.013)
Observations	11,460	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955	955
Panel B: Non-Head					
In-migration rate	0.010	-0.013	0.023	0.076***	-0.108***
	(0.024)	(0.019)	(0.018)	(0.014)	(0.019)
Observations	11,460	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955	955

Notes: This table shows origin-level SSIV coefficients on changes in employment (by sector), unemployment, and inactivity rates. Each observation is an origin municipality-year cell. The informal sector also includes self-employed workers. 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 are weighted by the working-age native population in 1991, include time dummies and control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water). Standard errors clustered at the origin municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table D3: Effects of predicted in-migration on employer-provided health insurance

	(1)	(2)	(3)	(4)	(5)
Predicted inflow	-0.015**	0.003	-0.004	-0.010**	-0.048**
	(0.007)	(0.002)	(0.003)	(0.005)	(0.022)
Mean of dep. var.	0.0158	0.0131	0.0448	0.0609	0.0758
Observations	4462346	4167842	138572	142100	13832
Municipalities	682	679	482	608	280
Firms	318739	297703	9898	10150	988
Time dummies	✓	✓	✓	✓	✓
Firm dummies	✓	✓	✓	✓	✓
Firm size weighted	All firms	1 to 50	51 to 100	101 to 1000	More than 1000

Notes: This table shows the reduced form coefficients of changes in the probability of a firm offering health insurance to its employees on the predicted inflow of migrants from the Semiarid region. Each observation is a firm-year cell. The dependent variable is the difference in the dummy variable that is equal to one for every year greater than or equal to the year when the health insurance contract was signed. The regressor is the predicted number of migrants from the Semiarid region in each destination municipality (excluding those in the Semiarid region), measured as a fraction of the native working-age population in 1991. Our sample comprises a balanced panel of all firms included in RAIS during the period. All the regressions are weighted by the number of employees in the firm in 1996. Standard errors clustered at the municipality level in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Appendix E Robustness

Table E1: **Shock balance test**

	(1)	(2)	(3)
Panel A: Lagged Δ log earnings	Overall	Formal	Informal
In-migration rate	-1.103***	-0.485**	-1.290***
	(0.187)	(0.197)	(0.125)
Panel B: Lagged Δ employment rate	Overall	Formal	Informal
In-migration rate	0.136***	-0.169***	0.305***
	(0.025)	(0.035)	(0.033)
Panel C: Lagged Δ benefits	Food	Transport	Health
In-migration rate	-0.306***	-0.438***	0.012
	(0.078)	(0.068)	(0.045)
Observations	10,505	10,505	10,505
Municipalities	955	955	955

Notes: This table shows origin-level SSIV coefficients of the lagged change in outcomes against the predicted in-migration rate. Each observation is a municipality-year cell. The informal sector also includes self-employed workers. All regressions control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) and are weighted by the working-age native population in 1991. Standard errors clustered at the municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table E2: **Effects of migration on employment and earnings - controlling for pre-trends**

	Employment			Earnings		
	Overall	Formal	Informal	Overall	Formal	Informal
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: SSIV - Without pre-trends						
In-migration rate	-0.017	-0.125***	0.108***	-0.867***	-0.591***	-0.744***
	(0.034)	(0.037)	(0.034)	(0.197)	(0.198)	(0.123)
Observations	11,460	11,460	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955	955	955
Panel B: SSIV - Controlling for pre-trends						
In-migration rate	-0.045	-0.147***	0.111**	-1.141***	-0.594**	-1.627***
	(0.037)	(0.041)	(0.048)	(0.266)	(0.251)	(0.205)
Observations	10,505	10,505	10,505	10,505	10,505	10,505
Municipalities	955	955	955	955	955	955

Notes: This table shows origin-level SSIV coefficients by sector. Each observation is a municipality-year cell. Panel A shows results from our baseline specification. Panel B uses annual data from 1998-1999 and 2001-2009, i.e., without the first year of our main sample, as we add the lagged outcome as an additional control. This specification uses annual data from 1998-1999 and 2001-2009. In Columns (1)-(3), the dependent variable is the change in the employment rate, while in Columns (4)-(6), it is the change in log earnings. The informal sector also includes self-employed workers. All regressions control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) and are weighted by the working-age native population in 1991. Standard errors clustered at the municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table E3: **Effects of migration on non-wage benefits - controlling for pre-trends**

