The Gains from Foreign Multinationals in an Economy with Distortions

Isabela Manelici, Mauricio Ulate, Jose P. Vasquez and Román D. Zárate*

March 2025

Link to the most up-to-date version†

Abstract: We study the local and aggregate welfare effects of foreign multinational enterprises (MNEs) in an economy with multiple distortions. First, we leverage rich microdata on *all* establishments in Mexico from 1994 to 2019, including domestic formal and informal ones, as well as foreign MNEs within or outside the maquila program. We construct *direct* measures of establishment-level distortions such as the burden of crime and corruption, labor taxes, and production subsidies. We find that, on average, domestic formal establishments face more distortions than informal ones, while foreign maquila MNEs are the least distorted among formal establishments. Distortions also vary significantly across commuting zones (CZs) in Mexico. Second, using an instrumental variable strategy, we show that an increase of one MNE worker per capita in a CZ raises domestic employment by 0.33 workers per capita. This growth mainly occurs in formal establishments, reducing the share of informal employment within local employment. Only non-maquila MNEs have positive effects on the local domestic economy. Third, we develop a quantitative model highlighting the ex-ante ambiguous role distortions can play in amplifying or attenuating the welfare effects of foreign MNEs. We use the IV estimates and the initial distortion measures to calibrate the model and quantify the GE effects of foreign MNEs. Simulating a 10% productivity shock of MNEs, we find that their expansion in Mexico (1994–2019) led to 6% aggregate welfare gains, with a quarter of the gains from reallocation effects and a fifth from local productivity spillovers.

JEL Codes: F23, F14, F16, D61, O17, J46, E26.

*Manelici: London School of Economics and CEPR (i.manelici@lse.ac.uk). Ulate: Federal Reserve Bank of San Francisco (mauricio.ulate@sf.frb.org). Vasquez: London School of Economics and CEPR (j.p.vasquez@lse.ac.uk). Zárate: World Bank, Development Research Group (rzaratevasquez@worldbank.org).

†We thank the STEG-CEPR programme, the Centre of Economic Performance (CEP/LSE), and the World Bank RSB for financial support. Diego Alvarez, Nicolás Fajardo, Maria del Mar Gómez, José Ignacio González, David Sosa and Alonso Venegas provided excellent research assistance. We are also grateful to Matías Busso, Juan Blyde, and Darío Romero for sharing the commuting zone data for Mexico, and to the National Institute of Statistics and Geography of Mexico (INEGI in Spanish), especially Dr. Natalia Volkow, for their hospitality and valuable institutional knowledge. The views expressed herein represent the authors' views and do not necessarily reflect those of the Federal Reserve Bank of San Francisco or the World Bank Group.

1 Introduction

A common policy prescription for catch-up development and growth in low- and middle-income countries (LMICs) is to attract foreign multinational enterprises (MNEs). MNEs often receive favorable treatment compared to domestic firms, including tax holidays, tariff exemptions, and streamlined regulations. The most common theoretical justification for departing from policy neutrality to favor MNEs is their potential to generate productivity spillovers in the host economy (Harrison and Rodríguez-Clare, 2010). Meanwhile, LMIC economies are characterized by multiple distortions (Hsieh and Klenow, 2009; Atkin and Donaldson, 2022), the severity of which varies both within and across locations, resulting in allocative inefficiency. Hence, a complete study of the welfare gains from MNEs in LMICs must account not only for their potential productivity spillovers but also for their impact on the allocation of resources across differently distorted establishments and locations. Such a study presents challenges, including the need for credible measures of distortions, comprehensive data on *all* establishments (including informal and non-manufacturing), reduced-form evidence of the reallocative effects of MNEs, and a general equilibrium framework to evaluate their aggregate welfare effects. We address these in four steps.

