Spatial Misallocation, Informality, and Transit Improvements: Evidence from Mexico City*

Román David Zárate†

UC San Diego

April 2026

Abstract

Developing countries face inefficiently high informality. Can transit infrastructure improve input allocation by reducing informality? This paper studies this question in Mexico City. I combine a rich collection of administrative microdata and exploit the construction of new subway lines. Transit improvements reduce informality by 8% in areas near the new subway stations. I develop a spatial general equilibrium model that accounts for the direct effects of transit infrastructure in efficient economies and on allocative efficiency in economies with distortions. Changes in allocative efficiency, driven by workers' reallocation to the formal sector, amplify welfare gains by approximately 20% and output gains by 30%-40%.

Keywords: Informality, allocative efficiency, urban transit infrastructure.

*I am extremely grateful to my advisors Ben Faber, Cecile Gaubert, and Andrés Rodríguez-Clare for their continuous support and guidance in this project. I also want to thank my discussants Clare Balboni, Gabriel Ulyssea, and Alejandro Molnar, and seminar participants at the Dallas Fed, McGill University, the World Bank, ITAM, Universidad de los Andes, Universidad del Rosario, PUC Rio, Vancouver School of Economics, Nottingham University, the Macroeconomics Conference at Oxford University, the Online Urban Economics Seminar, the NBER SI in Urban Economics, the Cities and Development Workshop, Hitotsubashi University, and the 2021 ASSA meetings for very useful suggestions. I also want to thank David Atkin, Kirill Borusyak, Manuel García-Santana, Andrés González-Lira, Marco González-Navarro, Patrick Kennedy, Chris Severen, Joaquín Klot, Isabela Manelici, Pablo Muñoz, Mathieu Pedemonte, Darío Tortarolo, Nick Tsivanidis, Jose P. Vásquez, Román Andrés Zárate, and Isabel Hincapie for very helpful comments and discussions. I thank the National Institute of Statistics and Geography (INEGI) and especially Dr. Natalia Volkow for granting me access to the data. Financial support from the Clausen Center at UC Berkeley is gratefully acknowledged. The views expressed are only those of the author and they do not reflect the views of INEGI.

†UC San Diego, School of Global Policy and Strategy. E-mail: rzaratevasquez@ucsd.edu.

1 Introduction

The growth of congestion and commuting times has become a pressing challenge for large cities in developing countries. Changes in commuting costs can have large effects on urban structure, wages, and welfare (e.g., Ahlfeldt et al. (2015)). In Mexico City, a low-skilled worker typically spends two to three hours commuting to the central business district (CBD), prompting large investments in public transportation.[1] A growing literature studies the gains from transit improvements, typically under the assumption of perfectly competitive economies. However, this framework abstracts from a central feature of developing countries: the presence of large economic distortions. In this paper, I study the economic effects of transit infrastructure in the presence of informality and input misallocation.

In developing economies, an important source of distortions arises from disparities in effective tax rates generated by costly regulations that are imperfectly enforced across establishments. Informal firms evade taxes and social security contributions, creating wedges in marginal products that reduce total factor productivity (Hsieh and Klenow, 2009). As a result, formal firms face higher distortions and operate at higher marginal revenue products than informal firms, inducing an inefficiently large informal sector. These distortions interact with commuting frictions, which limit workers' access to high-productivity formal jobs and may further amplify misallocation.

This paper studies whether reducing commuting frictions can mitigate misallocation in such an environment. I examine whether transit infrastructure improves allocative efficiency by enabling workers to reallocate from low-marginal-revenue-product informal establishments to high-marginal-revenue-product formal establishments. If so, the gains from these projects may exceed those predicted by urban models that abstract from distortions. The mechanism operates through workers' access to formal jobs: when commuting costs are high, workers in peripheral areas may prefer nearby informal jobs despite their lower productivity. By reducing these frictions, transit improvements expand access to formal employment and improve the allocation of labor across sectors and locations.

The paper makes two main contributions. First, I combine rich administrative microdata with a transit shock to provide empirical evidence on the relationship between informality and the geography of cities in developing countries. I estimate the effect of transit improvements on workers' reallocation across the formal and informal sectors. Second, I rationalize these results through the lens of a quantitative spatial model. I extend recent work (Ahlfeldt et al., 2015; Tsivanidis, 2023) and models of informality, such as Ulyssea (2018) and Dix Carneiro et al. (2018), by incorporating distortions, spatial frictions, misallocation, and endogenous sorting of firms and workers into the informal sector.

I study this question in Mexico City, a useful setting for three reasons. First, it has a high concentration of economic activity. Second, labor and firm informality are pervasive: more than 50% of the urban labor force and 70% of establishments are informal, generating substantial resource misallocation (Busso et al., 2012; Levy, 2018). Third, the construction of a major subway line in the early 2000s, connecting remote northern areas to the CBD, provides a plausible source of variation in access to formal employment.

[1]For example, according to Akbar et al. (2023), Mexico City is one of the 10 most congested cities in the world.

At the center of the analysis is a rich collection of unique administrative microdata. I observe the spatial distribution of jobs and workers for both the formal and informal sectors at the census tract level.[2] I use four primary sources of data: i) confidential microdata from the Economic Census, covering the universe of formal and informal business establishments in the city; ii) the Population Census to determine the location of both formal and informal workers based on having access to social security; iii) detailed information on the transportation network; and iv) transportation diaries.

The first part of the paper empirically investigates the relationship between job accessibility and informality. Using cross-sectional data, I document two key empirical findings. First, formal jobs are predominantly concentrated in the city center, and second, informal workers have shorter commutes than their formal counterparts. As a result, workers residing on the outskirts are more likely to engage in informal activities due to limited access to formal employment opportunities. Building on these findings, I then analyze the impact of transit improvements on informality.

The main finding suggests that transit improvements reduce informality rates. I provide evidence of this effect by analyzing the opening of a new subway line (Line B) that connected remote locations with the center of Mexico City. I estimate a series of difference-in-differences specifications using variation in access to the new transit line. These specifications control for initial census-tract characteristics and capture changes in informality following the transit shock in locations near the new subway line. The key identification assumption is that the opening dates of the new links were unrelated to other local demand- or supply-side shocks affecting nearby locations. This assumption is supported by the line's decades-long planning horizon and by several unexpected multi-year delays in its opening schedule due to the 1994 peso crisis. The results indicate that the ratio of formal to informal residents increases by approximately 7% to 9% in areas near the new stations.

I also perform two additional exercises to verify the robustness of the results. I use an expansion plan from 1980 and compare the new line with similar planned metro lines that were not completed over this period for unrelated reasons. Reassuringly, this robustness check yields estimates similar to those from the baseline specification. In addition, to alleviate remaining endogeneity concerns regarding infrastructure allocation, I construct an instrument based on the Least Cost Path in elevation connecting the main economic centers (Faber, 2014). The instrument is designed to capture the routes that planners would have built to minimize costs and connect these centers. The identification assumption is that cost-minimizing routes are unrelated to local shocks. The results are stronger than those from the baseline specification.

The reduced-form estimates suggest that transit improvements may generate welfare gains larger than those typically emphasized in the literature. This naturally raises two questions: are the gains indeed larger, and if so, by how much? To answer these questions, I develop a model with multiple sectors, endogenous firm location, and market distortions. The model incorporates several endogenous margins, including firms' and workers' selection into the formal and informal sectors. It also solves

[2] This is a unique characteristic of the Mexican data since I can observe directly which establishment is formal or informal. For instance, this differs from other datasets, such as India, where researchers use size as a proxy for informality. Throughout the paper, I use the standard definition of informality: a worker is informal if she does not receive social security benefits.

the "missing intercept" problem that arises in reduced-form approaches. This framework allows me to quantify the aggregate welfare effects of new infrastructure by explicitly accounting for changes in factor allocation. I derive a decomposition of welfare changes from commuting cost reductions into direct and allocation effects. Intuitively, allocation effects increase welfare when workers reallocate toward establishments that were initially more distorted.

A relevant statistic for assessing the effect of policy interventions on allocative efficiency is the wedge in labor and commercial floor space between formal and informal firms. Using information from Levy (2018), I construct wedges arising from multiple sources, including social insurance contributions (pensions, health insurance, disability, and other benefits), net payroll taxes, commercial floor space taxes, value-added and consumption taxes, and regulatory frictions. Combining labor and output distortions yields a labor wedge of approximately 0.92, while combining floor space and output distortions yields a floor space wedge of approximately 0.32.[3] These wedges imply sizable incentives for firms and workers to operate informally, generating misallocation that infrastructure investments can partially mitigate.

I estimate the model's key parameters using structural equations and a GMM procedure combined with indirect inference, matching the model's implied coefficients to their reduced-form counterparts. The estimates imply a commuting elasticity of approximately 6, indicating a high degree of substitutability across workplaces in Mexico; a parameter that transforms travel time to commuting costs around 0.09, consistent with values found in the literature; a sectoral labor supply elasticity of around 1.7, suggesting limited worker mobility between the formal and informal sectors, and a migration elasticity of approximately 0.4.

Armed with these estimates, I quantify and decompose the welfare gains from the new infrastructure by varying commuting costs in the model. Relative to a standard quantitative framework, incorporating the allocative efficiency margin increases welfare gains by 15%–20% and output gains by 30%–40%. In the preferred specification, the subway expansion raises aggregate welfare by approximately 0.75% and output by 0.85%, with improvements in allocative efficiency accounting for 16% of the welfare gains and 26% of the output gains. I also assess the model's fit and show that it replicates well the main empirical patterns observed in the data.

Overall, these results show that accounting for allocative efficiency amplifies the welfare and output benefits of transit investments. This insight has important implications for the optimal provision of infrastructure (Balboni, 2019; Fajgelbaum and Schaal, 2017; Santamaría, 2020), as existing planning frameworks typically abstract from infrastructure's effects on misallocation.

Related Literature: This paper relates to three strands of the literature: urban and spatial economics, informality in developing countries, and allocative efficiency in distorted economies.

A large literature in urban and spatial economics studies the effects of transit infrastructure in cities (Ahlfeldt et al., 2015; Baum-Snow, 2007; Gonzalez-Navarro and Turner, 2018; Heblich et al., 2018; Tsivanidis, 2023). Recent work combines quasi-experimental variation with spatial models to

[3]The magnitude of the labor wedge in Mexico is consistent with estimates for Brazil reported by Ulyssea (2018).

quantify the aggregate and spillover effects of urban interventions. For example, Tsivanidis (2023) studies the distributional effects of a bus rapid transit system in Bogotá, while Heblich et al. (2018) analyze the economic effects of the London Underground. Related work, such as Franklin et al. (2024), also emphasizes the importance of spillovers when evaluating public works. My paper builds on this literature by incorporating multiple sectors and by studying how commuting improvements interact with pre-existing distortions.

This paper also relates to the literature on informality in developing countries. Informality accounts for a substantial share of employment and output in low- and middle-income countries and has important implications for productivity, human capital, and welfare (Ulyssea, 2020). A growing body of work studies firms' and workers' decisions along different margins of informality (Alvarez and Ruane, 2024; Bosch and Esteban-Pretel, 2012; Bosch and Maloney, 2010; Ulyssea, 2018), while recent structural approaches emphasize its effects on aggregate productivity through misallocation and human capital accumulation (Bobba et al., 2022, 2021; Levy, 2018). My contribution is to show how distortions associated with informality interact with spatial frictions and transit infrastructure.

The paper also contributes to the literature on factor misallocation and aggregate productivity (Banerjee and Duflo, 2005; Hsieh and Klenow, 2009; Restuccia and Rogerson, 2008). This work shows that dispersion in distortions across establishments leads to inefficient input allocation and lower TFP. In the context of Mexico, Busso et al. (2012) show that reducing wedges between formal and informal firms could substantially increase productivity. Related work, including Albouy (2009), Fajgelbaum et al. (2019), and Hsieh and Moretti (2019), documents how taxes and housing regulations generate spatial misallocation in the United States. My contribution is to study how transit infrastructure affects misallocation in an economy where informality is a central source of distortions.

This paper is also related to the literature on how allocative efficiency shapes the gains from market integration. Reviews by Atkin and Khandelwal (2020) and Atkin and Donaldson (2021) emphasize the role of distortions in determining aggregate gains. A closely related contribution is Hornbeck and Rotemberg (2019), who study the expansion of the railroad network in the United States and its interaction with distortions. Related work examines how intersectoral distortions affect welfare (Święcki, 2017) and how trade shocks interact with informality (Dix Carneiro et al., 2018; McCaig and Pavcnik, 2018; McMillan and McCaig, 2019).[4] While most of this literature focuses on trade-related shocks, I study how distortions associated with informality shape the gains from reducing commuting frictions.

Related theoretical work, such as Moreno-Monroy and Posada (2018) and Suárez et al. (2016), studies commuting and informality using search models, arguing that high commuting costs raise informality by limiting access to formal jobs. This paper brings this idea to the data and to a quantitative spatial framework. Empirically, I show that transit infrastructure reduces informality. Quantitatively, I show that it also increases welfare and output by improving allocative efficiency.

The rest of the paper is organized as follows. Section 2 introduces the setting and describes the

[4] Other papers in this literature study how trade liberalization or infrastructure changes affect markups and firm behavior (Arkolakis et al., 2019; Asturias et al., 2019; Edmond et al., 2015).

transit shock. Section 3 presents the reduced-form evidence of the effect of commuting on informality. Section 4 develops the urban quantitative model. Section 5 estimates the main parameters. Section 6 quantifies and decomposes the welfare gains. Section 7 concludes.

2 Institutional Context

2.1 Transit System

In the second half of the 20th century, Mexico City experienced severe public transportation problems, characterized by congested main roads and highways. In 1967, the Government established a decentralized public office to build and operate an underground rapid transit system. The first line opened in September 1969. Today, the subway comprises 12 lines and 195 stations, spanning 128.4 miles, making it the largest system in Latin America and the second-largest in North America.

The Plan Maestro 1985-2010 guided the expansion of the subway. It established the mobility goals that the transport system needed to meet in the long term, based on best practices in urban development and the potential operational constraints. The Plan Maestro underwent modifications relative to the Government's initial plan. For example, Line B was originally Line 10 and experienced extensive changes (Ramírez et al., 2017). These modifications primarily responded to changing transportation demand patterns, which compelled the Government to redesign certain lines.

My empirical strategy exploits the construction of Metro Line B, which connects predominantly informal residential areas in the urban periphery to jobs in the central business district (CBD). Line B was inaugurated in 2000, but its route had largely been determined in the 1985 *Plan Maestro*. This reduces concerns that station placement responded to contemporaneous local economic conditions. In addition, the timing of construction was shaped by institutional and macroeconomic factors, including regulatory changes and the 1994 financial crisis. As a result, the project experienced substantial delays: although the original plan envisioned completion in 1997, the full line was only finished in 2002. These features imply that the timing of Line B was orthogonal to local shocks.

Line B is approximately 20 kilometers long and includes 21 stations. It connects the Mexico City metropolitan area with adjacent municipalities in the State of Mexico, including Ecatepec de Morelos and Ciudad Nezahualcoyotl. These areas are characterized by high poverty rates, low educational attainment, and high levels of informal employment. Line B is the fourth-busiest line in the metro network. The total cost of the project, including the net present value of operations and other overheads, amounted to \$2,230 million in 2014 USD, equivalent to 0.55% of the city's GDP.

Figure 1 presents a map of the Mexico City subway system in 2000. The black line denotes Line B, the blue lines correspond to the metro lines already in operation, and the green lines indicate the other feeder lines proposed in the 1985 *Plan Maestro*.

2.2 Informality

Following Busso et al. (2012); Kanbur (2009) and Levy (2018), I define informality based on firms' compliance with labor regulations. A worker is classified as informal if the firm does not pay mandatory social security contributions on their behalf. Similarly, a firm is considered informal if it does not make social security contributions for its workers. Social security benefits include health care, retirement savings, recreational social benefits, and disability insurance. Workers classified as informal may be either salaried or non-salaried.

As in most developing countries, informality in Mexico and the Latin America region is a significant problem. For instance, the average informality rate across the region is 50%, which is significantly higher than the OECD average of 17%. In Mexico, it affects 57% of the total workforce and 78% of firms (INEGI). Relative to other countries in the region, Mexico has one of the highest informality rates, and the difference is more pronounced when comparing Mexico to other countries with a similar income level, such as Argentina or Colombia (see Figure A1 in the Online Appendix).

The presence of the informal sector and the fact that informal firms avoid paying taxes create wedges across establishments. According to recent estimates, a firm that fully complies with salary regulations is expected to pay social security taxes amounting to 30%-40% of a worker's wage (Busso et al., 2012; Levy, 2018) and 25% on sales taxes. These wedges create distortions across firms that decrease welfare and TFP. Overall, most informal firms are small family businesses that are smaller and less productive than formal ones. The average formal firm is more than ten times as productive as the average informal firm (see Figure A2 in the Online Appendix).

In addition, informal workers are less educated, have lower incomes, work more hours, and are less likely to receive fixed salaries than their formal counterparts. For example, there is a negative correlation between informal workers and the probability of obtaining a technical or college degree, and informal workers get an average of 8% less income than formal workers (see Table B1). These differences also materialize across the locations where informal vs. formal workers live. A higher informality rate at the census tract level is associated with fewer public services, such as electricity, lower education, and lower average incomes (see Table B2).

2.3 Wedges between formal and informal firms

A central object for assessing the impact of infrastructure on allocative efficiency is the wedge between formal and informal firms. These wedges capture distortions in input costs that differentially affect firms by formality status. I quantify wedges for labor and commercial floor space using tax and regulatory information from Levy (2018). I define the labor and commercial floor space wedge faced by formal firms, respectively as:

$$t_L = \frac{1+\tau_L}{1-\tau_Y} - 1, \quad t_Z = \frac{1+\tau_Z}{1-\tau_Y} - 1,$$

where τ_L denotes differences in labor-related taxes and subsidies, τ_Y captures differences in output and value-added taxation, and τ_Z reflects differences in taxes on commercial floor space.

Multiple institutional and fiscal factors contribute to the wedge between formal and informal firms. Table B3 in the Appendix summarizes the main sources identified by Levy (2018).[5] These

[5]Table B4 summarizes the tax and benefits of formal and informal workers.

include contributions to social insurance programs (pensions, health insurance, disability, and other benefits), net payroll taxes paid by formal workers, subsidies received by informal workers through non-contributory programs, value-added taxes, as well as regulatory frictions such as contract enforcement, hiring and firing costs, and compliance burdens that disproportionately affect formal firms.

Quantifying these wedges is challenging because several distortions are not directly observed as statutory tax rates. I therefore rely on two complementary approaches. First, when Levy (2018) provides direct estimates of the relevant tax differentials, I use those values. Second, when distortions are reported as shares of GDP, I convert them into equivalent labor or output-tax rates using proportionality assumptions based on observed payroll taxes and factor shares.

Using this procedure, I estimate that the effective labor tax faced by formal firms is approximately $\tau_L \approx 0.504$. This estimate aggregates the following components:

- **Net tax from contributory programs (0.12):** Estimated by Levy as the difference between the total cost to firms of mandatory social contributions and the value of benefits received by formal workers.
- **Net state payroll taxes on salaried employment (0.03):** In Mexico City, payroll taxes average 3%. Levy estimates that this distortion accounts for roughly 0.39% of GDP in the country.
- **Net federal payroll taxes on salaried employment (0.192):** Federal payroll taxes paid by formal workers amount to approximately 2.5% of GDP. Mapping this figure to an equivalent labor tax rate yields 19.2%, based on the proportional relationship implied by the state payroll tax benchmark.
- **Net benefits to informal firms from non-contributory programs (0.162):** This number is taken directly from Levy's estimates, reflecting subsidies that reduce effective labor costs in the informal sector.

Formal firms also face higher output and value-added taxes than informal firms. I estimate the effective output tax wedge to be $\tau_Y \approx 0.217$, reflecting several institutional distortions:

- **Value-added taxes net of exemptions (0.12):** Based on Mexico's statutory VAT rate of 16%, adjusted for the fact that approximately 25% of GDP is produced in exempt sectors.
- **Special production taxes (IEPS) (0.05):** Applied to selected industries, including food and beverages, tobacco, and gasoline, and scaled by their aggregate GDP shares.
- **Corporate income taxes (0.047):** Levy (2018) estimates that differences in corporate taxation generate a gap of roughly 0.5% of GDP between formal and informal firms. Converting this magnitude into an equivalent output tax yields 0.047, using the same proportionality assumption and a labor share of 0.822.

Finally, I set $\tau_Z = 0.03$, based on the RESICO regime, which imposes a rental tax of approximately 3% on the floor space used by formal firms. I use these values in the quantification.

3 Data and Reduced Form Effects

3.1 Data

My primary unit of observation is the urban census tract (Area Geoestadística Básica in the Mexican micro-data). I use a sample of approximately 3,200 census tracts from 116 different neighborhoods.

The first source of information is standard GIS data on the evolution of the transportation network, including new transit subway lines, as well as data on roads and highways in Mexico City, which I use to calculate commuting times.

The second data source is the Mexican Economic Censuses. This is a unique establishment dataset that provides information on sales, value added, employment, social security, and other outcomes. The census is conducted every five years, starting in 1994. A unique characteristic of this dataset is that it allows observation of informal establishments at a very high level of granularity. Many self-employed individuals operating small establishments are recorded as establishments in this census, often as very small firms, so self-employment is partially captured in the dataset.

The third source of information is the Mexican Population Census. This census is conducted every 10 years, and INEGI has provided the data since 2000. Using this information, I calculate the numbers of informal, formal, and total residents for each location, where informality is defined as the absence of social security contributions, the standard definition in the Mexican context. This definition encompasses a broad range of informal arrangements, including self-employment without social contributions, unpaid family work, and salaried work without benefits. The 2000 Population Census also reported additional variables, such as household income and job characteristics, which I use as covariates in the empirical strategy.

I also use the 2015 Intercensal Survey and the 2017 Origin-Destination Survey to infer commuting flows across municipalities within the city and by transportation mode. Although these datasets were collected more than a decade after the shock, they help me estimate some of the key elasticities. The main assumption is that these elasticities are constant over time, which is a common assumption in the urban literature.[6]

Both datasets from the Census provide complementary measures of informality: the Economic Census captures the location of establishments, while the Population Census records workers' access to social security and therefore provides comprehensive coverage of informal employment regardless of establishment type. Concerns about misreporting are mitigated by institutional features of the data collection process. The Economic Census is conducted by INEGI under strict confidentiality guarantees, and the information collected cannot be used for tax enforcement purposes.

Table B6 in the Online Appendix reports summary statistics for the main variables used in the empirical analysis.

[6] Ahlfeldt et al. (2015) in one of their estimations also uses data in the post-period to estimate the commuting elasticity. The main assumption is that this parameter is constant over time, but depends on the context of the city.

3.2 Informality across Space

Before turning to the reduced-form analysis, I use cross-sectional variation to show that limited access to formal jobs is associated with higher informality.

Empirical Fact 1: *Formal jobs are concentrated in central areas of the city, while informal workers mainly reside in the outskirts.*

The first empirical fact is that formal jobs are primarily located in the center and west of Mexico City, whereas informal workers are more likely to reside in peripheral and less connected areas. High housing costs in central locations push many low-income workers to the outskirts, where access to formal employment is limited. Figure A4 shows a heat map of informality rates by job location and worker residence prior to the intervention.[7] Informality rates are substantially lower in central and western areas than in the east and the urban periphery.

A natural explanation for this spatial pattern is unequal access to formal jobs. Panel C plots the difference between market access indices for formal and informal employment in 1999, before the intervention. Eastern and peripheral neighborhoods exhibit markedly poorer access to formal jobs relative to informal ones. The construction of the subway line connected these areas to the city center, increasing relative access to formal employment in treated locations.[8]

Empirical Fact 2: *Informal workers commute less and work closer to home.*

The second empirical fact is that informal workers spend less time commuting than formal workers. To document this pattern, I use data from the 2015 Intercensal Survey, which reports individuals' municipality of residence, municipality of work, and time windows for average commuting duration. Exploiting cross-sectional variation, I compare commuting outcomes and workplace locations between formal and informal workers. I restrict the sample to individuals who worked in the week prior to the interview and estimate the following linear probability model:

$$y_i = \beta_0 + \beta_1 \text{Informal}_i + \gamma X_i + \gamma_{l(i)} + \gamma_{n(i)} + \gamma_{m(i)} + \epsilon_i, \tag{3.1}$$

where y_i is an indicator for whether individual i commutes to a different municipality than their residence, works in the central business district (CBD), or has an average commuting time within a given interval (e.g., 16 to 30 minutes). The vector X_i includes age, education, gender, household head relationship, and an indigenous background indicator. The terms $\gamma_{l(i)}$ and $\gamma_{n(i)}$ denote origin and destination fixed effects, and $\gamma_{m(i)}$ is a transportation-mode fixed effect, allowing comparisons between formal and informal workers using the same mode of transport. The error term is ϵ_i.

I focus on the distribution of commuting times by estimating the probability that a worker's average commute falls within different time windows. Informal workers consistently exhibit shorter commutes than formal workers (see Figure A3). Informal workers are more likely to work from home

[7] Consistent with this pattern, the west and center of Mexico City also exhibit the highest levels of economic activity (see Figure A7), a measure that is strongly correlated with the presence of formal employment.

[8] Figure 4 illustrates this change by plotting the variation in relative market access across sectors.

or commute less than 15 minutes, whereas formal workers are more likely to commute more than 1 or 2 hours. In addition, formal workers are more likely to commute across municipalities and to work in the CBD (see Table B5). Quantitatively, informal workers are about 15 percentage points less likely to commute to a different municipality and about 7 percentage points less likely to work in the CBD.

3.3 The Impact of Transit Infrastructure on Informality

The main finding suggests that informality rates decrease with transit improvements that improve market access to formal employment. To show this result, I exploit the construction of Line B and estimate a series of difference-in-differences specifications. I compare locations near the new subway line with locations in the rest of Mexico City and test whether areas that experienced larger improvements in market access also experienced changes in informality rates after the transit shock, while controlling for initial characteristics. The identification assumption is that the opening of the new stations is uncorrelated with local demand/supply shocks. The fact that most of the line was planned decades earlier makes this assumption plausible. Moreover, since infrastructure construction may be endogenous (Redding and Turner, 2015), I include a set of covariates as controls to compare similar areas and construct an instrument based on the least-cost path.

Using data from the Population Censuses between 2000 and 2010, I estimate the following specification relating the transit shock to the change in the ratio between formal and informal workers:[9]

$$\Delta^{2000-2010}\left(\ln L_{iF}-\ln L_{iI}\right)=\beta \mathrm{T}_i+\gamma X_i+\delta_{s(i)}+\epsilon_i, \tag{3.2}$$

where L_{is} is the number of individuals that live in census-tract i and sector s, T_i is one of two different treatment variables: log distance in meters, and a dummy variable whether the closest station is within 30 minutes, $\delta_{s(i)}$ are state or zone fixed effects,[10] and X_i is a vector of census-tract characteristics that include: the area, distance to other stations of public transit, a central business district dummy variable, distance to main roads, average income in the baseline year, the initial high-skilled share and some productivity measures in the baseline in which I include value-added per worker and the number of firms to capture local job density.[11] I estimate equation 3.2 for the pool of workers and for different skill-based groups.

Table 1 reports the results for different specifications of equation 3.2. The results suggest that locations near the new subway line experienced a decrease in informality rates. In particular, the ratio of formal to informal individuals increased by 5.0% to 8.7% after the shock. These results are robust to various specifications, including different definitions of the treatment variable and different sets of fixed effects or controls. In addition, in panels C and D, I control for changes in workers' skill composition and report results only for low-skilled workers. The estimates are very similar to the ones found for the entire pool of workers. The ratio of formal to informal low-skilled workers

[9] Since a large fraction of self-employed workers are informal, changes in informal employment should be interpreted as capturing both informal wage employment and informal self-employment.

[10] For the zone fixed effects specifications, I classify locations in the State of Mexico into three different groups: North, West, and East for a total of 19 zones that include these three areas and the sixteen municipalities of Mexico City.

[11] Equation 3.2 corresponds to a structural relationship that I will derive in the model (see section 5).

increased, on average, by 5.0%-9.0%. In panels E and F, I report the results for the sample restricted to areas outside the CBD. The effects should be larger in these locations, as more informal workers reside there and experience greater market access to formal jobs than to informal ones. I find larger effects under this specification; the ratio of formal to informal workers increased by almost 12.0%.[12]

I also estimate the effect of Line B on the log number of individuals and differentiate the impact between formal and informal workers. The point estimates indicate a modest effect on the number of individuals, with an increase of 1.8% for the pool of workers and 2.6% for low-skilled workers. The effect on informality is primarily driven by an increase in the number of formal workers, ranging from 3% to 7%, and a decrease in the number of informal workers, ranging from 2% to 3% (see Table B8).

3.4 Robustness checks

Comparison with other planned feeder lines: Line B was originally conceived as a feeder line within the 1985 *Plan Maestro* of the Mexico City metro system. The feeder lines considered in this exercise were selected simultaneously within a common planning framework, but only Line B was ultimately built during the study period. As a robustness check, I compare locations near Line B with areas near alternative feeder lines that were also planned but not contemporaneously built.[13] Figure 1 displays Line B, the metro network in place before 2000, and the set of feeder lines proposed in the 1985 *Plan Maestro*. I re-estimate the baseline difference-in-differences specification by defining treatment as census tracts whose centroids lie within a given buffer distance of Line B, and using as a control group census tracts located within the same buffer distance around the other feeder lines. I consider four buffer radii: 1,000, 1,500, 2,000, and 2,500 meters.

The estimated effects are similar to, and in some cases stronger than, those in the baseline specification. In particular, the log ratio of formal to informal workers increases by approximately 15% in treated areas relative to tracts near the other feeder lines. Consistent with spatial attenuation, the magnitude of the effect declines as the buffer radius increases (see Figure A6).

Least-cost path: To address additional endogeneity concerns of infrastructure placement, I construct an instrument for Line B using the least–cost–path approach of Faber (2014). I compute the minimum–cost routes connecting Line B's four main stations-Ciudad Azteca, Nezahualcóyotl, San Lázaro, and Buena Vista-and define an instrument equal to one if a census-tract centroid lies within the 10th–percentile corridor of these routes. The identifying assumption is that least-cost routes are orthogonal to local labor-market shocks.

Using this instrument yields larger effects (Table B9). For the pooled sample, the formal–to–informal ratio increases by 12%, compared to 8% in the baseline. For low-skilled workers and for peripheral areas, the effects reach roughly 19%, versus 8% and 10% in the baseline, respectively. These results indicate that areas gaining market access through cost-minimizing routes experience larger reductions in informality.

[12] Section D.1 in the Appendix describes the effects on employment.

[13] I exclude feeder line F from this exercise because it was designed for the northeastern areas of Mexico City and runs very close to Line B.

3.5 Discussion about worker sorting, household composition, and spillovers

A sufficient statistic to assess the effect of infrastructure on aggregate misallocation is the aggregate change in informality (see Baqaee and Farhi (2020) and the following subsection). A natural approach is to infer this change directly from reduced-form estimates. However, because these regressions rely on relative comparisons between treated and control areas, they do not identify the aggregate effect of the policy. In particular, the reduced-form coefficients are identified only up to a "missing intercept," since they do not reveal how outcomes in the control group adjust in equilibrium in response to the shock. Worker sorting across locations and spillovers to control areas are two mechanisms that can generate such adjustments, thereby complicating the interpretation of reduced-form estimates and any back-of-the-envelope calculations based on them.

Understanding how this missing intercept maps into aggregate outcomes requires clarifying the economic mechanisms underlying the reduced-form estimates. In the spatial equilibrium model developed in Section 4, infrastructure affects labor market outcomes through workers' behavioral responses, endogenous reallocation, and general equilibrium forces. Reduced-form coefficients, therefore, capture equilibrium responses at the location level, combining individual adjustments, worker sorting across space, and GE effects. The purpose of this section is to evaluate whether the empirical patterns are consistent with the model's key mechanisms and structural assumptions. Two issues are central for this assessment: first, whether worker sorting operates along the margins incorporated in the model; and second, whether spillovers to control areas affect the interpretation of relative estimates.

