Urban Sprawl and Productivity in Shrinking Cities: Causal Evidence from Brazil

Gustavo Castro* Carlos Azzoni† André Chagas‡

October 2, 2025

Abstract

We study how urban sprawl affects labor productivity in cities that are shrinking. This paradox, cities that lose population but continue to expand physically, is increasingly common in developing countries. Using data on over 5,500 Brazilian municipalities from 2010 to 2022, we estimate the causal effect of urban land expansion in shrinking cities on productivity growth. To address endogeneity, we use a SPEI-based shift-share instrument for population decline and implement a two-stage least squares model. We also apply a machine learning strategy to validate the results with high-dimensional controls. We find that urban sprawl significantly lowers productivity in shrinking cities across all major sectors. The results are robust to placebo tests. Our findings show that spatial expansion in shrinking areas creates economic inefficiencies, reinforcing the need for spatially-aware planning and policies that account for demographic contraction.
Keywords: Urban Shrinkage, Urban Sprawl, Labor Productivity

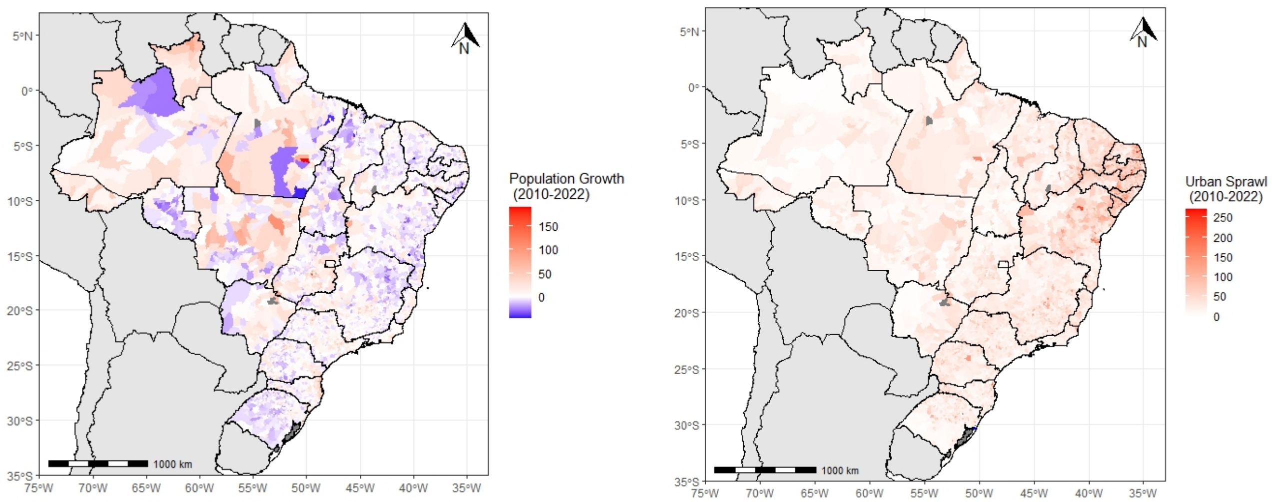
Abstract

Estudamos como a expansão urbana afeta a produtividade do trabalho em cidades em processo de encolhimento populacional. Esse paradoxo — municípios que perdem população mas continuam a se expandir fisicamente — é cada vez mais comum em países em desenvolvimento. Utilizando dados de mais de 5.500 municípios brasileiros entre 2010 e 2022, estimamos o efeito causal da expansão da mancha urbana em cidades em shrinkage sobre o crescimento da produtividade. Para lidar com endogeneidade, utilizamos um instrumento do tipo shift-share baseado no SPEI para declínio populacional e implementamos um modelo de mínimos quadrados em dois estágios. Também aplicamos uma estratégia de aprendizado de máquina para validar os resultados com controles de alta dimensão. Os resultados indicam que o espalhamento urbano reduz significativamente a produtividade em cidades em shrinkage em todos os grandes setores. As conclusões são robustas a testes placebo e reforçam a necessidade de políticas de planejamento urbano sensíveis à contração demográfica.
Palavras-chave: Shrinkage Urbano, Espraiamento Urbano, Produtividade do Trabalho

JEL Codes: R11, O18, C26

Área: 1 – Teoria, métodos e modelos de economia regional

*Departament of Economics, University of São Paulo, Brazil. E-mail: gustavocastro@usp.br
†Departament of Economics, University of São Paulo, Brazil. E-mail: cazzoni@usp.br
‡Departament of Economics, University of São Paulo, Brazil. E-mail: achagas@usp.br
This research was supported by the São Paulo Research Foundation (FAPESP), grant number 2025/00940-0. We thank Tiago Ferraz, and Pedro H. B. de Barros for valuable comments and suggestions.



A. Shrinking Cities

B. Urban Sprawl (Built-Up Area Growth)

Figure 1: Demographic Contraction in Brazil, 2010–2022 Notes: Panel A shows the rate of growth population between 2010 and 2022. Panel B displays the rate of growth in urban built-up area over the same period.

1 Introduction

Across the globe, a growing number of cities are facing population decline. Yet paradoxically, many of these shrinking cities continue to expand physically. This disconnect between demographic contraction and spatial growth, what we call ”shrinking but sprawling cities”, is especially pronounced in developing countries, where land-use regulation is limited and urban planning often lags behind demographic shifts (Rink et al., 2014; Mallach et al., 2017; Haase et al., 2017; Batunova and Gunko, 2018; Yang et al., 2022).

Brazil offers a particular case. Between 2010 and 2022, over 40% of its municipalities experienced population decline, even as urban built-up areas continued to grow (IBGE, 2023a; MapBiomas, 2023). Figure 1 shows how demographic contraction and spatial expansion are distributed across the country, revealing a clear mismatch. This expansion in the absence of population growth raises concerns about urban inefficiency. The expansion of urban footprints can increase infrastructure and service delivery costs (Behnisch et al., 2022), lead to underutilized land, and strain already limited public budgets. More importantly, such patterns can hinder local economic performance by reducing the benefits of agglomeration and misallocating resources Duranton and Puga (2004); Behrens et al. (2014).

Despite its relevance, there is limited empirical evidence on how this spatial-demographic mismatch affects economic performance. Most existing studies on urban shrinkage emphasize demographic, social, and innovation impacts (Jarzebski et al., 2021; Lv et al., 2024), while sprawl research has focused mainly on environmental degradation, infrastructure costs, or mobility issues (Ortiz-Moya, 2020; Rao et al., 2023; Li and Chen, 2023; Wang et al., 2023). Only a handful of studies attempt to identify the causal effects of urban form in shrinking contexts on productivity outcomes (Yang et al., 2022), and virtually none do so for developing countries in Latin America. To our knowledge, no prior research has causally estimated the productivity implications of urban sprawl in shrinking Brazilian cities.

This paper addresses this gap by asking: What are the labor productivity effects of urban sprawl in shrinking cities? We provide novel evidence based on a panel of 5,515 Brazilian municipalities from 2010 to 2022, combining satellite-derived measures of urban expansion, administrative labor market microdata, and municipal GDP figures (IBGE, 2023a; MapBiomas, 2023; RAIS, 2024).

A credible estimation of this relationship faces three major empirical challenges. First, urban shrinkage is endogenous to productivity: municipalities may lose population due to poor economic performance, raising reverse causality concerns. Second, urban sprawl is likely to be correlated with unobserved factors, such as local infrastructure investments or land market conditions, which also influence productivity. Third, the interaction effects are potentially nonlinear and high-dimensional: the interaction of demographic trends with urban form depends on a complex array of socioeconomic and geographic characteristics.

To overcome these issues, we implement a two-stage least squares (2SLS) approach, instrumenting both urban shrinkage and urban sprawl. For shrinkage, we construct a shift-share instrument based on long-term weather shocks [1](Bartik, 1991; Busso and Chauvin, 2025; Corbi et al., 2025a). This climate-based exposure captures persistent migratory responses to environmental stress while remaining orthogonal to current productivity levels. For sprawl, we exploit the spatial diffusion of urban form by using lagged urban expansion in neighboring municipalities as an instrument (Duranton et al., 2011; Koster and Rouwendal, 2017). Both strategies are theoretically grounded and empirically validated, allowing us to isolate exogenous variation in demographic and spatial dynamics.

We further complement this approach with Double Machine Learning (DML) using LASSO. This method orthogonalizes the endogenous regressor with respect to a high-dimensional set of controls, flexibly capturing nonlinearities and reducing overfitting (Chernozhukov et al., 2018). By combining 2SLS and DML, we ensure valid inference even under complex treatment assignment mechanisms and weak instrument settings.

Our results provide novel causal evidence that urban sprawl significantly depresses labor productivity in shrinking Brazilian cities. The interaction between shrinkage and sprawl is consistently negative and statistically significant across specifications, placebo tests, bootstrap inference, and estimation strategies. This adverse effect persists across all major sectors (agriculture, manufacturing, and services), highlighting a broad-based spatial inefficiency associated with demographic contraction and uncoordinated urban expansion [2].

These findings contribute to the urban economics literature by documenting a new mechanism of spatial inefficiency in developing contexts. When cities expand physically despite losing population, productivity suffers. Using high-resolution spatial data, a credible IV strategy, and machine learning for high-dimensional inference, this article bridges the gap between urban form and economic performance under shrinkage, an increasingly relevant dynamic for middle-income countries. Our evidence suggests that spatially disjointed contraction, if unmanaged, can exacerbate economic decline rather than cushion it.

This paper is organized as follows. Section 2 presents the theoretical framework. Section 3 presents the data and the identification strategy. Section 4 reports the main results and robustness checks. Section 5 presents the conclusion.

[1] Specifically, the historical incidence of droughts measured by the Standardized Precipitation-Evapotranspiration Index (SPEI) between 2000 and 2009.

[2] While our results point to negative productivity effects from the interaction between shrinkage and sprawl, it is worth noting that other urban outcomes may respond differently. A strand of the literature highlights potential benefits of urban contraction and decentralization, such as reductions in pollution, noise, and traffic congestion, especially in high-density contexts (Haase et al., 2014a; Martinez-Fernandez et al., 2012). Our analysis focuses on economic productivity, but future work could explore these complementary dimensions to provide a more holistic view of the trade-offs involved in urban decline.

2 Theoretical Framework

Urban economic theory has long emphasized the critical role of spatial structure in shaping local productivity. Cities exist because they facilitate agglomeration economies—productivity, enhancing interactions that arise when firms and individuals concentrate in space. These benefits include knowledge spillovers, better matching in labor and input markets, and shared use of infrastructure, all of which reduce per capita costs and raise the efficiency of economic activity (Rosenthal and Strange, 2004; Duranton and Puga, 2004). In this context, urban density emerges as a central determinant of productivity, acting as a proximate channel through which spatial proximity translates into economic gains (Glaeser et al., 1992).

A foundational tool for understanding the relationship between urban form and population is the canonical monocentric city model, developed by Alonso (1964), Muth (1969), and Mills (1972). In this framework, a city is organized around a central business district (CBD) to which all residents commute. Individuals face a trade-off between housing costs and transportation costs: as one moves farther from the CBD, land becomes cheaper, but commuting becomes more expensive. The equilibrium outcome is a downward-sloping bid-rent curve, with higher population densities and land prices near the center, declining with distance. Importantly, the model predicts that the spatial extent of a city adjusts endogenously with its population. When the number of residents increases, urban expansion accommodates demand for housing. Conversely, when the population falls, the built environment should, in theory, contract accordingly to preserve efficiency and spatial balance.