First, we build a unique collection of microdata from Mexico spanning 1994 and 2019, which includes (i) direct measures of initial establishment-level distortions; (ii) panel data on commuting-zone (CZ)-level foreign MNE employment, disaggregated by origin country and *maquiladora* status (maquila for short); and (iii) panel data on the CZ-level size of the domestic formal and informal economies. We then describe the four establishment types in our analysis (domestic formal, domestic informal, foreign MNE maquila, and foreign MNE non-maquila) and document the significant heterogeneity in distortions both within and across CZs in Mexico. Second, we estimate the effects of changes in the importance of foreign MNE employment within a CZ on the local domestic economy and the reallocation of resources between formal and informal domestic establishments. To provide plausibly causal estimates, we employ an instrumental variable (IV) strategy that predicts the expansion of foreign MNEs in a CZ based on historical spatial clustering by the MNEs' country of origin. Third, we develop a general equilibrium model with heterogeneous firms that incorporates the endogenous decision of domestic firms to operate (in)formally, based on their productivity, and the distortions and opportunities associated with each formality choice in their location. The model also includes MNEs whose expansions may differentially benefit formal and informal establishments, for example, through local productivity spillovers and demand linkages. This model helps us interpret the reduced-form effects on the local domestic economy, and enables a counterfactual analysis of the welfare effects of foreign MNEs in the absence of the observed initial distortions. Finally, the fourth step combines the work from previous steps to quantify the welfare gains from foreign MNEs in Mexico and highlight the gains from MNEs' resource reallocation across formal and informal establishments, relative to the gains from productivity spillovers.

Mexico provides a particularly well-suited context for this research. Following the implementation of the North American Free Trade Agreement (NAFTA) in 1994, the country began lib-

eralizing foreign direct investment (FDI) and trade. Over the subsequent twenty-five years, Mexico emerged as the tenth-largest recipient of FDI from various origin countries and the twelfth-largest exporter globally. Meanwhile, like most LMICs, Mexico remains severely afflicted by various distortions, leading to significant resource misallocation and reduced aggregate productivity (Hanson, 2010; Hsieh and Klenow, 2014; Levy, 2018; Misch and Saborowski, 2020; Bloom et al., 2025). In this project, we investigate how the presence of these distortions has affected the gains from foreign MNEs in Mexico, using a benchmark where potential productivity spillovers are the sole market failure justifying MNE-favoring policies. To this end, Mexico offers a uniquely valuable opportunity, as it allows us to track *all* establishments in the economy since 1994, including small informal establishments outside the manufacturing sector. Despite representing a large share of the overall economy in LMICs, these establishments are not captured in the administrative datasets that have been used to study the impact of MNEs on LMICs thus far. Moreover, several representative surveys enable us to construct direct measures of establishment-level "taxes" stemming from the weak rule of law, such as those imposed by crime and corruption. As Atkin and Khandelwal (2020) argue, this type of distortions are key to understanding the impact of globalization on developing countries but are typically either unavailable or challenging to extract from standard administrative datasets. Last, Mexico serves as a prominent example of an economy that embraced the maquila model of FDI as an early development strategy and a stepping-stone toward higher value-added investment.[1] Historically, the typical foreign MNE plant in the maquila program imported most of its inputs and did so tariff-free, processed them in a labor-intensive manner, and then exported the finished goods. Given notable differences between maquila and non-maquila foreign MNEs, both in terms of the distortions they face and their potential impact on the Mexican economy, we distinguish between them in our empirical analysis and quantification.

In the first step, we compile a uniquely rich microdata collection that enables a comprehensive study of the effects of foreign MNEs in a distortion-ridden economy. The first of two data pillars is the Economic Census, which tracks the outcomes of four types of establishments between 1994 and 2019: domestically-owned establishments, categorized by their choice to operate formally or informally, and foreign MNEs, distinguished by whether they operate in the maquila program or not. Establishments are classified as informal if they do not report paying their workers' legally required social security contributions. Foreign MNEs are establishments with at least 50% foreign ownership. Leveraging this data, we show that in 2019, informal establishments employed an average of 2.4 workers, compared to 23.6 workers in domestic formal establishments. While small, informal establishments are in the millions. Since 1994, the share of employment in informal establishments has risen nationwide from 31.7% to 36.8%, with Southern CZs experiencing informality rates above 80%. We then show that foreign MNEs have grown in importance since 1994, accounting for almost 10% of employment by 2019, with significant variation in their employment growth across CZs. Additionally, the composition of foreign MNE employment by