Worker sorting and household composition: The model allows workers to reallocate across locations in response to changes in commuting costs, and such reallocation may be a central channel through which infrastructure affects equilibrium outcomes. Workers differ in unobserved idiosyncratic shocks that govern both their location decision and their sectoral choice. As a result, changes in outcomes in treated areas may reflect both behavioral adjustments of incumbent residents and endogenous sorting of workers across space. The goal of this section is therefore not to rule out worker sorting, but to test empirical patterns that are consistent with the structure of the model.

Interpreting these equilibrium responses through the lens of the model relies on two assumptions. First, idiosyncratic location shocks are independent of observable characteristics, so improvements in job access should not generate systematic changes in observable household composition across locations. Second, the model features a nested structure in which, conditional on the location decision, workers choose between formal and informal employment based on a common distribution of idiosyncratic sectoral shocks. This distribution does not differ between incumbents and new residents. As a result, conditional on location, both groups have the same probability of choosing formal employment.

The empirical exercises in this subsection assess whether the data are consistent with the model's structural implications. I begin by evaluating the assumption that location-specific idiosyncratic shocks are independently and identically distributed across workers.[14] If the opening of Line B

[14] In principle, changes in observable household characteristics could arise even if workers' decisions depend only on unobservable shocks, provided that observables are correlated with these unobservables. Such an extension would not invalidate the interpretation of the model, but would require additional structure and parameters to map observables

induced selective inflows or outflows of particular worker types, this would be reflected in systematic changes in the observable characteristics of households residing in treated areas. To examine this, I estimate equation 3.2 using a broad set of covariates, including the shares of high-skilled (college) and medium-skilled (technical) workers, the average age of the household head, household size, number of children, gender composition, and years of education. Table B10 shows no evidence of systematic changes in household characteristics following the opening of Line B. The estimated coefficients are small, and most of them are precisely estimated around zero.[15] The absence of compositional changes therefore provides empirical support for the i.i.d. assumption embedded in the model.

To assess the second assumption, I re-estimate the main specifications restricting the sample to individuals who did not change their state of residence between census waves. This restriction focuses on individuals who remain in the State of Mexico, the peripheral areas of the metropolitan region directly affected by Line B, and excludes migrants from Mexico City, where the central business district is located. Since most treated locations are in the State of Mexico, this restriction holds the location decision fixed at the state level and isolates changes in sectoral choice within locations. The estimates from this non-migrant sample are nearly identical to the baseline results and remain statistically significant (see Table B11). This finding suggests that, conditional on the location decision, the choice between formal and informal employment is similar for incumbent and new residents, consistent with the model's structure.

Potential spillovers: In principle, reduced-form estimates can be used as inputs to back-of-the-envelope calculations to approximate the infrastructure's aggregate impact. However, this exercise assumes control areas are unaffected by the investment, implying that the "missing intercept" in relative comparisons is zero. These assumptions may fail if the infrastructure generates spillovers that affect locations classified as controls, thereby potentially violating SUTVA and complicating the interpretation of calculations based solely on reduced-form coefficients. Two types of spillovers are particularly relevant. First, general equilibrium effects may alter wages, employment, or prices throughout the metropolitan area, affecting both treated and control locations. Second, localized spillovers may arise if nearby areas benefit indirectly from improved access to the transit network.

Addressing these concerns requires an explicit general equilibrium framework, which the quantitative model in Section 4 provides. It incorporates spillovers by construction: reductions in commuting costs affect wages, sectoral allocation, and residential choices simultaneously across all locations. As a result, the model delivers the aggregate effect of the infrastructure, including its impact on the control areas in the reduced-form analysis. In this sense, the model resolves the missing-intercept problem inherent in relative comparisons and provides a consistent framework for welfare evaluation. The key question, then, is whether the model credibly captures the spatial adjustments observed in the data, which motivates the validation exercises.

into the choice process. The absence of systematic changes in household composition suggests that such extensions are not quantitatively important in this setting.

[15] The only outcome for which the coefficients are statistically different from zero is the student share. Nevertheless, the point estimates are small since they represent only 0.5 percentage points. Moreover, as shown by Alba-Vivar (2024), new infrastructure can also change education opportunities.

As a complementary exercise, I examine whether spillovers are primarily local. If nearby control areas are partially treated, excluding them should increase the estimated treatment effects. I therefore re-estimate the main specifications excluding census tracts located within 2.5 kilometers of treated areas. Under this restriction, the estimated effects increase (Table B12), indicating that nearby control locations were indeed affected by Line B. This pattern is consistent with local spillovers and further underscores the need for a general equilibrium framework to interpret aggregate effects.

3.6 Back-of-the-envelope calculation

Before presenting the full model, this section provides a back-of-the-envelope calculation to build intuition for how reallocation effects can generate first-order welfare changes. It also shows that changes in informal employment serve as a sufficient statistic for changes in allocative efficiency.

There are two different firm types: formal (F) and informal (I). Preferences take a general functional form: $U = f(q_F, q_I)$, where $f(\cdot,\cdot)$ is a function that is homogeneous of degree 1 and homothetic.[16] Firms produce with a technology that has constant or decreasing returns to labor. Without loss of generality, I assume that firms F pay a tax rate τ for each unit of labor, while firms I do not pay taxes. The wedge τ creates differences in the value of the marginal product of labor (VMPL), leading to misallocation. Workers have upward-sloping labor supply functions:

$$L_s = \frac{B_s w_s^{\kappa} \tilde{d}_s^{-\kappa}}{\sum_k B_k w_k^{\kappa} \tilde{d}_k^{-\kappa}} \bar{L}, \tag{3.3}$$

where w_s is the wage, B_s is a labor-supply shifter, $\tilde{d}_s$ corresponds to a commuting cost, and κ is the labor supply elasticity. To simplify the analysis, I assume that $\tilde{d}_F > 1$ and $\tilde{d}_I = 1$. We examine the welfare effects of a commuting shock to $\tilde{d}_F$. Figure 2 summarizes the deadweight loss due to the wedge and the economic effects of the shock.

Panel A shows the inefficiency created by wedges in this economy. The labor wedge faced by formal firms shifts the VMPL downward, shifting the market allocation from L_F^* to L_F' and creating a deadweight loss, captured by the vertical black area. Panel B illustrates the impact of a decrease in commuting costs. The labor supply for formal firms shifts to the right, and the total amount of labor increases from L_F' to L_F''. Defining $\lambda_F = L_F/\bar{L}$, the change in welfare to a first-order is (see Appendix Section E.1):

$$d\ln W = \underbrace{\lambda_F d\ln\tilde{d}_F}_{\text{Direct effects}} + \underbrace{\frac{t_F}{1 + t_F\lambda_F} d\lambda_F}_{\text{Reallocation effects}} . \tag{3.4}$$

This equation shows that, to a first-order, the change in welfare depends on only two terms, as shown in panel B. The first is the direct effect captured by the green rectangle. Its height corresponds to the change in commuting cost, and its width to the initial number of jobs affected by the shock, $\lambda_F = \frac{L_F}{\bar{L}}$. If the economy is not at its first-best, there is a second term, shown by the purple rectangle. Its height corresponds to the size of the wedge, and its width to the change in formal employment, $d\lambda_F$. Most

[16] For example, in the case of perfect substitutes, $f(q_F, q_I) = q_F + q_I$, and in the CES case, $f(q_F, q_I) = [q_F^{\rho} + q_I^{\rho}]^{1/\rho}$.

papers focus only on the first term and ignore distortions. The goal of this paper is to understand the contribution of the purple rectangle to the overall welfare gains.

Using this formula, I conduct a back-of-the-envelope calculation to gauge the magnitude of the welfare effects. The reduction in commuting costs induced by Line B is $d \ln \tilde{d}_F \approx -2.14\%$. The initial share of formal employment in Mexico City is $\lambda_F = 0.5267$, the formal–informal labor wedge is $\tau_L = 0.92$, and the reduced-form estimate implies a change in the formal employment share of $d\lambda_F = 0.0049$ (under the assumption that the "missing intercept" is zero). Plugging these values into the decomposition yields that Line B increases welfare by approximately 1.31%, with reallocation across the formal and informal sectors amplifying the direct gains by about 30%. While informative, this exercise provides only a first-order approximation. It relies on reduced-form estimates and therefore implicitly assumes that infrastructure has no impact on control areas, leaving the "missing intercept" issue unaddressed. It also abstracts from several features captured by the full quantitative model. For example, commuting costs affect only formal jobs, labor is the sole factor of production, firms do not endogenously determine their formal status or location, and there is a single consumption good.

4 Model

This section presents the full quantitative model, building on recent work in the spatial literature and classic models of informality. I extend the basic framework by adding wedges, firm location choice, and an endogenous decision for firms and workers on formal status. There are three groups of agents: workers denoted by L, house owners denoted by H, and commercial land owners denoted by Z.[17]

4.1 Preferences

There is a discrete set of locations, N, indexed by n and i. There is a mass of L_L workers that operate in 2 sectors indexed by $s \in I, F$, where I and F represent the informal and formal sectors, respectively. The utility function is Cobb-Douglas, and consumers derive utility from a composite consumption good and housing. The composite good is freely tradable, and I normalize its price to 1. The indirect utility of worker ω living in location n and working in sector s and location i is:

$$V_{nism\omega} = \frac{w_{is} \cdot \tilde{d}_{nim\omega}^{-1} \cdot \epsilon_{nis\omega} \cdot \exp(\xi_{n\omega}) \cdot (1 + \bar{t})}{P^{\alpha} r_{Hn}^{1-\alpha}}, \tag{4.1}$$

where w_{is} is the wage per efficiency unit in location i and sector s, $\tilde{d}_{nim\omega}$ is an iceberg commuting cost to move from location n to i using transportation mode m, $\epsilon_{nis\omega}$ is an idiosyncratic productivity shock that determines workplace and sector choice, $\xi_{n\omega}$ is a residential amenity shock, P is the price index of the consumption good (normalized to 1), r_{Hn} is the price for housing with $1-\alpha$ representing the expenditure share in housing, and $\bar{t}$ is a proportional tax rebate from the Government.[18] Workers

[17] The focus of the paper is efficiency. In the Appendix, I generalize the results to account for different worker groups, such as high- and low-skilled workers.

[18] I model the government as rebating fiscal revenues to households in proportion to labor income. This assumption is intended to capture the idea that revenues collected from formal labor are returned to households through Mexico's broader social protection system, while preserving tractability in the spatial equilibrium model. In Mexico,

first choose the transportation mode, then the sector and workplace, and finally the location to live. I solve the problem by backward induction.

Transportation mode choice: Conditional on choosing where to live and work, agents choose the transportation mode to commute from n to i. Individuals have idiosyncratic preferences across transportation modes that generate a nested logit demand across modes m: a nest of public modes $\Omega_{pub} \equiv \{Bus, Metro, Metrobus, Walk\}$ and a nest of private modes $\Omega_{priv} \equiv \{Car\}$. The disutility of commuting from n to i using mode m for agent ω is:

$$\tilde{d}_{nim\omega} = \delta\, time_{nim} - \alpha_m + \nu_{nim\omega}, \tag{4.2}$$

where $time_{nim}$ denotes travel time from residential location n to workplace i using transportation mode m, α_m is a mode-specific utility shifter, and $\nu_{nim\omega}$ is an idiosyncratic preference shock drawn from a nested extreme value distribution. The expected commuting disutility (or generalized commuting cost) from n to i, denoted $\tilde{d}_{ni} = \exp(\delta IVni)$, is determined by the inclusive value IV_{ni} that aggregates utility across transportation modes within the nested logit structure. When the private nest contains a single alternative (Car), the inclusive value simplifies to:

$$IV_{ni} = \frac{-1}{\delta} \ln\left[\exp(\delta\, IV_{ni,\text{pub}}) + \exp(\alpha_{\text{Car}} - \delta\, time_{ni,\text{Car}})\right], \tag{4.3}$$

where $IV_{ni,\text{pub}}$ is the inclusive value for the public nest:

$$IV_{ni,\text{pub}} = -\frac{\chi}{\delta} \ln\left(\sum_{m \in \Omega_{\text{pub}}} \exp\left(\alpha_m - \frac{\delta}{\chi} time_{nim} \right) \right), \tag{4.4}$$

and $\chi \in (0, 1]$ is the nesting parameter capturing the correlation across idiosyncratic preferences within the public modes. The variable $\tilde{d}_{ni}$ is the one that is shocked in the model.

Workplace and sector choices: The productivity shock $\epsilon_{nis\omega}$ is drawn from a extreme-value type II distribution $H(\cdot)$.[19] Workers who live in location n are ex-ante identical but become ex-post heterogeneous due to different draws across sectors and workplaces. Each worker receives a one-time shock and makes two decisions conditional on living in n: 2) sector (formal or informal), and 3) workplace. The parameters κ and θ measure the productivity dispersion across sectors and workplaces and capture the notion of comparative advantage.[20] The parameter θ corresponds to the

formal salaried workers receive contributory benefits, while workers outside formal salaried employment receive non-contributory benefits financed from general revenues. A proportional rebate is convenient because, unlike a location-specific lump-sum transfer, it does not distort migration decisions across space. It also provides a reasonable representation in this context, since payments and benefits in the contributory system are tied to wages, so higher-wage formal workers receive larger transfers in absolute terms. See Levy (2018) for a detailed discussion of contributory and non-contributory programs in Mexico.

[19] $H(\vec{\epsilon}) = \exp\left(\sum_s B_{ns} \left(\sum_i \epsilon_{nis}^{-\theta} \right)^{\frac{\kappa}{\theta}} \right)$ with $1 < \kappa < \theta$.

[20] Different articles have assumed a similar structure to analyze the allocation of workers across sectors. For example, Lagakos and Waugh (2013) studies selection in the agricultural sector; Hsieh et al. (2019) studies the allocation of talent in the US, and Galle et al. (2023) studies the distributional implications of trade.

commuting elasticity and captures how easy it is for workers to substitute jobs across locations, and the parameter κ represents the labor supply elasticity across sectors and measures how easy it is for workers to substitute formal for informal jobs. The parameters B_{ns} capture productivity shifters that attract residents to sector s in location n, which are fixed over time.

The probability of working in (i, s) conditional on living in n is:

$$\lambda_{nisL|n} = \underbrace{\left(\frac{B_{ns} W_{ns|n}^{\kappa}}{\sum_{s'} B_{ns'} W_{ns'|n}^{\kappa}} \right)}_{\lambda_{nsL|n}} \underbrace{\left(\frac{w_{is}^{\theta} \tilde{d}_{ni}^{-\theta}}{\sum_{i'} w_{i's}^{\theta} \tilde{d}_{ni'}^{-\theta}} \right)}_{\lambda_{nisL|ns}}, \tag{4.5}$$

where $W_{ns|n}^{\theta} = \sum_{i'} w_{is}^{\theta} \tilde{d}_{nis}^{-\theta}$ denotes a wage index from location n and sector s, which corresponds to the commuter market access measure for sector s in location n. This probability can be decomposed into two terms: the probability of working in s and the probability of working in i conditional on choosing sector s. The total amount of labor $\tilde{L}_{is}$ hired by (i, s) is equal to the amount of efficiency units supplied by all locations and is:

$$\tilde{L}_{is} = \sum_{n} W_n \lambda_{nis} \cdot \bar{L}_L. \tag{4.6}$$

where $W_n^{\kappa} = \sum_s W_{ns|n}^{\kappa}$ is a wage index that corresponds to the ex-ante average income in n.

Residential choice: The residential amenity shock $\xi_{n\omega}$ is drawn from a Gumbel distribution (Type-I Extreme Value distribution) with shape parameter $\frac{1}{\eta}$. Then, the share of workers that live in n is:

$$\lambda_n = \frac{B_n r_{Hn}^{-(1-\alpha)\eta} W_n^{\eta}}{\sum_{n'} B_{n'} r_{Hn'}^{-(1-\alpha)\eta} W_{n'}^{\eta}}, \tag{4.7}$$

where B_n is an amenity shifter that captures the attractiveness of location n and η is the migration elasticitiy. By the properties of extreme value type shocks, the log expected ex-ante utility is equal to the following constant:

$$\ln \bar{U}_L = \ln \left(\sum_{n'} B_{n'} r_{Hn'}^{-(1-\alpha)\eta} W_{n'}^{\eta} \right)^{\frac{1}{\eta}} + \gamma_\eta, \tag{4.8}$$

where γ_η is the Euler-Mascheroni constant. Preferences for the composite good C take a CES form of different varieties $q(\varphi)$:

$$C = \left(\sum_{i,s} \int_{\varphi \in \Omega_{is}} q_{is}(\varphi)^{\frac{\sigma-1}{\sigma}} d\varphi \right)^{\frac{\sigma}{\sigma-1}}, \tag{4.9}$$

where σ corresponds to the elasticity of substitution across all varieties.

4.2 Production of the Composite Good

On the production side, A mass of entrepreneurs M first decides whether to operate in the formal or informal sector and then chooses a production location within the city.[21] I solve the firm's problem by backward induction: I begin with the choice of location and then characterize the decision between operating formally or informally. The production function for each variety is:

$$q_{is}(\varphi) = A_{is}\varphi \left(\frac{\ell_{is}(\varphi)}{\beta}\right)^{\beta} \left(\frac{z_{is}(\varphi)}{1-\beta}\right)^{1-\beta},$$

where β represents the output elasticity with respect to labor, and $1-\beta$ is the output elasticity with respect to commercial floor space; ℓ and z are the total amount of labor and commercial floor space hired by firm φ, and A_{is} corresponds to the aggregate productivity of location i and sector s, φ is the total factor productivity of the firm. The firm's objective is to maximize profits. If the firm operates in the formal sector, it must pay a labor and a floor space wedge for the use of its inputs.

The problem of firm φ in location i and sector s is to maximize the following profit function:

$$\Pi^{op}_{is}(\varphi) = Q^{\frac{1}{\sigma}} q_{is}(\varphi)^{\frac{\sigma-1}{\sigma}} - (1+t_{Ls})w_{is}\ell_{is}(\varphi) - (1+t_{Zs})r_{Zi}z_{is}(\varphi),$$

where Q is the aggregate quantity of the composite good, determined in general equilibrium. The firm charges a constant markup $\tilde{\sigma} = \frac{\sigma}{\sigma-1}$. Optimal factor demands imply that labor hired by firm φ in (i,s) is:

$$\ell_{is}(\varphi) = \beta \left(\frac{\tilde{\sigma}^{-1} Q^{\frac{1}{\sigma}} (A_{is}\varphi)^{\frac{1}{\tilde{\sigma}}} \left[w_{is}(1+t_{Ls}) \cdot [(1+t_{Zs})r_{Zi}]^{-1}\right]^{\frac{1-\beta}{\tilde{\sigma}}}}{w_{is}(1+t_{Ls})} \right)^{\sigma}. \tag{4.10}$$

Labor demand is increasing in firm productivity φ and sector–location productivity A_{is}, and decreasing in the effective wage $(1+t_{Ls})w_{is}$. From the firm's optimality conditions, operating profits are:

$$\Pi^{op}_{is}(\varphi) = \hat{\sigma} \left(\frac{Q \cdot A_{is}\varphi}{[w_{is}(1+t_{Ls})]^{\beta} [r_{Zi}(1+t_{Zs})]^{1-\beta}} \right)^{\sigma-1}, \tag{4.11}$$

where $\hat{\sigma} = \frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}}$ is a constant. Let $c_{is} \equiv [w_{is}(1+t_{Ls})]^{\beta} [r_{Zi}(1+t_{Zs})]^{1-\beta}$ denote the unit input costs for location i and sector s. The previous expression implies that profits are increasing and convex in productivity φ. A key implication is that if $\frac{A_{iF}}{c_{iF}} \geq \frac{A_{iI}}{c_{iI}}$, the profit gap between the formal and informal sectors increases in φ.

Intuitively, when the formal sector offers higher productivity relative to its input costs, high-productivity firms benefit disproportionately from operating formally. In contrast, informal firms face a productivity penalty that becomes increasingly costly as firm productivity rises. As a result, more productive firms have stronger incentives to select in formality. This mechanism gives rise to a cutoff productivity level $\tilde{\varphi}$ at which firms are indifferent between formal and informal production,

[21] In the Online Appendix, I present two alternative formulations of the production environment: (i) a case with a representative firm for each sector and location operating under perfect competition and (ii) a setting with homogeneous firms in each sector and location under monopolistic competition.

consistent with standard models of informality.[22] In the quantification, I show that, on average, this condition holds in the Mexican data.

4.2.1 Firm location choice: For each firm with productivity φ, a mass of entrepreneurs decide where to locate in the city based on an idiosyncratic shock as in Suárez-Serrato and Zidar (2016) and Caliendo and Parro (2020). The utility of entrepreneur ω with productivity φ in sector s that locates in area i is:

$$\log V_{is\varphi\omega} = \log \pi_{is}(\varphi) + \phi_{is} + \zeta_{is\omega}(\varphi),$$

where ϕ_{is} is a firm-location shifter and $\zeta_{is\varphi\omega}$ is an idiosyncratic shock that is drawn from a Type-I Extreme value distribution with zero mean and dispersion parameter Ψ. Firms of productivity φ produce in location i if their value function is higher than in any other location j. The fraction of firms $\chi_{is}(\varphi)$ that decide to produce in location i is:

$$\chi_{is}(\varphi) = \frac{\Phi_{is}\pi_{is}(\varphi)^{1/\psi}}{\sum_j \Phi_{js}\pi_{js}(\varphi)^{1/\psi}}, \tag{4.12}$$

the parameter ψ has a similar interpretation to the migration elasticity in the case of households. It measures how responsive firm location choices are to changes in profits and how easy it is for firms to substitute locations in the city. A higher value implies that the productivity distribution is more dispersed across locations, and, as a result, it is more difficult to substitute across areas.

4.2.2 Firm decision between the formal and informal sector: Each entrepreneur receives an ex-ante productivity shock φ that determines the firm's overall productivity. This productivity shock is drawn from an $F(\cdot)$ distribution, which I assume is Pareto. Based on this productivity shock and the ex-ante expected profits in each sector, the entrepreneur decides whether to operate in the formal or informal sector.[23] An entrepreneur with productivity φ chooses to operate in the formal sector if:

$$\bar{\pi}_F(\varphi) - FC_F \geq \bar{\pi}_I(\varphi), \tag{4.13}$$

where $\bar{\pi}_s(\varphi)$ represents the ex-ante expected profits in each sector, and FC_F corresponds to a production fixed cost that firms need to pay to produce in the formal sector and is in terms of the numeraire. The expected profit of variety φ in sector s is a weighted average of profits across locations:

$$\bar{\pi}_s(\varphi) = \sum_i \chi_{is}(\varphi)\pi_{is}(\varphi). \tag{4.14}$$

As noted above, the formal sector is, on average, more productive than the informal sector. Then, the

[22] The model could be extended to include size-dependent costs of informality, but such costs would generate misallocation within the informal sector. To isolate the core selection mechanism emphasized in the informality literature, the model abstracts from these costs and focuses instead on the productivity penalty. Notably, the key prediction remains: more productive firms gain relatively more from operating formally than less productive firms.

[23] I model the choice of formality status as an ex-ante decision based on average profits, with location chosen conditional on that sectoral choice. This captures formality as a broader organizational and regulatory regime that affects firms' cost structures across locations, rather than as a purely location-specific margin.

difference between the expected profits in the formal vs. the informal sector increases monotonically in φ. As a result, a marginal firm $\tilde{\varphi}$ is indifferent between producing in the formal vs. the informal sector. All firms with a productivity level $\varphi \geq \tilde{\varphi}$ produce in the formal sector and pay the wedges and fixed costs, while all firms with a productivity level $\varphi < \tilde{\varphi}$ produce in the informal sector and face the productivity penalty.

4.3 Labor Market Clearing

With all these ingredients, the labor market clearing conditions imply that the total amount of labor hired by firms should equalize the total number of efficiency units supplied by workers. For the formal sector in location i, this condition implies that:

$$\tilde{L}_{iF} = M \int_{\varphi \in [\tilde{\varphi}, \bar{\varphi}]} \chi_{iF}(\varphi) \ell_{iF}(\varphi) \mu(\varphi) d\varphi, \tag{4.15}$$

where M represents the mass of firms, and μ is the density function of the $F(\cdot)$ distribution with an upper support $\bar{\varphi}$. Similarly, in the case of the informal sector, the labor market clearing condition is:

$$\tilde{L}_{iI} = M \int_{\varphi \in [\underline{\varphi}, \tilde{\varphi}]} \chi_{iI}(\varphi) \ell_{iI}(\varphi) \mu(\varphi) d\varphi, \tag{4.16}$$

where $\underline{\varphi}$ is the lower bound of the $F(\cdot)$ distribution. The left-hand side of equations 4.15 and 4.16 is an increasing function in wages, while the right-hand side is a decreasing function of wages. Hence, there is a fixed point at which supply and demand are equal.

4.4 Housing and Commercial Floor Space

I assume that there are two additional industries: $\tilde{H}$ and $\tilde{Z}$, which produce residential housing and commercial floor space, respectively. Both industries operate under perfect competition and the only factors of production are the agents H and Z, and there is no commuting. The production function for both sectors is linear in labor. Hence, the supply of residential and commercial floor space is perfectly inelastic. The prices are given by r_{Hi} and r_{Zi}, respectively. The equilibrium conditions for the commercial floor space and housing markets are:

$$r_{Zi} \tilde{Z}_i = \frac{(1-\beta)}{\beta} \sum_s w_{is} \tilde{L}_{is}, \quad r_{Hn} \tilde{H}_n = (1-\alpha) X_n, \tag{4.17}$$

where X_n is the total expenditure from location n and is given by the following equation:

$$X_n = \left(\bar{y}_n L_n + r_{Zn} Z_n + r_{Hn} H_n + \iota_n \Pi\right)(1+\bar{t}), \tag{4.18}$$

where ι_n corresponds to the share of profits assigned to location n and Π to the total profits obtained by firms in the economy. I assume that ι_n is proportional to the labor income of location n.

4.5 Government Budget Constraint

The Government collects taxes and gives a rebate to households captured by $\bar{t}$. I assume that the rebate is proportional to household income. The rebate is:

$$\sum_{i,s}\left(t_{Ls}w_{is}\tilde{L}_{is}+t_{Zs}q_i\tilde{Z}_{is}\right)=\bar{t}\cdot\sum_{n}X_n. \tag{4.19}$$

This equation implies that the Government's income from tax revenue equals its total expenditure.

4.6 Equilibrium

The equilibrium of the model is described by the following vector of endogenous variables:

$$x=\{w_{is},r_{Fi},r_{Hn},\bar{y}_n,W_{ns},P_{is},\tilde{L}_{is},\tilde{Z}_{is},L_n,\tilde{\varphi},\chi_{is},X_n\},$$

and the constant $\bar{U}$ given a set of exogenous parameters:

$$A=\{\tilde{d}_{ni},A_{is},B_n,T_{is},\bar{L},\bar{L}_H,\bar{L}_Z,\tilde{Z}_i,\tilde{H}_i,M,\tau,b_i,FC,\theta,\kappa,\eta,\sigma,\psi,\alpha,\beta\},$$

that solve the following system of equations: workplace and sector choice probabilities from equation 4.5; residence choice probabilities from equation 4.5; labor conditions in the formal and informal sectors 4.10; firm choice probabilities from equation 4.12; labor market clearing conditions from equations 4.15 and 4.16; commercial floor space and housing market clearing conditions from equation 4.17; total expenditure from equation 4.18; and the Government budget constraint from equation 4.19.

4.7 Welfare Decomposition

I assume a social planner who takes a utilitarian perspective. The aggregate welfare function is:

$$\bar{U}=\left(\omega_L\bar{U}_L+\omega_H\bar{U}_H+\omega_Z\bar{U}_Z\right), \tag{4.20}$$

where ω_g represents the weights for each group of agents in the economy.[24]

Let $\mathcal{L}$ denote an allocation of factors given fundamentals A, and let $\mathcal{U}(A,\mathcal{L})$ be the associated welfare level $\bar{U}$. The expression below applies under any of the following environments: (i) residential populations are fixed; (ii) workers choose where to live with $\eta\to\infty$; or (iii) a planner maximizes total output or a weighted sum of real income across locations, using initial income shares as weights.[25] A first-order approximation to the welfare effect of a commuting shock is:

[24] For the parametric case of this model, these weights solve the following expressions: $\frac{\omega_L\bar{U}_L}{\bar{U}}=\alpha\beta$, $\frac{\omega_Z\bar{U}_Z}{\bar{U}}=\alpha(1-\beta)$, and $\frac{\omega_H\bar{U}_H}{\bar{U}}=(1-\alpha)$.

[25] These assumptions isolate the economic impact of wedges in the absence of other market failures. Related work, such as Fajgelbaum and Gaubert (2020) and Donald et al. (2025) (DFM), studies the first-order effects of other market failures. For example, DFM shows that congestion arising from finite taste dispersion can generate ex-post differences in marginal utility across locations.

$$d\ln \bar{U} = \underbrace{\frac{\partial \ln \bar{\mathcal{U}}}{\partial \ln A} d\ln A}_{\text{“Direct” effect}} \quad + \quad \underbrace{\frac{\partial \ln \bar{\mathcal{U}}}{\partial L} dL}_{\text{Allocation effects}} \quad . \tag{4.21}$$

Equation 4.21 implies that the welfare effect of any shock can be decomposed into two components: a direct effect, reflecting changes in exogenous parameters such as iceberg commuting costs $\tilde{d}_{ni}$, and a first-order allocative term capturing reallocation effects. The change in welfare is:[26]

$$\text{“Direct” effect} = \sum_{n,i,s} -\alpha\beta \cdot \tilde{\lambda}_{nisL} \cdot d\ln \tilde{d}_{ni} \tag{4.22a}$$

$$\text{Allocation} = \alpha \left(\sum_{n,i,s} \beta \left(\frac{t_{Ls} - \bar{t}}{1+\bar{t}} \right) \tilde{\lambda}_{nisL} \cdot d\ln \tilde{L}_{nis} + \sum_{n,s} (1-\beta) \left(\frac{t_{Zs} - \bar{t}}{1+\bar{t}} \right) \tilde{\lambda}_{nsZ} \cdot d\ln \tilde{Z}_{ns} \right). \tag{4.22b}$$

The first term represents the “direct” effect of the commuting shock. In a perfectly efficient economy, this term implies that the cost-time-saving approach captures the welfare impact of any commuting shock to a first order. In such an economy, it is sufficient to know the value of jobs between locations n and i, since all nominal adjustments cancel out. This result follows from the envelope theorem: when the allocation is first-best, shocks to commuting costs have no first-order welfare effects through changes in the optimizers. This effect corresponds to the green area in Figure 2.

The presence of wedges implies that the economy is not at its first-best, leading to the second term, which captures changes in allocative efficiency. This term reflects whether workers reallocate toward sectors with higher or lower wedges. Firms facing greater distortions exhibit higher Total Factor Revenue Productivity (TFPR). If the commuting shock induces reallocation toward more distorted firms, TFPR dispersion falls, the allocation moves closer to the optimum, and welfare increases. This effect corresponds to the sum across input markets of the purple areas in Figure 2.

5 Estimation of the model parameters

This section describes the parameter estimation strategy. I first present the commuting parameters, then outline the GMM procedure, and finally discuss the parameters calibrated from cross-sectional moments. Table 3 summarizes the parameters and their identification.