While this model has been instrumental in shaping urban economic theory, more recent treatments emphasize its limitations in accounting for long-lived capital, institutional rigidities, and persistent sprawl in shrinking cities (Brueckner, 1987; Duranton and Puga, 2015).

However, this prediction assumes an idealized world of perfect flexibility in land use and instantaneous adjustment. In practice, cities exhibit substantial inertia. The physical footprint of urban areas often persists or even continues to expand despite declining populations, a phenomenon widely documented in shrinking cities across Europe, North America, and Asia (Rink et al., 2014; Mallach et al., 2017; Haase et al., 2017; Batunova and Gunko, 2018; Yang et al., 2022). This empirical reality reflects structural rigidities omitted by the Alonso-Muth-Mills model, such as sunk costs of infrastructure, long-lived housing stock, institutional zoning constraints, and political fragmentation, that prevent rapid or proportional spatial contraction.

The persistence or expansion of the urban area in the face of population decline has critical implications for productivity. In a well-functioning spatial equilibrium, the physical size of a city should scale with its population. When that relationship breaks down, effective urban density falls, weakening the intensity of agglomeration mechanisms. Infrastructure and public services become underutilized, increasing their per capita cost, while land consumption outpaces economic activity. These dynamics are especially problematic when low-density expansion leads to fragmented labor markets and diminished proximity between complementary firms (Yang et al., 2022).

To formalize the link between urban form and productivity, we adopt a standard production framework augmented by density-based agglomeration effects. Let the total output in municipality i be:

$$Y_i = A_i \left(\frac{N_i}{S_i} \right)^{\phi} K_i^{\alpha} L_i^{1-\alpha} \tag{1}$$

In this expression, Y_i denotes total output in municipality i, A_i is total factor productivity (TFP), capturing exogenous technological and institutional conditions, N_i represents population, and S_i is the urbanized land area. The variables K_i and L_i denote the aggregate stock of physical capital and the employed labor force, respectively, with $L_i = N_i$ under the assumption of full participation. The parameters $\alpha \in (0, 1)$ and $1 - \alpha$ correspond to the elasticities of output with respect to capital and labor, while $\phi > 0$ captures the elasticity of productivity with respect to population density (N_i/S_i), reflecting agglomeration economies.

This function exhibits constant returns to scale in capital and labor inputs, but increasing returns in density, reflecting spatial externalities such as knowledge spillovers and infrastructure efficiencies (Ciccone and Hall, 1993, 1996; Combes et al., 2011).

Focusing on labor productivity, we divide both sides by L_i to obtain [3] :

$$\frac{Y_i}{L_i} = A_i \left(\frac{L_i}{S_i}\right)^{\phi} \left(\frac{K_i}{L_i}\right)^{\alpha} \tag{2}$$

Taking logarithms, we derive the linear form:

$$log\left(\frac{Y_i}{L_i}\right) = \log A_i + \phi \log\left(\frac{L_i}{S_i}\right) + \alpha \log\left(\frac{K_i}{L_i}\right) \tag{3}$$

This formulation makes explicit how labor productivity depends on urban density and capital deepening. Specifically, if L_i decreases while S_i remains constant or increases, as is often observed in shrinking but sprawling cities, then the density term $\log(L_i/S_i)$ declines. Given that $\phi > 0$, this leads to a reduction in productivity.

To highlight the mechanism, consider the derivative of labor productivity with respect to urban area S_i, holding L_i and K_i constant. [4] :

$$\frac{\partial}{\partial S_i}\left(\frac{Y_i}{L_i}\right) = -\phi A_i \left(\frac{L_i}{S_i}\right)^{\phi} \left(\frac{K_i}{L_i}\right)^{\alpha} \cdot \frac{1}{S_i} \tag{4}$$

This expression confirms that, ceteris paribus, an increase in urban area (S_i) leads to a decrease in labor productivity whenever $\phi > 0$, due to dilution of density and thus of agglomeration economies.

The empirical specification used in our analysis is derived from this theoretical foundation, with density L_i/S_i serving as a central explanatory variable and capital per worker captured through controls.

This framework suggests several mechanisms through which spatial-demographic mismatch can harm productivity. First, infrastructure under-utilization: lower density increases the per capita cost of maintaining roads, public transit, utilities, and other fixed infrastructure. Second, reduced agglomeration: lower density and physical fragmentation weaken the intensity of knowledge diffusion, input-output linkages, and labor market pooling.

These mechanisms yield three main testable predictions: (1) sprawl reduces productivity more in shrinking cities than in growing ones; (2) the productivity penalty is larger in sectors with high agglomeration dependence; and (3) controlling for geography, pollution, infrastructure, and socioeconomic traits does not eliminate the effect. These hypotheses guide the empirical strategy developed in subsequent sections.

[3] An equivalent formulation is provided in Appendix A, where we rewrite the density and capital terms as separate power functions to clarify their individual elasticities: $\frac{Y_i}{L_i} = A_i \cdot S_i^{-\phi} \cdot K_i^{\alpha} \cdot L_i^{\phi-\alpha}$.

[4] This derivative follows from applying the chain rule to the term $(L_i/S_i)^{\phi}$, which is the only component of $\frac{Y_i}{L_i}$ that depends on S_i. Since L_i and K_i are held constant, we have:

$$\frac{\partial}{\partial S_i}\left(\frac{L_i}{S_i}\right)^{\phi} = -\phi\left(\frac{L_i}{S_i}\right)^{\phi} \cdot \frac{1}{S_i}.$$

Multiplying by $A_i\,(K_i/L_i)^{\alpha}$ yields the full expression in Equation (4). This confirms that urban expansion reduces productivity by lowering density, thereby weakening agglomeration effects.

3 Empirical Design

3.1 Data

Our main outcome is municipal labor productivity growth, measured as the change in log wages per formal worker between 2010 and 2022. Data are drawn from RAIS (Annual Social Information Report), which covers all formal-sector establishments in Brazil. Wages are deflated using the IBGE consumer price index (IPCA), expressed in December 2024 prices. To ensure comparability, we restrict the sample to private-sector workers and exclude public administration.

Shrinkage is a binary indicator equal to 1 if a municipality experienced population decline between the 2010 and 2022 Censuses (IBGE, 2023b), and zero otherwise. It captures urban contraction in demographic terms and serves as one of our endogenous regressors.

Urban sprawl is measured as the percentage increase in built-up urban area from 2010 to 2022. Data come from MapBiomas, which uses Landsat satellite imagery at 30-meter resolution to classify land cover. The built-up category includes impervious surfaces such as streets, buildings, and infrastructure.

To construct the shrinkage instrument, we use the Standardized Precipitation Evapotranspiration Index (SPEI) from Vicente-Serrano et al. (2010) and Beguería et al. (2019). SPEI accounts for both rainfall and evapotranspiration, making it a robust indicator of climatic water stress. he SPEI captures the climatic water balance by comparing observed precipitation with the amount of water needed to maintain stable surface moisture conditions. This water requirement is determined by evapotranspiration, which varies not only over time but also significantly across regions, as it is influenced by multiple atmospheric variables, particularly temperature [5]. Importantly for our empirical strategy, these characteristics make the SPEI a superior predictor of drought conditions compared to precipitation-only indicators such as the Standardized Precipitation Index (SPI). SPEI has been widely used in studies of weather-induced migration and climate vulnerability due to its ability to capture comprehensive hydrological stress (Kubik and Maurel, 2016; Albert et al., 2021; Busso and Chauvin, 2025).We computed the average share of months between 2000 and 2009 in which SPEI was below -1, following Busso and Chauvin (2025). To ease interpretation, the index is multiplied by -1, so that higher values denote greater dryness.

Socioeconomic controls include population, average household size, and formal employment in R&D-intensive occupations[6] (IBGE, 2023b; RAIS, 2024). Productivity levels in 2010 are included to capture initial economic conditions.

We use municipality-level PM2.5 emissions from EDGAR v8.1 (Crippa et al., 2024) as a proxy for environmental pressure. The data, originally expressed as $kg/m^2/s$, are spatially aggregated to total emissions (kg/s) and divided by population to obtain a per capita indicator. This variable serves as a control for pollution-driven productivity effects via health and operational channels (Zivin and Neidell, 2012; Yang et al., 2022).

Geographic controls include altitude (IPEA, 2024), latitude, longitude (IBGE shapefiles), and log distances to both the state capital and the Federal District (calculated via the Haversine formula). Infrastructure is proxied by log paved road area (DNIT, 2005).

All covariates are measured in or before 2010 to mitigate simultaneity. Summary statistics

[5] The SPEI data used in this study were originally provided as gridded monthly time series with a spatial resolution of 0.5° latitude by 0.5° longitude. To derive drought exposure at the municipal level, we computed monthly SPEI values using CRU TS climate variables (precipitation and potential evapotranspiration) over the period 1991–2010. The resulting grid-level SPEI values were converted into raster layers and spatially averaged within each municipality using exact raster-polygon extraction methods. This approach ensures precise alignment between gridded climate data and administrative boundaries without requiring interpolation.

[6] See Table B.1 in the Appendix B

are reported in Appendix Table B.2.

3.2 Identification Strategy

We aim to estimate the causal impact of urban sprawl in shrinking cities on local productivity growth. Our empirical focus is on the interaction between two spatial dynamics: population contraction and urban land expansion. The main specification is:

$$\Delta \log(Y_i) = \beta_0 + \beta_1(\text{Shrinkage}_i \times \text{UrbanSprawl}_i) + \theta^\top X_i + \varepsilon_i, \tag{5}$$

The dependent variable, $\Delta \log(Y_i)$, denotes the cumulative growth in labor productivity in municipality i between 2010 and 2022, computed as the difference in the log of output per formal worker. The variable $Shrinkage_i$ is a binary indicator equal to 1 if the population of municipality i declined between the 2010 and 2022 Brazilian Censuses, and 0 otherwise. $UrbanSprawl_i$ is defined as the percentage growth in built-up area in i between 2010 and 2022, based on high-resolution satellite imagery from MapBiomas. The term X_i is a vector of baseline controls, including altitude, latitude, longitude, pollution, paved road infrastructure, initial productivity levels, population size, household size, and distances to both the state capital and the Federal District, that capture initial socioeconomic and geographic characteristics. The coefficient β_1 captures the effect of the interaction between population contraction and urban expansion on productivity growth, and ε_i is an idiosyncratic error term.

Both regressors are likely endogenous due to distinct yet potentially overlapping threats to identification. First, urban shrinkage may suffer from reverse causality: municipalities experiencing negative productivity shocks, such as industrial decline or labor market deterioration, may subsequently lose population. Second, urban sprawl may reflect the influence of unobserved factors, such as infrastructure investments, fiscal incentives, or land market dynamics, that simultaneously shape productivity trajectories. To mitigate these concerns, we adopt a two-stage least squares (2SLS) strategy with separate and theoretically motivated instruments for each endogenous variable.