[1]Even though they are referred to by different names (e.g., export processing zones), incentives to attract foreign MNEs under the assembly-for-export model are widespread across LMICs.

country of origin has diversified, moving away from U.S. dominance. Last, maquila foreign MNEs employ over three times more workers on average than non-maquila MNEs. Despite their greater size, import and export intensity—traits often linked to higher productivity—maquila MNEs are less productive than non-maquila MNEs.

The second data pillar is a database we built containing direct measures of initial distortions in Mexico, with data at the CZ and establishment type level (domestic formal or informal, foreign MNE in the maquila program or not). The first category of distortions in this database includes taxes and subsidies, which can vary due to either government preferences for policy non-neutrality, differences in enforcement capacity across CZs or compliance across establishment types. Specifically, we consider value-added taxes, government subsidies, and output taxes (such as those faced by exporters) as output distortions, expressing them as a share of sales. Labor allocations are distorted by compliance with social security contributions (and other labor costs), which we express as a share of an establishment's wage bill. Informal establishments evade labor costs but face a size-dependent probability of being caught and penalized for non-compliance. The second category of distortions in our database pertains to weak institutions and the rule of law. We use two establishment-level representative surveys that track the burden of regulation and bureaucracy, as well as the burden of crime and corruption. These distortions are treated as output "taxes" and are expressed as a share of sales.

Using this second database, we establish three main facts about the severity of distortions in Mexico. First, our direct measurement of distortions allows us to document the severity of each distortion separately. We observe that the ranking of establishment types by the burden of distortions can be distortion-specific. For instance, informal establishments are the most "taxed" by crime and corruption; however, as they evade statutory output taxes and social security contributions, they appear as the least distorted in these dimensions. Second, when considering all distortions together, we observe significant variation in total distortions across CZs for the same establishment type, even after accounting for differences in establishment size and four-digit industry. This spatial heterogeneity in distortions underscores the importance of the multi-location aspect of our theoretical framework. Third and last, informal establishments are the least distorted on average,[2] followed by foreign maquila MNEs. Foreign non-maquila MNEs and domestic formal establishments rank closely as the most distorted establishment types.

In the second step, we examine the local effects of foreign MNEs in Mexico by leveraging variation in the importance of foreign MNE employment across time and CZs. We evaluate both their impact on the overall size of the local domestic economy and the reallocation of resources between formal and informal domestic establishments. This reallocation is of interest for two reasons. First, as outlined above, formal and informal establishments face distinct distortions, and MNE-driven reallocations between these establishment types have implications for allocative efficiency. Second, the unique characteristics of foreign MNEs—such as their skill intensity, product type and

[2]This finding aligns with Atkin and Donaldson (2022), who find a global pattern of greater distortions in the formal sector. Busso et al. (2012) show that lower distortions (proxied by lower TFPR, following Hsieh and Klenow, 2009) are often associated with establishments operating informally and evading payroll taxes in Mexico.

quality, and the responsible sourcing standards they impose in their supply chains—suggest that MNEs may affect the formal and informal sectors differently.

The key explanatory variable measures the change in employment of foreign MNEs in a CZ between two Economic Censuses, divided by the initial population in that CZ. An OLS estimation raises two primary concerns: (i) omitted variable bias if there are unobserved factors correlated with both the future performance of the domestic economy in a CZ and changes in the importance of foreign MNEs in that CZ (e.g., a new government commitment to invest in productive infrastructure); or (ii) attenuation bias if the explanatory variable suffers from measurement error. To address these concerns with OLS, we use an instrumental variable (IV) approach. Specifically, to construct an IV for the growth in foreign MNE employment from a particular country of origin in a CZ, we leverage (i) the historical spatial clustering of foreign MNE employment from that country of origin; and (ii) the growth of foreign MNE employment from that country of origin in all Mexican CZs *except* the CZ of interest.[3] The final formula for the IV sums across countries of origin to capture the supply-push component of the foreign MNE employment growth in a CZ. We find this IV to be a statistically significant predictor of foreign MNE employment growth. The IV satisfies the exclusion restriction if it affects the outcomes of a specific CZ solely through its impact on the change in foreign MNE employment in that CZ. To help in this direction, the IV does not capture CZ-specific changes in market conditions or other factors relevant to firm performance. Moreover, to alleviate remaining concerns with identification, we also use a rich set of CZ-level controls.