1. Commuting costs and elasticities: I first estimate the parameter δ, which measures the marginal disutility of travel time and converts minutes into utility units within the transportation-mode choice problem. This parameter is estimated directly from observed mode choices in the origin-destination (OD) survey using the nested logit structure. From equation 4.2, and using the properties of extreme value shocks, the share of workers commuting from municipality n to municipality i in

[26] The proof is provided in Appendix Section E.2. Section E.3 further generalizes this result to multiple worker-types and general utility and production functions by solving the social planner's problem.

sector s who choose transportation mode m is:

$$\lambda_{nism|nis} = \left(\frac{\exp\left(\alpha_m - \frac{\delta}{\chi_j} time_{nim}\right)}{\sum_{r\in\Omega_j} \exp\left(\alpha_r - \frac{\delta}{\chi_j} time_{nir}\right)} \right) \cdot \frac{\left(\sum_{r\in\Omega_j} \exp\left(\alpha_r - \frac{\delta}{\chi_j} time_{nir}\right)\right)^{\chi_j}}{\sum_h \left(\sum_{r'\in\Omega_h} \exp\left(\alpha_{r'} - \frac{\delta}{\chi_h} time_{nir'}\right)\right)^{\chi_h}}. \tag{5.1}$$

The travel-time disutility parameter δ and the mode-specific preference shifters α_m are estimated using individual-level mode choices from the OD survey through maximum likelihood. Identification relies on variation in travel times across transportation modes and origin-destination pairs, which disciplines substitution patterns across modes. The estimated value, $\delta \approx 0.0096$, is consistent with existing estimates in the literature (see Table B13 in the Appendix).

Commuting flows across municipalities are then constructed using the 2015 Intercensal Survey. For each municipality pair, I compute the inclusive value of travel times using equation 4.3 and map it to iceberg commuting costs. Travel times are computed separately for each transportation mode using mode-specific speeds; for car travel, speeds additionally depend on road type (see Table D24 in the Appendix).[27] The resulting iceberg commuting costs enter the following gravity equation:

$$\ln \lambda_{nis|ns} = -\theta \ln \tilde{d}_{ni} + \gamma_{is} + \gamma_{ns} + \epsilon_{nis}, \tag{5.2}$$

where $\lambda_{nis|ns}$ is the share of workers in sector s residing in municipality n who work in municipality i; $\tilde{d}_{ni}$ denotes iceberg commuting costs; γ_{ns} and γ_{is} are origin–sector and destination–sector fixed effects; and ϵ_{nis} captures measurement error. The parameter θ governs the elasticity of commuting flows with respect to commuting costs and summarizes the degree of spatial substitutability across job locations within the city. Equation 5.2 is estimated by Poisson pseudo–maximum likelihood (PPML) to include zero commuting flows across municipalities. I include origin-sector fixed effects and destination-sector fixed effects, which absorb origin-specific labor supply and destination-specific labor demand factors within each sector. Identification, therefore, comes from variation in bilateral commuting costs across destinations for a given origin and across origins for a given destination.[28]

Table 2 reports the estimates. Commuting flows decline sharply with commuting costs, with an elasticity of 6.02. This magnitude is comparable to estimates in other urban contexts and implies that workplaces within Mexico City are gross substitutes.[29]

2. Generalized method of moments and indirect inference: I estimate the parameter vector $\Theta \equiv \{\zeta, \kappa, \eta\}$ using a Generalized Method of Moments (GMM) procedure based on three empirical moments. Two parameters, κ and η, are identified through an indirect inference approach, matching the reduced-form treatment dummy effects estimated in the data to their model-implied counterparts,

[27] For example, walking speed is set to 5 km per hour (80 meters per minute), while metro speed is set to 35 km per hour (584 meters per minute). For metro trips, I additionally include five minutes to account for station entry and exit.

[28] Taken together, the nested logit model identifies how travel time maps into commuting costs through δ, while the gravity equation identifies how commuting flows respond to those commuting costs through θ. The two parameters are therefore identified from different behavioral margins in the data.

[29] Tsivanidis (2023) estimates an elasticity of 3.92 for Bogotá, and Khanna et al. (2022) a value of 4.94 for Medellín.

while ζ is identified using a direct cross-sectional moment.

The parameter κ is identified by matching the reduced-form coefficient on the treatment indicator from Table 1, Panel B, where the outcome is the log ratio of formal to informal workers at the residential level. This moment is informative about κ because the parameter governs the relative responsiveness of formal versus informal employment to changes in commuting costs from the labor supply side. The parameter η is identified by matching the reduced-form treatment coefficient where the outcome is the log number of residents, capturing the elasticity of residential location choice with respect to changes in market access (see Table B8).[30] The remaining parameter, ζ, which governs the shape parameter of the Pareto productivity distribution, is identified by matching the standard deviation of log value added in the data. In the model, higher values of ζ imply less dispersion in firm productivity and, consequently, lower dispersion in value added, yielding a monotonic mapping from this moment to ζ. The GMM estimator solves

$$\hat{\theta} = \arg\min_{\theta} \left[m^{\text{data}} - m^{\text{model}}(\theta)\right]' W \left[m^{\text{data}} - m^{\text{model}}(\theta)\right], \tag{5.3}$$

where the vector of moments is:

$$m = \left(\hat{\beta}^{\ln\left(\frac{L_{Fn}}{L_{In}}\right)}, \hat{\beta}^{\ln L_n}, \text{SD}(\log \text{VA})\right)', \tag{5.4}$$

and m corresponds to the moments used in estimation and W denotes the weighting matrix, which I set equal to the identity. Standard errors are computed using a bootstrap with 50 replications. Table 4 reports the point estimates and 95% confidence intervals for each parameter. The estimated productivity shape parameter is $\zeta = 2.9$, similar to the value found by Ulyssea (2018) in the Brazilian context. The sectoral labor supply elasticity, $\kappa = 1.7$, indicates frictions to worker reallocation across the formal/informal sectors and is consistent with recent estimates in the literature (Galle et al., 2023). The migration elasticity, $\eta = 0.4$, implies limited spatial mobility within the city relative to other contexts, as it suggests that reallocation across residential areas is more difficult in this context.

Figure 3 illustrates the relationship between the estimated parameters and the identifying moments in the spirit of Andrews et al. (2017). The figure also reports three curves corresponding to different values of the commuting elasticity, using the 95% confidence interval bounds. In the first three rows, the diagonal panels display a monotonic relationship between each parameter and its corresponding moment. Higher values of the Pareto shape parameter ζ reduce the standard deviation of log value added since there is less productivity dispersion, while larger values of κ and η increase the treatment dummy effects on the formal–informal employment ratio and population, respectively. The off-diagonal panels are largely flat, indicating limited cross-moment effects. This means that the different parameters do not affect other moments; they affect only their corresponding moments. The commuting elasticity has little effect in most moments and only mildly affects the formal-informal ratio since it enters directly into the wage indices, W_{ns}. The figure suggests that the moments provide

[30] I use the point estimates of my preferred specification, which includes state fixed effects and the full set of covariates.

local identification and that the objective function exhibits a well-defined minimum.[31]

3. Other parameters in the cross-section: I calibrate additional model parameters using simple data moments. For the housing expenditure share, I use household expenditure survey data and find $\alpha = 0.822$, implying a housing expenditure share of $(1 - \alpha) = 0.178$. Similarly, using aggregate data from the 1999 Economic Census, I find an average labor share value of $\beta = 0.85$. To calculate the total amount of housing $(\tilde{H})$ and commercial floor space $(\tilde{Z})$, I use the area in each tract from the Global Human Settlement Layer in 2000, weighted by the total number of employees and residents.

I calibrate the fixed cost for formal firms (FC_F) to match the fact that 83% of firms operate informally. I then calibrate the parameter ψ in equation 4.12, which governs how the number of firms responds to differences in profitability across locations, and obtain an estimate of $\psi = 13.5$.[32] To simulate the economy, I discretize the firm productivity distribution using a grid of 1,000 firms and calibrate the bounds $\underline{\varphi}$ and $\bar{\varphi}$ using the 10th and 90th percentiles of the residuals from a regression of log value added per worker on census-tract and sector fixed effects in the pre-period.

Moreover, I find that firms face a productivity penalty relative to their effective costs when operating in the informal sector. As a result, the difference between average profits in the formal and informal sectors increases with the ex-ante productivity, φ. Figure A8 plots informal profits, formal gross profits, and formal profits net of the fixed cost. As φ increases, the gap between the profit functions widens, generating a unique productivity cutoff $\tilde{\varphi}$ that determines selection into formality.

Additionally, following the spatial literature, I invert the model and recover productivity (A_{is}), residential shifters (B_n), sectoral shifters (B_{ns}), and firm location shifters (Φ_{is}) so that the model matches the observed baseline spatial distribution of employment, firms, and population. This is important because counterfactual exposure to changes in commuting costs depends on the initial allocation of economic activity. Section D.4 describes the methodology.

6 Counterfactual Analysis

This section evaluates the economic impact of Line B through the lens of the model. The counterfactual exercise is a shock to bilateral commuting costs $\tilde{d}_{ni}$. I compute travel times between all pairs of census tracts before and after the opening of Line B using the Spatial Analyst Network toolkit in ArcMap. I assign different speeds to each transportation mode (see Table D24). The pre-period network excludes Line B, while the post-period network includes it as the only change. The observed changes in travel times are then mapped into changes in iceberg commuting costs using equation 4.3, which combines travel-time reductions and the parameter δ. This step translates time savings into changes in commuting costs. I then solve for the general equilibrium before and after the shock.

Throughout the analysis, I assume that the city is closed, as the infrastructure improvement

[31] The figure also includes a 4th row, showing that the firm-reallocation elasticity, $\frac{1}{\psi}$, does not affect the identifying moments of the other parameters.

[32] Section D.2 in the Appendix describes the estimation in detail. I estimate a regression relating the number of firms in a census tract to the average value added. Because this regression may be endogenous due to firm selection and also suffer from mechanical division bias, I instrument the average value added with the average wage in the census tract.

consists of a single transit line. In the baseline counterfactual, workers can reallocate across residential locations within the city with a finite η, while firms' locations are fixed. Afterward, I explore alternative assumptions about the mobility of workers and firms.

The main objective is to assess whether distortions amplify the welfare gains from transit investment and, if so, by how much. To isolate the role of misallocation, I compare two economies: (i) a benchmark economy without wedges, in which formal firms do not face distortions, and (ii) a distorted economy in which wedges generate misallocation between formal and informal firms. In each case, I invert the model to recover initial productivity and amenity shifters, and then compute the counterfactual equilibrium. I compute confidence intervals for the counterfactual results using a bootstrap procedure with 50 repetitions.

6.1 Magnitude of the commuting shock

Before turning to aggregate outcomes, I characterize the magnitude and spatial heterogeneity of the commuting-cost shock. Following Tsivanidis (2023) and Donaldson and Hornbeck (2016), I construct a sector-specific commuter market access (CMA) measure that summarizes changes in job accessibility resulting from reduced commuting costs. Specifically, commuter and firm market access are computed by solving the following system of equations before and after the construction of Line B:

$$\text{CMA}_{ns} = \sum_i \frac{\tilde{L}_{is}\tilde{d}_{ni}^{-\theta}}{\text{FMA}_{is}}, \qquad \text{FMA}_{is} = \sum_n \frac{L_{ns}\tilde{d}_{ni}^{-\theta}}{\text{CMA}_{ns}}, \tag{6.1}$$

where $\tilde{L}_{is}$ denotes employment in location i and sector s, L_{ns} denotes the number of residents in location n employed in sector s, and FMA_{is} captures the ease with which firms attract workers.[33] Sector-specific measures are aggregated into a single index CMA_n using initial sectoral employment shares and the labor supply elasticity κ. Changes in CMA_n therefore summarize how access to jobs improves across space as a direct consequence of the reduction in commuting costs.

Panel A of Figure 4 illustrates substantial spatial heterogeneity in the magnitude of the commuting-cost shock. Locations closer to Line B experience larger reductions in commuting costs and correspondingly larger increases in commuter market access. Average commuting costs decline by about 7.0% in the most exposed areas, compared to roughly 2.5% in less exposed areas, implying a relative reduction of approximately 4.5 percentage points. This difference translates into an average increase in the commuter market access index that is 7.3% higher in "treated" census tracts relative to "control" tracts. Measured directly in terms of travel times between each census tract and all other locations in the city, average commuting times decline by 5.3% in treated areas and by 1.2% in control areas.[34]

[33]When computing market access, the spatial distribution of workers and residents is held fixed at its initial level, so changes in the index reflect only changes in commuting costs.

[34]For example, the metro commuting time using Line B from Ecatepec de Morelos to the Mexico City CBD is approximately 50 minutes, compared to around 80 minutes using alternative transportation modes.

6.2 Aggregate welfare and output effects

Table 5 reports the aggregate effects of Line B on welfare and output. The first four columns present welfare effects computed using equation (4.20), while the last four columns report changes in consumption-good output from equation (4.9). For each outcome, the table reports the mean, standard deviation, and 95% confidence interval obtained from the bootstrap procedure.

Panel A presents the baseline results under the assumption that workers can reallocate across residential locations. The first row reports outcomes in the benchmark economy without wedges. In this case, the reduction in commuting costs increases aggregate welfare by approximately 0.63%, with a 95% confidence interval ranging from 0.49% to 0.78%. Consumption-good output rises by a similar magnitude, with average gains of 0.63% and a confidence interval between 0.50% and 0.76%. The second row reports results for the distorted economy in which wedges generate misallocation between formal and informal firms. Introducing these wedges amplifies the gains from the commuting-cost shock. Welfare gains increase to 0.75%, with a confidence interval between 0.58% and 0.92%, while output gains rise to 0.85%, with a confidence interval between 0.67% and 1.03%. These differences reflect improvements in allocative efficiency induced by the infrastructure investment.

Figure 5 summarizes the additional gains from allocative efficiency, expressed as the percentage increase relative to the direct undistorted gains, along with 95% confidence intervals. In the baseline specification, allocative efficiency amplifies welfare gains by approximately 18% (95% CI: 14%-23%) and output gains by about 34% (95% CI: 31%-39%). These results indicate that reallocative effects are quantitatively meaningful and that ignoring distortions would substantially understate the benefits of transit investment in this setting.

I proceed to translate these effects into a cost-benefit analysis. According to official government documents, the net present value of Line B was approximately USD 2,230 million in 2014, corresponding to about 0.55% of Mexico City's GDP. In the benchmark economy, this implies a return of roughly 1.15 USD per dollar invested. Once allocative efficiency effects are incorporated, the return increases to approximately 1.37 USD per dollar spent, representing an 18% increase in total welfare gains. The difference is even bigger if we use as a benchmark the effects on the output gains. In the baseline, the gains would have been 1.15 USD per dollar spent, and 1.56 per dollar spent, considering allocative efficiency. These findings imply that infrastructure projects yield smaller welfare gains when implemented in areas where most workers are already formal, underscoring the importance of initial distortions in evaluating infrastructure investments.

Comparison with the first-order approximation: At the aggregate level, the transit improvement reduces informality by about 0.81% and increases formality by 0.73%. Using equation (4.22), I compare the full-model results with those implied by the first-order approximation. The direct effects are similar between the FOA and the full model, with aggregate welfare gains of approximately 0.63%. For allocative efficiency, however, the first-order approximation implies additional gains of about 0.19 percentage points, compared to 0.12 percentage points in the full welfare model and 0.22 percentage points for output.[35] The FOA may overstate welfare gains because it treats commuting shocks as infinitesimal and abstracts from endogenous wage adjustments, which dampen reallocative responses in the full equilibrium.[36]

Mobility scenarios: I also examine the sensitivity of these results to assumptions about worker and firm mobility. Panels B, C, and D of Table 5 report results for alternative scenarios in which (i) workers' residential locations are fixed at their 2000 distribution, (ii) both workers and firms can reallocate, and (iii) only firms can reallocate. Figure 5 reports the corresponding additional gains from allocative efficiency. Across all cases, the results are remarkably stable. Welfare gains remain close to 0.63% without wedges and 0.75% with wedges, while output gains range between 0.63–0.65% without wedges and 0.85–0.87% with wedges. These patterns indicate that the improvements in allocative efficiency are primarily driven by reductions in commuting costs rather than by large-scale spatial reallocation of workers or firms.[37]

City with higher baseline formality: To further assess the role of labor reallocation in a city with high informal employment, I implement the Line B counterfactual using as a baseline a city characterized by a higher initial level of employment formality. To construct this baseline, I first simulate an economy in which the productivity penalty faced by informal firms is increased tenfold, raising the baseline aggregate share of formal employment to 97%. I then evaluate the counterfactual by introducing the changes in travel times. In this environment, because formality is already very high, the scope for additional reallocation from the informal to the formal sector is limited, implying smaller gains from the allocative efficiency margin.

The overall welfare gains from the transportation investment are slightly larger, reflecting the higher initial concentration of formal jobs in central locations. At the same time, the additional gains from allocative efficiency are negligible: welfare and output gains are nearly identical in simulations with and without the wedge (see Table B17). These results indicate that when baseline formality is already high, transportation investments primarily operate through direct effects rather than through reallocation adjustments across sectors.

Alternative model specifications: I also consider alternative production-side specifications. Section C of the Online Appendix presents two additional formulations of the production environment: (i) a representative-firm model in each sector and location operating under perfect competition, and (ii) a setting with homogeneous firms under monopolistic competition and free entry.[38]

Overall, the results are qualitatively similar to those obtained in the heterogeneous-firm benchmark. In the representative-firm case, the absence of free entry and the lack of endogenous firm

[35] This implies a gain of 0.30 usd on allocative efficiency per dollar spent in the infrastructure.

[36] Geometrically, the change in allocative efficiency is a trapezoid, but the FOA captures it as a rectangle as it assumes that changes in commuting costs are small.

[37] This result is also consistent with the findings in Table B11 that show that the effects on the formal to informal ratio is similar in the baseline estimation and in the case in which workers did not change their residential state.

[38] For both specifications, I fully recalibrate the model using the indirect inference procedure to discipline the key elasticities and then re-compute the counterfactuals. This section also reports a set of robustness exercises that vary the main structural parameters.

transitions between the formal and informal sectors limit the available margins of adjustment, leading to smaller aggregate gains. Welfare gains in this specification are approximately 0.66%, with the allocative efficiency margin contributing an additional increase of about 15%. Output gains are around 0.71%, with allocative efficiency accounting for an additional increase of roughly 25%. By contrast, under monopolistic competition with free entry, the results closely resemble those of the heterogeneous-firm model. Welfare gains are approximately 0.78%, with the allocative efficiency margin amplifying gains by about 18%, while output gains are around 0.89%, with an amplification of roughly 30% due to allocative efficiency (see Table C21).

Other robustness checks: I also conduct a series of robustness checks by varying the model's key structural parameters. Specifically, I consider lower and higher values of the commuting elasticity, θ; the sectoral labor supply elasticity, κ; the migration elasticity, η; the firm reallocation elasticity, $\frac{1}{\psi}$; and the elasticity of substitution, σ. In all specifications, the qualitative and quantitative results are very similar, and the effects are stable across the different mobility scenarios (see Table B18).

Figure 6 plots the additional gains associated with allocative efficiency for each robustness specification. The orange bars represent additional welfare gains, and the blue bars the additional output gains. The black lines denote 95% confidence intervals. Allocative efficiency increases welfare gains by approximately 15–20% and output gains by 30–40% across most parameter values. The smallest gains arise when the sectoral labor supply elasticity, κ, is relatively low (1.2), reflecting limited worker reallocation from the informal to the formal sector. In this case, welfare gains increase by about 10% and output gains by roughly 25%. Conversely, when κ takes a higher value (2.4), reallocation is stronger, and the allocative efficiency channel becomes more important: welfare gains increase by more than 20% and output gains by over 40%. These results demonstrate that the main findings are robust to substantial variation across a range of plausible elasticities and confirm that allocative efficiency is an important channel.

6.3 Model Validation

This section evaluates the model's performance by comparing its predictions with the data. I first compare cross-sectional moments and then test whether the model's predictions replicate those observed in the data, using changes in market access as an instrument.

Cross-section: First, I show that the model performs very well in the cross-section by matching the initial equilibrium with the employment and population distribution by inverting the model. I start by correlating the productivity and amenity shifters, A_{is} and B_n, from the model's inversion with external productivity and amenity measures using the 1999 Economic Census and 2000 Population Census. These external measures are not used in the calibration.

The productivity measure for formal firms is positively associated with the log of the average value-added per worker and the average wage at the census tract level (see table B14). A 10% increase in value-added is associated with a 2.8% increase in the productivity parameter for formal firms and a 0.46% increase for informal firms. Additionally, the relationship is steeper for formal firms, indicating greater dispersion in productivity across locations in the formal sector. Regarding amenities, there is a positive correlation between the amenities recovered by the model and various measures of public services from the 2000 population census (see table B15). Specifically, a 1 percentage point increase in the share of households without electricity is associated with a 0.4% decrease in local amenities. Similar correlations are found for households without access to water and sanitation or living in poor-quality dwellings. The results are robust to different sets of fixed effects.

Moreover, I calibrate the model at the *census-tract level* using employment and residential information. Because the calibration does not use commuting flows at the municipality level, these flows provide an out-of-sample validation of the model. I construct the commuting flows across census tracts implied by the model and then aggregate them to the municipality level, which allows me to compare them with the flows observed in the 2015 Intercensal Census.[39] I then estimate a PPML regression of observed commuting flows on the flows predicted by the model, including origin fixed effects, destination fixed effects, or both. Across all specifications, the model closely matches the data. The estimated coefficient is close to one, and in none of the specifications I can reject the null hypothesis that the coefficient equals one (see Table B16). These results indicate that the model captures well the spatial structure of commuting patterns, despite not being calibrated to match these flows directly.

Changes in the data vs. the model: Finally, I assess whether the model can replicate observed changes in population and the formal-to-informal employment ratio.

First, I show that the model generates changes in the formal-to-informal employment ratio that are consistent with sector-specific changes in market access implied by the data. Figure 4 illustrates the mechanism behind the additional gains. Panel B plots the change in sector-specific market access following the construction of Line B. Treated locations experience larger increases in market access for formal employment than for informal jobs.[40] This differential improvement in job access induces reallocation toward the formal sector. Panel C plots the model-predicted change in the formal/informal ratio against changes in relative market access, revealing a positive relationship: locations that experience larger formal-sector access gains exhibit stronger shifts toward formality. These patterns provide a transparent mechanism that links the commuting-cost shock to allocative-efficiency gains and to workers' reallocation toward the formal sector.

Second, following Costinot and Donaldson (2012) and Adao et al. (2023, henceforth ACD), I evaluate whether the model reproduces the patterns observed in the data using the following regression:

$$\Delta y_i^{\text{data}} = \beta \, \Delta y_i^{\text{model}} + \gamma X_i + \delta_{s(i)} + \epsilon_i, \tag{6.2}$$

where Δy_i^{data} denotes the observed change in outcome y in location i, and $\Delta y_i^{\text{model}}$ is the corresponding change predicted by the model. The vector X_i includes initial characteristics, using the same covariates as in specification 3.2, and $\delta_{s(i)}$ denotes state fixed effects. I estimate this regression for

[39] Since the observed commuting data are from 2015, I use the commuting flows predicted by the model after the construction of Line B.

[40] To construct these market access measures, I vary only iceberg commuting costs while holding the spatial distribution of workers and residents fixed.

the baseline counterfactual. To isolate variation attributable to the infrastructure shock and reduce concerns about mechanical correlation between the estimation and validation moments discussed in ACD, I instrument the model-predicted changes using the change in commuter market access. Under the null hypothesis that the model is correctly specified, the coefficient satisfies $\beta = 1$.

Table 6 presents the results for the baseline counterfactual. Panel A reports estimates for changes in the formal-to-informal ratio, while Panel B reports estimates for changes in population. Differences across columns reflect alternative sets of control variables and fixed effects. Across specifications, the estimated coefficients are close to one. The third row in each panel reports p-values for tests of the null hypothesis $\beta = 1$. In all cases, the null cannot be rejected: the lowest p-value is 0.68 for the formal-to-informal ratio and 0.46 for population.

A potential concern is that the validation exercise may inherit variation already used in calibration, since the parameters are estimated to match the treatment dummy effects generated by the same infrastructure shock. Following the discussion in ACD, the relevant issue is the extent to which the testing moment is mechanically related to the estimation moments. In my setting, calibration targets average treatment effects, whereas the validation regression exploits continuous cross-sectional variation in market-access changes. While these objects are related, the overlap is limited: after residualizing both variables using the controls in equation (6.2), the treatment indicator explains approximately 24% of the variation in market-access changes. This suggests that the validation exercise is not simply a mechanical restatement of the calibration moments.

As an additional and more stringent validation exercise, I conduct two tests using samples not directly targeted in the estimation. First, I re-estimate equation (6.2) excluding locations within 2.5 to 5 kilometers of the new transit line (see Panels C and D in Table 6) to assess whether the model reproduces the local spillover patterns observed in the data. In these specifications, the estimated coefficients are statistically indistinguishable from 1 for both the formal-to-informal ratio and population. Second, I perform a stricter test by estimating the specification only on the subsample used in the empirical robustness exercises, comparing locations near Line B with locations near the feeder lines proposed in the 1985 Plan Maestro at buffer distances of 1000, 1500, 2000, and 2500 meters (see Table B19). This restriction reduces the sample and limits the comparison to locations used only in the empirical robustness exercises, which are not moments used in estimating the model parameters. Across these different samples, the estimated coefficients remain close to 1 and are statistically indistinguishable from 1, suggesting that the model reproduces the patterns observed in the data well. The results are also robust to the inclusion or exclusion of state fixed effects.

Overall, these findings indicate that the model's predictions align closely with the empirical responses in the data and capture well the change in informality associated with the infrastructure expansion.

6.4 Final Assessment:

The reduced-form results and back-of-the-envelope calculations show that Line B increased access to formal employment and reduced informality in more exposed areas, suggesting gains beyond the direct commuting-cost effect. However, those calculations suffer from the ‘missing intercept’ problem and do not incorporate general equilibrium adjustments or potential spillovers. As a result, they cannot quantify the aggregate reallocation required for a welfare decomposition or determine how much of the gains are explained by changes in allocative efficiency.

The quantitative model addresses these limitations by solving for wages, sectoral allocation, and spatial equilibrium, thereby delivering the aggregate change in formal employment implied by the commuting-cost shock. Structural counterfactuals, however, raise concerns about potential misspecification. Tables 6 and B19 validate the model by showing that the predicted changes in informality and population closely match those observed in the data, with coefficients statistically indistinguishable from one across different specifications and samples.

Taken together, the reduced-form evidence, the structural counterfactual, and the validation exercise provide a coherent assessment of the role of informality. The results indicate that distortions amplify the welfare gains from transit investment and that accounting for informality is an important mechanism when evaluating infrastructure improvements.

7 Conclusion

This paper examines the welfare implications of transit improvements in developing-country cities, highlighting a mechanism often overlooked in studies of infrastructure impacts: distortions that affect allocative efficiency. I show that commuting infrastructure can affect not only labor mobility and access to jobs but also the allocation of workers between the formal and informal sectors.

From an empirical perspective, I study the expansion of a transit system in Mexico City that connected peripheral, low-income areas to the urban core. The evidence shows that informality rates decrease by 8% in treated areas relative to the rest of the city, suggesting that workers reallocate toward higher-TFPR firms. This reallocation generates welfare gains beyond those predicted by standard models with fully efficient labor markets.

To quantify these effects, I develop a spatial equilibrium model with endogenous formal–informal sector choice, wedges, resource misallocation, and location decisions by firms and workers. The model includes a first-order welfare decomposition that isolates the role of allocative efficiency. The quantitative results indicate that improvements in allocative efficiency account for roughly 15%–20% of the total welfare gains and 30%-40% of the output gains from the transit investment.

These findings have relevant implications for policymakers. When evaluating transit projects, it is not enough to focus solely on transportation demand or time savings. In contexts where informality is widespread, transit expansions can enhance allocative efficiency by enabling informal workers to access formal employment opportunities. Thus, even in the absence of explicit redistributive objectives, connecting low-income and informal areas to high-efficiency zones can yield welfare benefits that would be missed under traditional cost-benefit analysis. This underscores the importance of incorporating allocative efficiency into infrastructure policy design, particularly in developing countries.

References

Adao, R., Costinot, A., and Donaldson, D. (2023). Putting quantitative models to the test: An application to trump's trade war. NBER Working Papers 31321, National Bureau of Economic Research, Inc.

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2015). The Economics of Density: Evidence From the Berlin Wall. *Econometrica*, 83:2127–2189.

Akbar, P. A., Couture, V., Duranton, G., and Storeygard, A. (2023). The Fast, the Slow, and the Congested: Urban Transportation in Rich and Poor Countries. NBER Working Papers 31642, National Bureau of Economic Research, Inc.

Alba-Vivar, F. (2024). Opportunity Bound: Transport and Access to College in a Megacity. Working paper, *Columbia University*.

Albouy, D. (2009). The unequal geographic burden of federal taxation. *Journal of Political Economy*, 117(4):635–667.

Allen, T. and Arkolakis, C. (2025). Quantitative regional economics. NBER Working Papers 33436, National Bureau of Economic Research, Inc.

Alvarez, F. and Lucas, R. J. (2007). General equilibrium analysis of the Eaton-Kortum model of international trade. *Journal of Monetary Economics*, 54(6):1726–1768.

Alvarez, J. A. and Ruane, C. (2024). Informality and aggregate productivity: The case of Mexico. *European Economic Review*, 167(C):None.

Andrews, I., Gentzkow, M., and Shapiro, J. M. (2017). Measuring the sensitivity of parameter estimates to estimation moments. *The Quarterly Journal of Economics*, 132(4):1553–1592.

Arkolakis, C., Costinot, A., Donaldson, D., and Rodríguez-Clare, A. (2019). The Elusive Pro-Competitive Effects of Trade. *Review of Economic Studies*, 86(1):46–80.

Asturias, J., García-Santana, M., and Ramos, R. (2019). Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India. *Journal of the European Economic Association*, 17(6):1881–1940.

Atkin, D. and Donaldson, D. (2021). The role of trade in economic development. Working Paper 29314, National Bureau of Economic Research.

Atkin, D. and Khandelwal, A. K. (2020). How distortions alter the impacts of international trade in developing countries. *Annual Review of Economics*, 12(1):213–238.

Balboni, C. (2019). In harm's way? infrastructure investments and the persistence of coastal cities. *Unpublished paper, MIT*.

Banerjee, A. V. and Duflo, E. (2005). Growth Theory through the Lens of Development Economics. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 1 of *Handbook of Economic Growth*, chapter 7, pages 473–552. Elsevier.

Baqaee, D. R. and Farhi, E. (2020). Productivity and Misallocation in General Equilibrium. *The Quarterly Journal of Economics*, 135(1):105–163.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics*, 122(2):775–805.

Bobba, M., Flabbi, L., and Levy, S. (2022). Labor market search, informality, and schooling investments. *International Economic Review*, 63(1):211–259.

Bobba, M., Flabbi, L., Levy, S., and Tejada, M. (2021). Labor market search, informality, and on-the-job human capital accumulation. *Journal of Econometrics*, 223(2):433–453.

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

Bosch, M. and Maloney, W. F. (2010). Comparative analysis of labor market dynamics using markov processes: An application to informality. *Labour Economics*, 17(4):621–631.

Busso, M., Fazio, M. V., and Algazi, S. L. (2012). (In)Formal and (Un)Productive: The Productivity Costs of Excessive Informality in Mexico. Research Department Publications 4789, Inter-American Development Bank, Research Department.

Caliendo, L. and Parro, F. (2020). The Quantitative Effects of Trade Policy on Industrial and Labor Location . Working paper, *Penn State University*.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45.

Costinot, A. and Donaldson, D. (2012). Ricardo's Theory of Comparative Advantage: Old Idea, New Evidence. *American Economic Review*, 102(3):453–458.

Dix Carneiro, R., Goldberg, P., Meguir, C., and Ulyssea, G. (2018). Trade and Informality in the Presence of Labor Market Frictions and Regulations. Working Paper, *Yale University*.

Donald, E., Fukui, M., and Miyauchi, Y. (2025). Unpacking aggregate welfare in a spatial economy. NBER Working Papers 34075, National Bureau of Economic Research, Inc.

Donaldson, D. and Hornbeck, R. (2016). Railroads and American Economic Growth: A "Market Access" Approach. *The Quarterly Journal of Economics*, 131(2):799–858.

Edmond, C., Midrigan, V., and Xu, D. Y. (2015). Competition, Markups, and the Gains from International Trade. *American Economic Review*, 105(10):3183–3221.