To instrument for shrinkage, we follow a shift-share design inspired by Bartik (1991) further developed in recent literature on migration and environmental shocks (Busso and Chauvin, 2025; Corbi et al., 2025a). The instrument interacts historical migrant inflow shares with plausibly exogenous climatic shocks in origin regions. Specifically, we define:

$$Z_i^{\text{Shrinkage}} = \sum_o s_{i,o} D_o, \tag{6}$$

Where $s_{i,o}$ is the share of migrants from origin o to municipality i, based on self-reported municipality of residence five years prior in the 2000 Brazilian Census, and D_o is the average monthly SPEI below -1 for municipality r from 2000 to 2009, reversed to reflect drought severity. To reinforce the rural nature of these shocks and strengthen the exclusion restriction, we compute D_r exclusively based on non-urban areas, excluding built-up zones in each origin municipality. This adjustment isolates droughts affecting agricultural land and rural populations, which are the main sources of climate-induced migration, consistent with Busso and Chauvin (2025).This instrument captures the intensity of historical climate-induced migration pressures.

Shift-share instrumental variables have a long tradition in applied economics, dating back to the foundational contributions of Bartik (1991) and Autor et al. (2013). Recent developments have clarified the conditions under which these instruments yield valid causal inference. Two main frameworks have emerged: one requires the exogeneity of the share component (Goldsmith-Pinkham et al., 2020), while the other, more relevant to our context, relies on the exogeneity of the shocks, treating shares as predetermined (Borusyak et al., 2022). We adopt the latter approach, leveraging historical exposure to droughts between 2000 and 2009, as measured by the SPEI, as our source of exogenous variation. Since these climatic shocks predate the outcome period and originate in rural sending regions, they are plausibly orthogonal to unobserved determinants of productivity in the destination municipalities between 2010 and 2022. This design allows us to capture migration-driven demographic pressure rather than endogenous shifts in local labor market conditions.

A key identifying assumption is that historical migration patterns, measured by $s_{i,r}$ from the 2000 Census, remain predictive of climate-induced flows during the 2000–2009 period. While strong, this assumption is supported by the persistence of internal migration channels in Brazil, particularly from drought-prone semiarid regions, which have long been embedded in stable economic networks (da Cunha, 2018; Corbi et al., 2025a). Moreover, by restricting the climate shocks to rural origin areas, we limit confounding from urban transformations that might otherwise disrupt historical trajectories.

Nonetheless, a potential concern is that migrants from drought-affected areas may differ systematically in unobserved traits, such as skills, income, or behavioral profiles, that could influence productivity at destination through channels unrelated to population size. To mitigate this, we include a rich set of baseline controls capturing initial productivity levels, demographic structure, and geographic access. These controls account for selective migration and potential confounding from historical migration dynamics.

Another concern is that persistent characteristics of origin regions, such as low institutional quality or educational attainment, might be correlated with the climate shocks and, through long-term influence, affect productivity in destination areas. Our empirical strategy addresses this by relying exclusively on rural-origin shocks and controlling for initial conditions in destination municipalities.

The validity of the instrument ultimately depends on two criteria: relevance and exogeneity. We confirm the former empirically by showing that historical migrant flows from drought-affected areas predict population decline in receiving municipalities. For exogeneity, we rely on the assumption that the SPEI-based shocks are orthogonal to unobserved determinants of productivity, conditional on controls and fixed effects. This aligns with the "shock exogeneity" framework formalized by Borusyak et al. (2025).

To reinforce the credibility of this assumption, we compute the inverse Herfindahl-Hirschman Index (HHI) of the migration shares. The value of 348 suggests that most municipalities are exposed to a broad set of origin regions, satisfying the law of large numbers condition required for valid shift-share identification. Together, these design choices ensure that the variation we exploit stems from exogenous environmental shocks rather than from economic processes endogenous to local productivity.

Urban sprawl is instrumented using the average growth in built-up area between 2000 and 2009 among spatial neighbors of municipality i, based on satellite imagery from MapBiomas. Specifically:

$$Z_i^{\text{Sprawl}} = \frac{1}{N_i} \sum_{j \in \mathcal{N}_i,\, j \neq i} \text{Sprawl}_{j,2000\text{–}2009}, \tag{7}$$

where $\mathcal{N}_i$ denotes the set of spatial neighbors of i, defined using a k-nearest neighbors (knn) spatial weights matrix with k=4, based on the geographic distance between municipal centroids. We explicitly exclude municipality i from its own neighborhood to ensure that the instrument captures only exogenous variation from surrounding land-use dynamics. This formulation reflects the idea that urban expansion in nearby areas may influence sprawl in i through spatial diffusion, policy imitation, or shared infrastructure networks while remaining orthogonal to unobserved productivity shocks once spatial controls are included.

The relevance of this instrument stems from spatial diffusion in land-use practices, infrastructure investments, and policy spillovers. Its exogeneity is supported by the use of urban expansion (2000–2010) in neighboring municipalities, excluding municipality i itself. This strategy builds on the literature on spatial instruments and spillover-based identification in urban economics (Duranton et al., 2011; Koster and Rouwendal, 2017).

We estimate the following first-stage regressions:

$$\text{Shrinkage}_i = \gamma_0 + \gamma_1 Z_i^{\text{Shrinkage}} + \delta_1^\top X_i + \eta_i \tag{8}$$

$$\text{UrbanSprawl}_i = \pi_0 + \pi_1 Z_i^{\text{Sprawl}} + \delta_2^\top X_i + \nu_i \tag{9}$$

The fitted values $\widehat{\text{Shrinkage}}_i$ and $\widehat{\text{UrbanSprawl}}_i$ are then interacted to generate the instrumented version of the interaction term:

$$\widehat{\text{Interaction}}_i = \widehat{\text{Shrinkage}}_i \times \widehat{\text{UrbanSprawl}}_i \tag{10}$$

A final identification concern relates to the use of the interaction term $ShrinkageSprawl$ as a treatment. Even with separate instruments for each component, causal inference for the interaction requires stronger assumptions. Specifically, the exclusion restriction must also hold for the product of the instruments, meaning that the interaction between the shift-share shock and the spatially lagged sprawl affects productivity only through the endogenous interaction term (Wooldridge, 2010; Bun and Harrison, 2019). While this assumption cannot be directly tested, it is plausible in our context. The climate-based shift-share instrument captures historical droughts in rural areas of origin, while the sprawl instrument reflects urban expansion in neighboring municipalities during the previous decade. These sources of variation are geographically and causally distinct, making it unlikely that their interaction influences productivity through channels other than the interaction of shrinkage and sprawl. Still, we acknowledge that this exclusion restriction is more demanding, and interpret our results with the appropriate caution. We further reinforce the credibility of our estimates through robustness checks and placebo tests.

The second stage replaces the endogenous regressors in Equation (1) with their fitted counterparts:

$$\Delta \log(Y_i) = \beta_0 + \beta_1 \widehat{\text{Interaction}}_i + \theta^\top X_i + \varepsilon_i \tag{11}$$

Standard errors are clustered at the municipality level to allow for arbitrary heteroskedasticity and spatial correlation [7]. The identification assumption is that, conditional on controls

[7] In shift-share research designs, standard errors require careful correction due to the potential correlation in the constructed instrument across observational units that share exposure to common shocks. This issue has been extensively documented by Adão et al. (2019), who develop the AKM estimator to account for such dependence, and by Goldsmith-Pinkham et al. (2020), who emphasize that identification in Bartik-type instruments arises from variation in the shift component, necessitating clustering at the level of the shock. More recently, Borusyak et al. (2022) propose a generalized approach to exposure-robust standard errors that remains valid even with high-dimensional shifts.

In our case, however, each observational unit (municipality i) is exposed to a unique set of idiosyncratic shocks $shift_s$—specifically, climate shocks (SPEI) measured at the municipality-of-origin level (s) between 2000 and 2009, weighted by pre-determined migration shares. Because the shift component is disaggregated at the municipal level and varies across i, there is no overlap in the shock exposure between municipalities. This structure reduces or eliminates the cross-sectional correlation in the instrument that typically motivates the use of exposure-robust inference, as shown by Borusyak et al. (2022).

Nevertheless, we conducted a robustness exercise by clustering standard errors at the level of the dominant region of origin for each municipality, defined as the region (*microregion*) that contributed the most to the shift-share instrument (i.e., the region with the highest cumulative $share_{is} \cdot shift_s$ for each i). This approach follows the practical strategy used by Busso and Chauvin (2025) to account for possible latent correlation in exposure across space. Our results show that clustering standard errors in this way yields minimal changes in standard errors and preserves the statistical significance of the estimated coefficients. We interpret this as and fixed effects, the instruments affect productivity only through their impact on shrinkage and urban sprawl, not directly [8].

3.3 Alternative Instruments for Shrinkage

To reinforce the causal validity of our empirical strategy, we follow recent contributions in the literature (Imbert et al., 2022; Imbert and Ulyssea, 2022; Corbi et al., 2025b) and construct an alternative instrument based on international agricultural price shocks. The idea is that exogenous fluctuations in global commodity markets affect agricultural income in rural regions of Brazil, thereby altering internal migration flows which, in turn, influence the probability of population decline or growth in receiving municipalities.

International monthly price data were obtained from the *World Bank Commodities Database*, covering the period 1970–2009. We consider 11 agricultural commodities: bananas, cocoa, coffee, cotton, maize, oranges, rice, soybeans, sugar, tobacco, and wheat. This set follows established practice in the literature, differing only from Imbert and Ulyssea (2022) by excluding wood. For each crop c and month m, we compute a price shock ε_{cm} as the residual of an AR(1) process.

The first step consists of computing a price-shock index at the level of each origin municipality o. For each crop c, we use the historical agricultural structure observed in the 1980 by Municipal Agricultural Survey. The aggregate price shock is defined as:

$$s_o^{prices} = \sum_m \sum_c \pi_{oc} \, \varepsilon_{cm}, \tag{12}$$

where π_{oc} denotes the share of crop c in the total agricultural value of municipality o in 1980, and ε_{cm} is the international price shock for crop c in month m. This index captures the historical exposure of each origin municipality to external agricultural price fluctuations.

In the second step, we combine these shocks with pre-determined migration networks observed in the 2000 Population Census. For each destination municipality i, the instrument is defined as:

$$Z_i^{prices} = \sum_o m_{io}^{2000} \, s_o^{prices}, \tag{13}$$

where m_{io}^{2000} is the fraction of migrants in municipality i who originated from municipality o in 2000. This shift-share design exploits (i) exogenous agricultural price shocks in origin regions and (ii) predetermined migration linkages that govern the relative exposure of each destination municipality.

An important caveat raised in the literature is that price shocks to certain commodities, most notably soybeans, may trigger broader structural transformations in local economies (Bustos et al., 2016). In such cases, the price shock may not only affect migration pressures but also capture changes in productive specialization. To address this concern, we conduct robustness exercises in which soybeans are excluded from the basket of commodities used to construct the price-shock index, following the approach suggested by Corbi et al. (2025b).

By adopting this strategy, we provide a source of exogenous variation that is distinct from climate-based shocks and grounded in global market dynamics. In line with the established literature on migration and regional economics (Imbert et al., 2022; Imbert and Ulyssea, 2022;

evidence that our main estimates are robust to concerns about inference in shift-share designs.