	Food	Transport	Health
	(1)	(2)	(3)
Panel A: SSIV - Without pre-trends			
In-migration rate	-0.222***	-0.106***	-0.109***
	(0.034)	(0.024)	(0.024)
Observations	11,460	11,460	11,460
Municipalities	955	955	955
Panel B: SSIV - Controlling for pre-trends			
In-migration rate	-0.357***	-0.171***	-0.117***
	(0.036)	(0.028)	(0.025)
Observations	10,505	10,505	10,505
Municipalities	955	955	955

Notes: This table shows origin-level SSIV coefficients of changes in the proportion of working-age native workers receiving non-wage benefits. Each observation is a municipality-year cell. Panel A shows results from our baseline specification. Panel B uses annual data from 1998-1999 and 2001-2009, i.e., without the first year of our main sample, as we add the lagged outcome as an additional control. In Columns (1)-(3), the dependent variable is the change in the employment rate, while in Columns (4)-(6), it is the change in log earnings. The informal sector also includes self-employed workers. All regressions control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) and are weighted by the working-age native population in 1991. Standard errors clustered at the municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table E4: **Correlation between predicted migration from the Semiarid and other regions**

	(1)	(2)
	Non-Semiarid migration	
Predicted inflow	0.080	0.080
	(2.847)	(2.847)
Observations	8,190	8,190
Municipalities	684	684
Time dummies	✓	✓
Municipality dummies	✓	✓
Baseline × time		✓

Notes: This table shows destination-level regression coefficients of the observed inflow of migrants from other regions on the predicted number of migrants from the Semiarid, both measured as a fraction of the working-age native population in 1991. Each observation is a destination municipality-year cell. All regressions are weighted by the working-age native population in 1991 and include municipality and time dummies. Column (2) controls for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) interacted with time dummies. Standard errors clustered at the destination municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table E5: **Effects of migration on earnings, by industry**

	Agriculture	Manufacturing	Services
	(1)	(2)	(3)
In-migration rate	-0.470	-0.251	-0.830***
	(0.375)	(0.188)	(0.187)
Observations	11,447	11,460	11,460
Municipalities	955	955	955

Notes: This table shows origin-level SSIV coefficients on changes in log earnings by industry. Each observation is an origin municipality-year cell. All regressions are weighted by the working-age native population in 1991, include time dummies and control for destination-level 1991 characteristics (log of the working-age native population; shares of population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) . Standard errors clustered at the origin municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Appendix F Heterogeneity

Table F1: **Effects of migration on employment and earnings, by level of MW bite**

	High MW Bite			Low MW Bite		
	Overall	Formal	Informal	Overall	Formal	Informal
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Δ log earnings						
In-migration rate	-1.293***	-0.695	-1.604***	-0.653***	-0.717***	-0.177
	(0.445)	(0.497)	(0.234)	(0.176)	(0.159)	(0.135)
Panel B: Δ employment rate						
In-migration rate	0.001	-0.302***	0.304***	-0.025	-0.078***	0.053
	(0.064)	(0.098)	(0.061)	(0.035)	(0.024)	(0.046)
Observations	9,840	9,840	9,840	11,460	11,460	11,460
Municipalities	820	820	820	955	955	955

Notes: This table shows origin-level SSIV coefficients on changes in employment rate and log earnings for each sector and by level of minimum wage bite. Each observation is an origin municipality-year cell. The informal sector also includes self-employed workers. All regressions include time dummies, control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) , and are weighted by the working-age native population in 1991. Standard errors clustered at the origin municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table F2: **Effects of migration on non-wage benefits, by level of MW bite**

	High MW Bite			Low MW Bite		
	Food	Transport	Health	Food	Transport	Health
	(1)	(2)	(3)	(4)	(5)	(6)
In-migration rate	-0.096*	-0.004	0.032	-0.238***	-0.127***	-0.170***
	(0.055)	(0.041)	(0.028)	(0.033)	(0.028)	(0.023)
Observations	9,840	9,840	9,840	11,460	11,460	11,460
Municipalities	820	820	820	955	955	955

Notes: This table shows origin-level SSIV coefficients on changes in the proportions of working-age native population who receive health insurance, food, or transport subsidies by level of minimum wage bite. Each observation is an origin municipality-year cell. All regressions include time dummies, control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) , and are weighted by the working-age native population in 1991. Standard errors clustered at the origin municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table F3: **Effects of migration on employment and earnings, by level of MW bite and education**