Faber, B. (2014). Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. *Review of Economic Studies*, 81(3):1046–1070.

Fajgelbaum, P. and Schaal, E. (2017). Optimal Transport Networks in Spatial Equilibrium. Working Paper, *UC Los Angeles*.

Fajgelbaum, P. D. and Gaubert, C. (2020). Optimal Spatial Policies, Geography, and Sorting*. *The Quarterly Journal of Economics*, 135(2):959–1036.

Fajgelbaum, P. D., Morales, E., Serrato, J. C. S., and Zidar, O. (2019). State Taxes and Spatial Misallocation. *Review of Economic Studies*, 86(1):333–376.

Franklin, S., Imbert, C., Abebe, G., and Mejia-Mantilla, C. (2024). Urban public works in spatial equilibrium: Experimental evidence from ethiopia. *American Economic Review*, 114(5):1382–1414.

Galle, S., Rodríguez-Clare, A., and Yi, M. (2023). Slicing the Pie: Quantifying the Aggregate and Distributional Effects of Trade. *The Review of Economic Studies*, 90(1):331–375.

Gonzalez-Navarro, M. and Turner, M. A. (2018). Subways and urban growth: Evidence from earth. *Journal of Urban Economics*, 108(C):85–106.

Heblich, S., Redding, S. J., and Sturm, D. M. (2018). The Making of the Modern Metropolis: Evidence from London. NBER Working Papers 25047, National Bureau of Economic Research, Inc.

Hornbeck, R. and Rotemberg, M. (2019). Railroads, Reallocation and the Rise of American Manufacturing. Working paper, *University of Chicago*.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The allocation of talent and u.s. economic growth. *Econometrica*, 87(5):1439–1474.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Hsieh, C.-T. and Moretti, E. (2019). Housing Constraints and Spatial Misallocation. *American Economic Journal: Macroeconomics*, 11(2):1–39.

Kanbur, R. (2009). Conceptualizing Informality: Regulation and Enforcement. Working Papers 48926, Cornell University, Department of Applied Economics and Management.

Khanna, G., Medina, C., Nyshadham, A., Ramos-Menchelli, D., Tamayo, J., and Tiew, A. (2022). Spatial Mobility, Economic Opportunity, and Crime. Working paper, *UC San Diego*.

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy*, 99(3):483–499.

Lagakos, D. and Waugh, M. E. (2013). Selection, agriculture, and cross-country productivity differences. *American Economic Review*, 103(2):948–80.

Levy, S. (2018). *Under-Rewarded Efforts: The Elusive Quest for Prosperity in Mexico*. Interamerican Development Bank.

McCaig, B. and Pavcnik, N. (2018). Export Markets and Labor Allocation in a Low-Income Country. *American Economic Review*, 108(7):1899–1941.

McMillan, M. S. and McCaig, B. (2019). Trade liberalization and labor market adjustment in botswana. Working Paper 26326, National Bureau of Economic Research.

Moreno-Monroy, A. I. and Posada, H. M. (2018). The effect of commuting costs and transport subsidies on informality rates. *Journal of Development Economics*, 130(C):99–112.

Ramírez, S. B., Rojo, M. H., and Gault, D. A. (2017). Decisiones e Implementación en la Construcción de las Primeras Once Líneas de la Red del Metro en La Ciudad de México Hacia la Desorganización del Metro (1967-2000). Working Paper, *Centro de Investigaci ón y Docencia Económicas CIDE.*

Redding, S. J. and Rossi-Hansberg, E. (2017). Quantitative Spatial Economics. *Annual Review of Economics*, 9(1):21–58.

Redding, S. J. and Turner, M. A. (2015). Transportation Costs and the Spatial Organization of Economic Activity. In Duranton, G., Henderson, J. V., and Strange, W. C., editors, *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics*, chapter 0, pages 1339–1398. Elsevier.

Restuccia, D. and Rogerson, R. (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Plants. *Review of Economic Dynamics*, 11(4):707–720.

Santamaría, M. (2020). The Gains from Reshaping Infrastructure: Evidence from the division of Germany. Working paper, *University of Warwick*.

Suárez, M., Murata, M., and Campos, J. D. (2016). Why do the poor travel less? Urban structure, commuting and economic informality in Mexico City. *Urban Studies*, 53(12):2548–2566.

Suárez-Serrato, J. C. and Zidar, O. (2016). Who benefits from state corporate tax cuts? a local labor markets approach with heterogeneous firms. *American Economic Review*, 106(9):2582–2624.

Święcki, T. (2017). Intersectoral distortions and the welfare gains from trade. *Journal of International Economics*, 104(C):138–156.

Tsivanidis, N. (2023). Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá's Transmilenio. Conditionally accepted at the *American Economic Review*, *UC Berkeley*.

Ulyssea, G. (2018). Firms, Informality, and Development: Theory and Evidence from Brazil. *American Economic Review*, 108(8):2015–47.

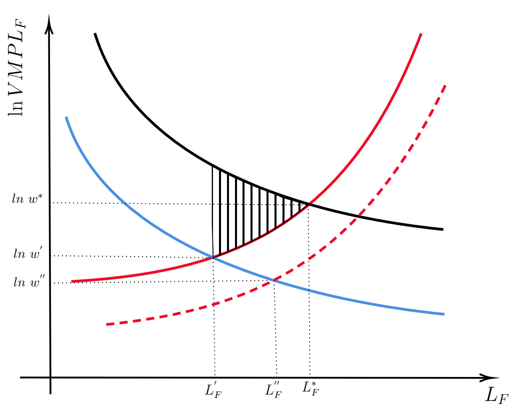
Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics*, 12(1):525–546.

Figures

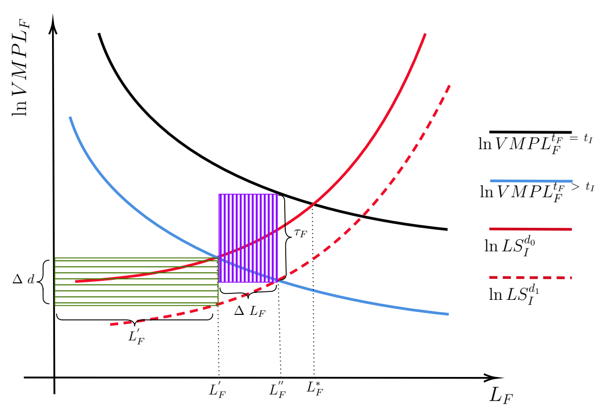
Figure 1: Transit System and Plan Maestro 1985

Notes: This figure shows the public transportation network in Mexico City. The black line highlights Line B, which is exploited in the main empirical specification. The green lines correspond to feeder lines proposed in the Government's 1985 *Plan Maestro*, while the blue lines represent the subway lines operating in 2000.

Figure 2: Effect of a labor wedge and a shock on allocative efficiency



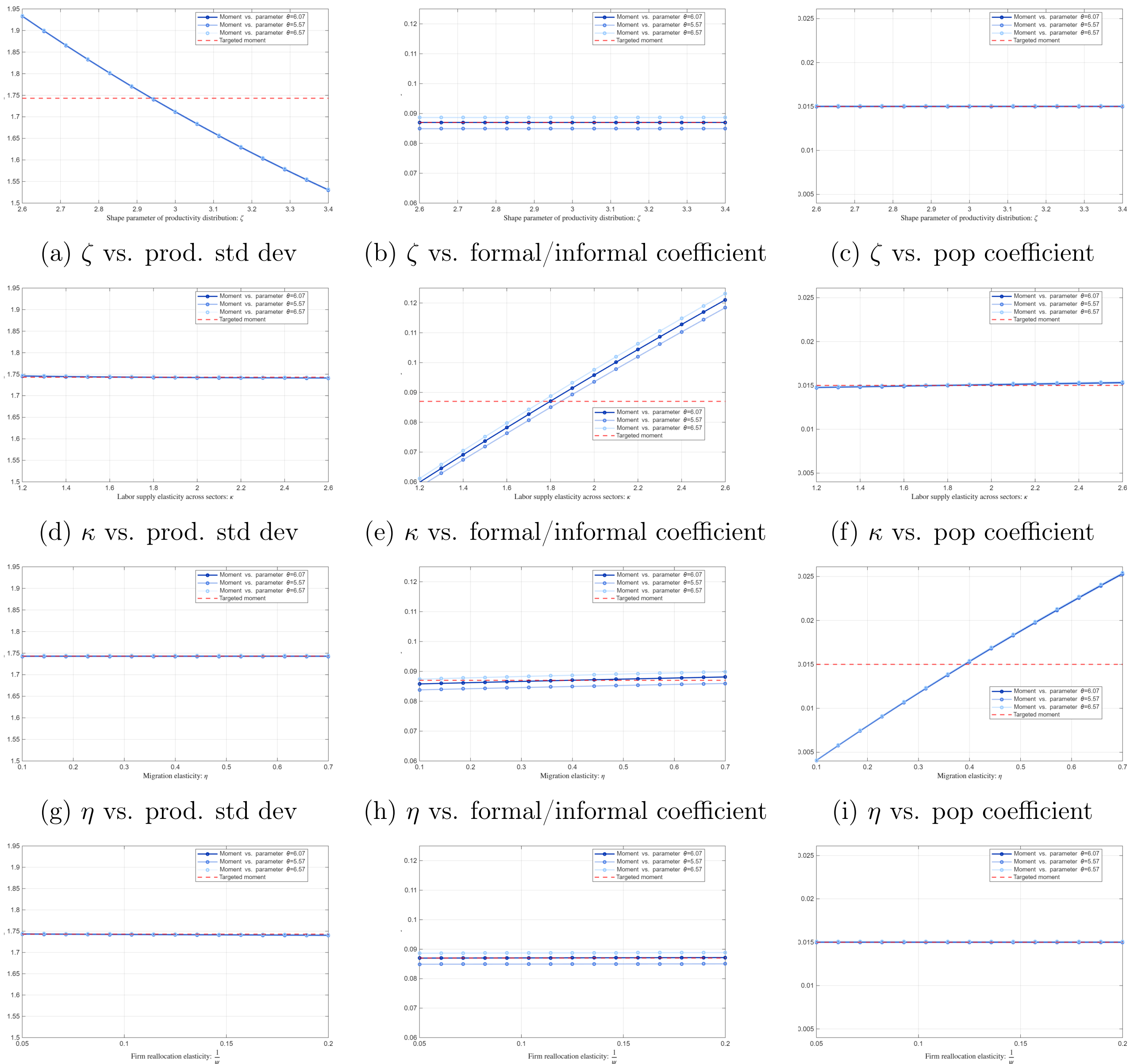
a) DWL due to the wedge



b) Reallocation effects of a shock on welfare

Notes: This figure plots the relationship between the VMPL for formal and informal firm types. Panel A illustrates the efficiency loss resulting from the labor wedge. With no distortions ($\tau = 0$), total welfare is maximized. The curve representing this case is the black line, while the market solution corresponds to the blue solid line. The area of the grey triangle approximates the DWL due to the differences in the wedge between formal and informal firms. Panel B displays the results of the first-order approximation in the event of an economic shock. In this case, if the shock reallocates workers to the formal firm, the purple area in panel B represents the change in welfare resulting from the reallocation effects. The height of the rectangle is the wedge, and the width corresponds to the change in formal labor.

Figure 3: Model parameters vs. moment conditions



(a) ζ vs. prod. std dev (b) ζ vs. formal/informal coefficient (c) ζ vs. pop coefficient

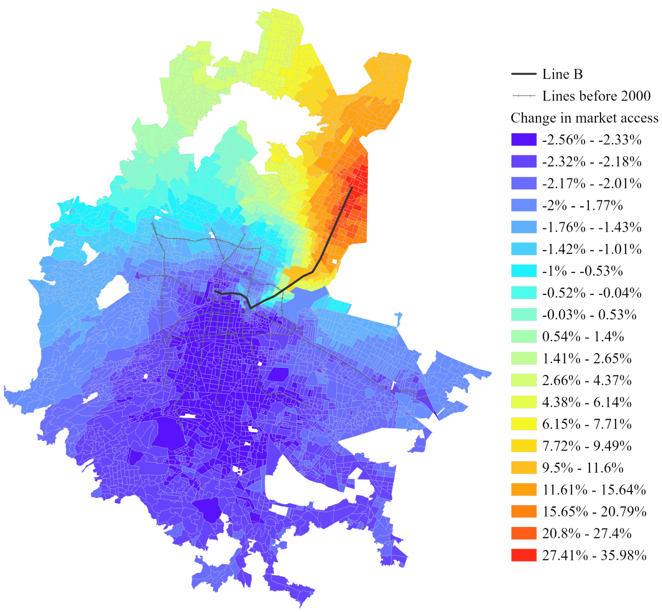
(d) κ vs. prod. std dev (e) κ vs. formal/informal coefficient (f) κ vs. pop coefficient

(g) η vs. prod. std dev (h) η vs. formal/informal coefficient (i) η vs. pop coefficient

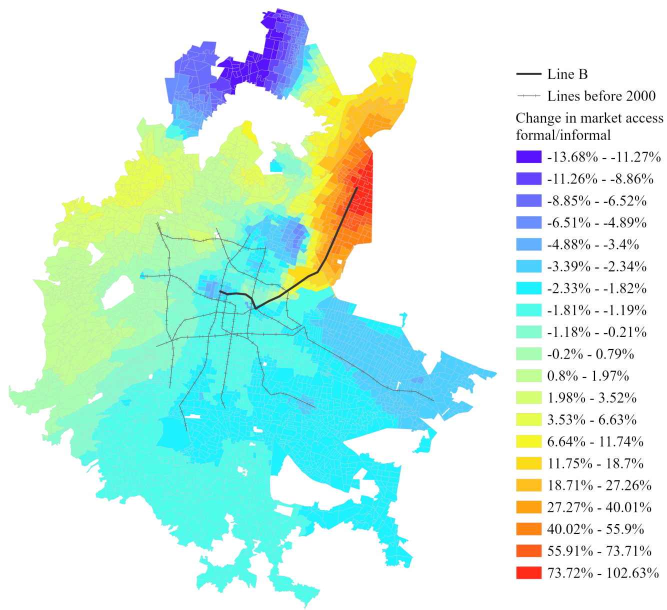
(j) $\frac{1}{\psi}$ vs. prod. std dev (k) $\frac{1}{\psi}$ vs. formal/informal coefficient (l) $\frac{1}{\psi}$ vs. pop coefficient

Notes: This figure illustrates how each model parameter maps into the moment conditions used in the GMM and indirect inference procedure. Each row corresponds to one parameter: the shape parameter of the productivity distribution (ζ), the labor supply elasticity across sectors (κ), the migration elasticity (η), and the firm reallocation elasticity ($\frac{1}{\psi}$). Each column represents a different empirical moment: (i) the standard deviation of log value added, (ii) the treatment effect on the log ratio of formal to informal workers, and (iii) the treatment effect on the number of individuals. The diagonal panels in the first three rows display the sensitivity of each moment to its targeted parameter: ζ governs dispersion in log value added, κ aligns with the treatment effect on the formal-informal employment ratio, and η determines population responses. The off-diagonal panels and the fourth row plot the cross-relationships between parameters and other targeted moments; as expected, these curves are almost flat, reflecting weaker sensitivity. Each figure also shows the sensitivity with respect to the commuting elasticity using three different values that correspond to the point estimate and 95% confidence interval.

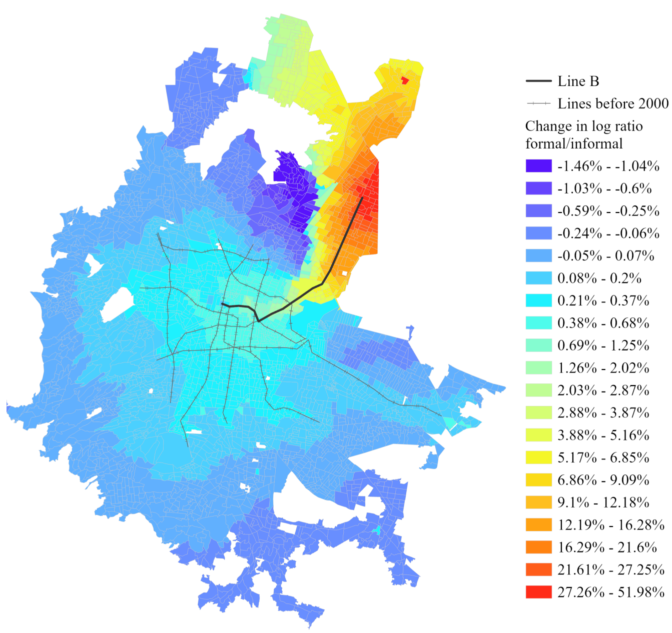
Figure 4: Spatial impact of Line B



a) Change in local CMA



b) Change in CMA across sectors



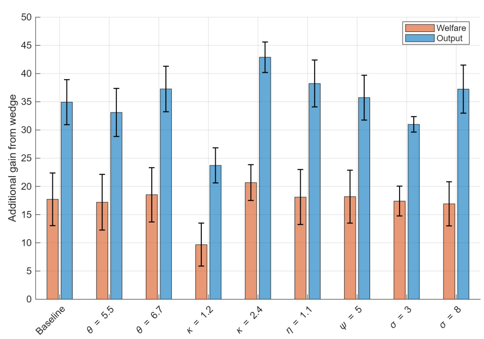
c) Δ in the log ratio formal/informal

Notes: This figure plots a heat map of Mexico City. Panel A plots the spatial distribution of the change in commuter market access (CMA) holding the number of workers and residents fixed. Panel B plots the spatial distribution of the change in CMA across sectors after the transit shock, holding the number of workers and residents fixed. Panel C plots the change in the log ratio of formal vs. informal workers from the model. I construct natural breaks across locations. Each color represents one of the natural breaks categories. Blue colors represent a very small change, while red color a very large change. From the figure, census tracts close to the new line got better access to formal employment relative to the informal sector. Thus, workers reallocate to the formal sector.

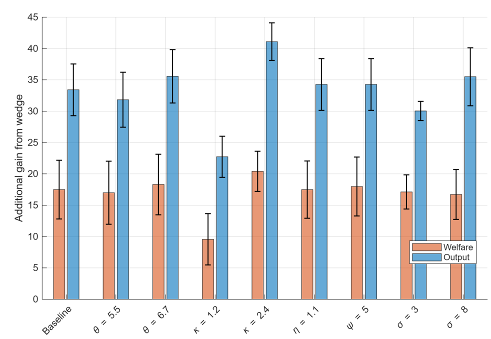
Figure 5: Additional gains: Reallocation from the informal to the formal sector

Notes: This figure plots the additional gains from the infrastructure that arise due to the reallocation of workers from the informal to the formal sector for 50 different samples using the bootstrap procedure. The orange bars report the mean and 95% CI for the additional welfare gains, and the blue bar for the additional output gains.

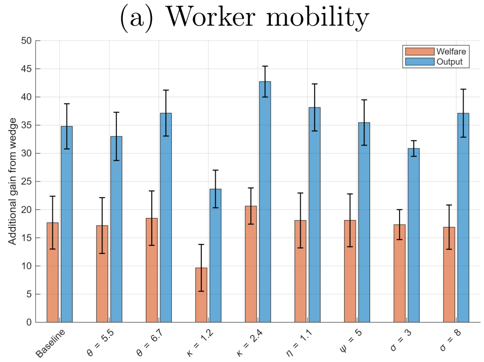
Figure 6: Robustness checks: Additional gains from the reallocation



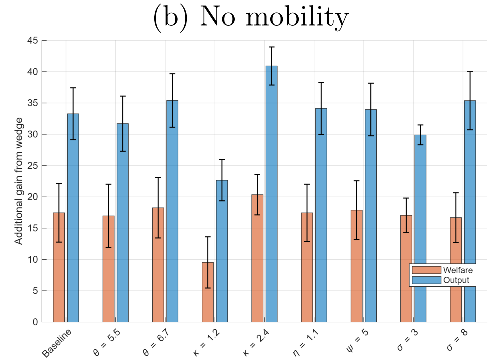
(a) Worker mobility



(b) No mobility



(c) Worker and firm mobility



(d) Firm mobility

Notes: This figure plots the additional gains from the infrastructure that arise due to the reallocation of workers from the informal to the formal sector for the different robustness-check scenarios. The orange bars report the mean and 95% CI for the additional welfare gains, and the blue bar for the additional output gains. The black lines represent the 95% confidence interval estimated using a bootstrap procedure with 50 repetitions for each mobility scenario. Panel A reports the results for the worker mobility scenario, panel B for the no-mobility scenario, panel C for the worker and firm mobility scenario, and panel D for the firm mobility scenario.

Tables

Table 1: Difference-in-Difference - Ratio of Formal to Informal Residents

	(1)	(2)	(3)	(4)	(5)	(6)
Outcomes:	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$
Panel A: Continuous treatment measure-Pool of residents						
-ln dist	0.058***	0.062***	0.062***	0.064***	0.030***	0.034***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)
Obs	3156	3156	3156	3156	3156	3156
R2	0.267	0.279	0.215	0.223	0.223	0.232
Panel B: Treatment dummy variable-Pool of residents						
T_i	0.070***	0.078***	0.083***	0.087***	0.061***	0.068***
	(0.019)	(0.019)	(0.018)	(0.018)	(0.020)	(0.021)
Obs	3156	3156	3156	3156	3156	3156
R2	0.258	0.270	0.205	0.212	0.223	0.231
Panel C: Continuous treatment measure-Low-skilled workers						
-ln dist	0.061***	0.064***	0.064***	0.065***	0.032***	0.036***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.011)
Obs	3156	3156	3156	3156	3156	3156
R2	0.242	0.248	0.175	0.179	0.191	0.197
Panel D: Treatment dummy variable-Low-skilled workers						
T_i	0.080***	0.086***	0.087***	0.090***	0.064***	0.070***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.021)	(0.022)
Obs	3156	3156	3156	3156	3156	3156
R2	0.233	0.239	0.165	0.169	0.190	0.196
Panel E: Continuous treatment measure-Outskirts areas						
-ln dist	0.090***	0.095***	0.092***	0.094***	0.044***	0.048***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Obs	2128	2128	2128	2128	2128	2128
R2	0.289	0.306	0.209	0.220	0.209	0.222
Panel F: Treatment dummy variable-Outskirts areas						
T_i	0.125***	0.140***	0.138***	0.145***	0.087***	0.096***
	(0.024)	(0.025)	(0.024)	(0.025)	(0.027)	(0.027)
Obs	2128	2128	2128	2128	2128	2128
R2	0.272	0.288	0.190	0.202	0.207	0.220
Pop Controls	X	X	X	X	X	X
Dist Controls	X	X	X	X	X	X
Prod Controls		X		X		X
State FE			X	X		
Zone FE					X	X

Notes: This table reports the results of a regression relating changes in the ratio of formal to informal residents in each location to the opening of Line B of the subway. Panel A reports the results for the continuous treatment measures and the pool of residents, panel B for the treatment dummy variables and the pool of residents, panel C for the continuous treatment measure and low-skilled workers, panel D for the treatment dummy variables and low-skilled workers, panel E for the continuous treatment measure on the locations that are not in the CBD, and panel F for the treatment dummy variable on the locations that are not in the CBD. In the first two columns, I do not include geographic fixed effects; in columns 3 and 4, I include state fixed effects, and in columns 5 and 6, zone fixed effects. The regressions are weighted using the initial population shares. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses with a bandwidth of 500 meters (Conley, 1999). $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table 2: Gravity Equations-Commuting Elasticities

	(1) Overall	(2) Overall	(3) Formal	(4) Informal
Outcome:	$\ln \lambda_{ni}$	$\ln \lambda_{ni}$	$\ln \lambda_{niF}$	$\ln \lambda_{niI}$
$\ln \tilde{d}_{ni}$	-6.040***	-6.094***	-5.074***	-7.377***
	(0.292)	(0.285)	(0.249)	(0.377)
Observations	1152	1152	576	576
R-squared	0.341	0.344	0.289	0.408
Origin FE	X		X	X
Destination FE	X		X	X
Origin-sector FE		X		
Destination-sector FE		X		

Notes: This table reports the results of a gravity equation relating commuting flows at the municipality level with the inclusive value of travel times. I estimate this regression via PPML to include the zeros. The first column presents the results for the overall elasticity, the second column for the overall elasticity with sector fixed effects, the third column for the formal sector, and the fourth column for the informal sector. Standard errors are clustered at the municipality of origin level and reported in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table 3: Model Parameters and Identification Strategy

	Symbol	**Value**	**Source / Identification**
			Panel A: Parameters calibrated from the cross-section or other sources
Expenditure share on housing	α	0.822	ENOE: Housing expenditure share
Labor share	β	0.855	Economic Census: Labor share in value added
Residential floor space	$\tilde{H}$	–	GHSL 2000: Residential built-up area
Commercial floor space	$\tilde{Z}$	–	GHSL 2000: Commercial built-up area
Fixed cost for formal firms	FC_F	–	Calibrated to match share of formal firms (83%)
Firm reallocation elasticity	ψ	13.5	Firm location response to profits
Elasticity of substitution across varieties	σ	6.0	Standard values from trade literature
Elasticity of substitution across modes	δ	0.0096	Nested logit estimation
Commuting elasticity	θ	≈ 6	Gravity equation: commuting flows
			Panel B: Parameters identified via GMM and indirect inference
Shape parameter-Pareto distribution	ζ	2.939	GMM: dispersion (sd) of log value added across firms
Migration elasticity	η	0.391	Indirect inference: migration response reduced-form
Sectoral labor supply elasticity	κ	1.798	Indirect inference: sectoral response reduced-form
			Panel C: Parameters recovered via model inversion
Productivity shifters	A_{is}	–	Model inversion: labor demand by location and sector
Sectoral amenity shifters	B_{ns}	–	Model inversion: sectoral labor supply by location
Residential amenity shifters	B_n	–	Model inversion: residential population distribution
Firm amenity shifters	Φ_{is}	–	Model inversion: firm location distribution

Notes: This table reports the parameters of the model and their identification strategy. Panel A presents parameters calibrated using cross-sectional moments, including expenditure shares, factor shares, and aggregate stocks. Panel B reports structural elasticities identified through GMM and indirect inference, matching reduced-form responses of employment and firm outcomes to the model. Panel C lists location- and sector-specific shifters recovered by model inversion, ensuring that the model exactly matches observed labor demand, labor supply, and firm location decisions in the baseline equilibrium.

Table 4: Estimation of the Main Model Parameters

Parameter	Point est.	s.d.	95% CI
η (migration elasticity)	0.391	0.040	[0.313, 0.469]
κ (labor supply elasticity)	1.798	0.209	[1.388, 2.207]
ζ (Pareto shape parameter)	2.939	0.004	[2.930, 2.947]

Notes: This table reports GMM estimation results. The first column reports the point estimates from the baseline sample. The second column reports standard deviations obtained from a bootstrap procedure with 50 replications, and the third column reports 95% confidence intervals based on a normal distribution. The migration elasticity, η, and the labor supply elasticity across sectors, κ, are estimated using an indirect inference approach that matches the treatment dummy coefficients. The shape parameter of the Pareto distribution is estimated by matching the standard deviation of the log value added.

Table 5: Effects on Welfare: Counterfactual Results $\hat{X} = X'/X$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome measure:		1. Welfare				2. Output		
	Mean	sd	p05	p95	Mean	sd	p05	p95
			Panel A: Worker mobility (Baseline)					
% Δ at baseline	0.635	0.072	0.494	0.777	0.630	0.065	0.504	0.757
% Δ-wedge	0.749	0.086	0.582	0.917	0.850	0.092	0.669	1.031
			Panel B: No worker and no firm mobility					
% Δ at baseline	0.636	0.072	0.494	0.778	0.649	0.067	0.518	0.779
% Δ-wedge	0.749	0.085	0.581	0.916	0.866	0.093	0.684	1.048
			Panel C: Worker and firm mobility					
% Δ at baseline	0.637	0.073	0.494	0.779	0.632	0.065	0.505	0.760
% Δ-wedge	0.750	0.086	0.582	0.919	0.852	0.093	0.670	1.033
			Panel D: Firm mobility					
% Δ at baseline	0.637	0.073	0.495	0.780	0.651	0.067	0.519	0.782
% Δ-wedge	0.750	0.086	0.582	0.918	0.868	0.094	0.684	1.051

Notes: This table reports the results of the main counterfactuals after running a bootstrap with 50 replications. The first four columns report results for workers' welfare, and the last four columns report the effect on total output. For each measure, the table reports the mean, the standard deviation, and the 5th and 95th percentiles of the bootstrap distribution.

Table 6: Data vs. Model Predictions

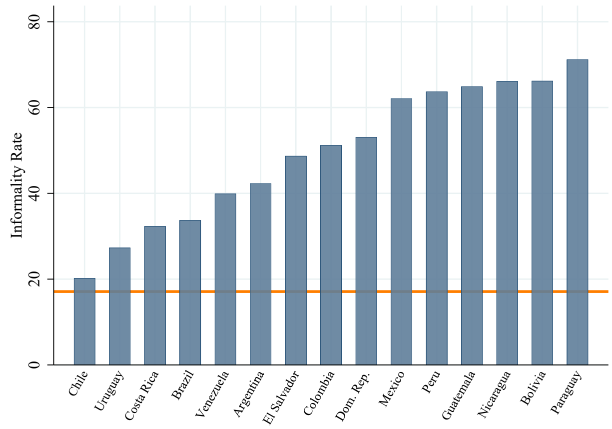
	(1)	(2)	(3)	(4)
	Worker mobility	Worker mobility	Worker mobility	Worker mobility
	Panel A: log formal/informal data vs. log formal/informal model			
Outcome:	$\Delta \ln L_{iF}/L_{iI}$	$\Delta \ln L_{iF}/L_{iI}$	$\Delta \ln L_{iF}/L_{iI}$	$\Delta \ln L_{iF}/L_{iI}$
$\Delta \ln L_F/L_I$ model	0.942***	1.014***	0.971***	1.010***
	(0.145)	(0.147)	(0.143)	(0.146)
p-value: $\beta = 1$	0.689	0.922	0.840	0.948
Obs	3156	3156	3156	3156
R2	0.266	0.280	0.289	0.296
	Panel B: log pop data vs. log pop model			
Outcome:	$\Delta \ln L_i$	$\Delta \ln L_i$	$\Delta \ln L_i$	$\Delta \ln L_i$
$\Delta \ln pop$ model	0.736*	0.918**	0.679	0.930**
	(0.432)	(0.423)	(0.438)	(0.428)
p-value: $\beta = 1$	0.542	0.846	0.463	0.869
Obs	3156	3156	3156	3156
R2	0.295	0.312	0.303	0.320
	Panel C: log formal/informal data vs. log formal/informal model (spillover)			
Outcome:	$\Delta \ln L_{iF}/L_{iI}$	$\Delta \ln L_{iF}/L_{iI}$	$\Delta \ln L_{iF}/L_{iI}$	$\Delta \ln L_{iF}/L_{iI}$
$\Delta \ln L_F/L_I$ model	0.919***	1.014***	0.950***	1.006***
	(0.153)	(0.157)	(0.151)	(0.156)
p-value: $\beta = 1$	0.595	0.928	0.741	0.967
Obs	2829	2829	2829	2829
R2	0.270	0.286	0.296	0.304
	Panel D: log pop data vs. log pop model (spillover)			
Outcome:	$\Delta \ln L_i$	$\Delta \ln L_i$	$\Delta \ln L_i$	$\Delta \ln L_i$
$\Delta \ln pop$ model	0.700	0.831*	0.639	0.850*
	(0.461)	(0.450)	(0.469)	(0.458)
p-value: $\beta = 1$	0.515	0.708	0.441	0.743
Obs	2829	2829	2829	2829
R2	0.304	0.324	0.313	0.333
State FE			X	X
Pop controls	X	X	X	X
Dist controls	X	X	X	X
Prod controls		X		X
Instrument	Δ CMA	Δ CMA	Δ CMA	Δ CMA

Notes: This table reports the results of the main model predictions for the case of worker mobility. Panel A reports the results of a regression relating the ratio between formal and informal observed in the data vs. the predictions implied by the model. Panel B reports the results of a regression relating the change in population from the data to the one from the model. Panels C and D report the same regressions excluding locations between 2.5 and 5.0 km of the new transit line. The predictions of the model are instrumented with the commuter market access measures of Line B. The regressions are weighted using the initial population shares. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses with a bandwidth of 500 meters (Conley, 1999). $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

For Online Publication: Appendix
Spatial Misallocation, Informality, and Transit Improvements

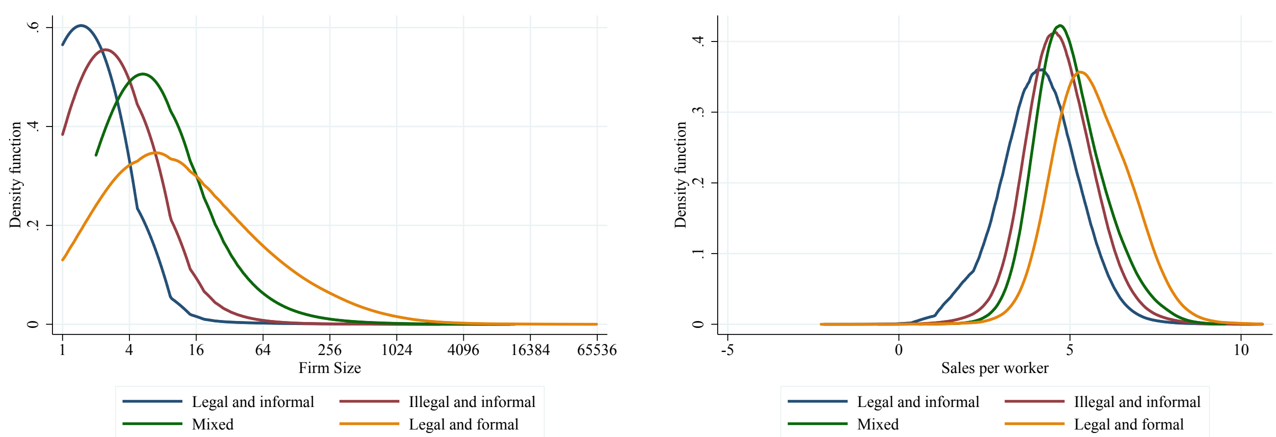
A Additional Figures

Figure A1: Informality Rates-Latin America and the Caribbean



Notes: This figure plots informality rates across countries from Latin America and the Caribbean. The data source is the online appendix from Ulyssea (2018) that uses data from *SEDLAC*, an initiative from the World Bank and Universidad Nacional de la Plata. Informal workers are defined as those without social security. The orange line represents the average informality rate of countries from the OECD. The figure shows that informality rates in LAC are very high, and even within the region, Mexico is one of the countries with the highest informality rates.