[8] This is a preliminary version of a working paper. Ongoing work includes robustness checks using economics-based shift-share instruments, including employment and price shocks, to further evaluate the strength and validity of our identification strategy.

Corbi et al., 2025b), combining alternative instruments strengthens the credibility of our identification strategy and ensures that our results are not driven by a specific set of complier municipalities.

3.4 Double Machine Learning (DML)

As a robustness check to our 2SLS estimates, we adopt a Double Machine Learning (DML) strategy with LASSO regularization, following Chernozhukov et al. (2018). This method is particularly useful when dealing with a high-dimensional set of controls, where the risk of overfitting or functional form misspecification in the first stage is non-negligible.

Rather than specifying the first-stage equations parametrically, we allow for flexible model selection by estimating them via LASSO. In particular, we estimate the first stage of shrinkage using a LASSO-penalized regression and that of urban sprawl using LASSO in a linear model. The LASSO estimator selects a sparse subset of variables that best predict endogenous regressors by minimizing a penalized objective function. For the shrinkage equation, the LASSO solves the following:

$$\hat{\beta}^{LASSO} = \arg\min_{\beta} \left\{ \frac{1}{n} \sum_{i=1}^{n} \log\left(1 + e^{-y_i x_i^{\top}\beta}\right) + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \tag{14}$$

where y_i is the shrinkage or urban sprawl indicator , x_i includes the drought-based instrument and all controls, and λ is the penalty parameter selected via cross-validation. This procedure allows us to identify the most predictive covariates and avoid overfitting, particularly in the presence of many potentially irrelevant controls.

We then use the fitted values from these LASSO regressions in the second stage, maintaining the same identification strategy as in the baseline 2SLS. The goal is not to improve point estimates per se, but to assess the robustness of the causal effect to a more flexible, data-driven model selection procedure in the first stage.

This approach enhances credibility by ensuring that the exclusion restriction is not violated due to misspecified or omitted first-stage covariates. It is particularly relevant in urban and regional applications, where spatial, socioeconomic, and infrastructural covariates often interact in complex, high-dimensional ways.

4 Results

4.1 Descriptive facts

To first characterize the urban economies in our study, we begin our empirical analysis by examining the summary statistics of the main variables used in the regression models. Table B2 in Appendix B reports descriptive statistics for the full sample of 5,515 Brazilian municipalities. The dependent variable, the formal variation in labor productivity between 2010 and 2022, averages 1.09 with a standard deviation of 0.26. Approximately 43% of the municipalities experienced a population decline during the period, as captured by the shrinkage dummy. On average, municipalities exhibited high urban sprawl and experienced mild climatic stress, with an average reverse SPEI of <0.03, indicating slightly drier than normal conditions. The mean of neighboring urban expansion is similar in magnitude. Additional controls such as initial productivity, population size, infrastructure, and environmental conditions exhibit substantial variation throughout space, underscoring the importance of conditioning these covariates in subsequent analysis. In particular, per capita exposure to fine particulate matter (pm25 per capita) is extremely low in absolute terms, reflecting the fact that values are expressed in scientific notation and normalized by population. However, it still displays considerable cross-sectional variation that may influence health-related productivity effects.

4.2 First-stage Results

Panel A of Table 1 reports the first-stage estimates where the dependent variable is an indicator for whether a municipality experienced population shrinkage between 2010 and 2022. The coefficient on the SPEI-based shift-share instrument is negative and statistically significant across all specifications. This instrument combines exogenous variation in drought conditions in migrant-sending municipalities with historical migration linkages to each municipality of destination. A higher value of the instrument thus reflects a stronger exposure to potential inflows of climate-displaced migrants.

The negative sign suggests that municipalities with stronger historical ties to drought-affected regions were less likely to experience population decline, plausibly due to their role as preferred destinations for displaced individuals. This interpretation is consistent with the migration literature emphasizing that climate shocks often trigger selective migration flows toward urban or economically connected regions (Cattaneo and Peri, 2016; Busso and Chauvin, 2025). In particular, Busso and Chauvin (2025) shows that droughts in Brazil increase emigration rates significantly, especially toward municipalities with preexisting migratory channels. Our findings support the view that receiving municipalities benefit demographically from such displacement patterns, thereby attenuating their risk of shrinkage.

To further assess the validity of the shift-share instrument, we examine its correlation with a set of pre-treatment covariates. As shown in Appendix C (Table 8), the instrument is moderately correlated with economic variables such as employment in R&D-intensive sectors, sectoral GDP participation, and total population. These associations reflect the historical geography of migratory patterns and climate exposure. Importantly, the instrument shows weak or negligible correlation with locational, geographic, and environmental characteristics,such as road density, pollution levels, or distance to Brasília, reducing concerns about omitted variable bias. This result support the plausibility of the exclusion restriction and suggest that variation in the instrument arises primarily through climate-induced migratory pressure rather than structural differences in municipal characteristics.

Importantly, the F-statistics across specifications range from 4.1 to over 20, with the full model (column 5) yielding an F-statistic of 20.7, well above the conventional threshold of 10, indicating that weak instrument concerns are mitigated. Moreover, the use of lagged migration shares and pre-treatment climate shocks ensures that the identifying variation is plausibly exogenous to contemporaneous unobservables affecting population dynamics. This aligns with recent best practices in the shift-share literature (Borusyak et al., 2025).

Panel B shows the first-stage estimates for urban sprawl, instrumented using a spatial lag of urban expansion in neighboring municipalities. The coefficient on $\log(\text{W\_sprawl}_{t-1})$ is consistently large, positive, and highly significant across all specifications, with F-statistics well above conventional thresholds (minimum of 165), confirming the strength of the instrument. This result is consistent with the idea that urban expansion is spatially contagious, with local land-use changes being influenced by dynamics in nearby areas.

Taken together, these results provide strong support for the relevance and validity of our instruments. The sources of variation are both theoretically grounded and empirically powerful, lending credibility to the subsequent second-stage estimates.

Table 1: First-Stage Estimates for Instrumental Variables

	(1)	(2)	(3)	(4)	(5)
Panel A: Effects of Weather Shocks on Shrinkage					
SPEI-based shift-share IV	−0.027***	−0.012**	−0.023**	−0.045***	−0.045***
	(0.010)	(0.006)	(0.009)	(0.009)	(0.010)
States FE	No	Yes	Yes	Yes	Yes
Socioeconomics	No	No	Yes	Yes	Yes
Geography and Environment	No	No	No	Yes	Yes
Infrastructure	No	No	No	No	Yes
F-statistic	6.80***	4.1*	5.5**	20.4***	20.7***
Observations	5,515	5,515	5,515	5,515	5,515
Panel B: Effects of Spatio-Temporal Lag Shocks on Urban Sprawl					
$\log(\text{W\_sprawl}_{t-1})$	0.349***	0.216***	0.212***	0.208***	0.208***
	(0.019)	(0.016)	(0.016)	(0.016)	(0.016)
States FE	No	Yes	Yes	Yes	Yes
Socioeconomics	No	No	Yes	Yes	Yes
Geography and Environment	No	No	No	Yes	Yes
Infrastructure	No	No	No	No	Yes
F-statistic	324.4***	177.3***	170.9***	165.3***	168.1***
Observations	5,515	5,515	5,515	5,515	5,515

Notes: Each column reports estimates from a first-stage regression corresponding to the two endogenous variables in the 2SLS specification. Standard errors are clustered at the municipality level and reported in parentheses. All specifications include the shift-share instrument in Panel A and the spatial lag of urban sprawl in Panel B. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2 presents the second-stage estimates examining the causal effects of population shrinkage, urban sprawl, and their interaction on labor productivity growth across municipalities between 2010 and 2022. The OLS estimates in column (1) suggest a modest but significant negative interaction between shrinkage and sprawl. However, the IV estimates reveal substantially larger effects once endogeneity concerns are addressed.

Table 2: Second-Stage Estimates: Effects of Shrinkage and Urban Sprawl on Productivity Growth

	OLS	IV (1)	IV (2)	IV (3)	IV (4)	IV (5)
Shrinkage	0.099***	6.377***	4.984***	1.062***	0.570***	0.534***
	(0.029)	(2.110)	(1.867)	(0.183)	(0.134)	(0.131)
Urban Sprawl	0.014	3.588***	2.158*	0.591***	0.398***	0.340***
	(0.024)	(1.132)	(1.120)	(0.155)	(0.100)	(0.104)
Interaction	−0.130***	−8.073***	−4.443*	−1.196***	−0.798***	−0.661***
	(0.036)	(2.636)	(2.372)	(0.272)	(0.156)	(0.167)
State FE	No	No	Yes	Yes	Yes	Yes
Socioeconomic	No	No	No	Yes	Yes	Yes
Geographs and Env.	No	No	No	No	Yes	Yes
Infrastructure	No	No	No	No	No	Yes
R-squared	0.008	0.006	0.314	0.308	0.319	0.320
Observations	5,515	5,515	5,515	5,515	5,515	5,515

Notes: This table reports OLS and 2SLS estimates of the effects of shrinkage, urban sprawl, and their interaction on formal labor productivity growth between 2010 and 2022. All IV models use the SPEI-based shift-share and the spatial lag of urban sprawl as instruments. Standard errors clustered at the municipality level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The coefficient on *Shrinkage* is consistently positive and significant across all specifications. While this may appear counterintuitive at first glance, it likely reflects a conditional selection effect. When population declines, outmigration tends to involve less productive individuals or firms, mechanically raising average productivity among those who remain. This result aligns with mechanisms observed in shrinking regions, where consolidation and selection can intensify measured efficiency (Duranton and Puga, 2004).

Similarly, the positive coefficient on *Urban Sprawl* may capture land reallocation and congestion relief in municipalities not undergoing demographic pressure. In stable or growing areas, physical expansion can facilitate access to land, reduce density-related frictions, and support economic decentralization, especially when supported by infrastructure (Glaeser and Kahn, 2004; Burchfield et al., 2006).

However, it is important not to interpret these main effects in isolation. The core mechanism we identify is inherently interactive. When shrinkage and sprawl co-occur, their individual advantages are offset by spatial inefficiencies. This is captured by the consistently negative and statistically significant coefficient on the interaction term, which suggests that uncoordinated expansion in shrinking cities imposes a net productivity cost. Our interpretation thus emphasizes the joint dynamics of demographic contraction and land-use change, rather than the marginal effects of each process on its own.

To further examine the selection-based interpretation of the positive shrinkage coefficient, we calculate the proportion of municipalities that experienced above-median productivity growth, conditional on shrinkage status (Table 3). Among shrinking municipalities, 48.2% recorded above-median productivity growth, compared to 51.3% among non-shrinking municipalities. While the difference is small, it suggests that shrinkage is not systematically associated with economic decline and may reflect selective outmigration in some regions.