	High MW Bite			Low MW Bite		
	Overall	Formal	Informal	Overall	Formal	Informal
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Δ log earnings (low-education)						
In-migration rate	-0.807***	-0.078	-1.157***	-1.320***	-1.322***	-0.866***
	(0.169)	(0.242)	(0.193)	(0.131)	(0.148)	(0.171)
Panel B: Δ log earnings (high-education)						
In-migration rate	-2.287***	-1.278**	-3.190***	-0.364**	-0.591***	0.274
	(0.614)	(0.592)	(0.474)	(0.185)	(0.178)	(0.169)
Panel C: Δ employment rate (low-education)						
In-migration rate	0.184***	-0.095***	0.279***	0.017	-0.073***	0.090***
	(0.055)	(0.033)	(0.065)	(0.038)	(0.019)	(0.031)
Panel D: Δ employment rate (high-education)						
In-migration rate	-0.183**	-0.207***	0.024	-0.042	-0.005	-0.037*
	(0.086)	(0.075)	(0.020)	(0.029)	(0.027)	(0.020)
Observations	9,840	9,840	9,840	11,460	11,460	11,460
Municipalities	820	820	820	955	955	955

Notes: This table shows origin-level SSIV coefficients on changes in employment rate and log earnings for each sector and by the minimum wage bite and education levels. Each observation is an origin municipality-year cell. The informal sector also includes self-employed workers. All regressions include time dummies, control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to

Table F4: **Effects of migration on non-wage benefits, by level of MW bite**

	High MW Bite			Low MW Bite		
	Food	Transport	Health	Food	Transport	Health
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Δ benefits (low-education)						
In-migration rate	-0.033	0.030	0.042***	-0.109***	-0.057***	-0.032***
	(0.030)	(0.024)	(0.015)	(0.008)	(0.012)	(0.008)
Panel B: Δ benefits (high-education)						
In-migration rate	-0.062**	-0.034	-0.010	-0.129***	-0.070***	-0.137***
	(0.028)	(0.031)	(0.017)	(0.031)	(0.022)	(0.018)
Observations	9,840	9,840	9,840	11,460	11,460	11,460
Municipalities	820	820	820	955	955	955

Notes: This table shows origin-level SSIV coefficients on changes in the proportions of working-age native population who receive health insurance, food, or transport subsidies by the minimum wage bite and education levels. Each observation is an origin municipality-year cell. All regressions include time dummies, control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) , and are weighted by the working-age native population in 1991. Standard errors clustered at the origin municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table F5: **Effects of migration on employment and earnings, by baseline level of informality**

	High informality			Low informality		
	Overall	Formal	Informal	Overall	Formal	Informal
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Δ log earnings						
In-migration rate	-2.948***	-1.797***	-3.390***	-0.571**	-0.693***	-0.104
	(0.234)	(0.335)	(0.217)	(0.247)	(0.258)	(0.151)
Panel B: Δ employment rate						
In-migration rate	0.004	-0.290***	0.294***	-0.066	-0.127**	0.062
	(0.047)	(0.055)	(0.046)	(0.046)	(0.051)	(0.042)
Observations	9,696	9,696	9,696	11,460	11,460	11,460
Municipalities	808	808	808	955	955	955

Notes: This table shows origin-level SSIV coefficients on changes in employment rate and log earnings for each sector and by baseline level of informality. Each observation is an origin municipality-year cell. The informal sector also includes self-employed workers. All regressions include time dummies, control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) , and are weighted by the working-age native population in 1991. Standard errors clustered at the origin municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table F6: **Effects of migration on non-wage benefits, by baseline level of informality**

	High informality			Low informality		
	Food	Transport	Health	Food	Transport	Health
	(1)	(2)	(3)	(4)	(5)	(6)
In-migration rate	-0.251***	-0.387***	-0.080***	-0.259***	-0.062*	-0.122***
	(0.034)	(0.049)	(0.027)	(0.050)	(0.033)	(0.037)
Observations	9,696	9,696	9,696	11,460	11,460	11,460
Municipalities	808	808	808	955	955	955