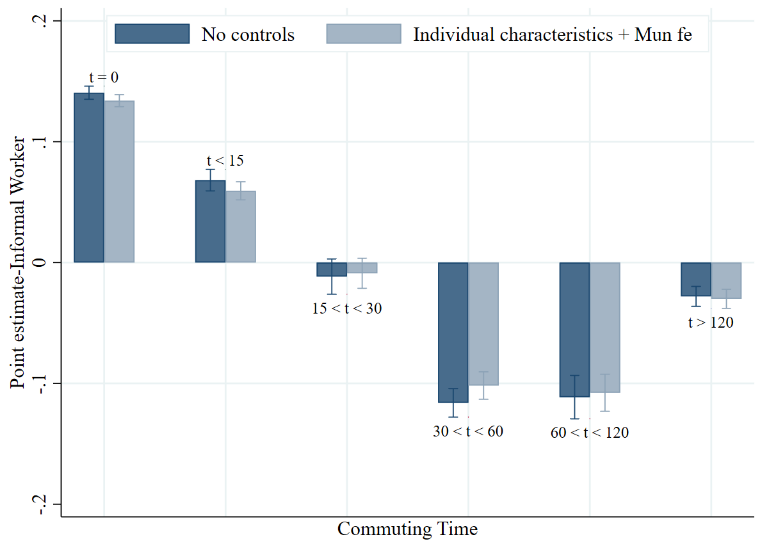
Figure A2: Firm size and Productivity Distribution-Economic Census 1999



(a) Firm size (b) Productivity

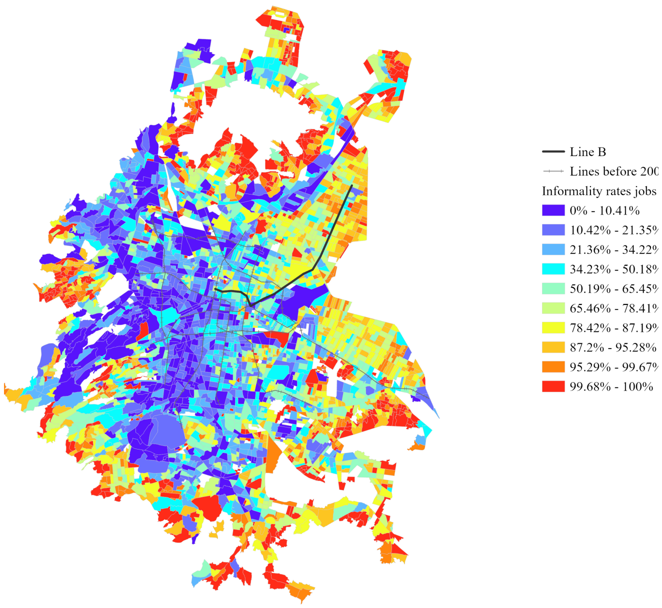
Notes: This figure plots the firm size and productivity distribution for the four different categories of firms: 1) Legal and informal 2) Illegal and informal, 3) Mixed, and 4) Legal and formal. I use the 2004 economic census. Panel (a) plots the firm size distribution and panel (b) the productivity distribution. Firm size is measured as the number of workers, and productivity as the logarithm of sales per worker.

Figure A3: Commuting Time- Informal vs. Formal

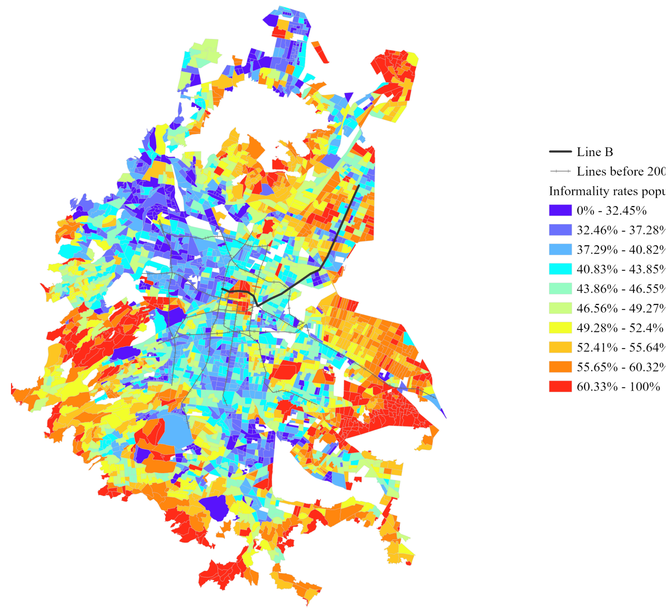


Notes: This figure plots the point estimate and 95% confidence interval of a regression that relates the probability of commuting within some window of time with an informal dummy variable. The first bar reports the results for the category of non-commuting, the second bar if the worker spends on average between 1 to 15 minutes, the third bar between 16 to 30 minutes, the fourth bar between 30 to 60 minutes, the fifth bar between 60 to 120 minutes, and the sixth bar more than 120 minutes. The dark-blue bar does not include controls, while the light-blue bar includes individual controls and municipality fixed effects. Standard errors are computed with clusters at the municipality level.

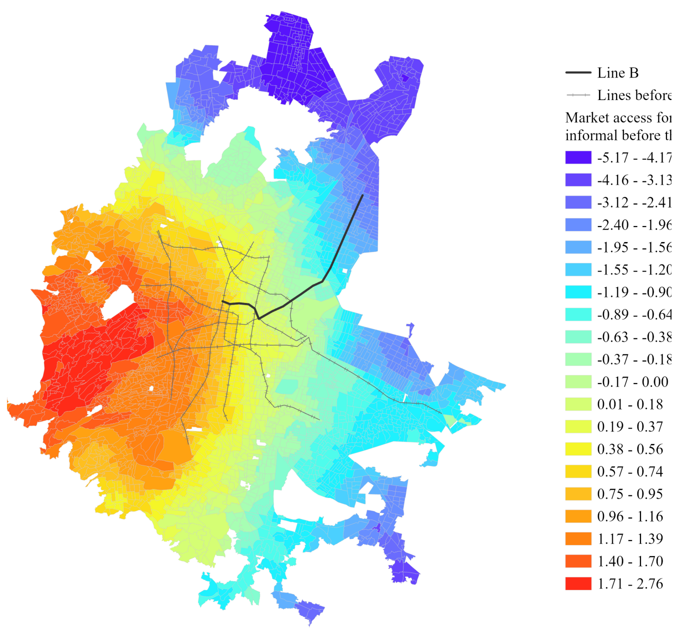
Figure A4: Spatial distribution of informality



(a) Informal workers



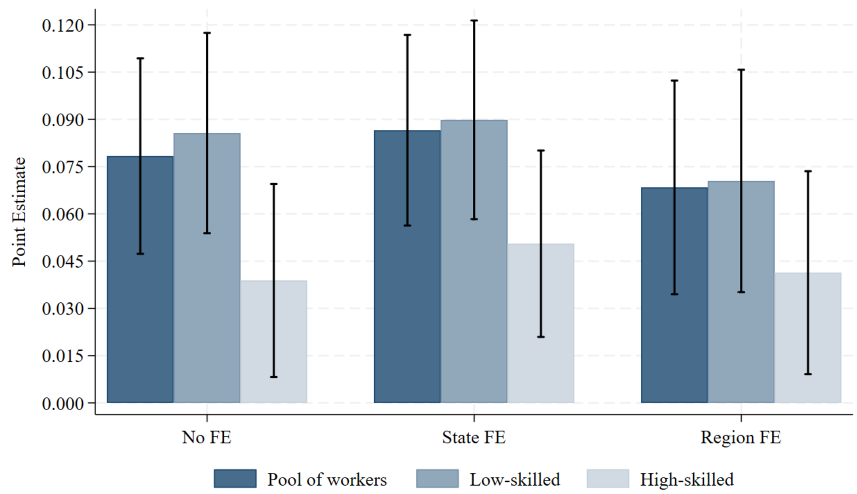
(b) Informal Residents



c) Relative market access between formal/informal employment before the shock

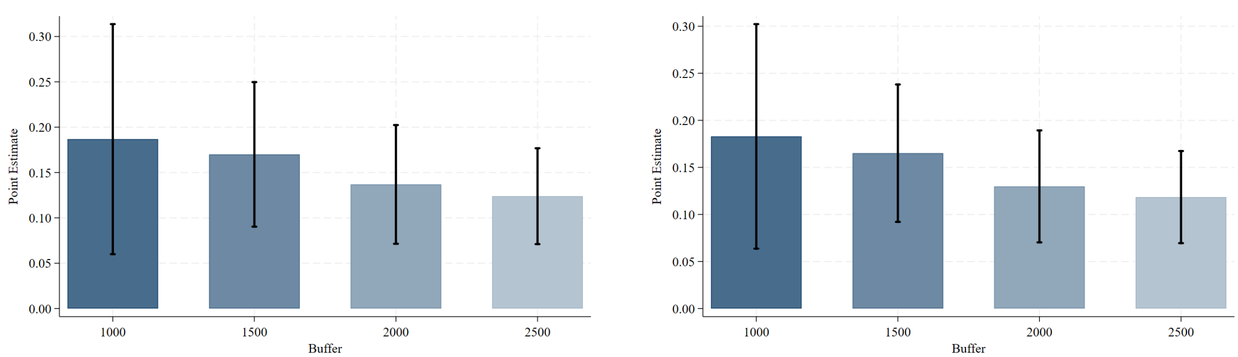
Notes: This figure plots a map of Mexico City with the spatial distribution of informality rates. Panel A plots a heat map of workers' informality rates by deciles in 1999. Panel B plots a heat map of residents' informality rates by deciles in 2000 and panel C plots a heat map of the relative commuter market access between formal and informal employment before the shock. The main takeaway of this map is that in the middle-west and center of the city informality rates are lower than on the boundaries and east of Mexico City. As a result, informal workers that live in the outskirts have poor access to most of the formal employment, which is usually located in the center of the city.

Figure A5: Difference in Difference Results-Ratio Formal to Informal



Notes: This figure depicts the point estimate and 95% confidence interval of a regression that relates the change over time in the log of the ratio between formal and informal residents with the transit shock. The treatment variable takes a value of 1 for census tracts with a centroid within a 30-minute walking range. The first three bars show the results of a regression including distance and population controls with no fixed, the second three bars state fixed effects, and the third three bars report the results with zone fixed effects. The regressions are weighted using the initial population shares. Standard errors are clustered at the census tract level. The dark-blue bar reports the results for the pool of workers, the middle-blue bar for low-skilled workers and the light-blue bar for high skilled workers. Line B increased the ratio of formal to informal residents on approximately 8% when I compare treated areas vs. the rest of Mexico City.

Figure A6: Robustness Checks-Ratio Formal to Informal

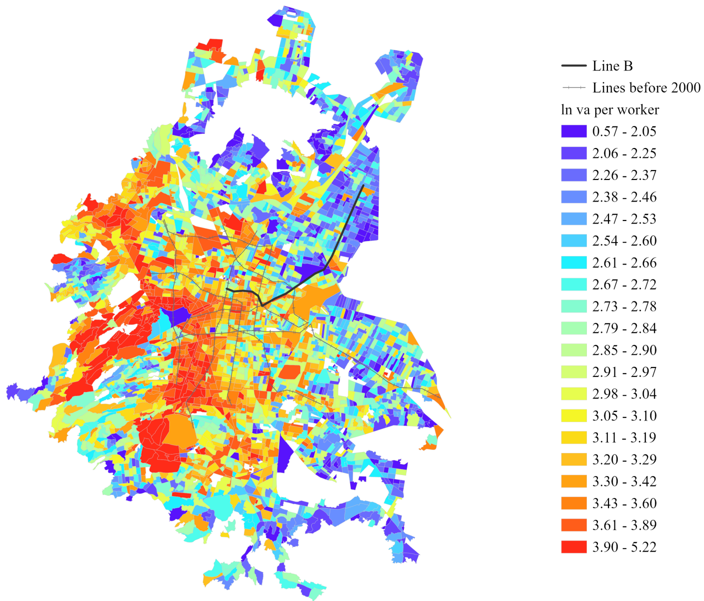


(a) No state fixed-effects

(b) with state fixed-effects

Notes: This figure depicts the point estimate and 95% confidence interval of a regression that relates the change over time in the log of the ratio between formal and informal residents with the transit shock. The treatment variable takes the value 1 for census tracts with a centroid within a buffer zone of the new subway line (Line B). The control group comprises locations within the buffer zone of the feeder lines in the Plan Maestro in 1985. The Government planned to build these lines in the 1980s, but they were not constructed during my period of analysis. I exclude Line F since it is very close to Line B. I use different buffer zones: 1000, 1500, 2000, and 2500 meters. Panel A is the specification without fixed effects, and Panel B plots the point estimates with state fixed effects. The regressions are weighted using the initial population shares. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors with a bandwidth of 500 meters are estimated to construct the 95% confidence interval (Conley, 1999).

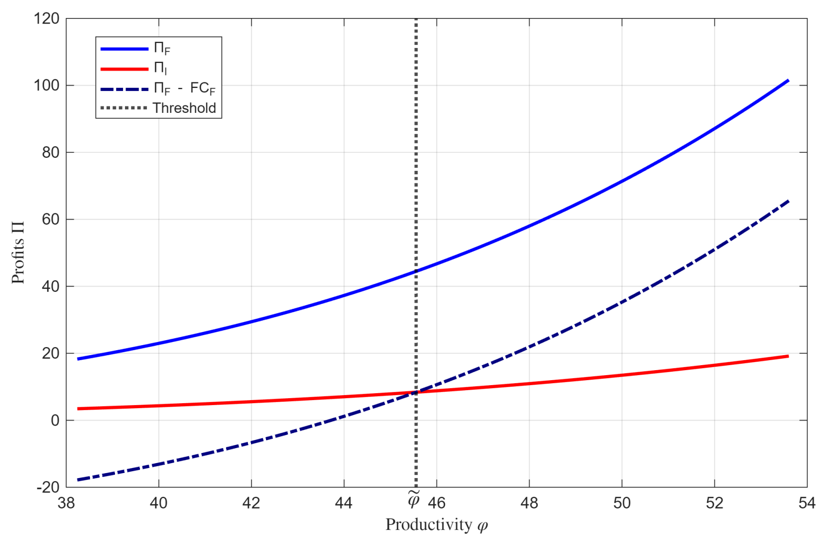
Figure A7: Spatial Distribution of Productivity



(a) Productivity - log value added per worker

Notes: This figure plots a map of Mexico City with the spatial distribution of productivity measured as value added per worker. I construct ventiles across locations after aggregating value-added measures and the total number of workers. Each color represents one of the quantile categories. Census tracts in central areas have higher productivity measures.

Figure A8: Profits in the formal and informal sectors



Notes: This figure illustrates the relationship between firm productivity φ and profits in the formal and informal sectors implied by the model. The blue solid line shows gross profits in the formal sector, while the red solid line shows gross profits in the informal sector. The dark blue dotted line represents net formal-sector profits after subtracting the fixed cost to operate in the formal sector. The marginal firm, $\tilde{\varphi}$, is indifferent between operating formally and informally and is defined by the intersection of informal profits and net formal profits.

B Additional Tables

Table B1: Characterization of Informality at the Individual level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Low-skilled		Medium-skilled		High-skilled		Fixed salary	
Informal	0.175***	0.171***	-0.066***	-0.066***	-0.084***	-0.081***	-0.526***	-0.531***
	(0.009)	(0.007)	(0.004)	(0.003)	(0.006)	(0.005)	(0.006)	(0.007)
Obs	751160	751160	751160	751160	769625	769625	722274	722274
R2	0.036	0.076	0.011	0.021	0.010	0.044	0.333	0.339
	Log income		Hours of work <35		35<Hours of work<48		Hours of work>48	
Informal	-0.432***	-0.391***	0.130***	0.129***	-0.238***	-0.227***	0.108***	0.098***
	(0.021)	(0.024)	(0.005)	(0.005)	(0.008)	(0.007)	(0.008)	(0.007)
Obs	710825	710825	686645	686645	686645	686645	686645	686645
R2	0.083	0.143	0.033	0.043	0.060	0.085	0.017	0.041
Quarter FE	X	X	X	X	X	X	X	X
Municipality FE		X		X		X		X

Notes: This table reports the results of a regression correlating different individual characteristics with a dummy variable that takes the value of 1 if the worker is informal. I use household survey data for the entire country from 2000, which is the year before the opening of Line B. All columns include quarter fixed-effects and the even columns also include municipality fixed-effects. Standard errors are clustered at the municipality level and reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B2: Characterization of informality at the census-tract level

	(1)	(2)	(3)	(4)	(5)	(6)
	High-skilled share		Labor force participation share		log avg income	
Average informality rate	-0.822***	-0.786***	-0.112***	-0.097***	-0.182***	-0.069***
	(0.035)	(0.033)	(0.011)	(0.010)	(0.036)	(0.017)
Obs	3491	3491	3487	3487	3491	3491
R2	0.331	0.331	0.078	0.069	0.020	0.014
	log avg hours worked		Share with electricity		Share with water	
Average informality rate	0.264***	0.248***	0.015***	0.014***	0.227***	0.230***
	(0.019)	(0.018)	(0.002)	(0.002)	(0.024)	(0.024)
Obs	3491	3491	3487	3487	3487	3487
R2	0.292	0.298	0.026	0.024	0.064	0.065
State FE		X		X		X

Notes: This table reports the results of a regression correlating different census-tract characteristics with the informality share at the census-tract level. I use the data at the census-tract level from the 2000 population census. The even columns include state fixed-effects. Conley standard errors with a bandwidth of 500 meters are reported in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table B3: Institutional factors and wedges between formal and informal firms

Category	**Policy/Institution**	**Observations / Impact**
Contributory Social Insurance	Benefits paid by firms and workers proportional to wages Low-quality services by IMSS and Infonavit Imperfect enforcement by IMSS and Infonavit	Applies only to salaried workers Implicit tax of ∼12% on salaried contracts despite 0.5% of GDP in subsidies Discriminates against medium and large firms with salaried workers
Regulations on Dismissal and Reinstatement	Contingent firing costs and legal risks from "unjustified dismissals" Imperfect adjudication by JCAs	Applies only to salaried workers Implicit tax beyond 12%; costly adjustment to negative shocks; high legal fees and delays
Non-Contributory Social Insurance	Free benefits to workers regardless of earnings Fully funded by the government (∼1.7% of GDP)	Applies to non-salaried and de facto to illegally hired salaried workers Implicit subsidy of ∼16% of earnings to non-salaried and illegal salaried contracts
Labor Taxation	State payroll taxes (2–3% of wages) Federal employment subsidy Federal income taxes Tax enforcement by SAT	Collect 0.39% of GDP; applies only to salaried workers 0.24% of GDP; applies to workers earning up to 3 minimum wages Withheld for salaried workers; self-reported for non-salaried Salaried workers taxed at 2.5% of GDP vs. 0.1% for non-salaried (implicit subsidy of 0.4% of GDP)
Firm Competition	93% of firms qualify for Repeco regime Imperfect enforcement of Repeco by states Weak enforcement of general regime by SAT	Covers 52% of labor and 25% of capital; favors small, low-productivity firms Implicit subsidy to small firms (0.5% of GDP) Marginal tax rates increase with firm size
Consumption Taxation	42% of VAT base in special regimes VAT exemptions on 25% of GDP	Discriminates across sectors Leads to 1.5% of GDP in foregone revenue; fosters informal supply chains
Enforcement of Contracts	Contract enforceability varies by state Weak property rights and costly collateral seizure	Imperfect enforcement reduces firm size and client base Lowers value of assets as collateral; reduces credit access
Insufficient Competition	Market concentration in banking and low-exposure sectors	Lowest firm credit among OECD; most firms lack commercial credit; lending favors large firms

Notes: This table reports the different sources that generate differences in wedges between formal and informal firms based on Levy (2018).

Table B4: Illustrative accounting of taxes and benefits across worker types

Panel A: Labor contract distortions

Worker type	Cost to firm	Income / benefits to worker	Parameter in the model
All informal workers	w_I	w_I + non-contributory benefits	baseline 0
All formal workers	$w_F(1+\tau_L)$	w_F + contributory benefits	labor taxes $\tau_L = 0.504$
Marginal informal workers	w_I	w_I + non-contributory benefits	baseline 0 (informal sector)
Marginal formal workers	$w_F(1+\tau_L)$	w_F + contributory benefits	switching margin (informal $\rightarrow$ formal) - labor taxes $\tau_L = 0.504$

Panel B: Output and value-added distortions

Firm type	Firm revenue	Tax component	Parameter in the model
Informal firms	$p_I q_I$	Limited VAT compliance, simplified regimes	baseline 0
Formal firms	$p_F q_F(1-\tau_Y)$	VAT, IEPS, corporate taxation	$\tau_Y = 0.217$

Notes: This table reports the taxes and benefits of formal and informal workers. Panel A summarizes labor contract distortions across worker types following the logic of Levy (2018). The "marginal" workers correspond to the subset of workers whose sectoral choice is affected by the reduction in commuting costs generated by the subway expansion. They are not a different institutional category and therefore face the same tax schedule as other formal workers and informal workers, respectively. In the institutional system, formal workers receive contributory benefits, while informal workers receive non-contributory benefits. In the model, these transfers are represented through the proportional rebate of fiscal revenues to households. The effective labor tax differential is $\tau_L = 0.504$. Panel B summarizes output and value-added tax distortions across firms. The effective output tax differential is $\tau_Y = 0.217$. In the quantitative model, the labor wedge relevant for the welfare decomposition is $t_L = (1+\tau_L)/(1-\tau_Y) - 1 = 0.9208$.

Table B5: Informality and Commuting Patterns

	(1)	(2)	(3)	(4)	(5)
	Panel A: Probability of working in the same municipality of residence				
Outcome:	Workplace municipality	Workplace municipality	Workplace municipality	Workplace municipality	Workplace municipality
Informal	-0.265***	-0.231***	-0.231***	-0.132***	-0.079***
	(0.009)	(0.008)	(0.008)	(0.006)	(0.008)
Observations	577,041	577,039	577,039	517,354	516,931
R-squared	0.069	0.098	0.123	0.215	0.465
	Panel B: Probability of working in the CBD of Mexico City				
Outcome:	Workplace-CBD	Workplace-CBD	Workplace-CBD	Workplace-CBD	Workplace-CBD
Informal	-0.086***	-0.056**	-0.059***	-0.037***	-
	(0.026)	(0.024)	(0.018)	(0.011)	-
Observations	577,041	577,039	577,039	517,354	-
R-squared	0.007	0.042	0.468	0.444	-
Individual Characteristics		X	X	X	X
Origin FE			X	X	X
Transportation Mode FE				X	X
Destination FE					X

Notes: This table reports the results of a linear probability model relating the probability of working in the same municipality as the one in which the worker resides, and the probability of working in the CBD with a dummy variable that takes the value of 1 if the worker is informal. Panel A reports the results for working in the same municipality, and panel B whether the individual works in the CBD. Standard errors are clustered at the residence municipality level and reported in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table B6: Descriptive Statistics 1999 and 2000

Panel A: Main Variables

Variable	Mean	Median	Sd	Min	Max
Share informal residents (Population census)	46.58	46.52	10.46	1.09	91.98
Share informal high-skilled residents (Population census)	35.59	35.07	7.40	1.43	76.99
Share informal low-skilled residents (Population census)	50.23	50.56	10.37	1.07	93.01
Share informal workers (Economic census)	59.97	66.31	33.08	0.00	100.00
Share informal firms (Economic census)	84.27	91.43	17.85	0.00	100.00

Panel B: Treatment Variables

Variable	Mean	Median	Sd	Min	Max
Walking Distance to new stations (minutes)	138.53	135.20	82.64	1.21	410.49
Log distance to new stations	4.65	4.91	0.88	0.19	6.02
Dummy variable (minutes < 30)	0.11	0.00	0.31	0.00	1.00
Observations	3156				

Notes: This table reports summary statistics of the main variables. Panel A presents the statistics for the main variables: workers' informality rates from the Economic Census in 1999 and residents' informality rates from the Population Census in 2000. Panel B reports the treatment variables, including the network walking distance, log distance, and a dummy variable for whether the centroid of the census tract is within a 30-minute walking range. The descriptive statistics of shares are presented in %.

Table B7: Results: Census tract characteristics 1999 and 2000 vs. Treatment

	(1)	(2)	(3)	(4)
Outcome:	ln Income	High Skill Share	Occupation share	Informality Rates
T_i	-0.010***	-0.035***	-0.012***	0.029***
	(0.003)	(0.008)	(0.002)	(0.006)
Obs	3156	3156	3156	3156
R2	0.849	0.196	0.330	0.123
Distance controls	X	X	X	X
State FE	X	X	X	X

Notes: This table reports the results of a regression relating census tract characteristics with a dummy variable whether the centroid of the census tract is within a 30-minute walking range of Line B. The first column reports the results for the log of income, the second column for the share of high-skilled workers, the third column for occupation, and the fourth column for the informality rate. Conley standard errors with a bandwidth of 500 meters are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B8: Difference-in-Difference - Log individuals

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Outcome:	log individuals			log informal residents			log formal residents		
				Panel A: Pool of workers					
T_i	0.018*	0.015	-0.007	-0.015	-0.023	-0.036**	0.064***	0.063***	0.032**
	(0.010)	(0.011)	(0.010)	(0.015)	(0.015)	(0.016)	(0.012)	(0.012)	(0.013)
Obs	3156	3156	3156	3156	3156	3156	3156	3156	3156
R2	0.312	0.313	0.303	0.182	0.183	0.197	0.433	0.404	0.390
				Panel B: Low-skilled workers					
T_i	0.030***	0.023**	-0.000	-0.004	-0.013	-0.027*	0.076***	0.076***	0.043***
	(0.010)	(0.010)	(0.010)	(0.015)	(0.015)	(0.016)	(0.013)	(0.013)	(0.013)
Obs	3156	3156	3156	3156	3156	3156	3156	3156	3156
R2	0.422	0.463	0.417	0.224	0.227	0.217	0.474	0.474	0.438
				Panel C: High-skilled workers					
T_i	0.001	0.004	-0.016	-0.022	-0.029	-0.042**	0.022	0.022	-0.001
	(0.017)	(0.017)	(0.015)	(0.020)	(0.020)	(0.020)	(0.017)	(0.017)	(0.016)
Obs	3156	3156	3156	3156	3156	3156	3156	3156	3156
R2	0.431	0.432	0.452	0.375	0.351	0.360	0.443	0.443	0.447
Pop Controls	X	X	X	X	X	X	X	X	X
Dist Controls	X	X	X	X	X	X	X	X	X
Prod Controls	X	X	X	X	X	X	X	X	X
State FE		X			X			X	
Zone FE			X			X			X

Notes: This table reports the results of a regression relating changes in the log of the number of individuals in each location and sector with Line B of the subway. Panel A reports the results for the pool of workers, panel B for low-skilled workers, and panel C for high-skilled workers. The first three columns report the results for the overall number of individuals, the fourth to sixth columns for individuals in the informal sector, and the seventh to ninth columns for workers in the formal sector. The regressions are weighted by the population in 2000. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses with a bandwidth of 500 meters (Conley, 1999). $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table B9: Least Cost Path IV - Ratio of Formal to Informal Residents

	(1)	(2)	(3)	(4)	(5)	(6)
Outcomes:	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$
	Panel A: Continuous treatment measure-Pool of residents					
-ln dist	0.053***	0.058***	0.060***	0.063***	0.052***	0.057***
	(0.011)	(0.012)	(0.011)	(0.012)	(0.014)	(0.014)
Obs	3156	3156	3156	3156	3156	3156
R2	0.267	0.279	0.215	0.223	0.222	0.230
	Panel B: Treatment dummy variable-Pool of residents					
T_i	0.114***	0.125***	0.130***	0.135***	0.108***	0.117***
	(0.025)	(0.026)	(0.026)	(0.026)	(0.029)	(0.030)
Obs	3156	3156	3156	3156	3156	3156
R2	0.257	0.269	0.203	0.211	0.222	0.230
	Panel C: Continuous treatment measure-Low-skilled workers					
-ln dist	0.062***	0.066***	0.066***	0.068***	0.058***	0.062***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.015)	(0.015)
Obs	3156	3156	3156	3156	3156	3156
R2	0.242	0.248	0.175	0.179	0.189	0.195
	Panel D: Treatment dummy variable-Low-skilled workers					
T_i	0.133***	0.141***	0.142***	0.146***	0.119***	0.128***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.031)	(0.032)
Obs	3156	3156	3156	3156	3156	3156
R2	0.231	0.237	0.163	0.167	0.188	0.194
	Panel E: Continuous treatment measure-Outskirts areas					
-ln dist	0.083***	0.090***	0.091***	0.093***	0.067***	0.071***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)	(0.017)
Obs	2128	2128	2128	2128	2128	2128
R2	0.289	0.306	0.209	0.220	0.208	0.220
	Panel F: Treatment dummy variable-Outskirts areas					
T_i	0.210***	0.226***	0.230***	0.234***	0.168***	0.176***
	(0.040)	(0.041)	(0.042)	(0.042)	(0.041)	(0.042)
Obs	2128	2128	2128	2128	2128	2128
R2	0.269	0.285	0.186	0.198	0.204	0.216
Pop Controls	X	X	X	X	X	X
Dist Controls	X	X	X	X	X	X
Prod Controls		X		X		X
State FE			X	X		
Zone FE					X	X