Table 3: Share of Municipalities with Above-Median Productivity Growth by Shrinkage Status

Shrinkage Status	Total Municipalities	Above Median	Share (%)
Not Shrinking	3,138	1,611	51.3
Shrinking	2,377	1,146	48.2

However, the key insight from our estimates lies in the negative and robust coefficient on the *Interaction* term. Across all IV specifications, the interaction between shrinkage and urban sprawl is strongly negative and statistically significant, with magnitudes far exceeding those from OLS. This result indicates that the benefits of urban expansion do not materialize, and may in fact reverse, in the context of demographic decline. In shrinking municipalities, sprawl appears to dilute density without generating offsetting gains, thereby undermining agglomeration economies and raising per capita infrastructure burdens (Yang et al., 2022).

This finding validates the theoretical prediction derived in Section 2, where urban expansion in the presence of population decline reduces effective density and thus productivity, holding capital intensity constant. As Equation (4) shows, the marginal effect of expanding urban area (S_i) is negative when population (L_i) is stagnant or falling. The empirical results thus confirm the presence of spatial inefficiencies: in shrinking cities, continued expansion leads to productivity losses through under-utilization of infrastructure, fragmentation of economic activity, and erosion of agglomeration benefits.

The consistency of the interaction effect across specifications also reinforces the argument that the spatial-demographic mismatch is a first-order determinant of urban productivity. Even after controlling for geography, infrastructure, and socioeconomic traits, the penalty of sprawling under shrinkage persists. This suggests that the interaction is not merely capturing omitted heterogeneity, but reflects a fundamental tension in spatial equilibrium dynamics, i.e., cities that expand without growing population-wise are prone to inefficient urban forms.

The coefficient of –0.66 on the interaction term implies an economically meaningful effect. Average productivity growth between 2010 and 2022 was 1.09 log-points (about 8.6 percent).

The interaction effect corresponds to a reduction of 61 percent relative to this mean, i.e. a loss of 5.2 percentage points. In practice, while the typical municipality experienced productivity gains of about 8.6 percent, shrinking but sprawling municipalities achieved growth of only 3.4 percent over the period.

4.3 Placebo Test

To assess the robustness of our main results, we conduct a placebo exercise to test whether the observed negative interaction effect between urban shrinkage and sprawl could arise spuriously due to model misspecification or functional form artifacts.

In this placebo test, we redefine the shrinkage variable by assigning a value of one to municipalities that experienced population growth between 2010 and 2022, rather than shrinkage. This reversed coding allows us to evaluate whether the negative productivity effects previously attributed to demographic decline also emerge when applied to expanding cities. If our original findings are driven by a true causal mechanism associated with shrinkage, we expect the interaction term to lose significance or change sign in this placebo scenario.

Table 9 reports the results. In stark contrast to the main specification, the interaction term is now positive and statistically significant at the 1 percent level, while both main effects, shrinkage (now redefined as growth) and sprawl, enter with negative signs. These patterns suggest that the negative interaction observed in the main specification is not mechanically induced by model structure or variable scaling. Rather, the productivity penalty appears to be uniquely associated with sprawl occurring under demographic contraction.

This placebo result strengthens the causal interpretation of our findings. It supports the notion that urban expansion is not inherently detrimental to productivity; indeed, it may even be beneficial in growing municipalities, but its effects become counterproductive when combined with population decline. Such a pattern is precisely what one would expect from a framework in which declining density weakens agglomeration economies and raises inefficiencies in spatially overextended urban environments.

4.4 Bootstrap Simulation

As a robustness check, we assess the sampling distribution of the interaction coefficient using 1,000 bootstrap replications of the 2SLS procedure. Figure 2 shows the empirical distribution of the estimated coefficient on *Shrinkage* $\times$ *Sprawl*, which is centered around -0.540 and remains statistically significant across replications.

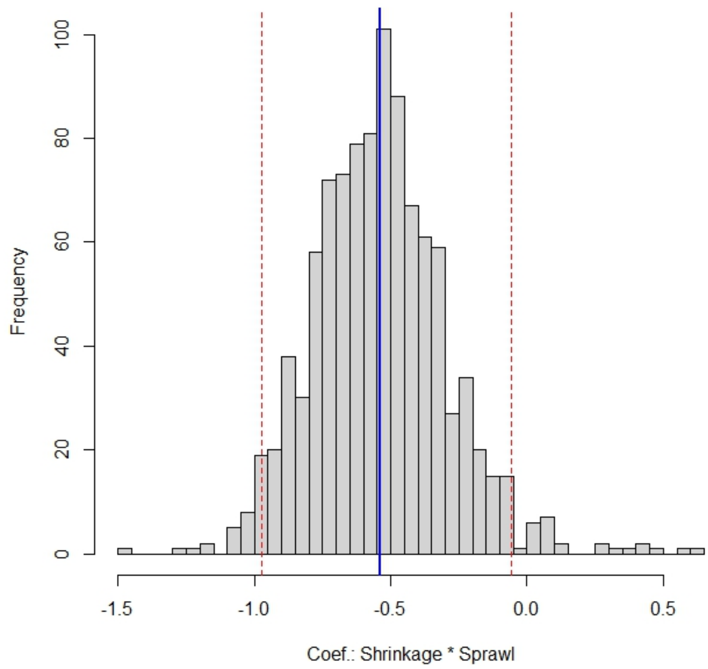


Figure 2: **Bootstrap Distribution of the Interaction Effect (*Shrinkage* $\times$ *Sprawl*)**
Distribution based on 1,000 bootstrap replications of the 2SLS estimator. In each iteration, the two-stage model is re-estimated on a resampled dataset (N = 5,515). The average coefficient is **-0.540**, with a bootstrap standard error of **0.245**, and a 95% confidence interval of [**-0.972, -0.058**].

4.5 Heterogeneous Effects by Sector

To investigate whether the interaction between urban shrinkage and urban sprawl affects sectors differently, we estimate Equation (2) separately for agriculture, manufacturing, and services. Table 4 presents the results.

Across all sectors, both Shrinkage and Urban Sprawl display positive and statistically significant main effects, suggesting that, when considered separately, each process is associated with higher labor productivity. However, the interaction term *Shrinkage* $\times$ *Urban Sprawl* is negative and significant at the 1% level in all cases, indicating that the productivity-enhancing effects of either dynamic are nullified or reversed when they co-occur. This finding reinforces our baseline results and highlights the importance of spatial-demographic coordination in urban policy.

The magnitude of the interaction is particularly large in the agricultural sector ($\beta = -2.52$). While both shrinkage and sprawl individually increase productivity, likely due to land reallocation and improved market access, their joint presence undermines these gains. A plausible explanation is that spatial expansion in shrinking areas leads to inefficient land fragmentation and higher per capita infrastructure costs, limiting scale economies in agri-logistics and input delivery. This pattern aligns with evidence from China, where Yang et al. (2022) show that built-up area expansion in shrinking cities leads to resource misallocation and declining productivity, particularly in peripheral zones with low density.

In manufacturing, the negative interaction effect ($\beta = -1.23$) also reflects spatial inefficiency. As industries rely on dense input-output networks and labor pools, the combination of demographic contraction and spatial dispersal erodes agglomeration economies. This interpretation is consistent with the findings of Fallah et al. (2011) , who document that low-density sprawl reduces productivity in U.S. manufacturing hubs, especially when decoupled from population growth.

The services sector presents a slightly attenuated, but still significant, interaction effect ($\beta = -0.93$). Unlike agriculture or industry, services can partially decouple from physical infrastructure and may benefit from digital connectivity. However, even in this context, spatial mismatch in shrinking cities appears to hinder efficiency. Dispersed urban form likely inflates travel times, reduces face-to-face interactions, and fragments consumer markets, mechanisms that are important to productivity in knowledge-intensive and client-based services (Duranton and Puga, 2004; Glaeser and Gottlieb, 2009).

These sectoral results underscore a central mechanism of our paper, that the detrimental effects of urban sprawl in shrinking cities operate through the erosion of spatial efficiency and density-dependent externalities. When cities lose population but continue to expand physically, they not only incur higher per capita infrastructure costs but also dilute the economic forces that sustain productivity. This is especially problematic in contexts where land-use and population dynamics are misaligned due to governance failures or path-dependent urban planning.

These findings add a new dimension to the literature on urban decline by highlighting the sector-specific channels through which spatial-demographic misalignment undermines productivity. They suggest that effective spatial containment policies and adaptive reuse strategies are essential to prevent long-run economic decline in shrinking urban systems (Haase et al., 2014b; Pallagst et al., 2017).

Table 4: Heterogeneous Effects of Shrinkage and Urban Sprawl on Labor Productivity, by Sector

	Agriculture	Manufacturing	Services
Shrinkage	1.949***	0.977***	0.792***
	(0.410)	(0.184)	(0.218)
Urban Sprawl	1.658***	0.564***	0.521***
	(0.316)	(0.143)	(0.169)
Shrinkage × Urban Sprawl	-2.516***	-1.225***	-0.929***
	(0.528)	(0.238)	(0.287)
State Fixed Effects	Yes	Yes	Yes
Socioeconomic	Yes	Yes	Yes
Geographic and Environmental	Yes	Yes	Yes
Infrastructure	Yes	Yes	Yes
Adjusted R^2	0.448	0.215	0.135
N	5,515	5,515	5,515

Note: Clustered standard errors at the municipality level in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Robustness checks further support these sectoral findings. In Appendix C.3 (Table 10), we present placebo regressions in which the treatment variable is redefined as population growth rather than decline. The results confirm that the negative interaction effects observed in shrinking areas do not arise in growing municipalities. Additionally, bootstrap simulations for each sector (Appendix C.4) validate the statistical significance of the interaction terms, with all estimated distributions centering around the original coefficients and displaying narrow confidence intervals (Figures C.2, C.3 and C.4).

4.6 Robustness Check: DML

As an additional robustness check, we implement a DML strategy based on LASSO, following the framework developed by Chernozhukov et al. (2018). This approach addresses two critical concerns in our empirical setting. First, the high dimensionality of control variables, including socioeconomic, geographic, and infrastructural factors. Second, the risk of model misspecification in conventional IV frameworks.

In particular, the determinants of urban shrinkage are likely to involve complex interactions among a large set of observables, which standard parametric specifications cannot fully capture. While the 2SLS estimator used in the baseline analysis is consistent under correct specification, its performance can deteriorate in high-dimensional settings or when functional form assumptions are violated. The LASSO-DML estimator mitigates these concerns by performing data-driven covariate selection and regularization in the first stage, thus allowing for a more flexible and robust estimation of treatment effects.

Importantly, the DML framework provides valid post-selection inference even when the number of controls is large relative to the sample size, and it guards against overfitting. This is particularly relevant in our case, where a rich set of contextual factors may jointly influence both urban shrinkage and urban sprawl.

To address concerns about transparency and interpretability, we report the variables selected by the LASSO procedure in the first-stage regressions for both endogenous regressors (shrinkage and sprawl). Appendix Table 10 shows that the algorithm retains a rich set of covariates, including labor market indicators (e.g., R& D employment, GDP shares), demographic factors (e.g., household size, population), and spatial variables (e.g., distance to capitals, altitude, road infrastructure). This selection provides evidence that the model controls for multiple potential confounders, reinforcing the credibility of the identification strategy and helping illuminate the mechanisms behind population decline and urban expansion.