Notes: This table shows origin-level SSIV coefficients on changes in the proportions of working-age native population who receive health insurance, food, or transport subsidies by baseline level of informality. Each observation is an origin municipality-year cell. All regressions include time dummies, control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) , and are weighted by the working-age native population in 1991. Standard errors clustered at the origin municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table F7: **Effects of migration on employment and earnings, by education level**

	Employment			Earnings		
	Overall	Formal	Informal	Overall	Formal	Informal
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Low-Education						
In-migration rate	0.085***	-0.056***	0.141***	-1.271***	-0.910***	-1.173***
	(0.030)	(0.015)	(0.029)	(0.109)	(0.126)	(0.117)
Observations	11,460	11,460	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955	955	955
Panel B: High-education						
In-migration rate	-0.102***	-0.070**	-0.033**	-0.846***	-0.714***	-0.665***
	(0.039)	(0.034)	(0.014)	(0.235)	(0.230)	(0.178)
Observations	11,460	11,460	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955	955	955

Notes: This table shows origin-level SSIV coefficients on changes in employment rate and log earnings for each sector and by education level. Each observation is an origin municipality-year cell. The informal sector also includes self-employed workers. All regressions are weighted by the working-age native population in 1991, include time dummies and control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water). Standard errors clustered at the origin municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table F8: **Effects of migration on non-wage benefits, by education level**

	Food	Transport	Health
	(1)	(2)	(3)
Panel A: Low-Education			
In-migration rate	-0.090***	-0.040***	0.003
	(0.009)	(0.010)	(0.007)
Observations	11,460	11,460	11,460
Municipalities	955	955	955
Panel B: High-education			
In-migration rate	-0.132***	-0.066***	-0.111***
	(0.028)	(0.018)	(0.019)
Observations	11,460	11,460	11,460
Municipalities	955	955	955

Notes: This table shows origin-level SSIV coefficients on changes in the proportion of the working-age native population receiving non-wage benefits, by education level. Each observation is an origin municipality-year cell. The informal sector also includes self-employed workers. All regressions are weighted by the working-age native population in 1991, include time dummies and control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water). Standard errors clustered at the origin municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Appendix G Weather shocks and predicted migration

In this section, we discuss the weather data and provide further details about how we construct our instrument. We also show that our results are robust to an alternative measure of weather shocks.

G.1 Weather data

Our main source for weather data comes from the CRUTS v4, a gridded dataset produced by the Climatic Research Unit at the University of East Anglia (Harris et al., 2020). It provides information on monthly precipitation and temperature covering the whole globe (except Antarctica) from 1901 to 2018. The grid resolution is $0.5° \times 0.5°$ (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 georeference the coordinates from each municipality's centroid. Then, for each municipality, we find the grid's four points that are closest to its centroid and calculate the average level of precipitation and temperature from these 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 from 1901 to 2010, which we aggregate in yearly measures. Precipitation is defined as the sum of monthly levels, and temperature is the average. For each municipality in the Semiarid region, we calculate the historical mean from both variables and take the natural logarithm of these variables (both levels and long-term averages).

Finally, our weather shock variables are defined as

$$Rainfall_{ot} = \ln\left(\sum_{\tau \in \{GS\}} r_{o\tau t}\right) - \ln(\bar{r}_o) \tag{G1}$$

where $r_{o\tau t}$ is the rainfall in the 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. In our main specifications, we use data from the Semiarid's growing season (from November to April), but results are very similar when we use the full year (see Table G1).

Table G1: **Migration outflows induced by weather shocks (12 months)**

	(1)	(2)	(3)	(4)	(5)	(6)
Rainfall$_{t-1}$	-0.126***	-0.107***	-0.106***	-0.110***	-0.112***	-0.109***
	(0.033)	(0.033)	(0.034)	(0.034)	(0.033)	(0.033)
Rainfall$_{t-2}$			0.029	0.047		
			(0.037)	(0.039)		
Rainfall$_{t-3}$				0.046		
				(0.033)		
Rainfall$_{t}$					-0.014	
					(0.039)	
Rainfall$_{t+1}$						-0.068*
						(0.037)
Observations	14,400	14,400	14,400	14,400	14,400	14,400
Municipalities	960	960	960	960	960	960
R-Squared	0.461	0.465	0.465	0.466	0.465	0.466
Time dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Temperature shocks	✓	✓	✓	✓	✓	✓
Demographics		✓	✓	✓	✓	✓