The IV estimates suggest that an increased exposure to foreign MNEs leads to expansion of the domestic economy. We estimate that one additional foreign MNE worker per capita in a CZ increases domestic establishment employment by 0.33 workers and sales by about $35,700 (2018 USD) per capita.[4] Value-added, assets, and wage bills also increase by around $15,230, $6,230, and $3,530 per capita (2018 USD), respectively. Additionally, we find that only non-maquila foreign MNEs have positive effects on the local domestic economy, consistent with their higher productivity and greater integration into Mexico's economy. Next, we separately analyze the effects of increased foreign MNE exposure on the formal and informal segments of the local economy. We learn that the formal sector is the main beneficiary of the local positive impact of foreign MNEs and that most of this impact comes from exposure to non-maquila foreign MNEs.

In the third step, we develop the theoretical framework for our general equilibrium analysis. Our model combines insights from the informality literature with recent advances in spatial economics. It features multiple locations connected by trade and migration linkages and three establishment types: foreign MNEs (to be further divided in the model quantification into those participating in the maquila program and those not), domestic formal, and domestic informal. Domestic establishments in each location are heterogeneous and endogenously decide whether to

[3] A similar IV was recently used by Setzler and Tintelnot (2021) to study the local effects of foreign MNEs in the United States. Its rationale mirrors that of a classic IV from the migration literature, which instruments immigrant inflows from a specific source country to a location by using the historical settlement patterns of immigrants from that source country and the total number of newly arriving immigrants from that country (Card, 2001).

[4] This corresponds to an extreme change in the exposure variable of 1. The difference between the 10th and 90th percentiles of the exposure variable during the period is 0.07, implying a yearly increase in sales per capita of 120 USD.

operate in the formal or informal sector based on their individual productivity, the productivity of their location, distortions they face given their location and formality choice, and general equilibrium forces. All establishment types face location- and establishment-type-specific "tax-like" input and output distortions that affect the allocation of resources. Informal establishments additionally face a size-dependent distortion.[5]

In our model, foreign MNEs repatriate their profits abroad and interact with the local economy through two MNE-specific channels: (i) Local productivity spillovers: MNEs can increase the productivity of both formal and informal local domestic establishments, albeit to different extents. These spillovers are modeled as external economies of scale resulting from the local size of MNE employment. (ii) Input demand linkages: MNEs source inputs exclusively from formal domestic establishments (reflecting MNEs' preference for supplier compliance with VAT and labor taxes and regulations, as in Gadenne et al., 2024; Alfaro-Ureña et al., 2022a). We show that the contribution of the allocative efficiency margin to the aggregate gains from foreign MNEs hinges on whether MNEs reallocate workers toward the more distorted segments of the economy. This reallocation depends on the local productivity spillovers from foreign MNEs and their input demand linkages.

In the fourth and last step, we calibrate and quantify the model. We simulate a productivity shock to foreign MNEs that aligns, through the lens of the model, with the observed expansion of MNE employment across CZs in Mexico during the study period. We then calibrate the two spillovers elasticities to match our plausibly causal reduced-form estimates for the effects of MNE expansions on total local domestic employment and the reallocation between formal and informal employment. We calibrate all other elasticities using their estimated values in the literature. We infer an elasticity of formal and informal establishments' productivity to foreign MNE employment equal to 0.12 and 0.025, respectively. With the calibrated model in hand, we conduct counterfactual comparisons of the Mexican economy with and without distortions or spillovers, and quantify the aggregate welfare changes under these scenarios.