Results from the DML estimation (Table 5) confirm our main findings. The interaction term between *Shrinkage* and *Urban Sprawl* remains negative and statistically significant, with a point estimate of -0.960 (standard error: 0.243). While slightly attenuated compared to the 2SLS estimate, the magnitude and direction are consistent, reinforcing the interpretation that sprawling in shrinking areas depresses local productivity.

Table 5: Double Machine Learning (DML) Estimates and First-Stage Relevance Tests

	ML
Labor Productivity	
Shrinkage	9.905***
	(2.519)
Urban Sprawl	0.450***
	(0.131)
Shrinkage × Urban Sprawl	-0.960***
	(0.243)
Controls	Yes
R^2	0.320
N	5,515
Effects of Weather Shocks on Shrinkage	
SPEI-based shift-share IV	-0.044***
	(0.017)
Controls	Yes
N	5,515
Effects of Spatio-temporal Lag Shocks on Urban Sprawl	
$\log(\text{W\_sprawl}_{t-1})$	0.205***
	(0.016)
Controls	Yes
N	5,515

Note: Clustered standard errors at the municipality level in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

We also test the predictive performance of the first-stage models using standard out-of-sample validation metrics (Table 6). For both treatment variables, the LASSO and conventional estimators yield nearly identical values for MSE, RMSE, and MAE. This suggests that model selection does not significantly distort the first-stage predictions. However, the explanatory power for *Urban Sprawl* is notably higher ($R^2 = 0.38$) than for *Shrinkage* ($R^2 = 0.11$), which is consistent with our conceptual framework where sprawl is more spatially persistent and easier to predict via lagged spillovers.

To enhance transparency, we examined the variables selected by the LASSO procedure in the first-stage regression for *shrinkage*. Using the value of $\lambda_{\min} = 4.68 \times 10^{-5}$ that minimizes the cross-validated mean squared error, the model selects a rich set of covariates including the main instrument (*shift-share*), sectoral indicators (e.g., *R&D employees*, *GDP share*), demographic and geographic controls (e.g., *Population in 2010*, *Latitude*, *Distances*), and multiple state fixed effects. These results confirm that the selection captures both structural economic characteristics and spatial heterogeneity. A full list of selected covariates and coefficients is reported in Appendix Table 11.

Overall, the DML results corroborate our baseline conclusions and rule out the possibility that our findings are driven by omitted variable bias or overfitting. They reinforce the view that spatial inefficiency arises when physical urban expansion occurs in the absence of demographic growth, and they strengthen the causal interpretation of our core estimates.

Table 6: Model Validation Metrics – First-Stage Predictions

Métricas	Shrinkage		Urban Sprawl	
	OLS	Lasso	OLS	Lasso
MSE	0.2172	0.2172	0.0071	0.0072
RMSE	0.4661	0.4661	0.0848	0.0848
MAE	0.4391	0.4392	0.0606	0.0607
R^2	0.1142	0.1140	0.3808	0.3800

4.7 Alternative Instrument - Price Shocks

Table 7 reports the estimates obtained when using international agricultural price shocks as an alternative instrument for shrinkage, following the robustness strategy described in Section 3.3. Panel A shows the second-stage results for labor productivity, while Panel B presents the first-stage relationship between price shocks and shrinkage.

The results in Panel A confirm the main findings obtained with our baseline specification. The interaction term between shrinkage and sprawl remains strongly negative and statistically significant across all specifications. This reinforces the interpretation that urban expansion in shrinking cities erodes agglomeration economies and generates spatial inefficiencies, thereby lowering productivity. The magnitude of the interaction effect is stable and close to baseline estimates, indicating that the adverse impact of sprawling under demographic decline is not driven by a specific choice of instrument.

Panel B provides evidence on the first-stage relationship between agricultural price shocks and shrinkage. Overall price shocks at $t-1$ are positively associated with shrinkage, with first-stage F-statistics exceeding conventional thresholds for instrument relevance. When decomposed, negative shocks significantly reduce the likelihood of shrinkage, while positive shocks increase it. This pattern is consistent with the mechanism whereby adverse price shocks in origin regions trigger out-migration toward destination municipalities, attenuating their risk of demographic contraction, whereas favorable price shocks in origins retain population and thereby increase the probability of shrinkage in destinations. As an additional robustness check, we exclude soybeans from the commodity basket given their well-documented role in structural transformation in Brazil (Bustos et al., 2016; Corbi et al., 2025a). The results remain stable, with the instrument retaining strong predictive power.

Taken together, these results demonstrate that the negative productivity effects of urban sprawl in shrinking cities are robust to the use of agricultural price shocks as an alternative instrument. This triangulation of exogenous sources, climatic shocks and international price shocks, strengthens the credibility of our causal interpretation.

Table 7: Labor Productivity and Price Shocks as Alternative Instrument for Shrinkage

	(1)	(2)	(3)	(4)
Panel A: Labor Productivity				
Shrinkage	0.462***	0.460***	0.462***	0.463***
	(0.113)	(0.125)	(0.120)	(0.113)
Urban Sprawl	0.348***	0.353***	0.352**	0.349***
	(0.097)	(0.091)	(0.098)	(0.097)
Shrinkage × Urban Sprawl	-0.668***	-0.675***	-0.674***	-0.670***
	(0.150)	(0.154)	(0.152)	(0.150)
State FE	Yes	Yes	Yes	Yes
Socioeconomic controls	Yes	Yes	Yes	Yes
Geographic & environmental	Yes	Yes	Yes	Yes
Infrastructure	Yes	Yes	Yes	Yes
R-squared	0.321	0.321	0.321	0.321
Observations	5,539	5,539	5,539	5,539
Panel B: Effects of Price Shocks on Shrinkage				
Price Shocks	0.014***			
	(0.004)			
Negative Shocks		-0.006***		
		(0.001)		
Positive Shocks			0.004***	
			(0.001)	
Price Shocks (no soy)				0.015***
				(0.005)
State FE	Yes	Yes	Yes	Yes
Socioeconomic controls	Yes	Yes	Yes	Yes
Geographic & environmental	Yes	Yes	Yes	Yes
Infrastructure	Yes	Yes	Yes	Yes
F-statistic	10.8	15.6	15.4	10.4
Observations	5,539	5,539	5,539	5,539

Notes: This table reports IV estimates using a placebo outcome measured prior to treatment. Standard errors clustered at the municipality level are shown in parentheses. All models include both instruments and the full set of controls. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5 Conclusion

This paper examines how urban sprawl affects labor productivity in the context of population decline. Using a novel panel of Brazilian municipalities based on Census data from 2010 to 2022, we find that the interaction between urban expansion and demographic contraction significantly reduces productivity growth.

To address endogeneity, we implement a 2SLS strategy, instrumenting shrinkage with a climate-based shift-share variable derived from the SPEI and sprawl with spatial lags of urban expansion. These instruments are grounded in theory and satisfy the exclusion restrictions under plausible identifying assumptions. The results consistently indicate that urban sprawl undermines productivity, but only in municipalities experiencing population decline. Neither shrinkage nor sprawl alone generates comparable effects.

Our conclusions are supported by a series of robustness checks, including placebo tests, sector-specific regressions, an alternative estimation strategy using DML, and the use of international agricultural price shocks as an alternative instrument for shrinkage. The DML approach, based on LASSO, confirms the significance and stability of the main interaction term while providing transparency on variable selection. The price-shock strategy confirms the relevance of the instrument in the first stage and shows that our results hold under a distinct source of exogenous variation, thereby reinforcing the robustness of our causal interpretation. Together, these exercises validate our interpretation of the interaction as a structural inefficiency associated with uncoordinated spatial expansion under demographic trends.

This work contributes to the emerging literature on urban decline by identifying a key mechanism through which spatial form interacts with population dynamics. When cities continue to expand physically despite shrinking demographically, they dilute agglomeration economies, overextend infrastructure, and generate land-use mismatches. These inefficiencies result in persistent productivity penalties.

From a policy perspective, our findings call for a shift away from growth-oriented urban paradigms. In places where demographic decline is structural or irreversible, spatial containment becomes critical. Policymakers should prioritize adaptive reuse, infrastructure optimization, and compact urban design. Tools such as land banking, adaptive zoning, and infrastructure decommissioning can help realign spatial structures with population trajectories. Additionally, regional coordination is essential to curb spillover effects and prevent inefficient expansion from propagating across declining urban systems.

References

Adão, R. R., Kolesár, M., and Morales, E. M. (2019). Shift-share designs: Theory and inference. *American Economic Review*, 109(4):1393–1429.

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

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent*. Harvard University Press, Cambridge, MA.

Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The china syndrome: Local labor market effects of import competition in the united states. *American economic review*, 103(6):2121–2168.

Bartik, T. J. (1991). *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute for Employment Research, Kalamazoo, MI.

Batunova, E. and Gunko, M. (2018). Urban shrinkage: An unspoken challenge of spatial planning in russian small and medium-sized cities. *European Planning Studies*, 26(8):1580–1597.

Beguería, S., Latorre, B., Reig, F., and Vicente-Serrano, S. (2019). Global spei database. *Recuperado de https://spei. csic. es/database. html*.

Behnisch, M., Krüger, T., and Jaeger, J. A. (2022). Rapid rise in urban sprawl: Global hotspots and trends since 1990. *PLOS Sustainability and Transformation*, 1(11):e0000034.

Behrens, K., Duranton, G., and Robert-Nicoud, F. (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, 122(3):507–553.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *Review of Economic Studies*, 89(1):1–44.

Borusyak, K., Hull, P., and Jaravel, X. (2025). A practical guide to shift-share instruments. *Journal of Economic Perspectives*, 39(1):181–204.

Brueckner, J. K. (1987). The structure of urban equilibria: A unified treatment of the muth-mills model. In Mills, E. S., editor, *Handbook of Regional and Urban Economics*, volume 2, pages 821–845. Elsevier.

Bun, M. J. and Harrison, T. D. (2019). Ols and iv estimation of regression models including endogenous interaction terms. *Econometric Reviews*, 38(7):814–827.

Burchfield, M., Overman, H. G., Puga, D., and Turner, M. A. (2006). Causes of sprawl: A portrait from space. *The Quarterly Journal of Economics*, 121(2):587–633.

Busso, M. and Chauvin, J. P. (2025). Long-term effects of weather-induced migration on urban labor and housing markets. *Journal of Urban Economics*, 146:103739.

Bustos, P., Caprettini, B., and Ponticelli, J. (2016). Agricultural productivity and structural transformation: Evidence from brazil. *American Economic Review*, 106(6):1320–1365.

Cattaneo, C. and Peri, G. (2016). The migration response to increasing temperatures. *Journal of development economics*, 122:127–146.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., and Newey, W. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68.

Ciccone, A. and Hall, R. E. (1993). Productivity and the density of economic activity.

Ciccone, A. and Hall, R. E. (1996). Productivity and the density of economic activity. *American Economic Review*, 86(1):54–70.

Combes, P.-P., Duranton, G., and Gobillon, L. (2011). The identification of agglomeration economies. *Journal of Economic Geography*, 11(2):253–266.

Corbi, R., Ferraz, T., and Narita, R. (2025a). Internal migration and labor market adjustments in the presence of non-wage compensation. *Journal of Development Economics*, page 103534.

Corbi, R., Ferraz, T., and Sanches, F. M. (2025b). Internal migration and religious change: Evidence from brazil's evangelical rise. Technical Report 2025-10, Working Paper Series.