Notes: Each observation is a municipality-year cell. The 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 the historical average. All specifications include controls for temperature shocks, municipality, and year fixed effects. Columns (2)-(6) control for municipality-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) . Standard errors are clustered at the grid level. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

G.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 use the Standardized Precipitation Evapotranspiration Index (SPEI) developed by Vicente-Serrano et al. (2010). The SPEI is a measure that accounts for both precipitation and potential evapotranspiration, providing a measure of water balance in a given period. This index also captures deviations from the historical average (1905-2018) in the net water need for a given location. An SPEI value of -1 means that the precipitation level is one standard deviation below the historical level needed to maintain the balance given the potential evapotranspiration. According to Vicente-Serrano et al. (2010), the SPEI is particularly useful for detecting, monitoring, and exploring the consequences of global warming on drought conditions. We repeat the first step in our procedure to construct an instrument for in-migration, using the SPEI instead of rainfall and temperature shocks. Once again, we calculate the average SPEI for the 6-month growing season. Table G2 shows that we can also use this measure to predict the out-migration rate from the origin municipalities in the Semiarid region, although the estimates are noisier than those in Table 4.

We estimate the same specification of column 3 from Tables 5 and 6 using this new instrument and show in Tables G3 and G4 that the results are very similar.

Table G2: **Migration outflows induced by weather shocks: Standardized Precipitation Evapotranspiration Index (SPEI)**

	(1)	(2)	(3)	(4)	(5)	(6)
SPEI_{t-1}	-0.033**	-0.041**	-0.037**	-0.042**	-0.045***	-0.043***
	(0.016)	(0.016)	(0.016)	(0.017)	(0.017)	(0.016)
SPEI_{t-2}			0.041**	0.043**		
			(0.019)	(0.019)		
SPEI_{t-3}				0.032*		
				(0.019)		
SPEI_{t}					-0.020	
					(0.016)	
SPEI_{t+1}						-0.003
						(0.017)
Observations	14,400	14,400	14,400	14,400	14,400	14,400
Municipalities	960	960	960	960	960	960
R-Squared	0.461	0.465	0.465	0.466	0.465	0.465
Time dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Temperature shocks	✓	✓	✓	✓	✓	✓
Demographics		✓	✓	✓	✓	✓

Notes: Each observation is a municipality-year cell. The dependent variable is the number of individuals who left the origin municipality divided by the total population in the 1991 Census. All specifications include controls for temperature shocks, municipality, and year fixed effects. Columns (2)-(6) control for municipality-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) . Standard errors are clustered at the grid level. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table G3: **Effects of migration on earnings and employment, using SPEI to predict out-migration**

	Employment			Earnings		
	Overall	Formal	Informal	Overall	Formal	Informal
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
In-migration rate	-0.084	-0.032	-0.052	0.146	0.296	-0.231
	(0.084)	(0.089)	(0.075)	(0.405)	(0.438)	(0.424)
Observations	8,190	8,190	8,190	8,190	8,162	8,179
Municipalities	684	684	684	684	684	684
Panel B: SSIV						
In-migration rate	-0.017	-0.125***	0.108***	-0.867***	-0.591***	-0.744***
	(0.034)	(0.037)	(0.034)	(0.197)	(0.198)	(0.123)
Observations	11,460	11,460	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955	955	955

Notes: This table shows origin-level destination-level OLS and SSIV coefficients of regressions on log earnings and on the employment rate for native workers. Each observation is an origin municipality-year cell. The instrument for the migrant inflow is calculated using the SPEI to predict out-migration distributed by the 1991 share of Semiarid migrants. All regressions include time dummies and destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water). All regressions are weighted by the working-age native population in 1991. Standard errors clustered at the origin municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table G4: **Effects of migration on non-wage benefits, using SPEI to predict out-migration**

	Food	Transport	Health
	(1)	(2)	(3)
Panel A: OLS			
In-migration rate	-0.063	0.018	-0.049
	(0.088)	(0.077)	(0.076)
Observations	8,190	8,190	8,190
Municipalities	684	684	684
Panel B: SSIV			
In-migration rate	-0.222***	-0.106***	-0.109***
	(0.034)	(0.024)	(0.024)
Observations	11,460	11,460	11,460
Municipalities	955	955	955