Notes: This table reports the results of a regression relating changes in the ratio of formal to informal residents in each location to the opening of Line B of the subway using as an instrument the least cost path based on the routes that minimize the cost of building the metro. Panel A reports the results for the continuous treatment measures and the pool of residents, panel B for the treatment dummy variables and the pool of residents, panel C for the continuous treatment measure and low-skilled workers, panel D for the treatment dummy variables and low skilled workers, panel E for the continuous treatment measure on the locations that are not in the CBD, and panel F for the treatment dummy variable on the locations that are not in the CBD. In the first two columns, I do not include geographic fixed effects, in columns 3 and 4, I include state fixed effects, and in columns 5 and 6, zone fixed effects. The regressions are weighted using the initial population shares. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses with a bandwidth of 500 meters (Conley, 1999). *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table B10: Change in Covariates after the Transit Shock

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Outcome:			Number of kids						Household size			
-ln dist	0.039		0.034		0.053		-0.198		-0.439		-3.973	
	(0.049)		(0.048)		(0.059)		(2.857)		(3.022)		(4.720)	
T_i		0.057		0.033		0.049		6.492		5.255		5.558
		(0.108)		(0.102)		(0.123)		(4.622)		(4.358)		(4.715)
Obs	3156	3156	3156	3156	3156	3156	3156	3156	3156	3156	3156	3156
R2	0.038	0.038	0.041	0.041	0.049	0.048	0.078	0.078	0.079	0.079	0.085	0.085
Control mean (pre)	3.170	3.170	3.170	3.170	3.170	3.170	8.613	8.613	8.613	8.613	8.613	8.613
Outcome:			Household Head gender: Male						Age			
-ln dist	0.000		0.000		-0.000		0.014		0.014		0.053	
	(0.000)		(0.000)		(0.000)		(0.028)		(0.028)		(0.033)	
T_i		-0.001		-0.000		-0.002**		0.012		0.012		0.101
		(0.001)		(0.001)		(0.001)		(0.060)		(0.060)		(0.070)
Obs	3156	3156	3156	3156	3156	3156	3156	3156	3156	3156	3156	3156
R2	0.226	0.226	0.266	0.266	0.201	0.202	0.140	0.140	0.114	0.114	0.066	0.065
Control mean (pre)	0.473	0.473	0.473	0.473	0.473	0.473	32.532	32.532	32.532	32.532	32.532	32.532
Outcome:			High-skilled share						Student share			
-ln dist	-0.001		-0.000		-0.001		-0.003***		-0.002***		-0.002	
	(0.001)		(0.001)		(0.001)		(0.001)		(0.001)		(0.001)	
T_i		-0.004		-0.003		-0.002		-0.005**		-0.005**		-0.003
		(0.002)		(0.002)		(0.002)		(0.002)		(0.002)		(0.002)
Obs	3156	3156	3156	3156	3156	3156	3156	3156	3156	3156	3156	3156
R2	0.195	0.196	0.219	0.220	0.146	0.146	0.143	0.142	0.141	0.140	0.145	0.145
Control mean (pre)	0.268	0.268	0.268	0.268	0.268	0.268	0.155	0.155	0.155	0.155	0.155	0.155
Outcome:			Years of education						Medium-skilled share			
-ln dist	-0.006		-0.005		-0.016		-0.000		-0.000		-0.000	
	(0.009)		(0.009)		(0.010)		(0.001)		(0.001)		(0.001)	
T_i		-0.039*		-0.030		-0.027		0.001		-0.000		0.001
		(0.021)		(0.021)		(0.021)		(0.002)		(0.002)		(0.002)
Obs	3156	3156	3156	3156	3156	3156	3156	3156	3156	3156	3156	3156
R2	0.039	0.040	0.063	0.064	0.033	0.032	0.601	0.601	0.603	0.602	0.563	0.563
Control mean (pre)	9.529	9.529	9.529	9.529	9.529	9.529	0.456	0.456	0.456	0.456	0.456	0.456
Pop Controls	X	X	X	X	X	X	X	X	X	X	X	X
Dist Controls	X	X	X	X	X	X	X	X	X	X	X	X
Prod Controls	X	X	X	X	X	X	X	X	X	X	X	X
State FE			X	X					X	X		
Zone FE					X	X					X	X

Notes: This table reports the results of a difference-in-difference specification relating changes in household composition and covariates with the transit shock. The odd columns report the results for the continuous treatment variable, and the even columns for the treatment dummy variable. The regressions are weighted using the initial population shares. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses with a bandwidth of 500 meters (Conley, 1999). $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table B11: Sample that did not change the state of residence between the two census waves

	(1)	(2)	(3)	(4)	(5)	(6)
Outcomes:	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$
			Panel A: Continuous treatment measure-Pool of residents			
-ln dist	0.057***	0.061***	0.062***	0.063***	0.029***	0.033***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)
Obs	3157	3157	3157	3157	3157	3157
R2	0.266	0.280	0.217	0.225	0.226	0.235
			Panel B: Treatment dummy variable-Pool of residents			
T_i	0.068***	0.076***	0.083***	0.086***	0.062***	0.069***
	(0.019)	(0.019)	(0.018)	(0.019)	(0.020)	(0.021)
Obs	3157	3157	3157	3157	3157	3157
R2	0.258	0.271	0.207	0.215	0.226	0.235
			Panel C: Continuous treatment measure-Low-skilled workers			
-ln dist	0.061***	0.063***	0.064***	0.065***	0.031***	0.035***
	(0.010)	(0.009)	(0.010)	(0.009)	(0.011)	(0.011)
Obs	3157	3157	3157	3157	3157	3157
R2	0.242	0.250	0.176	0.181	0.193	0.200
			Panel D: Treatment dummy variable-Low-skilled workers			
T_i	0.078***	0.084***	0.087***	0.090***	0.064***	0.071***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.021)	(0.022)
Obs	3157	3157	3157	3157	3157	3157
R2	0.234	0.241	0.166	0.171	0.193	0.200
			Panel E: Continuous treatment measure-Outskirts areas			
-ln dist	0.092***	0.098***	0.094***	0.096***	0.044***	0.048***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)
Obs	2128	2128	2128	2128	2128	2128
R2	0.286	0.303	0.211	0.223	0.213	0.226
			Panel F: Treatment dummy variable-Outskirts areas			
T_i	0.129***	0.144***	0.143***	0.149***	0.089***	0.098***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.027)	(0.028)
Obs	2128	2128	2128	2128	2128	2128
R2	0.269	0.286	0.193	0.205	0.211	0.224
Pop Controls	X	X	X	X	X	X
Dist Controls	X	X	X	X	X	X
Prod Controls		X		X		X
State FE			X	X		
Zone FE					X	X

Notes: This table reports the results of a regression relating changes in the ratio of formal to informal residents in each location to the opening of Line B of the subway for the sample that did not change the state of residence between 2000 and 2010. Panel A reports the results for the continuous treatment measures and the pool of residents, panel B for the treatment dummy variables and the pool of residents, panel C for the continuous treatment measure and low-skilled workers, panel D for the treatment dummy variables and low skilled workers, panel E for the continuous treatment measure on the locations that are not in the CBD, and panel F for the treatment dummy variable on the locations that are not in the CBD. In the first two columns, I do not include geographic fixed effects, in columns 3 and 4, I include state fixed effects, and in columns 5 and 6, zone fixed effects. The regressions are weighted using the initial population shares. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses with a bandwidth of 500 meters (Conley, 1999). $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table B12: Results controlling for spillovers: Dropping observations between 2.5-5.0 km to the transit shock

	(1)	(2)	(3)	(4)	(5)	(6)
Outcomes:	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$	$\Delta(\ln L_{Fi} - \ln L_{Ii})$
	Panel A: Continuous treatment measure-Pool of residents					
-ln dist	0.055***	0.059***	0.058***	0.060***	0.021**	0.026**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.011)	(0.011)
Obs	2829	2829	2829	2829	2829	2829
R2	0.270	0.285	0.215	0.224	0.225	0.235
	Panel B: Treatment dummy variable-Pool of residents					
T_i	0.089***	0.102***	0.109***	0.115***	0.075***	0.087***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.024)	(0.024)
Obs	2829	2829	2829	2829	2829	2829
R2	0.265	0.280	0.211	0.220	0.227	0.237
	Panel C: Continuous treatment measure-Low-skilled workers					
-ln dist	0.059***	0.062***	0.061***	0.063***	0.021*	0.026**
	(0.010)	(0.010)	(0.010)	(0.009)	(0.011)	(0.011)
Obs	2829	2829	2829	2829	2829	2829
R2	0.251	0.259	0.179	0.185	0.197	0.204
	Panel D: Treatment dummy variable-Low-skilled workers					
T_i	0.102***	0.112***	0.115***	0.120***	0.071***	0.082***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.025)	(0.025)
Obs	2829	2829	2829	2829	2829	2829
R2	0.246	0.254	0.174	0.180	0.198	0.206
	Panel E: Continuous treatment measure-Outskirts areas					
-ln dist	0.082***	0.088***	0.082***	0.085***	0.035***	0.040***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)
Obs	2039	2039	2039	2039	2039	2039
R2	0.290	0.308	0.213	0.224	0.213	0.226
	Panel F: Treatment dummy variable-Outskirts areas					
T_i	0.150***	0.170***	0.167***	0.177***	0.111***	0.123***
	(0.025)	(0.026)	(0.025)	(0.026)	(0.027)	(0.028)
Obs	2039	2039	2039	2039	2039	2039
R2	0.281	0.300	0.205	0.218	0.216	0.230
Pop Controls	X	X	X	X	X	X
Dist Controls	X	X	X	X	X	X
Prod Controls		X		X		X
State FE			X	X		
Zone FE					X	X

Notes: This table reports the results of a regression relating changes in the ratio of formal to informal residents in each location to the opening of Line B of the subway after dropping the census tracts within 2.5 to 5.0 km close to Line B. The idea consists of assuming that the potential spillovers are local in nature. This regression compares the "treated" areas with more distant locations. Panel A reports the results for the continuous treatment measures and the pool of residents, panel B for the treatment dummy variables and the pool of residents, panel C for the continuous treatment measure and low-skilled workers, panel D for the treatment dummy variables and low skilled workers, panel E for the continuous treatment measure on the locations that are not in the CBD, and panel F for the treatment dummy variable on the locations that are not in the CBD. In the first two columns, I do not include geographic fixed effects; in columns 3 and 4, I include state fixed effects, and in columns 5 and 6, zone fixed effects. The regressions are weighted using the initial population shares. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses with a bandwidth of 500 meters (Conley, 1999). $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table B13: Nested Logit - Iceberg Costs

	(1)
Costs	*Commuting: Trips to workplace*
Minutes	-0.010***
	(0.000)
Bus	0.524***
	(0.088)
Metro	0.033
	(0.094)
Metrobus	-0.364***
	(0.120)
Car	-0.145
	(0.107)
λ public	0.476***
	(0.025)
Observations	76,585
Trips	15,317
Iceberg cost before (simple mean)	4.098
Iceberg cost after (simple mean)	3.983

Notes: This table reports estimates from a nested logit model using the 2017 Originâ€“Destination survey, restricting the sample to trips that use a single transportation mode. The sample includes commuting trips between home and work in either direction, departing between 6:00 and 10:00 AM and 5:00 and 9:00 PM. The regressions include controls for a male indicator, age-group fixed effects, and time-of-day fixed effects based on departure time from home.

Table B14: Correlation between Scale Parameters and Productivity Measures

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	*Panel A: log productivity formal firms* $\ln A_{iF}$					
ln va pw	0.283***		0.241***		0.090***	
	(0.027)		(0.026)		(0.028)	
ln w		0.563***		0.491***		0.282***
		(0.046)		(0.046)		(0.047)
Obs	3221	3143	3221	3143	3221	3143
R2	0.055	0.094	0.040	0.072	0.005	0.023
Outcome:	*Panel B: log productivity informal firms* $\ln A_{iI}$					
ln va pw	0.046***		0.034***		-0.010	
	(0.012)		(0.012)		(0.012)	
ln w		0.046**		0.023		-0.050***
		(0.019)		(0.019)		(0.018)
Obs	3221	3143	3221	3143	3221	3143
R2	0.008	0.004	0.004	0.001	0.000	0.004
State FE			X	X		
Zone FE					X	X

Notes: This table reports the results of the correlation between the scale parameters derived after inverting the model and traditional observed productivity measure. Panel A reports the correlation for formal firms and panel B for informal firms. The odd columns report the results for the log of the average value added per worker and the even columns for the average wage in the commuting zone. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses with a bandwidth of 500 meters (Conley, 1999). *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table B15: Correlation between Scale Parameters for Amenities and Deficit of Public Services

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	*Panel A: log amenities $\ln B_n$ vs. water and electricity*					
% No elect	-0.403***		-0.416***		-0.287*	
	(0.146)		(0.148)		(0.156)	
% No water		-0.055***		-0.056***		-0.018***
		(0.011)		(0.012)		(0.006)
Obs	3312	3312	3312	3312	3189	3189
R2	0.030	0.052	0.032	0.053	0.011	0.006
Outcome:	*Panel B: log amenities $\ln B_n$ vs. floor quality and sanitation*					
% No floor	-0.157***		-0.164***		-0.074***	
	(0.030)		(0.030)		(0.021)	
% No sanitation		-0.178***		-0.178***		-0.127**
		(0.041)		(0.041)		(0.058)
Obs	3312	3312	3312	3312	3189	3189
R2	0.059	0.038	0.064	0.038	0.011	0.022
State FE			X	X		
Zone FE					X	X

Notes: This table reports the results of the correlation between the scale parameters derived after inverting the model and traditional observed amenity measure. Panel A reports the correlation between amenities and the share of dwelling at the census tract with a deficit of water and electricity and panel B with the quality of floors and sanitation. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses with a bandwidth of 500 meters (Conley, 1999). *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table B16: Model Validation: Municipality-Level Commuting Flows

	(1)	(2)	(3)
Outcome:		*Observed commuting flows*	
log commuting flows-model	1.014***	0.953***	1.005***
	(0.052)	(0.040)	(0.038)
Obs	576	576	576
p-value: $\beta = 1$	0.787	0.241	0.891
Origin FE	X		X
Destination FE		X	X

Notes: This table reports Poisson Pseudo-Maximum Likelihood (PPML) estimates relating observed municipality-level commuting flows from the 2015 Intercensal Census to commuting flows predicted by the model. The independent variable is the log of predicted commuting flows, $\log \lambda_{nis}$ from the model. Columns (1)-(3) include origin fixed effects, destination fixed effects, or both. Standard errors are clustered two-way at the origin and destination municipality level. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table B17: Effects on Welfare: Counterfactual Results $\hat{X} = X'/X$ - Low Informality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome measure:		1. Welfare				2. Output		
	Mean	sd	p05	p95	Mean	sd	p05	p95
			Panel A: Worker mobility (Baseline)					
% Δ at baseline	0.839	0.097	0.649	1.029	0.828	0.092	0.648	1.009
% Δ-wedge	0.836	0.098	0.644	1.029	0.843	0.094	0.659	1.028
			Panel B: No worker and no firm mobility					
% Δ at baseline	0.842	0.095	0.656	1.028	0.860	0.091	0.682	1.038
% Δ-wedge	0.841	0.094	0.656	1.026	0.876	0.093	0.694	1.058
			Panel C: Worker and firm mobility					
% Δ at baseline	0.839	0.098	0.647	1.031	0.829	0.092	0.649	1.009
% Δ-wedge	0.837	0.100	0.642	1.032	0.844	0.094	0.660	1.028
			Panel D: Firm mobility					
% Δ at baseline	0.839	0.096	0.651	1.028	0.858	0.092	0.678	1.039
% Δ-wedge	0.837	0.096	0.648	1.026	0.875	0.094	0.690	1.059

Notes: This table reports the results of the counterfactuals in which initial informality is low in the city after running a bootstrap with 50 replications. The first four columns report results for workers' welfare, and the last four columns report the effect on total output. For each measure, the table reports the mean, the standard deviation, and the 5th and 95th percentiles of the bootstrap distribution..

Table B18: Effects on Welfare: Counterfactual Results $\hat{X} = X'/X$-Robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Outcome measure:		1. Welfare				2. Output			Robustness scenario
	Mean	sd	p05	p95	Mean	sd	p05	p95	Alternative value of the parameter
					Worker mobility (Baseline)				
% Δ at baseline	0.635	0.072	0.494	0.777	0.630	0.065	0.504	0.757	Baseline
% Δ-wedge	0.749	0.086	0.582	0.917	0.850	0.092	0.669	1.031	Baseline
% Δ at baseline	0.637	0.071	0.497	0.776	0.646	0.067	0.515	0.777	Lower commuting elasticity ($\theta = 5.5$)
% Δ-wedge	0.746	0.086	0.577	0.915	0.858	0.096	0.670	1.045	Lower commuting elasticity ($\theta = 5.5$)
% Δ at baseline	0.637	0.075	0.490	0.783	0.616	0.064	0.491	0.742	Higher commuting elasticity ($\theta = 6.7$)
% Δ-wedge	0.753	0.088	0.581	0.925	0.841	0.092	0.662	1.021	Higher commuting elasticity ($\theta = 6.7$)
% Δ at baseline	0.607	0.071	0.467	0.747	0.588	0.065	0.461	0.716	Lower labor supply elasticity ($\kappa = 1.2$)
% Δ-wedge	0.671	0.083	0.510	0.833	0.727	0.087	0.557	0.897	Lower labor supply elasticity ($\kappa = 1.2$)
% Δ at baseline	0.643	0.068	0.509	0.777	0.643	0.059	0.527	0.759	Higher labor supply elasticity ($\kappa = 2.4$)
% Δ-wedge	0.778	0.076	0.629	0.927	0.915	0.084	0.751	1.080	Higher labor supply elasticity ($\kappa = 2.4$)
% Δ at baseline	0.648	0.075	0.501	0.795	0.602	0.063	0.477	0.726	Higher migration elasticity ($\eta = 1.1$)
% Δ-wedge	0.763	0.089	0.588	0.938	0.824	0.094	0.641	1.008	Higher migration elasticity ($\eta = 1.1$)
% Δ at baseline	0.635	0.072	0.494	0.777	0.630	0.065	0.504	0.757	Higher firm location elasticity (lower $\psi = 5$)
% Δ-wedge	0.749	0.086	0.582	0.917	0.850	0.092	0.669	1.031	Higher firm location elasticity (lower $\psi = 5$)
% Δ at baseline	0.903	0.065	0.776	1.031	0.930	0.056	0.820	1.039	Lower elasticity of substitution across varieties ($\sigma = 3$)
% Δ-wedge	1.069	0.065	0.942	1.196	1.220	0.073	1.077	1.362	Lower elasticity of substitution across varieties ($\sigma = 3$)
% Δ at baseline	0.582	0.067	0.451	0.712	0.577	0.058	0.464	0.690	Higher elasticity of substitution across varieties ($\sigma = 8$)
% Δ-wedge	0.683	0.074	0.537	0.829	0.793	0.081	0.634	0.952	Higher elasticity of substitution across varieties ($\sigma = 8$)

Notes: This table reports the results for the robustness checks of the main counterfactuals after running a bootstrap with 50 replications. The first four columns report results on workers' welfare, and the last four columns report the effects on output. The table reports the mean, standard deviation, and the 5th and 95th percentiles of the bootstrap distribution for each measure. The results are reported for the worker mobility scenario.

Table B19: Data vs. Model Predictions Across Different Buffer Areas

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Buffer:	1,000 meters	1,000 meters	1,500 meters	1,500 meters	2,000 meters	2,000 meters	2,500 meters	2,500 meters
	Panel A: log formal/informal data vs. log formal/informal model							
Outcome:	$\Delta \ln L_{iF}/L_{iI}$	$\Delta \ln L_{iF}/L_{iI}$	$\Delta \ln L_{iF}/L_{iI}$	$\Delta \ln L_{iF}/L_{iI}$	$\Delta \ln L_{iF}/L_{iI}$	$\Delta \ln L_{iF}/L_{iI}$	$\Delta \ln L_{iF}/L_{iI}$	$\Delta \ln L_{iF}/L_{iI}$
$\Delta \ln L_F/L_I$ model	0.766*	0.649**	0.967***	0.866***	0.860***	0.767***	0.810***	0.756***
	(0.417)	(0.312)	(0.297)	(0.232)	(0.237)	(0.190)	(0.203)	(0.169)
p-value: $\beta = 1$	0.574	0.260	0.913	0.562	0.554	0.221	0.348	0.149
Obs	476	476	695	695	902	902	1098	1098
R2	0.224	0.152	0.237	0.158	0.217	0.125	0.224	0.131
	Panel B: log pop data vs. log pop model							
Outcome:	$\Delta \ln L_i$	$\Delta \ln L_i$	$\Delta \ln L_i$	$\Delta \ln L_i$	$\Delta \ln L_i$	$\Delta \ln L_i$	$\Delta \ln L_i$	$\Delta \ln L_i$
$\Delta \ln pop$ model	2.132***	1.807**	1.221**	0.910	1.307**	0.994**	1.414***	1.121**
	(0.766)	(0.709)	(0.615)	(0.598)	(0.509)	(0.501)	(0.465)	(0.463)
p-value: $\beta = 1$	0.139	0.255	0.720	0.881	0.547	0.990	0.373	0.794
Obs	476	476	695	695	902	902	1098	1098
R2	0.301	0.290	0.288	0.269	0.304	0.285	0.312	0.291
Pop controls	X	X	X	X	X	X	X	X
Dist controls	X	X	X	X	X	X	X	X
Prod controls	X	X	X	X	X	X	X	X
State FE		X		X		X		X
Instrument	Δ CMA	Δ CMA	Δ CMA	Δ CMA	Δ CMA	Δ CMA	Δ CMA	Δ CMA

Notes: This table reports the results of a regression testing the changes observed in the data vs. the model predictions comparing locations close to Line B vs. locations within a buffer zone to the other feeder lines in the Plan Maestro in 1985. Panel A reports the results of a regression relating the change in the ratio of formal to informal workers who live in those areas observed in the data vs. the predictions implied by the model. Panel B reports the results for the change in populations. The model's predictions are instrumented with the change in commuter market access. The regressions are weighted using the initial population shares. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses with a bandwidth of 500 meters (Conley, 1999). $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

C Alternative Model Specifications

This section considers two alternative production-side specifications from the main text. These alternatives differ in their market structure and in how production is organized within each location–sector pair. In both models, preferences, technologies, factor intensities, and wedges are identical, and expenditure shares π_{is} are endogenous equilibrium objects. The key distinction lies in whether production in a given (i, s) pair is carried out by a single representative firm operating under perfect competition or by a continuum of monopolistically competitive firms with free entry.

Importantly, the factor demand conditions take the same accounting form in both models conditional on revenues. However, revenues differ across specifications because the determination of π_{is} differs: under the single representative firm case, π_{is} depends only on relative prices, whereas under monopolistic competition it also depends on the endogenous number of firms operating in each location–sector pair.

C.1 Representative firm

In the representative firm benchmark, there is a firm in each sector s and location i. Preferences over freely tradable goods take a standard CES form:

$$C = \left(\sum_{i,s} x_{is}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}},$$

where σ denotes the elasticity of substitution across varieties. The associated price index is

$$P = \left(\sum_{i,s} p_{is}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}. \tag{C.1}$$

Production in each location–sector pair takes the following form:

$$Q_{is} = A_{is} \left(\frac{\tilde{L}_{is}}{\beta} \right)^{\beta} \left(\frac{\tilde{Z}_{is}}{1-\beta} \right)^{1-\beta}, \tag{C.2}$$

where A_{is} denotes total factor productivity, $\tilde{L}_{is}$ is labor, and $\tilde{Z}_{is}$ is commercial floor space. Cost minimization implies that prices equal marginal cost:

$$p_{is} = \frac{(w_{is}[1+t_{sL}])^{\beta}(q_i[1+t_{sZ}])^{1-\beta}}{A_{is}}, \tag{C.3}$$

where w_{is} denotes the wage per efficiency unit in location i and sector s, q_i is the price of commercial floor space in location i, t_{sL} and t_{sZ} capture sector-specific wedges on labor and floor space, respectively, and β is the expenditure share on labor. The CES demand properties imply that the expenditure share on goods produced in (i, s) is

$$\pi_{is} = \frac{p_{is}^{1-\sigma}}{\sum_{r,k} p_{rk}^{1-\sigma}}, \tag{C.4}$$

which is endogenous through prices and productivity.

Let Y_{is} denote total revenue in location i and sector s, $Y_{is} = \pi_{is} \sum_n \alpha X_n$ where X_n is total expenditure in location n and α is the expenditure share on the composite good. Then, aggregate factor demands satisfy the following equations:

$$w_{is}(1+t_{sL})\tilde{L}_{is} = \beta Y_{is}, \qquad q_i(1+t_{sZ})\tilde{Z}_{is} = (1-\beta)Y_{is}. \tag{C.5}$$

In equilibrium, wages and commercial floor space prices adjust so that factor demands equal the corresponding supplies described in the main text.

C.2 Monopolistic Competition

We now consider an alternative market structure in which a continuum of monopolistically competitive firms produce differentiated varieties within each location–sector pair. Preferences take the form:

$$C = \left(\sum_{i,s} \int_j x_{isj}^{\frac{\sigma-1}{\sigma}} dj \right)^{\frac{\sigma}{\sigma-1}},$$

with price index

$$P = \left(\sum_{i,s} \int_j p_{isj}^{1-\sigma} dj \right)^{\frac{1}{1-\sigma}}, \tag{C.6}$$

where p_{isj} denotes the price charged by firm j in (i,s). The production of each good follows the model from Krugman (1991). Firms compete monopolistically and to produce a variety a firm must incur both a constant variable cost and a fixed cost. Both costs use labor and commercial floor space with the same intensity factor for all firms, which implies that the production function is homothetic. The variable cost varies with the productivity from location i and sector s, and is represented by A_{is}.

The total cost of producing x_{ij} units of variety j in location i and sector s is:

$$\Gamma_{isj} = \left(F_s + \frac{x_{isj}}{A_{is}} \right) (w_{is}[1+t_{sL}])^{\beta} (q_i[1+t_{sZ}])^{1-\beta}, \tag{C.7}$$

where w_{is} is the wage per efficiency unit in (i,s), q_i is the price of commercial floor space, β is the output elasticity with respect to labor and F_s is a fixed cost that varies by sector. The variables t_{sL} and t_{sZ} represent the wedges in each sector. Profit maximization implies that the equilibrium price is the standard constant mark-up over marginal cost. The price charged by firms in location i and sector s is:

$$p_{isj} = \left(\frac{\sigma}{\sigma-1} \right) \frac{(w_{is}[1+t_{sL}])^{\beta} (q_i[1+t_{sZ}])^{1-\beta}}{A_{is}}. \tag{C.8}$$

The zero-profit condition implies that the equilibrium output of each variety is constant across firms in the same sector and location and is:

$$x_{isj} = \bar{x}_{is} = A_{is} F_s (\sigma - 1). \tag{C.9}$$

Aggregate payments to labor and commercial floor space (inclusive of taxes) represent constant shares of total revenue in location i and sector s. These cost shares are given by β for labor and $1-\beta$ for commercial floor space.

Total revenue in location i and sector s is given by $Y_{is} = \pi_{is} \sum_n \alpha X_n$, where X_n denotes expenditure originating in location n, α is the expenditure share on the composite good, and π_{is} is the aggregate expenditure share on sector-s goods produced in location i:

$$\pi_{is} = \frac{M_{is} p_{is}^{1-\sigma}}{\sum_{r,k} M_{rk} p_{rk}^{1-\sigma}} \tag{C.10}$$

where M_{is} denotes the total number of firms in location i and sector s. All firms within a given location–sector pair choose identical quantities of labor and commercial floor space. As a result, the total number of firms is pinned down by the aggregate amounts of labor and commercial floor space employed in that location and sector. Imposing the zero-profit condition implies that the equilibrium number of firms is given by

$$M_{is} = \frac{\tilde{\beta}\, \tilde{L}_{is}^{\beta}\, \tilde{Z}_{is}^{1-\beta}}{\sigma F_s}, \tag{C.11}$$

where $\tilde{\beta}$ is a sector-specific constant and F_s denotes the fixed cost of production in sector s. This implies, that the love of variety combined with free entry generates agglomeration externalities, which are summarized by the elasticity $\frac{1}{\sigma-1}$.

Factor demands for labor and commercial floor space satisfy the following equations:

$$w_{is}(1+t_{sL})\tilde{L}_{is} = \beta_s Y_{is}, \qquad q_i(1+t_{sZ})\tilde{Z}_{is} = (1-\beta_s) Y_{is}, \tag{C.12}$$

In equilibrium, market-clearing conditions require that factor demands equal their corresponding supplies, which are described in the main text. Then, the difference between the monopolistic competitive model and perfect competition with a representative firm relies on the number of firms and agglomeration forces.

C.3 Results

I perform the full calibration for both models. The full algorithm includes the indirect inference approach to estimate the labor supply elasticity across sectors, κ, the migration elasticity, η, and the counterfactual analysis. Table C20 reports the estimates of the main model parameters for both models:

Table C20: Estimation of the main model parameters-Alternative model specifications

Parameter	Point est.	s.d.	95% CI
	Panel A: Single representative firm		
η (migration elasticity)	0.390	0.040	[0.312, 0.468]
κ (labor supply elasticity)	1.799	0.209	[1.389, 2.209]
	Panel B: Monopolistic competition		
η (migration elasticity)	0.389	0.040	[0.311, 0.467]
κ (labor supply elasticity)	1.764	0.197	[1.379, 2.150]

Notes: This table reports GMM estimation results. The first column reports the point estimates from the baseline sample. The second column reports standard deviations obtained from a bootstrap procedure with 50 replications, and the third column reports 95% confidence intervals based on a normal distribution. The migration elasticity, η, and the labor supply elasticity across sectors, κ, are estimated using an indirect inference approach that matches the treatment dummy coefficients. Panel A presents the results for the perfectly competitive model with a representative firm, and Panel B presents the results for the monopolistic competitive model.

Overall, the results are very similar to the main estimates reported in Table 4. The migration elasticity is nearly identical across the three models. Estimates of the sectoral labor supply elasticity are also very similar under the single representative firm case and the model with heterogeneous firms (1.79), and only slightly lower under monopolistic competition (1.76).

Table C21 reports the estimates for the welfare and output gains for the two alternative models. As in the main table, the first four columns report the results for welfare and the second four columns for output. It reports the mean, standard deviation and confidence interval. Overall, the results are similar to the ones in the main text. For the representative-firm specifications, the estimated gains

Table C21: Effects on Welfare: Counterfactual Results $\hat{X} = X'/X$
Alternative models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome measure:		1. Welfare				2. Output		
	Mean	sd	p05	p95	Mean	sd	p05	p95
Panel A: Worker mobility-Representative firm								
% Δ at baseline	0.573	0.067	0.442	0.703	0.561	0.057	0.450	0.672
% Δ-wedge	0.659	0.075	0.511	0.806	0.714	0.075	0.567	0.861
Panel B: No worker mobility-Representative firm								
% Δ at baseline	0.573	0.067	0.443	0.704	0.581	0.059	0.465	0.696
% Δ-wedge	0.659	0.075	0.511	0.806	0.731	0.076	0.581	0.881
Panel C: Worker mobility-Monopolistic competition								
% Δ at baseline	0.666	0.075	0.519	0.812	0.684	0.069	0.548	0.819
% Δ-wedge	0.779	0.087	0.609	0.950	0.881	0.093	0.698	1.064
Panel D: No worker mobility-Monopolistic competition								
% Δ at baseline	0.669	0.076	0.521	0.817	0.704	0.072	0.563	0.846
% Δ-wedge	0.781	0.087	0.611	0.952	0.898	0.094	0.713	1.084

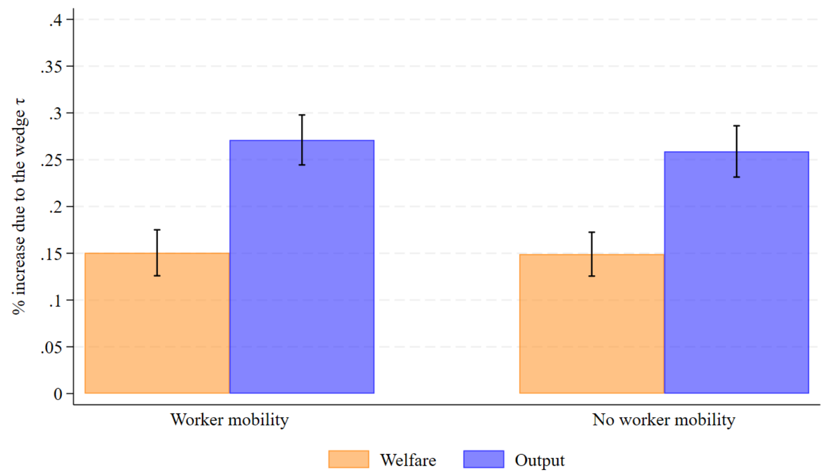
Notes: This table reports the results of the counterfactuals after running a bootstrap with 50 replications for the models that assume perfect or monopolistic competition. The first four columns report results on workers' welfare, and the last four columns report the results on output. The table reports the mean, the standard deviation, and the 5th and 95th percentiles of the bootstrap distribution. Panels A and B report the results for the perfectly competitive case, and panels C and D for the monopolistic competitive case.

are generally smaller, reflecting the absence of free entry and the lack of firm transitions from the informal to the formal sector that are present in the baseline model. In this case, average welfare gains in the absence of the wedge are about 0.57%, with a confidence interval of [0.44%, 0.70%], while output gains are 0.56% (CI: [0.45%, 0.67%]). When the wedge is introduced, welfare gains rise to 0.66% (CI: [0.51%, 0.80%]) and output gains to 0.71% (CI: [0.57%, 0.86%]).