Crippa, M., Guizzardi, D., Pagani, F., Schiavina, M., Melchiorri, M., Pisoni, E., Graziosi, F., Muntean, M., Maes, J., Dijkstra, L., et al. (2024). Insights into the spatial distribution of global, national, and subnational greenhouse gas emissions in the emissions database for global atmospheric research (edgar v8. 0). *Earth System Science Data*, 16(6):2811–2830.

da Cunha, J. M. P. (2018). Internal migration in brazil over the past 50 years: (dis)continuities and ruptures. In Arretche, M., Lavinas, L., and Ferreira, F. H. G., editors, *Paths of Inequality in Brazil: A Half-Century of Changes*, pages 209–231. Springer International Publishing, Cham.

DNIT (2005). Departamento nacional de infraestrutura de transportes (dnit). paved roads network – georeferenced shapefile. Georeferenced shapefile of Brazil's paved highway network. Accessed July 2025.

Duranton, G., Gobillon, L., and Overman, H. G. (2011). Assessing the effects of local taxation using microgeographic data. *The economic journal*, 121(555):1017–1046.

Duranton, G. and Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Handbook of regional and urban economics*, volume 4, pages 2063–2117. Elsevier.

Duranton, G. and Puga, D. (2015). Urban land use. In Duranton, G., Henderson, J. V., and Strange, W. C., editors, *Handbook of Regional and Urban Economics*, volume 5B, pages 467–560. Elsevier.

Fallah, B. N., Partridge, M. D., and Olfert, M. R. (2011). Urban sprawl and productivity: Evidence from us metropolitan areas. *Papers in Regional Science*, 90(3):451–472.

Glaeser, E. L. and Gottlieb, J. D. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the united states. *Journal of Economic Literature*, 47(4):983–1028.

Glaeser, E. L. and Kahn, M. E. (2004). Sprawl and urban growth. *Handbook of Regional and Urban Economics*, 4:2481–2527.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in cities. *Journal of political economy*, 100(6):1126–1152.

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

Haase, A., Nelle, A., and Mallach, A. (2017). Representing urban shrinkage—the importance of discourse as a frame for understanding conditions and policy. *Cities*, 69:95–101.

Haase, A., Rink, D., Grossmann, K., Bernt, M., and Mykhnenko, V. (2014a). Conceptualizing urban shrinkage. *Environment and Planning A*, 46(7):1519–1534.

Haase, A., Rink, D., Grossmann, K., Bernt, M., and Mykhnenko, V. (2014b). Conceptualizing urban shrinkage. *Environment and Planning A*, 46(7):1519–1534.

IBGE (2023a). Censo demográfico 2010 e 2022. Accessed: 2024-03-01. Instituto Brasileiro de Geografia e Estatística.

IBGE (2023b). Instituto brasileiro de geografia e estatística (ibge) - population censuses of 2010 and 2022. Microdata from the Brazilian Population Census. Accessed July 2025.

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

Imbert, C. and Ulyssea, G. (2022). Firm dynamics and the labor market effects of rural-urban migration. Working paper.

IPEA (2024). Instituto de pesquisa econômica aplicada (ipea) - regional statistics and municipal indicators database. Socioeconomic and geographic indicators at the municipal level. Accessed July 2025.

Jarzebski, M. P., Elmqvist, T., Gasparatos, A., Fukushi, K., Eckersten, S., Haase, D., Goodness, J., Khoshkar, S., Saito, O., Takeuchi, K., et al. (2021). Ageing and population shrinking: implications for sustainability in the urban century. *Npj Urban Sustainability*, 1(1):17.

Koster, H. R. and Rouwendal, J. (2017). Historic amenities and housing externalities: evidence from the netherlands.

Kubik, Z. and Maurel, M. (2016). Weather shocks, agricultural production and migration: Evidence from tanzania. *The Journal of Development Studies*, 52(5):665–680.

Li, Y. and Chen, Z. (2023). Does transportation infrastructure accelerate factor outflow from shrinking cities? an evidence from china. *Transport Policy*, 134:180–190.

Lv, X., Wu, Z., and Sui, Y. (2024). Understanding the dynamics of urban shrinkage and the impact on innovation in china: A comprehensive analysis. *Journal of the Knowledge Economy*, pages 1–34.

Mallach, A., Haase, A., and Hattori, K. (2017). The shrinking city in comparative perspective: Contrasting dynamics and responses to urban shrinkage. *Cities*, 69:102–108.

MapBiomas (2023). Projeto mapbiomas – coleção 8 de mapas de cobertura e uso da terra do brasil. Satellite-based land use dataset. Accessed via: https://mapbiomas.org.

Martinez-Fernandez, C., Audirac, I., Fol, S., and Cunningham-Sabot, E. (2012). Shrinking cities: Urban challenges of globalization. In *OECD Regional Outlook 2012: Building Resilient Regions for Stronger Economies*, pages 215–267. OECD Publishing, Paris.

Mills, E. S. (1972). *Studies in the Structure of the Urban Economy*. Johns Hopkins University Press, Baltimore.

Muth, R. F. (1969). *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*. University of Chicago Press, Chicago.

Ortiz-Moya, F. (2020). Green growth strategies in a shrinking city: Tackling urban revitalization through environmental justice in kitakyushu city, japan. *Journal of Urban Affairs*, 42(3):312–332.

Pallagst, K., Wiechmann, T., and Martinez-Fernandez, C. (2017). *Shrinking Cities: International Perspectives and Policy Implications*. Routledge, London.

RAIS (2024). Annual Report of Social Information (RAIS) - Ministry of Labor and Employment of Brazil. Microdata from 2010 to 2022. Accessed via Brazilian Ministry of Labor and Employment (MTE).

Rao, Y., Wu, C., and He, Q. (2023). The antagonistic effect of urban growth pattern and shrinking cities on air quality: Based on the empirical analysis of 174 cities in china. *Sustainable Cities and Society*, 97:104752.

Rink, D., Couch, C., Haase, A., Krzysztofik, R., Nadolu, B., and Rumpel, P. (2014). The governance of urban shrinkage in cities of post-socialist europe: Policies, strategies and actors. *Urban Research & Practice*, 7(3):258–277.

Rosenthal, S. S. and Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In *Handbook of regional and urban economics*, volume 4, pages 2119–2171. Elsevier.

Sobrinho, E. M. G. and Azzoni, C. R. (2016). Potencial inovativo da indústria nas regiões brasileiras. *Revista Brasileira de Inovação*, 15(2):275–304.

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

Wang, L., Wang, J., and Qian, X. (2023). Does the bullet train exacerbate urban shrinkage? lessons from japan. *Papers in Regional Science*, 102(1):187–213.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Yang, Z., Zhu, Y., and Zhang, Y. (2022). Does urban shrinkage lower labor productivity? the role of spatial expansion. *Regional Science Policy & Practice*, 14:106–117.

Zivin, J. G. and Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review*, 102(7):3652–3673.

A Appendix

A.1: Derivation of the Labor Productivity Expression

We begin with a standard Cobb-Douglas production function augmented with agglomeration economies, where total output in municipality i is given by:

$$Y_i = A_i \left(\frac{L_i}{S_i}\right)^{\phi} K_i^{\alpha} L_i^{1-\alpha} \tag{15}$$

where: Y_i is total output; A_i is total factor productivity (TFP); L_i is the labor force; K_i is the capital stock; S_i is the urbanized area;$\phi > 0$ is the elasticity of agglomeration with respect to density; $\alpha \in (0, 1)$ is the elasticity of output with respect to capital.

The term $(L_i/S_i)^{\phi}$ captures the productivity gains from urban density. Our goal is to derive the expression for labor productivity, i.e., output per worker Y_i/L_i.

First, express the density component as a product of powers:

$$\left(\frac{L_i}{S_i}\right)^{\phi} = L_i^{\phi} \cdot S_i^{-\phi} \tag{16}$$

Substitute this into the original production function:

$$Y_i = A_i \cdot L_i^{\phi} \cdot S_i^{-\phi} \cdot K_i^{\alpha} \cdot L_i^{1-\alpha} \tag{17}$$

Combine the exponents of L_i:

$$L_i^{\phi} \cdot L_i^{1-\alpha} = L_i^{\phi+1-\alpha} \tag{18}$$

so that:

$$Y_i = A_i \cdot S_i^{-\phi} \cdot K_i^{\alpha} \cdot L_i^{\phi+1-\alpha} \tag{19}$$

Divide both sides of the equation by L_i to obtain labor productivity:

$$\frac{Y_i}{L_i} = A_i \cdot S_i^{-\phi} \cdot K_i^{\alpha} \cdot L_i^{\phi+1-\alpha-1} \tag{20}$$

$$= A_i \cdot S_i^{-\phi} \cdot K_i^{\alpha} \cdot L_i^{\phi-\alpha} \tag{21}$$

We arrive at the following expression for labor productivity:

$$\frac{Y_i}{L_i} = A_i \cdot S_i^{-\phi} \cdot K_i^{\alpha} \cdot L_i^{\phi-\alpha} \tag{22}$$

This formulation makes explicit the marginal effects of urban land use and population on productivity. In particular, a decrease in density due to urban sprawl (i.e., an increase in S_i) lowers productivity when $\phi > 0$, while capital deepening and agglomeration from larger L_i may offset such effects depending on the relative magnitudes of ϕ and α.

A.2: Equivalence with Equation (2) in the main text

We now show that the expression derived above is algebraically equivalent to Equation (2) presented in the main text:

$$\frac{Y_i}{L_i} = A_i \left(\frac{L_i}{S_i}\right)^{\phi} \left(\frac{K_i}{L_i}\right)^{\alpha} \tag{23}$$

To prove this, recall that:

$$\left(\frac{L_i}{S_i}\right)^{\phi} = L_i^{\phi} \cdot S_i^{-\phi} \quad \text{and} \quad \left(\frac{K_i}{L_i}\right)^{\alpha} = K_i^{\alpha} \cdot L_i^{-\alpha}$$

Substituting these identities into Equation (2), we obtain:

$$\begin{aligned}
\frac{Y_i}{L_i} &= A_i \cdot \left(\frac{L_i}{S_i}\right)^{\phi} \cdot \left(\frac{K_i}{L_i}\right)^{\alpha} \\
&= A_i \cdot \left(L_i^{\phi} \cdot S_i^{-\phi}\right) \cdot \left(K_i^{\alpha} \cdot L_i^{-\alpha}\right) \\
&= A_i \cdot S_i^{-\phi} \cdot K_i^{\alpha} \cdot L_i^{\phi-\alpha}
\end{aligned}$$

which exactly matches the expression derived earlier. This confirms that Equation (2) and the expanded form obtained in this appendix are fully equivalent representations of labor productivity under density-based agglomeration economies.

B Appendix

B.1 R&D Occupations

Table B.1: R&D Occupations

CBO Code	**Occupation Description**
1237	Research and Development Directors
1426	Research and Development Managers and Related Occupations
2030	Researchers in Biological Sciences
2031	Researchers in Natural and Exact Sciences
2032	Researchers in Engineering and Technology
2033	Researchers in Health Sciences
2034	Researchers in Agricultural Sciences
2035	Researchers in Social Sciences and Humanities
3951	R&D Support Technicians

Source: Sobrinho and Azzoni (2016).