Notes: This table shows destination-level OLS and origin-level SSIV coefficients of regressions on the proportion of working-age native population receiving non-wage benefits. Each observation is an origin municipality-year cell. The instrument for the migrant inflow is calculated using the SPEI to predict out-migration distributed by the 1991 share of Semiarid migrants. All regressions include time dummies and destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water). All regressions are weighted by the working-age native population in 1991. Standard errors clustered at the origin municipality level in parenthesis. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Appendix H Shift-share instrument (SSIV)

In this section, we derive the origin-level SSIV estimator and present and discuss the identifying assumptions needed to produce a consistent estimator of the effects of the inflow of migrants from the Semiarid region on labor markets in the destination municipalities.

We start from the structural equation 5. To simplify notation, we omit the time subscript t. By the Frisch-Waugh-Lovell Theorem, we can re-write it as

$$y_d^{\perp} = \beta m_d^{\perp} + \varepsilon_d^{\perp} \tag{H1}$$

where all $y_d^{\perp}$ is the vector of outcomes, $m_d^{\perp}$[33] is the observed number of Semiarid's migrants who entered the destination municipality d and $\varepsilon_d^{\perp}$ is a structural residual. All variables are residualized to remove the effects from the covariates.

In equation 9, we defined the shift-share instrumental variable (SSIV) as

$$\widehat{m}_d = \sum_{o=1}^{O} s_{od} \frac{\widehat{M}_o}{N_d} \tag{H2}$$

where $s_{od} = \dfrac{M_{od}}{\sum_d M_{od}}$ is the share of migrants from origin municipality o who lived in the destination area d in 1991 and $\widehat{M}_o$ is the predicted number of migrants leaving the Semiarid region driven by weather shocks.

The more traditional approach would be to estimate β using $\widehat{m}_d$ as an instrument for the endogenous migrant inflow $m_d^{\perp}$. In such a case, we would have

$$\hat{\beta} = \frac{\sum_d \widehat{m}_d y_d^{\perp}}{\sum_d \widehat{m}_d m_d^{\perp}} \tag{H3}$$

By the definition of $\widehat{m}_d$ in equation H2 and switching the order of the summation,

$$\hat{\beta} = \frac{\sum_d \left(\sum_o s_{od} \frac{\widehat{M}_o}{N_d} \right) y_d^{\perp}}{\sum_d \left(\sum_o s_{od} \frac{\widehat{M}_o}{N_d} \right) m_d^{\perp}} = \frac{\sum_o \widehat{M}_o \left(\sum_d s_{od} \frac{y_d^{\perp}}{N_d} \right)}{\sum_o \widehat{M}_o \left(\sum_d s_{od} \frac{m_d^{\perp}}{N_d} \right)} = \frac{\sum_o s_o \widehat{M}_o \bar{y}_o}{\sum_o s_o \widehat{M}_o \bar{m}_o} \tag{H4}$$

where $\bar{y}_o = \dfrac{\sum_d s_{od} \frac{y_d^{\perp}}{N_d}}{\sum_d s_{od}}$ is a weighted average of the residualized outcome, normalized by the native population, which uses as weights the destination's average exposure to the shocks $s_o = \sum_d s_{od}$. The same result is valid for the endogenous variable $m_d^{\perp}$, meaning that we can estimate the following IV regression at the origin municipality level:

$$\bar{y}_o = \beta \bar{m}_o + \bar{\varepsilon}_o \tag{H5}$$

[33] In order to facilitate the interpretation of the coefficients we normalize this measure dividing by the working-age native population in 1991, which means $m_d = \dfrac{M_d}{N_d}$

using the predicted number of migrants from the Semiarid region, $\widehat{M_o}$, as instrumental variable and weighting by the average exposure s_o.

This derivation is almost identical to that presented by Borusyak et al. (2021), except for the fact that we need to divide both variables by the predetermined native population. Their equivalence result shows that the parameter β can be estimated at the level of the identifying variation, which in our case is the origin municipality hit by weather shocks.