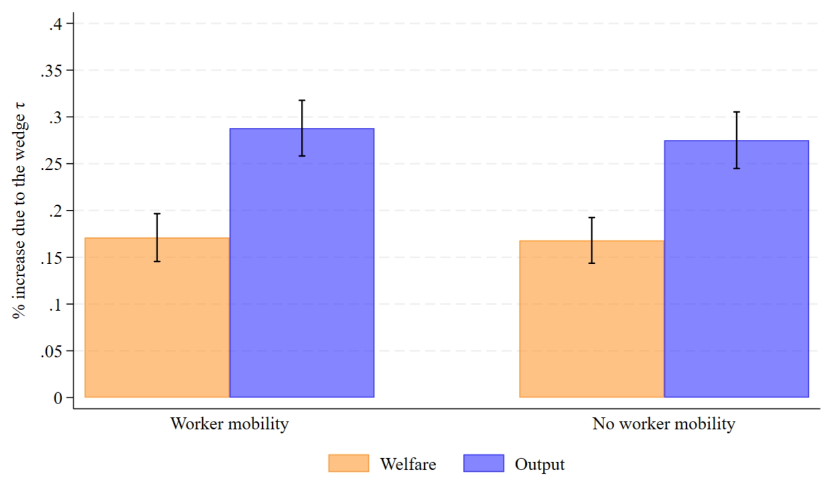
Under monopolistic competition, the presence of free entry leads to results that are very similar to those in the heterogeneous-firm model since the number of firms in the formal and informal sector respond to the change in economic conditions. Welfare gains without the wedge are approximately 0.67% (CI: [0.52%, 0.82%]), and output gains are about 0.68% (CI: [0.55%, 0.82%]). Accounting for the wedge further increases welfare gains to 0.78% (CI: [0.61%, 0.95%]) and output gains to 0.90% (CI: [0.70%, 1.06%]), corresponding to increases of 18% and 30%, respectively, closely mirroring the magnitudes obtained in the baseline heterogeneous-firm specification in the main text.

Figure C9 plots the additional gains for the two models and their 95% confidence intervals due to the wedge between formal and informal firms and the reallocation mechanism. In the representative-

Figure C9: Additional gains: Reallocation from the informal to the formal sector
Alternative models



(a) Representative firm



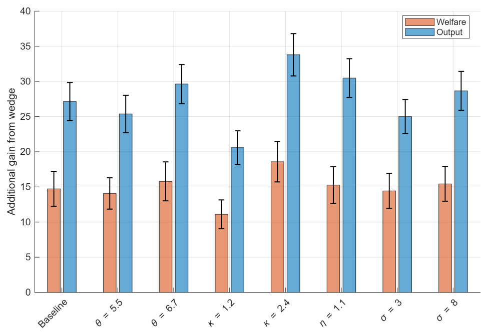
(b) Monopolistic competition

Notes: This figure plots the additional gains from the infrastructure that arise due to the reallocation of workers from the informal to the formal sector for 50 different samples using the bootstrap procedure for the alternative models. The orange bars report the mean and 95% CI for the additional welfare gains, and the blue bar for the additional output gains.

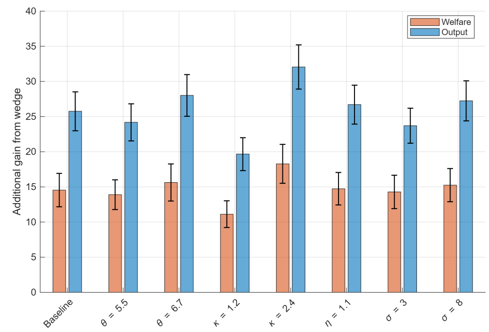
firm case, welfare gains increase by about 15% and output gains by approximately 25%. As discussed above, the absence of a free-entry condition and the lack of endogenous firm transitions between the formal and informal sectors imply a smaller response of formal employment. As a result, gains from the allocative efficiency margin are attenuated relative to the model with heterogeneous firms. By contrast, under monopolistic competition, the results closely mirror those of the main specification. In this case, the allocative efficiency margin accounts for increases of roughly 17% in welfare and 28% in output, consistent with the magnitudes obtained in the heterogeneous-firm model.

Figure C10 illustrates the additional gains across robustness exercises that vary key model parameters, including lower and higher values of the commuting elasticity, the sectoral labor supply elasticity, and the migration elasticity. The orange bars represent the additional welfare gains, the blue bars the additional output gains, and the black lines the 95% confidence interval. Overall, the

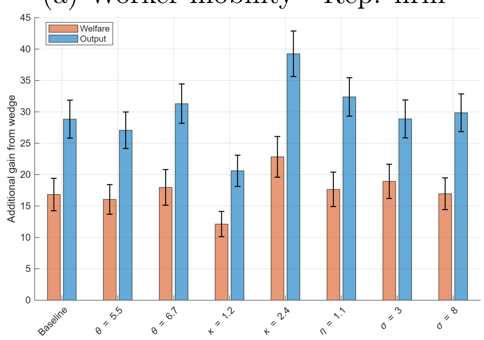
Figure C10: Robustness checks: Additional gains from the reallocation
Alternative models



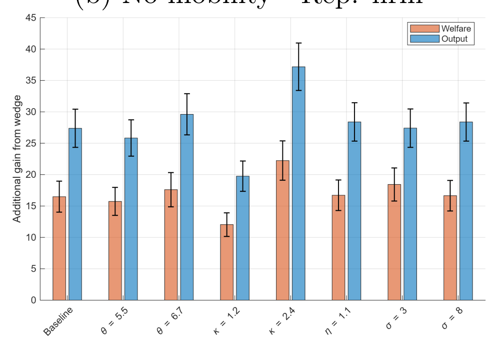
(a) Worker mobility - Rep. firm



(b) No mobility - Rep. firm



(a) Worker mobility - Monopolistic comp.



(b) No mobility - Monopolistic comp.

Notes: This figure plots the additional gains from the infrastructure that arise due to the reallocation of workers from the informal to the formal sector for the different robustness-check scenarios, changing the main model parameters. The orange bars report the mean and 95% CI for the additional welfare gains, and the blue bar for the additional output gains. The black lines represent the 95% confidence interval estimated using a bootstrap procedure with 50 repetitions for each mobility scenario. Panels A and B consider the model with a single representative firm that operates under perfect competition and Panels C and D the case of monopolistic competition

results are very similar to those obtained under the baseline parameterization. In the representative-firm case, the allocative efficiency (misallocation) channel increases welfare gains by approximately 15% and output gains by about 25%-30%. Under monopolistic competition, this margin raises welfare gains by around 18% and output gains by roughly 30%, closely matching the results from the heterogeneous-firm model. Consistent with the main text, the smallest gains arise when the labor supply elasticity across sectors is low, reflecting more limited reallocation of workers from the informal to the formal sector. In this scenario, welfare gains increase by only 10%-12% in both models, while output gains rise by around 20%. By contrast, when the labor supply elasticity is higher, the allocative efficiency margin becomes substantially stronger, with welfare gains increasing by more than 20% and output gains by approximately 35%. For different values of the other parameters, the gains are similar to the baseline. The welfare range is between 15%-20% and the output range between 30%-40%.

D Data and Quantification Appendix

D.1 Effects on employment

I use establishment-level data from the Economic Censuses (1994–2009), which are collected every five years, to test whether the shock reduced job informality in the "treated" areas. If improved transit access encourages workers to seek formal jobs in the city center, we should observe a decline in informal employment in areas close to the new subway line since jobs reallocate to the center. To examine this, I estimate a difference-in-differences specification using 1994 as the baseline year, comparing changes in informality rates for jobs near the new line to those farther away:

$$y_{it} = \sum_{\tau \neq 1994} \beta_\tau T_i + \delta_i + \delta_{s(i),t} + \alpha_t X_i + \epsilon_{i,t}, \tag{D.1}$$

where $y_{i,t}$ is one of the outcomes of interest, δ_i is a census-tract fixed effect, $\delta_{s(i),t}$ is a state-time fixed effect, and I also include a vector of covariates interacted with time dummies where I include the distance and population controls. The coefficients of interest are β_τ. Table D22 reports the point estimates for the primary outcome, the share of informal workers. Jobs' informality rates decrease in locations near Line B after the transit shock. The evidence also shows parallel trends, as the point estimate is small and insignificant in 1999. On average, informality rates for jobs decrease between 2.0 and 3.0 p.p. in locations that experienced the shock.

Table D22: Difference-in-difference: Employment

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	Share informal workers		log informal workers		log formal workers	
Ti x 1999	-0.009	-0.012	-0.034	-0.023	0.020	0.029
	(0.009)	(0.009)	(0.029)	(0.029)	(0.065)	(0.065)
Ti x 2004	-0.020	-0.022*	-0.101**	-0.070*	0.100	0.119
	(0.013)	(0.013)	(0.041)	(0.041)	(0.082)	(0.084)
Ti x 2009	-0.024*	-0.026*	-0.089**	-0.048	0.099	0.129
	(0.014)	(0.014)	(0.043)	(0.043)	(0.093)	(0.095)
Obs	12929	12929	12914	12914	10715	10715
R2	0.249	0.056	0.614	0.151	0.047	0.026
Distance controls	X	X	X	X	X	X
Pop controls	X	X	X	X	X	X
Census-tract FE	X	X	X	X	X	X
State-time FE		X		X		X

Notes: This table reports the results of a regression running a difference-in-difference specification relating changes in the share of informal workers, the log of informal workers, and the log of formal workers in each location with Line B of the subway. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses with a bandwidth of 500 meters (Conley, 1999). $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

I use the same specification to examine changes in the number of formal and informal workers. The results show that the number of informal workers decreases by approximately 7%- 10%, whereas the number of formal workers remains stable. This pattern is consistent with the idea that the new infrastructure induces a reorganization of economic activity: peripheral areas near the new stations increasingly specialize in residential uses, while central areas specialize more in commercial activity. Because the treated peripheral locations initially had a higher share of informal employment, this reallocation leads to a larger decline in informal jobs.

D.2 Estimation of firm location elasticity

This section describes the estimation of the firm-location elasticity parameter ψ, which governs the responsiveness of firm location decisions to changes in profits across locations in the model.

Model motivation In the model, firms choose their location based on idiosyncratic shocks. The share of firms of sector s that locate in area i is given by:

$$\chi_{is}(\varphi) = \frac{\Phi_{is}\pi_{is}(\varphi)^{1/\psi}}{\sum_j \Phi_{js}\pi_{js}(\varphi)^{1/\psi}}, \tag{D.2}$$

where $\pi_{is}(\varphi)$ denotes profits in location i, Φ_{is} is a location-specific shifter, and ψ governs the dispersion of location choices. A lower value of ψ implies that firms are more responsive to profit differences across locations, while a higher value implies that location choices are less responsive to profits.

Taking logs of the expression above yields:

$$\ln \chi_{is} = \ln \Phi_{is} + \frac{1}{\psi} \ln \pi_{is}. \tag{D.3}$$

In the model, profits are proportional to value-added:

$$\pi_{is}(\varphi) = \frac{\text{VA}_{is}(\varphi)}{\sigma}, \tag{D.4}$$

which implies that firm location shares depend on the log value-added. This relationship motivates the empirical specification used to estimate the firm location elasticity.

Empirical specification To estimate this relationship, I use variation across census tracts and time and estimate the following specification:

$$\ln M_{it} = \gamma_i + \gamma_{s(i),t} + \beta \ln \text{VA}_{it} + \epsilon_{it}, \tag{D.5}$$

where M_{it} is the total number of firms in census tract i at time t, VA_{it} is average value-added per firm in the census tract, γ_i are census-tract fixed effects that control for characteristics that are common across time, and $\gamma_{s(i),t}$ are state-year fixed effects that control for shocks within the state in each period of time. The coefficient β corresponds to $1/\psi$ in the model.

The estimation is conducted at the census-tract level rather than the census-tract sector level. The parameter ψ governs how the total mass of firms reallocates across locations in response to differences in profitability in the model. Therefore, the relevant empirical relationship is how the total number of firms in a location responds to changes in profitability in those areas. Using aggregate firm counts and average value-added at the census-tract level provides a direct empirical analog to this relationship and reduces measurement error relative to more disaggregated specifications.

Endogeneity and instrument Average value-added is endogenous for two reasons. First, unobserved local shocks may simultaneously affect both profitability and firm presence, generating selection bias. Second, the regression suffers from division bias because average value-added is constructed as total value-added divided by the number of firms, and the number of firms is the dependent variable in the regression, creating a mechanical relationship between the regressor and the outcome.

To mitigate these concerns, I instrument the log average value added with log average wages at the tract level. In the model, wages affect firms' operating profits and, therefore, shift location incentives. Conditional on census-tract fixed effects, wages do not enter the firm location decision directly, but operate through profits. The identifying assumption is that conditional on the fixed effects, tract-level wage variation provides exogenous shifts in profitability, allowing me to identify $\frac{1}{\psi}$. For example, higher wages may reflect higher productivity or stronger local demand conditions that increase profits, thereby attracting firms.

Results Table D23 reports the results. The estimates imply that when profits increase by 10%, the number of firms increases by approximately 0.67% to 0.76%, which implies a value of ψ of approximately 13.5. This is the value used in the baseline calibration of the model.

Table D23: Firm location elasticity

	(1)	(2)	(3)	(4)
Dependent variable	ln VA	ln VA	ln Firms	ln Firms
ln w	0.327***	0.327***		
	(0.014)	(0.014)		
ln VA			0.067***	0.076***
			(0.024)	(0.024)
Implied ψ			14.989	13.142
Obs.	12604	12604	12604	12604
Census-tract FE	X	X	X	X
State-year FE	X		X	
Year FE		X		X

Notes: This table reports the results of regressions relating the number of firms to the average value-added in a location, instrumenting value-added with average wages. The first two columns show the results of the first stage, and columns 3 and 4 show the results of the 2SLS specification. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses with a bandwidth of 500 meters (Conley, 1999). $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The parameter ψ governs how responsive firm location is to profit differences across locations. Therefore, it mainly affects the degree of spatial reallocation of firms across locations in counterfactual exercises. Importantly, the estimation of the main structural parameters and the model's quantitative results are not very sensitive to this parameter. In Section 6, I show that the counterfactual results are robust to alternative values of ψ, including lower values that allow for more firm reallocation.

Overall, these results suggest that the paper's quantitative findings are not driven by the specific value of ψ, but rather by the interaction among commuting costs, productivity differences, and distortions across sectors and locations.

D.3 Calibration of Speeds

This section describes the calibration of speeds across different transportation modes. I use different sources of information. For the transportation network in Mexico City, I use data from the Government of the city.[41] For the network of roads, I use information from the New York University digital archive, in which they report different types of roads for each census tract in the commuting zone of Mexico City. The different roads include: autopistas, calles, etc. I calculate an average speed for each road. With this information I compute commuting times across census tracts in Mexico City using the Network analysis toolkit from Arcmap. I compute these times for four different modes of transportation: walking, car, traditional buses, and the subway. I add 5 minutes to each station when computing times for the public transit network. I construct a matrix of approximately 13 million observations across census tracts.

[41] The data can be found here.

Table D24: Calibration of speeds

Type	Speed
Panel A: Public transit system	
Subway Lines	584.00 m/min
Metrobus	266.67 m/min
Bus	200.00 m/min
Walking	80.00 m/min
Panel B: Types of roads for cars	
Autopista	724.83 m/min
Avenida	336.67 m/min
Boulevard	609.70 m/min
Calle	282.06 m/min
Calzada	259.19 m/min
Carretera	621.91 m/min
Cerrada	221.92 m/min
Circuito	366.95 m/min
Corredor	251.80 m/min
Diagonal	282.06 m/min
Eje vial	342.39 m/min
Pasaje	315.77 m/min
Privada	248.38 m/min
Viaducto	443.20 m/min
Other	10.73 m/min

Notes: This table reports the calibration of speeds for different street types using trips from Google Maps. The calibration is based on random trips in 2019. I computed these times between 8 am and 11 am and 5 pm and 8 pm under different traffic scenarios.

To calibrate speeds for each mode and each type of road, I use random trips from Google Maps. I downloaded 4000 random trips in 2019 between 8 am-11 am and 5 pm-8 pm using the R command *gmaps* distance, which uses the Google Maps Distance Matrix API. I use as an origin and destination, the closest vertex of each type of road or metro line. This tool has the feature that you can calculate times for different modes under several traffic scenarios: pessimistic, optimistic, or none, and modes such as walking, car, or the public transit network. Using this information, I calibrate speeds for each road and each line using the average time spent to move from one vertex to the other. Table D24 reports the average speed for each one of the roads and the public transit system.

D.4 Model Inversion

This section details the procedure used to recover the shifters B_n and B_{ns}, which capture differences in amenities attracting residents to each location and sector; the firm location amenity shifters Φ_{is}; and the productivity scale parameters A_{is}. The idea is that, given the key elasticities and the observed distribution of workers, residents, and firms across locations and sectors, I can invert the model from Section 4 and recover these shifters. With these parameters in hand, I recover trade and commuting flows and solve counterfactuals starting from the observed initial equilibrium and changing commuting costs in the model.

Following the quantitative spatial literature (e.g., Allen and Arkolakis (2025)), I recover location-sector fundamentals so that the model exactly reproduces the observed spatial distribution of workers, firms, and residents in the baseline equilibrium. This ensures that counterfactual changes in commuting costs operate through the observed spatial allocation of economic activity rather than through differences between the model and the data in the baseline equilibrium.

I proceed in four steps. In the first step, I jointly recover the wage distribution and the sector amenity parameter shifters, B_{ns}, via a fixed-point algorithm equating labor supply to observed worker counts. In the second step, I recover productivity levels A_{is} by equating labor demand to observed worker counts, iterating jointly over productivities and the aggregate output index Q. In the third step, I recover the firm location amenity parameters Φ_{is} analytically from observed firm location shares and model-implied profits. In the fourth step, I recover location amenity parameters B_n by equating model-implied residential shares to the ones observed in the data.

Step 1: Wages and sectoral labor supply shifters

I jointly recover the wage vector $\{w_{is}\}$ and the sectoral amenity shifters $\{B_{ns}\}$ via a simultaneous fixed-point iteration, normalizing $B_{nI} = 1$ for all n without loss of generality.

Given wages, the commuting costs $\tilde{d}_{ni}$ and the Fréchet dispersion parameter θ, market access in sector s is defined as:

$$W_{ns} = \left(\sum_i w_{is}^{\theta} \, \tilde{d}_{ni}^{-\theta} \right)^{1/\theta} . \tag{D.6}$$

The share of residents in location n choosing sector s is:

$$\lambda_{ns|n} = \frac{B_{ns} \, W_{ns}^{\kappa}}{\sum_{s'} B_{ns'} \, W_{ns'}^{\kappa}}, \tag{D.7}$$

and the share choosing to commute from n to workplace i conditional on sector s is:

$$\lambda_{ni|ns} = \frac{w_{is}^{\theta} \, \tilde{d}_{ni}^{-\theta}}{\sum_j w_{js}^{\theta} \, \tilde{d}_{nj}^{-\theta}}. \tag{D.8}$$

Total labor supply in location i and sector s aggregates over all residential locations:

$$L_{is} = \bar{L}_L \sum_n \lambda_n \cdot \lambda_{ns|n} \cdot \lambda_{ni|ns}, \tag{D.9}$$

where $\bar{L}_L$ is the total mass of workers and λ_n is the residential share of location n. Wages are then updated by inverting the labor supply condition. matching the model with the data. Since $L_{is} \propto w_{is}^{\theta}$, the update rule is

$$w_{is}^{\text{upd}} = \left(\frac{L_{is}^{\text{data}}}{L_{is}/w_{is}^{\theta}} \right)^{1/\theta} . \tag{D.10}$$

In the second step, given the vector of wages, w_{is}, B_{nF} is updated analytically by inverting the sectoral share condition observed in the data. This means:

$$B_{nF}^{\text{upd}} = \frac{L_{nF}^{\text{data}} \, W_{nI}^{\kappa}}{L_{nI}^{\text{data}} \, W_{nF}^{\kappa}}, \tag{D.11}$$

where L_{ns}^{data} denotes the observed number of workers residing in n and working in sector s. Both objects, the vector of wages, w_{is}, and the shifters, B_{nF} are updated jointly until the algorithm reaches convergence.

Step 2: Productivity shifters

Using the converged wages from Step 1, I recover the productivity scale parameters A_{is} by equating labor demand to $\tilde{L}_{is} = \sum_i \lambda_{nis|ns} W_n L_{ns}^{\text{data}}$, the observed number of efficiency units of labor supplied in each location i and sector s. In particular, given data on wages, w_{is}, the shares of formal and informal firms, and total employment, L_{is}, I use the following algorithm: Firms produce a differentiated variety using labor and commercial floor space according to the production function $q_{is}(\varphi) = A_{is}\varphi(\ell_{is}/\beta)^{\beta}(z_{is}/(1-\beta))^{1-\beta}$, and face a downward-sloping demand curve from the CES aggregator with elasticity σ. Profit maximization yields optimal labor demand for firm φ in location i and sector s:

$$\ell_{is}(\varphi) = \beta \left(\frac{\tilde{\sigma}^{-1} Q^{\frac{1}{\sigma}} (A_{is}\varphi)^{\frac{1}{\tilde{\sigma}}} \left[w_{is}(1+t_{Ls}) \cdot [(1+t_{Zs}) r_{Zi}]^{-1} \right]^{\frac{1-\beta}{\tilde{\sigma}}} }{ w_{is}(1+t_{Ls}) } \right)^{\sigma}, \tag{D.12}$$

where $\tilde{\sigma} = \sigma/(\sigma-1)$ is the markup and r_{Zi} is the commercial floor space price in location i. Labor demand is increasing in firm-level productivity φ, location-sector productivity A_{is}, and the aggregate output index Q, and decreasing in the effective wage $(1+t_{Ls})w_{is}$.

Firm productivity φ is drawn from a Pareto distribution with shape ζ, lower bound x_m, and upper bound $x_{\max}$, with a mass M of potential entrants. Formal firms operate above the endogenous productivity cutoff $\tilde{\varphi}$, while informal firms operate below it. Aggregating over the productivity distribution and the firm location probabilities $\chi_{is}(\varphi)$ from equation 4.12, the labor market clearing conditions are:

$$\tilde{L}_{iF} = M \int_{\tilde{\varphi}}^{\bar{\varphi}} \chi_{iF}(\varphi)\, \ell_{iF}(\varphi)\, \mu(\varphi)\, d\varphi, \tag{D.13}$$

$$\tilde{L}_{iI} = M \int_{\underline{\varphi}}^{\tilde{\varphi}} \chi_{iI}(\varphi)\, \ell_{iI}(\varphi)\, \mu(\varphi)\, d\varphi, \tag{D.14}$$

where $\mu(\varphi)$ is the density of the productivity distribution. Because A_{is} and the aggregate output index Q enter labor demand jointly, Q determines equilibrium output prices and hence the scale of factor demands. I iterate over both simultaneously. Specifically, A_{is} is updated in each iteration in proportion to the gap between model-implied and observed employment:

$$A_{is}^{\text{upd}} = A_{is} \cdot \left(1 + \delta \cdot \frac{L_{is}^{D} - \tilde{L}_{is}^{\text{data}}}{\tilde{L}_{is}^{\text{data}}} \right)^{-1}, \tag{D.15}$$

and Q is updated using the CES aggregator evaluated at the current firm output decisions:

$$Q^{\text{upd}} = \left(M \int \left[q_{iF}(\varphi)^{\frac{\sigma-1}{\sigma}} + q_{iI}(\varphi)^{\frac{\sigma-1}{\sigma}} \right] \mu(\varphi)\, d\varphi \right)^{\frac{\sigma}{\sigma-1}}. \tag{D.16}$$

The algorithm iterates until both the location-sector labor market clearing- conditions D.13-D.14 and output market clearing is jointly satisfied. I can then compute operating profits for each firm if they operate under the formal or informal status:

$$\Pi_{is}^{op}(\varphi) = \hat{\sigma} \left(\frac{Q \cdot A_{is}\varphi}{[w_{is}(1+t_{Ls})]^{\beta} \, [r_{Zi}(1+t_{Zs})]^{1-\beta}} \right)^{\sigma-1}, \tag{D.17}$$

where $\hat{\sigma} = \frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}}$ is a constant.

Step 3: Firm location shifters and formal fixed-cost

With profits $\pi_{is}(\varphi)$ in hand from Step 2, I recover the firm location amenity shifters Φ_{is} analytically by inverting the firm location probabilities from equation 4.12. Taking the ratio of observed firm shares in the data to model-implied profits, and applying a geometric mean normalization to remove the scale indeterminacy, I obtain:

$$\Phi_{is}(\varphi) = \frac{\left(\chi_{is}^{\text{data}}(\varphi) \,/\, \tilde{\chi}_{is}\right)^{1/\psi}}{\pi_{is}(\varphi) \,/\, \tilde{\pi}_{is}}, \tag{D.18}$$

where $\tilde{\chi}_{is}$ and $\tilde{\pi}_{is}$ denote the geometric means of observed firm shares and profits across locations, respectively. This normalization ensures that Φ_{is} captures only the residual variation in firm location decisions not explained by profit differentials. Then, I can compute average profits across locations for each firm if they operate formally and informally:

$$\bar{\pi}_s(\varphi) = \sum_i \chi_{is}^{\text{data}}(\varphi)\pi_{is}(\varphi). \tag{D.19}$$

The fixed cost FC_F is identified from the indifference condition of the marginal firm. Since 83% of firms in the data operate informally, I identify FC_F by equating the profit gap at the 83rd percentile of the productivity distribution φ_{p83} to zero:

$$FC_F = \bar{\pi}_F(\varphi_{p83}) - \bar{\pi}_I(\varphi_{p83}), \tag{D.20}$$

so that φ_{p83} is the marginal firm $\tilde{\varphi}$ indifferent between sectors, consistent with the observed formality rate in the data. Then, I use this fixed cost value to compute the counterfactuals.

Step 4: Residential location amenity shifters

With wages, productivities, and sector amenities recovered, I obtain the location amenity parameters B_n by equating model-implied residential shares implied by the model to the ones observed in the data. Housing rents are first recovered from equilibrium expenditure shares and housing supply H_n:

$$r_n = \frac{(1-\alpha)\,X_n}{H_n}, \tag{D.21}$$

where X_n is the total expenditure in location n, itself computed from labor income, profits, and government transfers. The welfare index is then:

$$\mathcal{W}_n = \frac{\Phi_n}{r_n^{1-\alpha}}, \tag{D.22}$$

where $W_n = \left(\sum_s B_{ns}\, W_{ns}^{\kappa}\right)^{1/\kappa}$ is the overall wage index or market access measure. The model-implied residential share is:

$$\lambda_n = \frac{(B_n\,\mathcal{W}_n)^{\eta}}{\sum_{n'} (B_{n'}\,\mathcal{W}_{n'})^{\eta}}, \tag{D.23}$$

where η is the migration elasticity. Applying a geometric mean normalization to remove the scale indeterminacy, I recover B_n analytically as:

$$B_n \propto \left(\frac{\lambda_n^{\text{data}}}{\tilde{\lambda}_n}\right)^{1/\eta} \cdot \frac{1}{\mathcal{W}_n/\tilde{\mathcal{W}}}, \tag{D.24}$$

where tildes again denote geometric means. With all parameters recovered, I compute initial commuting and trade flows from the model's gravity equations, and then, I can solve for counterfactuals using the algorithm described in Section D.5.

D.5 Algorithm

In this section, I explain the main algorithm to solve the general equilibrium model. The system of equations is described in Section 4. The sub-index t represents simulation iterations. The algorithm is based on Alvarez and Lucas (2007) and Redding and Rossi-Hansberg (2017), incorporating firm heterogeneity. The algorithm functions as a contraction mapping and is as follows:

1. In an outer loop, for the grid of firms, compute the productivity values φ drawn from a bounded Pareto productivity distribution with bounds $\underline{\varphi}$ and $\bar{\varphi}$ and shape parameter ζ. I use a grid of 1000 firms.

2. Define the set of firms that are formal or informal by determining the marginal firm $\tilde{\varphi}$ that is indifferent between operating in the formal or informal economy. Initially, I match that 83% of the firms are informal.

3. In the first inner loop, guess an initial vector of wages $\vec{w}^0$, the number of residents in each location $\vec{L}^0$, the share of firms in each location χ_{is}, and total output Q^0.