B.2 Summary of Main Variables

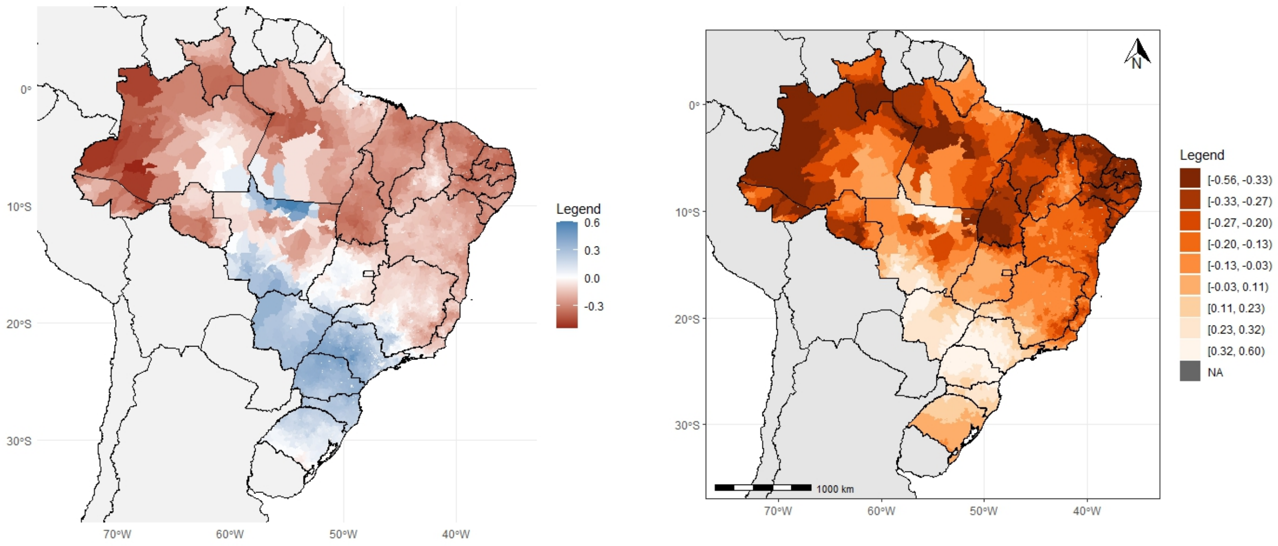
Table B.2:Summary Statistics of Main Variables

Variable	**Mean**	**Std. Dev.**	**Min**	**Max**
Δ Labor productivity [a]	1.090	0.264	0.000	6.356
Shrinkage	0.430	0.495	0.000	1.000
Δ Urban sprawl (ha)	1.333	0.279	0.999	3.702
SPEI x -1	-0.029	0.254	-0.551	0.604
Log Δ W urban sprawl	0.851	0.097	0.704	1.687
Labor productivity (2010)	2,266.223	923.908	0.000	1,7794.230
R&D employees	9.809	202.618	0.000	13,309.0
GDP share	0.000	0.002	0.000	0.123
Household size	3.389	0.430	2.560	6.890
Population (2010)	33,597	200,189	805	11,253,500
Altitude (m)	408.813	294.613	0.000	1,628.000
Latitude	-16.467	8.266	-33.653	4.685
Longitude	-46.254	6.434	-73.439	-34.842
Dist State Capital (km)	235.161	164.000	0.000	1415.789
Dist Federal District (km)	1074.166	446.610	42.076	2937.558
PM2.5 per capita (kg/s)	4.324	8.922	5.413	2.490
Paved road area (ha)	26.736	49.044	0.000	974.768

*Note:*To address potential biases from extreme values, we construct a dummy variable identifying municipalities with labor productivity above the 99th percentile of the national distribution. This binary variable is included in all specifications to control for outliers that could disproportionately influence the regression estimates.

C Appendix

C.1 Spatial Distribution of Climate Shocks



Panel A. Mean of $-1\times$ SPEI (2000–2009) **Panel B.** Percentile Distribution of $-1\times$ SPEI

Figure 3: **Spatial Distribution of Drought Exposure (Rural Areas Only)**
Notes: This figure presents two measures of drought exposure based on the Standardized Precipitation-Evapotranspiration Index (SPEI). Panel A shows the municipality-level average of $-1\times$ SPEI over the 2000–2009 period, with darker tones indicating more severe dryness. Panel B displays the percentile rank of this measure to facilitate comparison across municipalities. Calculations restricted to rural areas of sending municipalities.

C.2 Shift-Share Instrument and Baseline Covariates

Table 8: Correlation Between Shift-Share Instrument and Baseline Covariates

Variable	**SPEI-based shift-share IV**
Labor Productivy	0.11
R and D employees	0.66
GDP share	0.63
Household size	-0.12
Population	0.48
Altitude	0.14
Latitude	-0.22
Longitude	-0.14
Dist. State Capital	-0.09
Dist. Federal District	0.02
PM2.5 per capita (kg/s)	0.04
Paved road area	-0.02

C.3 Placebo Tests – Main Estimates

Table 9: Placebo Test: Instrument Validity

	Placebo (1)
Labor Productivity	
Shrinkage	−0.534***
	(0.131)
Urban Sprawl	−0.321***
	(0.105)
Interaction	0.661***
	(0.167)
State FE	Yes
Socioeconomic Controls	Yes
Geography and Environmental Controls	Yes
Infrastructure Controls	Yes
R-squared	
Observations	5,515
First Stage: Effects of Weather Shocks on Shrinkage	
SPEI-based shift-share IV	0.045***
	(0.009)
F-statistic	20.7***
First Stage: Effects of Spatial Lag on Urban Sprawl	
log(W_sprawl$_{t-1}$)	0.208***
	(0.016)
F-statistic	168.1***
State FE	Yes
Socioeconomic Controls	Yes
Geography and Environmental Controls	Yes
Infrastructure Controls	Yes
Observations	5,515

Notes: This table reports IV estimates using a placebo outcome measured prior to treatment. Standard errors clustered at the municipality level are shown in parentheses. All models include both instruments and the full set of controls. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

C.4 Sectoral Placebo Tests

Table 10: Placebo Test: Sectoral Effects of Shrinkage and Urban Sprawl on Labor Productivity

	Agriculture	Manufacturing	Services
Shrinkage	-1.949***	-0.977***	-0.792***
	(0.410)	(0.184)	(0.218)
Urban Sprawl	-1.658***	-0.564***	-0.521***
	(0.316)	(0.143)	(0.169)
Shrinkage × Urban Sprawl	2.516***	1.225***	0.929***
	(0.528)	(0.238)	(0.287)
State Fixed Effects	Yes	Yes	Yes
Socioeconomic Controls	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes
Infrastructure Controls	Yes	Yes	Yes
Adjusted R^2	0.448	0.215	0.135
N	5,515	5,515	5,515

Note: Clustered standard errors at the municipality level in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

C.5 Bootstrap – Sectoral Interaction Effect

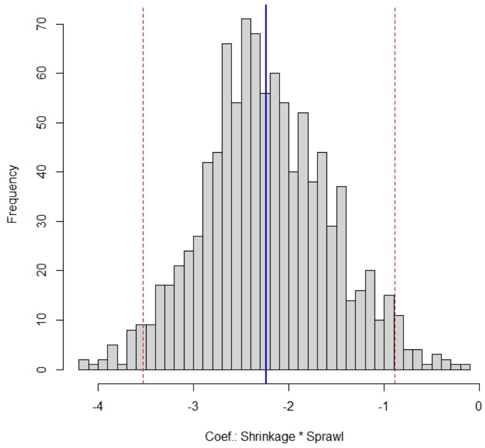


Figure C.2: Bootstrap Distribution of the Interaction Effect (*Shrinkage* $\times$ *Sprawl*) – Agriculture Sector
Distribution based on 1,000 bootstrap replications of the 2SLS estimator for the manufacturing sector. In each iteration, the two-stage model is re-estimated on a resampled dataset (N = 5,515). The average coefficient is **-2.236**, with a bootstrap standard error of **0.664**, and a 95% confidence interval of [**-3.525, -0.885**].

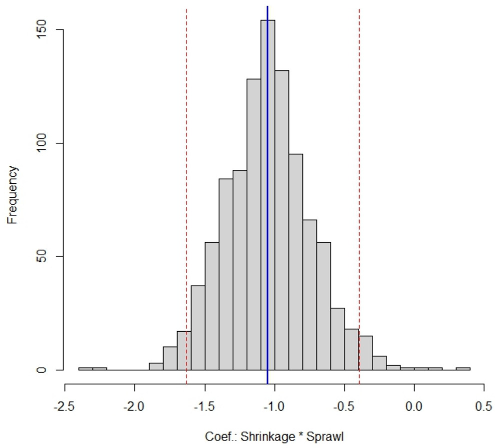


Figure C.3: Bootstrap Distribution of the Interaction Effect (*Shrinkage* $\times$ *Sprawl*) – Manufacturing Sector
Distribution based on 1,000 bootstrap replications of the 2SLS estimator for the manufacturing sector. In each iteration, the two-stage model is re-estimated on a resampled dataset (N = 5,515). The average coefficient is **-1.049**, with a bootstrap standard error of **0.311**, and a 95% confidence interval of [**-1.629, -0.395**.

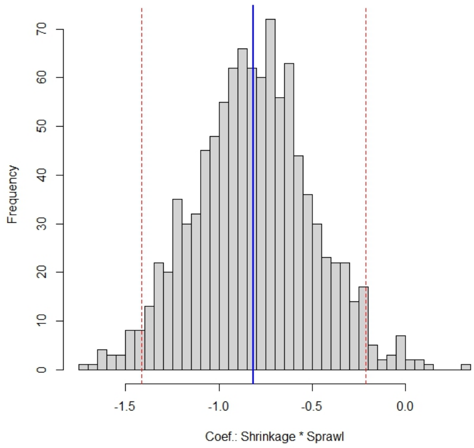


Figure C.4: Bootstrap Distribution of the Interaction Effect (*Shrinkage* $\times$ *Sprawl*) – Services Sector

Distribution based on 1,000 bootstrap replications of the 2SLS estimator for the manufacturing sector. In each iteration, the two-stage model is re-estimated on a resampled dataset (N = 5,515). The average coefficient is **-0.815**, with a bootstrap standard error of **0.311**, and a 95% confidence interval of [**-1.415, -0.213**].

Table 11: LASSO-Selected Covariates for First-Stage Shrinkage and Sprawl Regression

Variable	**shrinkage**	**sprawl**
(Intercept)	1.132	0.943
SPEI Shift-Share	-0.044	
IV Urban Sprawl		0.205
Labor productivy	0.001	-0.007
R and G employees	-0.046	-0.000
GDP shares	30.056	- 0.003
Population	-0.062	-0.013
Household size	0.003	0.055
Altitude	0.004	0.003
Latitude	0.017	0.000
Longitude	0.011	0.003
Dist. State Capital	0.071	0.007
Dist. Federal District	0.049	0.000
PM2.5 per capita (kg/s)	-0.008	
Paved road area	-0.006	-0.001