Table H1: **SSIV First Stage**

	(1)	(2)	(3)
Panel A: OLS			
mnrate	0.910***	0.909***	0.926***
	(0.015)	(0.015)	(0.020)
F-statistic	3527.26	3528.89	2172.53
Observations	11,460	11,460	11,460
Municipalities	955	955	955
Effective sample size	7,281	7,281	7,281
Time dummies		✓	✓
Baseline			✓

Notes: This table shows the SSIV first stage coefficients of the origin-level weighted average of the endogenous inflow of migrants at the destinations against the predicted number of migrants from the Semiarid region. Each observation is an origin municipality-year cell. The F-statistic is calculated as the square of the coefficient t-statistic (see Borusyak et al., 2021). The effective sample size is the inverse of the HHI of the origin-level exposure. Column (2) includes time dummies while Column (3) also controls for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) . Regressions are weighted by the working-age native population in 1991. Standard errors cluster by the municipality of origin in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Appendix I Spatial correlation in weather shocks

Weather events are likely correlated across space. Figure 1 shows that precipitation levels in the Semiarid are similar among nearby municipalities. Potentially, this could invalidate the consistency of our estimator given by *Assumption 2 (Many uncorrelated shocks)* discussed in Appendix H. Here, we investigate this issue by re-constructing our instrument according to different degrees of aggregation of regions of origin of a migrant - such as a microregion or a mesoregion - instead of a municipality. IBGE (1990) defines microregions as "groups of economically integrated municipalities sharing borders and structure of production.". Mesoregions are collections of microregions of which not all municipalities share borders.[34] Brazil has 5,565 municipalities, 361 micro, and 87 mesoregions overall. The Semiarid has 960 municipalities, 137 micro, and 35 mesoregions.

The intuition behind this exercise is that even if weather shocks are spatially correlated among contiguous municipalities, such a correlation should decrease as we consider larger areas. Table I2 displays Moran's index of spatial correlation of rainfall shocks for each of the three geographic aggregates in columns 1-3.[35] As expected, neighboring municipalities display a correlation above 0,24, but it decreases rapidly as we aggregate up to micro and meso regions, to 0,16 and 0,07, respectively.

Table I2 also shows the association between rainfall shocks and migration outflows. Column 1 is identical to Table 4 for reference. Columns 2 and 3 report almost identical point estimates and precision, indicating that we do not lose any significant information by aggregating origin areas. Next, we estimate our main specification from Column (3) in Tables 5 and 6 using instruments corresponding to micro and mesoregion-level aggregation. Tables I3-I4 show that our results associating migration and earnings, employment, and non-wage benefits are very similar to the municipality-level estimates, although standard errors increase substantially, as one would expect considering that there are fewer units from which we can leverage variation. All those results indicate that spatial correlation among rainfall shocks in origin municipalities is not a source of relevant bias in our setting.

[34] Table I1 reports summary statistics of our main variables for both levels of aggregation.

[35] Moran's I is calculated according to the following formula:

$$I = \frac{1}{\sum_i \sum_j w_{ij}} \times \frac{\sum_i \sum_j w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\frac{1}{N}(y_i - \bar{y})^2} \tag{I1}$$

Essentially, it is a correlation coefficient weighted by an appropriate matrix that models how different units are related across space. We use a contiguity matrix with the queen criterion, meaning that two localities i and j sharing either borders or vertices are considered 'neighbors', and the entry w_{ij} has a positive value. Non-adjacent pairs receive a zero weight. As discussed by Beenstock et al. (2019), Moran's I can be calculated for each period and averaged out with panel data.

Table I1: **Summary statistics: Micro- and meso-regions in the Semiarid**

Panel A: Micro-regions	Mean	Std. Dev.	Min	Max	Obs
Rainfall shock	-0.01	0.20	-0.70	0.47	2,055
Temperature shock	0.01	0.01	-0.01	0.05	2,055
Out-migration	1,500.70	1,371.95	6.00	9,685.00	2,055
Out-migration rate (p.p.)	1.08	0.41	0.12	3.12	2,055
Population	149,800.70	129,018.99	4,968.00	752,718.94	2,055
Area	7,150.16	7,857.60	84.94	55,358.33	2,055
Number of municipalities	8.20	4.56	2.00	26.00	2,055
Panel B: Meso-regions	**Mean**	**Std. Dev.**	**Min**	**Max**	**Obs**
Rainfall shock	-0.02	0.20	-0.69	0.44	525
Temperature shock	0.01	0.01	-0.01	0.05	525
Out-migration	5,874.18	5,766.16	51.00	34,800.00	525
Out-migration rate (p.p.)	1.08	0.37	0.24	2.32	525
Population	586,362.76	527,013.45	15,499.00	2,349,152.25	525
Area	27,986.83	30,649.61	84.94	124,505.71	525
Number of municipalities	37.20	21.51	10.00	118.00	525

Notes: Rainfall is measured in mm. Temperature is measured in degrees Celsius. Migration outflow (inflow) rate is the share of migrants over the local (native) population. The area is measured in km^2.