4. Given a vector $\vec{w}^t$, $\vec{L}^t$, $\vec{\chi}^t$ and Q^t compute the following equations:

   - Labor supply equations:

$$\lambda_{nisL|ns} = \frac{w_{is}^{\theta} \tilde{d}_{ni}^{-\theta}}{\sum_{i'} w_{i's}^{\theta} \tilde{d}_{ni'}^{-\theta}} \tag{D.25}$$

$$\lambda_{nsL|n} = \frac{W_{ns|n}^{\kappa}}{\sum_{s'} W_{ns'|n}^{\kappa}}, \quad W_{ns|n}^{\theta} = \sum_{i'} w_{is}^{\theta} \tilde{d}_{ni'}^{-\theta} \tag{D.26}$$

$$\tilde{L}_{is} = \sum_{n} \lambda_{nis} \cdot \bar{L}_{L}. \tag{D.27}$$

   - Solve the problem for informal and formal firms:

$$\ell_{is}(\varphi) = \beta \left(\frac{\tilde{\sigma}^{-1} Q^{\frac{1}{\sigma}} (A_{is}\varphi)^{\frac{1}{\tilde{\sigma}}} \left[w_{is}(1+t_{Ls}) \cdot [(1+t_{Zs}) r_{Zi}]^{-1} \right]^{\frac{1-\beta}{\tilde{\sigma}}}}{w_{is}(1+t_{Ls})} \right)^{\sigma}. \tag{D.28}$$

   - Compute the average income in each location:

$$\bar{y}_n \equiv \left(\sum_{s} \left(\sum_{i} w_{is}^{\theta} \tilde{d}_{ni}^{-\theta} \right)^{\frac{\kappa}{\theta}} \right)^{\frac{1}{\kappa}} \tag{D.29}$$

   - Compute commercial floor space prices:

$$q_i \tilde{Z}_i = \sum_{s} \frac{(1-\beta)(1+t_{isL}) w_{is} \tilde{L}_{is}}{\beta(1+t_{isZ})}, \tag{D.30}$$

- Compute profits for each firm in each location and sector:

$$\Pi_{is}^{op}(\varphi) = \hat{\sigma} \left(\frac{Q \cdot A_{is}\varphi}{[w_{is}(1+t_{Ls})]^{\beta} [r_{Zi}(1+t_{Zs})]^{1-\beta}} \right)^{\sigma-1}, \tag{D.31}$$

- Using the initial firm shares in each location, calculate expected profits subtracting the fixed cost (expressed in terms of the numeraire good):

$$\bar{\pi}_s(\varphi) = \sum_i \chi_{is}(\varphi)\pi_{is}(\varphi) - FC_s \tag{D.32}$$

- Compute total expenditure X_n:

$$X_n = (\bar{y}_n L_n + r_{Fn} Z_n + r_{Hn} H_n + \iota_n \Pi)(1+\bar{t}), \tag{D.33}$$

- Using the government budget constraint, calculate the rebate:

$$\sum_{i,s} \frac{1}{\beta} \left(t_{isL} w_{is} \tilde{L}_{is} \right) = \bar{t} \cdot \sum_n X_n. \tag{D.34}$$

- Compute housing prices:

$$r_{Hn}\tilde{H}_n = (1-\alpha)X_n, \tag{D.35}$$

5. Add the labor demand across firms that were initially defined as formal and informal to compute total demand in each sector:

$$\tilde{L}_{iF} = M \int_{\varphi \in [\tilde{\varphi}, \bar{\varphi}]} \chi_{iF}(\varphi)\ell_{iF}(\varphi)\mu(\varphi)d\varphi, \tag{D.36}$$

$$\tilde{L}_{iI} = M \int_{\varphi \in [\underline{\varphi}, \tilde{\varphi}]} \chi_{iI}(\varphi)\ell_{iI}(\varphi)\mu(\varphi)d\varphi, \tag{D.37}$$

6. Compute the new share for both formal and informal firms using the operating profits function and the firm reallocation elasticity $\frac{1}{\psi}$:

$$\tilde{\chi}_{is}^{t+1}(\varphi) = \frac{\Phi_{is}\pi_{is}(\varphi)^{1/\psi}}{\sum_j \Phi_{js}\pi_{js}(\varphi)^{1/\psi}}, \tag{D.38}$$

7. Compute the new total output $\tilde{Q}^{t+1}$:

$$\tilde{Q}^{t+1} = \left(M \sum_{i,s} \int \chi_{is}(\varphi) q_{is}(\varphi)^{\frac{\sigma-1}{\sigma}} \mu(\varphi) d\varphi \right)^{\frac{\sigma}{\sigma-1}} \tag{D.39}$$

8. Calculate the difference between labor demand and labor supply, the number of residents, the share of firms in each location, and total output of the composite good:

$$z_w = \frac{L_{is}^D - \tilde{L}_{is}}{\tilde{L}_{is}} \tag{D.40}$$

$$\tilde{L}_n^{t+1} = \left(\frac{B_n P_n^{-\alpha\eta} r_n^{-(1-\alpha)\eta} W_n^{\eta}}{\sum_{n'} B_{n'} P_{n'}^{-\alpha\eta} r_{n'}^{-(1-\alpha)\eta} W_{n'}^{\eta}} \right) \bar{L}_L \tag{D.41}$$

$$z_\chi = \chi_{is}^{t+1} - \chi_{is}^{t} \tag{D.42}$$

$$z_Q = Q^{t+1} - Q^t \tag{D.43}$$

9. If $||(z_w, \tilde{L}_i^t, z_\chi, z_Q) - (0, L_i^t, 0, 0)|| < \epsilon^{\text{tol}}$, then the algorithm of the inner loop stops. Otherwise, update:

$$w_{is}^{t+1} = w_{is}^t(1 + \nu_w z_w) \tag{D.44a}$$

$$L_n^{t+1} = \nu_L \tilde{L}_n^{t+1} + (1 - \nu_L) L_n^t, \tag{D.44b}$$

$$\chi_{is}^{t+1} = \nu_\chi \tilde{\chi}_{is}^{t+1} + (1 - \nu_\chi)\chi_{is}^t \tag{D.44c}$$

$$Q^{t+1} = \nu_Q \tilde{Q}^{t+1} + (1 - \nu_Q) Q^t \tag{D.44d}$$

where ν_w, ν_L, ν_χ, and ν_Q are convergence parameters and ϵ^{tol} is a tolerance value.

10. For the outer loop, solve for the new marginal firm $\tilde{\tilde{\varphi}}^{t+1}$ that is indifferent between operating in the formal or informal economy.

11. Compute the difference between the new and previous marginal firm:

$$z_\varphi = \tilde{\tilde{\varphi}}^{t+1} - \tilde{\varphi}^t \tag{D.45}$$

12. Update the new marginal firm:

$$\tilde{\varphi}^{t+1} = \nu_\varphi \tilde{\tilde{\varphi}}^{t+1} + (1 - \nu_\varphi)\tilde{\varphi}^t \tag{D.46}$$

13. Continue the algorithm until there is no significant difference in the number of the marginal firm. Otherwise, repeat the whole algorithm.

E Theoretical Appendix

E.1 Welfare Decomposition: Back-of-the-envelope calculation

Let's consider that the utility function is $U = f(q_F, q_I)$ where q_F corresponds to the quantity consumed from formal firm types and q_I from informal firm types. The production functions are:

$$q_F = A_F g_F(\lambda_F)$$

$$q_I = A_F g_I(1 - \lambda_F)$$

where the function g_s is a decreasing function in labor. I normalize the total labor force to 1. Formal firms have to pay a labor tax τ_F per efficiency unit of labor, while informal firms do not pay taxes.

On the other hand, the labor supply function takes the same form as in the main text. The share of workers who decide to work in firm type s is:

$$\lambda_s = \frac{B_s w_s^\kappa d_s^{-\kappa}}{\sum_k B_k w_k^\kappa d_k^{-\kappa}}, \tag{E.1}$$

where B_s is a productivity shifter, w_s is the wage paid per efficiency unit by firms s and d_s is an iceberg commuting cost. In this simple case, we assume that $d_F > 1$ and $d_I = 1$. The ex-ante average income for workers is:

$$W = \left(\sum_k B_k w_k^\kappa d_k^{-\kappa}\right)^{\frac{1}{\kappa}}$$

Under homotheticity, the indirect utility function is a function of real income:

$$V = \frac{W(1+\bar{t})}{P(p_F, p_I)},$$

where $P(p_F, p_I)$ is the price index from the expenditure minimization problem, W is the wage index and $\bar{t} = \lambda_F t_F$ is the proportional rebate in this economy. The goal is to understand the effect of a small perturbation in the commuting cost d_F. The change in welfare is:

$$d\ln V = d\ln W + d\ln(1+\bar{t}) - d\ln P, \tag{E.2}$$

where:

$$d\ln W = \lambda_F d\ln w_F - \lambda_F d\ln d_F + \lambda_I d\ln w_I \tag{E.3}$$

and,

$$d\ln P(p_F, p_I) = \frac{p_F e'_{p_F}(p_F, p_I, 1)}{e} d\ln p_F + \frac{p_I e'_{q_I}(p_F, p_I, 1)}{e} d\ln p_I \tag{E.4}$$

$$= \alpha_F d\ln p_F + \alpha_I d\ln p_I \tag{E.5}$$

where $e(p_F, p_I, 1)$ corresponds to the expenditure function and $e'_{p_s}(p_F, p_I, 1) = q_s$ by the Shephard's lemma. The equilibrium conditions imply that:

$$\lambda_F(1+t_F) = \alpha_F(1+\bar{t}) \tag{E.6}$$

$$\lambda_I = \alpha_I(1+\bar{t}) \tag{E.7}$$

Plugging these expressions into the change in the indirect utility, the change in welfare is:

$$\begin{aligned}
d\ln V &= d\ln W - d\ln P \\
&= \lambda_F d\ln w_F - \lambda_F d\ln d_F + \lambda_I d\ln w_I - \alpha_F d\ln w_F - \alpha_I d\ln w_I + d\ln(1+\bar{t}) \\
&= -\lambda_F d\ln d_F + \lambda_F\left(1 - \frac{t_F}{1+\bar{t}}\right) d\ln w_F + \lambda_I\left(1 - \frac{1}{1+\bar{t}}\right) d\ln w_I + d\ln(1+\bar{t}) \\
&= -\lambda_F d\ln d_F - \lambda_F\left(\frac{t_F - \bar{t}}{1+\bar{t}}\right) d\ln w_F + \lambda_I\left(\frac{\bar{t}}{1+\bar{t}}\right) d\ln w_I + d\ln(1+\bar{t}) \\
&= -\lambda_F d\ln d_F - \lambda_F\left(\frac{t_F - \bar{t}}{1+\bar{t}}\right) d\ln w_F + \lambda_I\left(\frac{\bar{t}}{1+\bar{t}}\right) d\ln w_I + \frac{\lambda_F t_F}{1+\bar{t}} d\ln\lambda_F \\
&= -\lambda_F d\ln d_F - \lambda_F\left(\frac{t_F - \bar{t}}{1+\bar{t}}\right) d\ln w_F + \lambda_I\left(\frac{\bar{t}}{1+\bar{t}}\right) d\ln w_I + \frac{\lambda_F t_F}{1+\bar{t}} d\ln\lambda_F \\
&= -\lambda_F d\ln d_F - \lambda_F\left(\frac{t_F - \bar{t}}{1+\bar{t}}\right) d\ln w_F + \frac{\lambda_F t_F}{1+\bar{t}}(d\ln w_F + d\ln\tilde{\lambda}_F) \\
&= -\lambda_F d\ln d_F - \frac{t_F}{1+\bar{t}} d\tilde{\lambda}_F
\end{aligned}$$

where $\tilde{\lambda}_F$ is the total amount of efficiency units, and this is the expression from section 3.6.

E.2 Welfare Decomposition: Full Model

This section derives the formula for the welfare decomposition. For this derivation, I work under one of the following three assumptions: i) the number of workers who decide to live in location n is fixed, ii) the planner maximizes an aggregate welfare function of real income across locations where the initial weights correspond to the initial income levels, or iii) the migration elasticity, $\eta \to \infty$. The idea is to abstract from other market failures, such as congestion forces, to focus on the wedges. I also assume that profits are rebated to households based on their workplace. As in the text, there are three groups of agents: workers denoted by L, commercial floor space owners denoted by Z, and residential floor space owners denoted by H. The latter two groups do not commute.

Indirect utility function: Recall that the indirect utility of agent ω based on the ex-ante inclusive value of commuting costs is:

$$V_{nis\omega} = \frac{w_{is} \cdot \tilde{d}_{ni\omega}^{-1} \cdot \epsilon_{nis\omega} \cdot \exp(\xi_{n\omega}) \cdot (1+\bar{t})}{P^{\alpha} r_{Hn}^{1-\alpha}}, \tag{E.8}$$

where w_{is} is the wage per efficiency unit in location i and sector s, $\epsilon_{nis\omega}$ is an idiosyncratic shock drawn from a nested Fréchet distribution with dispersion parameters θ in the lower nest, and κ in the upper nest, P is the price index that is normalized to 1, and r_{Hn} is the housing price in location n. By the properties of the Fréchet, the total amount of efficiency units $\tilde{d}_{ni}^{-1}\tilde{L}_{nis}$ net of commuting costs provided by location n to location i-sector s is:

$$\tilde{d}_{ni}^{-1}\tilde{L}_{nis} = w_{is}^{-1}\lambda_n\lambda_{ns|n}\lambda_{nis|ns}\bar{y}_n\bar{L}, \tag{E.9}$$

where $\bar{y}_n \equiv \left(\sum_s B_{ns}W_{ns}^{\kappa}\right)^{\frac{1}{\kappa}}$, and $W_{ns} \equiv \left(\sum_i w_{is}^{\theta}\tilde{d}_{ni}^{-\theta}\right)^{\frac{1}{\theta}}$. This expression characterizes the labor supply function.

Market clearing conditions: Under these assumptions, the market clearing conditions are:

$$\sum_n \lambda_n\lambda_{ns|n}\lambda_{nis|ns}\bar{y}_n\bar{L} = \left(\frac{1+\bar{t}}{1+t_{Ls}}\right)\alpha\beta\sum_j\int_{\varphi\in\Omega_{is}}\pi_{is}(\varphi)\left(\bar{y}_j\lambda_j\bar{L} + r_{Zj}\lambda_{jZ}Z + r_{Hj}\lambda_{jH}H\right)d\varphi \tag{E.10a}$$

$$r_{Zi}\lambda_{iZ}Z = \sum_s\left(\frac{1+\bar{t}}{1+t_{Zs}}\right)\alpha(1-\beta)\sum_j\int_{\varphi\in\Omega_{is}}\pi_{is}(\varphi)\left(\bar{y}_j\lambda_j\bar{L} + r_{Zj}\lambda_{jZ}Z + r_{Hj}\lambda_{jH}H\right)d\varphi, \tag{E.10b}$$

$$r_{Hn}\lambda_{nH}H = (1-\alpha)\left(\bar{y}_n\lambda_n\bar{L} + r_{Zn}\lambda_{nZ}Z + r_{Hn}\lambda_{nH}H\right)(1+\bar{t}) \tag{E.10c}$$

where r_{Zn} is the commercial floor space price, r_{Hn} is the residential floor space price, and $\lambda_{nZ}Z$ and $\lambda_{nH}H$ are the supplies of commercial and residential floor space in location n, respectively. Let's also define:

$$\pi_{is} = \int_{\varphi\in\Omega_{is}}\pi_{is}(\varphi)d\varphi$$

Aggregate welfare: The aggregate welfare function is:

$$\bar{U} = \underbrace{\omega_L\sum_n\delta_n\bar{U}_n}_{\bar{U}_L} + \underbrace{\omega_Z\sum_n\delta_{nZ}\bar{U}_{nZ}}_{\bar{U}_Z} + \underbrace{\omega_H\sum_n\delta_{nH}\bar{U}_{nH}}_{\bar{U}_H}, \tag{E.11}$$

where ω and δ are weights based on the initial allocation. I am interested in a shock to commuting costs. Then, by a first-order approximation, the effect of an infinitesimal change in $\vec{\tilde{d}}$ is:

$$d\ln\bar{U} = \frac{\omega_L \bar{U}_L}{\bar{U}} d\ln\bar{U}_L + \frac{\omega_Z \bar{U}_Z}{\bar{U}} d\ln\bar{U}_Z + \frac{\omega_H \bar{U}_H}{\bar{U}} d\ln\bar{U}_H + d\ln(1+\bar{t}), \tag{E.12a}$$

Under homothetic and homogeneous-of-degree-one preferences, and under any of the three assumptions above, the weights are: $\frac{\omega_L \bar{U}_L}{\bar{U}} = \alpha\beta$, $\frac{\omega_Z \bar{U}_Z}{\bar{U}} = \alpha(1-\beta)$, and $\frac{\omega_H \bar{U}_H}{\bar{U}} = 1-\alpha$.[42]

$$\begin{aligned} d\ln\bar{U} = {} & \alpha\beta \sum_n \tilde{\lambda}_{nL} \left(d\ln\bar{y}_n - \alpha d\ln P - (1-\alpha) d\ln r_{Hn}\right) && \text{(E.13a)} \\ & + \alpha(1-\beta) \sum_n \tilde{\lambda}_{nZ} \left(d\ln r_{Zn} - \alpha d\ln P - (1-\alpha) d\ln r_{Hn}\right) && \text{(E.13b)} \\ & + (1-\alpha) \sum_n \tilde{\lambda}_{nH} \left(d\ln r_{Hn} - \alpha d\ln P - (1-\alpha) d\ln r_{Hn}\right) && \text{(E.13c)} \\ & + d\ln(1+\bar{t}), && \text{(E.13d)} \end{aligned}$$

where $\frac{\delta_n \bar{U}_n}{\bar{U}_L} \equiv \tilde{\lambda}_{nL} = \frac{\bar{y}_n \lambda_n}{\sum_{n'} \bar{y}_{n'} \lambda_{n'}}$ is the share of total labor income in location n, and similarly, $\tilde{\lambda}_{nZ}$ is the share of total income of commercial floor space in location n, and $\tilde{\lambda}_{nH}$ is the share of total income of residential floor space in location n. Then, the change in the average income and the price index in each location is:

$$d\ln\bar{y}_n = \sum_{i,s} \lambda_{ns|n} \lambda_{nis|ns} d\ln w_{is} - \sum_{i,s} \lambda_{ns|n} \lambda_{nis|ns} d\ln\tilde{d}_{ni} \tag{E.14a}$$

$$d\ln P = \sum_{i,s} \pi_{is} (\beta d\ln w_{is} + (1-\beta) d\ln r_{Zi}), \tag{E.14b}$$

Rewriting the goods market clearing condition in terms of the income shares, we obtain:

$$\sum_n \tilde{\lambda}_{nL} \lambda_{ns|n} \lambda_{nis|ns} = \left(\frac{1+\bar{t}}{1+t_{Ls}}\right) \pi_{is} \tag{E.15}$$

$$\tilde{\lambda}_{isZ} = \left(\frac{1+\bar{t}}{1+t_{Zs}}\right) \pi_{is} \tag{E.16}$$

$$\tilde{\lambda}_{nH} = (1+\bar{t})(\alpha\beta\tilde{\lambda}_{nL} + \alpha(1-\beta)\tilde{\lambda}_{nZ} + (1-\alpha)\tilde{\lambda}_{nH}), \tag{E.17}$$

where $\tilde{\lambda}_{isZ} = \frac{r_{Zi} \lambda_{isZ}}{\sum_{n,s} r_{Zn} \lambda_{nsZ}}$. Then, from these expressions and after some algebraic manipulation, we get that:

$$\begin{aligned} d\ln\bar{U} = {} & -\alpha\beta \sum_{n,s,i} \left(\tilde{\lambda}_{nL} \lambda_{ns|n} \lambda_{nis|ns}\right) d\ln\tilde{d}_{ni} && \text{(E.18a)} \\ & - \alpha\beta \sum_{n,i,s} \tilde{\lambda}_{nL} \lambda_{ns|n} \lambda_{nis|ns} \left(\frac{t_{Ls} - \bar{t}}{1+\bar{t}}\right) d\ln w_{is} && \text{(E.18b)} \end{aligned}$$

[42] This is shown in section E.3.

$$- \alpha(1-\beta)\sum_{n,s}\tilde{\lambda}_{nsZ}\left(\frac{t_{Zs}-\bar{t}}{1+\bar{t}}\right)d\ln r_{Zn} \tag{E.18c}$$

$$- (1-\alpha)\sum_{n}\tilde{\lambda}_{nH}\left(\frac{-\bar{t}}{1+\bar{t}}\right)d\ln r_{Hn} \tag{E.18d}$$

$$+ d\ln(1+\bar{t}). \tag{E.18e}$$

Change in rebate: From the Government budget constraint,

$$\bar{t} = \sum_{i,n,s}\alpha\beta\tilde{\lambda}_{nL}\lambda_{ns|n}\lambda_{nis|ns}t_{Ls} + \sum_{n,s}\alpha(1-\beta)t_{Zs}\tilde{\lambda}_{nsZ}.$$

Then, we obtain that

$$d\ln(1+\bar{t}) = \frac{d\bar{t}}{1+\bar{t}} \tag{E.19a}$$

$$= \alpha\beta\sum_{n,i,s}\tilde{\lambda}_{nL}\lambda_{ns|n}\lambda_{nis|ns}\left(\frac{t_{Ls}}{1+\bar{t}}\right)(d\ln w_{is} + d\ln\tilde{L}_{nis}) \tag{E.19b}$$

$$+ \alpha(1-\beta)\sum_{n,s}\tilde{\lambda}_{nsZ}\left(\frac{t_{Zs}}{1+\bar{t}}\right)(d\ln r_{Zn} + d\ln\tilde{Z}_{ns}). \tag{E.19c}$$

Normalizing total expenditure, the following condition holds

$$\alpha\beta\sum_{n,i,s}\tilde{\lambda}_{nL}\lambda_{ns|n}\lambda_{nis|ns} + \alpha(1-\beta)\sum_{n,s}\tilde{\lambda}_{nsZ} + (1-\alpha)\sum_{n}\tilde{\lambda}_{nH} = 1, \tag{E.20}$$

which implies that:

$$\alpha\beta\sum_{n,i,s}\tilde{\lambda}_{nL}\lambda_{ns|n}\lambda_{nis|ns}(d\ln w_{is} + d\ln\tilde{L}_{nis}) \tag{E.21a}$$

$$+\alpha(1-\beta)\sum_{n,s}\tilde{\lambda}_{nsZ}(d\ln r_{Zn} + d\ln\tilde{Z}_{ns}) \tag{E.21b}$$

$$+(1-\alpha)\sum_{n}\tilde{\lambda}_{nH}(d\ln r_{Hn} + d\ln\tilde{H}_{n}) = 0 \tag{E.21c}$$

Plugging this expression multiplied by $\bar{t}$ into the aggregate change in welfare, the total change in welfare is:

$$d\ln\bar{U} = -\alpha\beta\sum_{n,s,i}\left(\tilde{\lambda}_{nL}\lambda_{ns|n}\lambda_{nis|ns}\right)d\ln\tilde{d}_{ni} \tag{E.22a}$$

$$+ \alpha\beta\sum_{n,i,s}\tilde{\lambda}_{nL}\lambda_{ns|n}\lambda_{nis|ns}\left(\frac{t_{Ls}-\bar{t}}{1+\bar{t}}\right)d\ln\tilde{L}_{nis} \tag{E.22b}$$

$$+ \alpha(1-\beta)\sum_{n,s}\tilde{\lambda}_{nsZ}\left(\frac{t_{Zs}-\bar{t}}{1+\bar{t}}\right)d\ln\tilde{Z}_{ns} \tag{E.22c}$$

$$+ (1-\alpha)\sum_{n}\tilde{\lambda}_{nH}\left(\frac{-\bar{t}}{1+\bar{t}}\right)d\ln\tilde{H}_{n}. \tag{E.22d}$$

Since residential floor space is fixed, we have that $d\ln\tilde{H}_n = 0$. Rearranging terms and defining

$\tilde{\lambda}_{nisL} = \tilde{\lambda}_{nL}\lambda_{nis|ns}\lambda_{ns|n}$, we obtain:

$$d\ln \bar{U} = -\alpha\beta \sum_{n,s,i} \tilde{\lambda}_{nisL} d\ln \tilde{d}_{ni} \tag{E.23a}$$

$$+ \alpha\beta \sum_{n,i,s} \tilde{\lambda}_{nisL} \left(\frac{t_{Ls} - \bar{t}}{1+\bar{t}}\right) d\ln \tilde{L}_{nis} + \alpha(1-\beta) \sum_{n,s} \tilde{\lambda}_{nsZ} \left(\frac{t_{Zs} - \bar{t}}{1+\bar{t}}\right) d\ln \tilde{Z}_{ns}, \tag{E.23b}$$

which corresponds to the equation in the main text. This equation suggests that the first-order change in welfare depends on two terms: i) the direct effects of changes in commuting costs, and ii) the reallocation of resources towards sectors depending on the relative difference between the wedge and the average in the economy.

E.3 The problem of the social planner

In this section, I find the equilibrium conditions for the problem of the social planner. The idea is to show that if $\eta \to \infty$ there are two important results. First, in the case in which the economy operates under perfect competition, the market allocation coincides with the efficient allocation. Second, the variable $\bar{U}$ is proportional to the aggregate total expenditure or income of the economy, which is the main assumption from the previous section.

There are different groups of workers indexed by g, sectors indexed by s, and a mass of locations $\mathcal{N}$ indexed by n and i. Each group has a utility function $U_g(c_{ng}, h_{ng})$, where c_{ng} represents the average consumption of a composite good in location n and h_{ng} is the average amount of housing in location n. The utility function is homogeneous of degree one and homothetic. In the optimal allocation, workers are indifferent across locations, $\eta \to \infty$, and workers face iceberg commuting costs. The problem of the planner is to maximize the following welfare function:

$$\bar{U} = \lambda_g \cdot U_g$$

subject to:

i) spatial mobility constraints

$$U_{ng}L_{ng} \leq \bar{U}_g \;\; \forall n, g$$

ii) composite and housing feasibility constraints

$$\sum_{n,s} \tau_{ni} Q_{is} \leq Y_{is}(E_{gis}) \;\; \forall i, s$$

$$L_{ng} \cdot c_{ng} \leq C(Q_{n11g}, ..., Q_{nSNg}) \;\; \forall n$$

$$L_{ng} \cdot h_{ng} \leq H_n(\tilde{E}_{nhg}) \;\; \forall n$$

iii) labor supply constraints

$$\tilde{E}_{isg} \leq \sum_{n} \tilde{d}_{ni}^{-1} E_{nisg} \;\; \forall i, s, g \text{ including the sector that produces housing}$$

$$E_g(E_{n11g}, ..., E_{nSNg}) \leq L_{n,g} \;\; \forall n, g$$

iv) non-negativity constraints of commuting flows and labor.

v) Labor Market clearing implies that:

$$\sum_{n,g} L_{n,g} \leq \bar{L}_g$$

where Y is the production function, $C(\cdot)$ is a composite good aggregator across locations and sectors, in my case the nested CES; $E(\cdot)$ is a efficiency units aggregator, in my case the nested Fréchet; and E_{nisg} are efficiency units provided from location n to i, s by group g.[43] The other parameters represent the same variables as in section 4.

The Lagrangian of the planning problem, omitting the non-negative constraints, is:

$$\begin{aligned}
\mathcal{L} = L_g U_g - \sum_{n,g} \omega_{ng} L_{ng} \left(U_g - U_g(c_{ng}, h_{ng})\right) \\
- \sum_{i,s} p^*_{is} \left(\sum_n \tau_{ni} Q_{nis} - Y_{is}(\tilde{E}_{isg}) \right) \\
- \sum_n P^*_n \left(\sum_g L_{ng} c_{ng} - C(Q(Q_{n11g}, ..., Q_{nSNg})) \right) \\
- \sum_{i,s,g} w^*_{isg} \left(\tilde{E}_{isg} - \sum_n \tilde{d}_{ni}^{-1} E_{nisg} \right) \\
- \sum_{n,g} \bar{y}^*_{n,g} \left(E_g(E_{n11g}, ..., E_{nSNg}) - L_{n,g} \right) \\
- \sum_n r^*_n \left(\sum_g L_{ng} h_{ng} - H_n(\tilde{E}_{nhg}) \right) \\
- \sum_g \Psi_g (\sum_n L_{ng} - \bar{L}_g) + ...
\end{aligned}$$

The planner chooses c_{ng}, h_{ng}, Q_{nis}, E_{nisg}, $\tilde{E}_{isg}$, $\tilde{E}_{ihg}$, L_{ng}, and U_g to maximize welfare. I proceed in two parts. First, I show the relationship between $\bar{U}$ and aggregate expenditure, and then, I show that the market allocation coincides with the efficient allocation. Then, I generalized the formula from the previous section using the goods market-clearing condition.

Utility and Total Expenditure

The F.O.C with respect to c_{ng} and h_{ng} is:

$$\omega_{ng} c_{ng} \frac{\partial U_g}{\partial c} \leq P^*_n c_{ng} \quad \forall g$$

$$\omega_{ng} h_{ng} \frac{\partial U_g}{\partial h} \leq r^*_n h_{ng} \quad \forall g$$

Since $U_g(\cdot)$ is homogeneous of degree one, then,

$$L_{ng} \left(P^*_n c_{ng} + r^*_n h_{ng} \right) = L_{ng} \omega_{ng} U_{ng} \tag{E.24}$$

[43] Recall that the efficiency units aggregator is $E_n^{\frac{\kappa}{\kappa-1}} \equiv \sum_s E_{ns}^{\frac{\kappa}{\kappa-1}}$, where $E_{ns}^{\frac{\theta}{\theta-1}} = \sum_i E_{nis}^{\frac{\theta}{\theta-1}}$.

The LHS of equation E.24 is the aggregate expenditure X_{ng} of group g who lives in location n. The F.O.C with respect to U_g is:

$$\sum_n \omega_{ng} L_{ng} = L_g$$

Combining this equation with equation E.24, and the fact that in equilibrium $U_{ng} = U_g$ for all the locations in which $L_{ng} > 0$ yield that:

$$L_g U_g = \sum_n X_{ng}$$

Recall that $\bar{U} = \sum_g L_g U_g$, thus,

$$\bar{U} \propto X$$

where $X \equiv \sum_{n,g} X_{ng}$ is aggregate expenditure. At the aggregate level, total expenditure is equal to total income then in the previous section $\bar{U} = \sum_n \bar{y}_n L_n + q_n \tilde{Z}_n + r_n H_n$, which was the assumption for the theoretical result of the first-order approximation if $\eta \to \infty$.

Efficient Allocation

Now, I show that the market allocation coincides with the efficient allocation if there are no distortions. The F.O.C with respect to other variables is:

$$[Q_{nis}] : P_n^* \frac{\partial C}{\partial Q_{nis}} \leq p_{is}^* \tau_{ni} \tag{E.25a}$$

$$[\tilde{E}_{isg}] : p_{is}^* \frac{\partial Y}{\partial \tilde{E}_{isg}} \leq w_{isg}^* \tag{E.25b}$$

$$[E_{nisg}] : w_{isg}^* \tilde{d}_{ni}^{-1} \leq \bar{y}_{ng} \frac{\partial E_g}{\partial E_{nisg}} \tag{E.25c}$$

$$[\tilde{E}_{nhg}] : r_n^* \frac{\partial H}{\partial \tilde{E}_{nhg}} \leq w_{nhg} \tag{E.25d}$$

$$[L_{ng}] : \bar{y}_n \leq \Psi_g \tag{E.25e}$$

Equations E.25a to E.25d are the same as the utility and profit maximization conditions of the consumer's and firm's problem. In the particular case in which the function $C(\cdot)$ is the nested CES utility function from section 4, $E(\cdot)$ is the nested Frechet, and assuming that $Y(\cdot)$ is homogeneous of degree one, I can rewrite these conditions as:

$$\lambda_{nsg|n} \lambda_{nisg|ns} \bar{y}_{ng}^* L_{ng} = w_{isg}^* \tilde{d}_{ni}^{-1} E_{nisg}$$

$$w_{isg}^* \tilde{E}_{isg} = \beta_{isg} p_{is}^* Y_{is}$$

$$p_{is}^* Y_{is} = \sum_{n,g} \alpha_{ng} \pi_{nis} \bar{y}_{ng}^* L_{ng}$$

$$r_n^* H_n = \sum_{n,g} (1 - \alpha_{ng}) \bar{y}_{ng}^* L_{ng},$$

where

$$\beta_{isg} \equiv \frac{E_{isg} \frac{\partial Y}{\partial E_{isg}}}{Y_{is}}$$

$$\alpha_{ng} \equiv \frac{c_{ng} \frac{\partial U_g}{\partial c_{ng}}}{U_{ng}}$$

These are the same conditions as the market allocation from the previous section. Then, the market allocation is efficient in the case in which there are no wedges.

We can generalize the welfare decomposition from the previous section for different groups of labor under the assumptions where the utility and production function is homogeneous of degree one. In particular, we can rewrite the change in $\bar{U}$ as:

$$d\ln \bar{U} = \underbrace{-\sum_{n,i,s,g} \alpha_{ng}\beta_{isg}\left(\tilde{\lambda}_{ng}\lambda_{nsg|n}\lambda_{nisg|ns}\right) \cdot d\ln \tilde{d}_{ni}}_{\text{“Pure” effect commuting costs}} \tag{E.26a}$$

$$+ \underbrace{\left(\sum_{n,i,s,g} \alpha_{ng}\beta_{isg}\tilde{\lambda}_{ng}\lambda_{nsg|n}\lambda_{nisg|ns}\left(\frac{t_{isg}-\bar{t}}{1+\bar{t}}\right) d\ln \tilde{L}_{nisg}\right)}_{\text{Allocative efficiency}} \tag{E.26b}$$

This result is similar to the one obtained by Baqaee and Farhi (2020) in GE models. However, this expression is in the context of an urban model in which workers face i) commuting costs, and ii) are indifferent to live across locations within the city.