Table I2: **Migration outflows induced by weather shocks according to different aggregation levels**

	(1)	(2)	(3)
	Municipality	Micro-region	Meso-region
Rainfall_$t-1$	-0.092***	-0.090***	-0.096***
	(0.028)	(0.031)	(0.024)
Observations	14,400	2,055	525
Regions	960	137	35
R squared	0.465	0.772	0.875
Moran's I	0.235	0.158	0.075

Notes: Each observation is a region-year cell. The dependent variable is the number of individuals who left the origin region divided by the total population in the 1991 Census. Rainfall is measured as log-deviation from the historical average. All specifications control for temperature shocks and control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water) interacted with time dummies, and include municipality, and year fixed effects. Moran's I show the spatial correlation in rainfall shocks among origin regions. Standard errors are clustered at the respective region level. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table I3: **Effects of migration on earnings according to different aggregation levels**

	Employment			Earnings		
	Overall	Formal	Informal	Overall	Formal	Informal
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Municipality						
In-migration rate	-0.017	-0.125***	0.108***	-0.867***	-0.591***	-0.744***
	(0.034)	(0.037)	(0.034)	(0.197)	(0.198)	(0.123)
Observations	11,460	11,460	11,460	11,460	11,460	11,460
Municipalities	955	955	955	955	955	955
Panel B: Micro-region						
In-migration rate	-0.003	-0.117**	0.114**	-0.844***	-0.556*	-0.743***
	(0.058)	(0.055)	(0.057)	(0.301)	(0.290)	(0.201)
Observations	1,644	1,644	1,644	1,644	1,644	1,644
Municipalities	137	137	137	137	137	137
Panel C: Meso-region						
In-migration rate	0.008	-0.125	0.132	-0.869	-0.554	-0.767**
	(0.091)	(0.098)	(0.095)	(0.550)	(0.527)	(0.343)
Observations	420	420	420	420	420	420
Municipalities	35	35	35	35	35	35

Notes: This table shows origin-level SSIV coefficients on changes in log earnings by sector, for different levels of aggregation. Each observation is an origin region-year cell. The informal sector also includes self-employed workers. All specifications include time and control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water). Panel A replicates the same results from Table 5. In Panels B and C, we aggregate the origin-level shocks at the micro- and meso-region levels, respectively. All regressions are weighted by native working-age population in 1991. Standard errors clustered at the respective aggregation level in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table I4: **Effects of migration on employment according to different aggregation levels**

	Food	Transport	Health
	(1)	(2)	(3)
Panel A: Municipality			
In-migration rate	-0.222***	-0.106***	-0.109***
	(0.034)	(0.024)	(0.024)
Observations	11,460	11,460	11,460
Municipalities	955	955	955
Panel B: Micro-region			
In-migration rate	-0.210***	-0.083**	-0.100***
	(0.053)	(0.041)	(0.035)
Observations	1,644	1,644	1,644
Municipalities	137	137	137
Panel C: Meso-region			
In-migration rate	-0.221**	-0.080	-0.101*
	(0.087)	(0.065)	(0.060)
Observations	420	420	420
Municipalities	35	35	35

Notes: This table shows origin level SSIV coefficients of change in the proportions of working-age native population who receive health insurance, food, or transport subsidies. Each observation is an origin region-year cell. The informal sector also includes self-employed workers. All specifications include time and control for destination-level 1991 characteristics (log of the working-age native population; shares of the population aged 15-25, 26-50, 51-65 and older than 65; share of the non-white population; share of the population with a college education; share of women in the total and employed populations; shares of employment in agriculture and manufacturing; logs of the average household income and size; and the shares of households with access to electricity and piped water). Panel A replicates the same results from Table 6. In Panels B and C, we aggregate the origin-level shocks at the micro- and meso-region levels, respectively. All specifications use the same set of controls defined in Table 6. All regressions are weighted by native working-age population in 1991. Standard errors clustered at the respective aggregation level in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.