Our quantification implies that the expansions of foreign MNEs increased aggregate welfare in Mexico by 6.3%. In the absence of differential distortions and productivity spillovers, this effect would have been 4.1%, indicating that differential distortions and spillovers together account for roughly one-third of the total gains. Existing distortions amplify the welfare effects of MNE expansions by driving a reallocation of resources toward the formal sector, which is the most constrained by the burden of distortions. Notably, the presence of distortions contributes slightly more to the overall amplification of the aggregate welfare gains than spillovers. This effect is particularly pronounced in locations that did not experience the shock directly, as distortions play a more significant role in these regions, and workers also reallocate to the formal sector there due to trade linkages with other CZs. By contrast, spillovers have a more prominent role in CZs directly affected by the shock, as their effects tend to be more localized.

[5]The size-dependent distortion represents an additional cost of informality. Its micro-foundation captures the higher likelihood of detection by authorities as informal establishments grow larger (Ulyssea, 2018).

Related literature This article builds upon and contributes to three key strands of literature. First, the longstanding interest in the developmental potential of FDI—primarily through the attraction of foreign MNE affiliates—has generated an extensive body of (mostly empirical) research evaluating its impacts on recipient LMIC economies (for a comprehensive VoxDev literature review, see Alviarez et al., 2025).[6] Most of this literature has focused on productivity spillovers from foreign MNEs to domestic firms as the primary market failure justifying policy interventions in favor of MNEs (Harrison and Rodríguez-Clare, 2010).[7] More recent exceptions examine instead the effects of MNEs in economies where inefficiencies stem from product market power (Asturias et al., 2022), frictions in access to capital (Bau and Matray, 2022), labor market power (Alfaro-Ureña et al., 2021; Setzler and Tintelnot, 2021; Méndez and Van Patten, 2022) or distortive firm-specific policies (such as "responsible sourcing" policies by MNEs, as in Alfaro-Ureña et al., 2022a).

We contribute to this body of work in three ways. First, to examine the local and aggregate effects of foreign MNEs in an economy with distortions, we leverage microdata on nearly all establishments in a large LMIC economy, Mexico, including informal establishments and those outside the manufacturing sector.[8] To our knowledge, prior research on foreign MNEs has focused exclusively on their effects on formal establishments (often further restricted to the manufacturing sector), even in contexts where informal or non-manufacturing establishments employ significant shares of workers.[9] Second, we focus on a set of empirically relevant output and input distortions that are directly observable across establishments and locations in Mexico, rather than inferred through a model. While this set is not exhaustive, it represents an important first step in studying the effects of foreign MNEs in economies where, in practice, *multiple* distortions operate simultaneously. Furthermore, the direct measurement of each distortion enables us to identify which distortions have the greatest impact on the welfare effects of foreign MNEs, making them important policy targets. Third, we develop a general equilibrium framework in which foreign MNEs can affect the welfare of recipient economies through both local productivity spillovers and the reallocation of resources across segments of the economy that differ in their initial output and input

[6] There is a smaller but complementary literature interested in the welfare gains from multinational production, in which Northern MNEs with superior technologies benefit the Southern countries that receive their affiliates and their accompanying technologies (Ramondo and Rodríguez-Clare, 2013; Tintelnot, 2017; Arkolakis et al., 2018).

[7] Prior research has either produced empirical evidence suggestive of productivity spillovers (e.g., Javorcik, 2004; Poole, 2013; Alfaro-Ureña et al., 2022b; Abebe et al., 2022) or assumed them as the driving force behind productivity growth among domestic firms located near foreign MNEs (Setzler and Tintelnot, 2021). To our knowledge, none of these studies quantifies the welfare implications of foreign MNEs in a setting with productivity spillovers.

[8] Prior research on foreign MNEs in Mexico has focused on the link between the maquila sector, the skill-premium and incentives to acquire education (Feenstra and Hanson, 1997; Verhoogen, 2008; Atkin, 2016). Atkin et al. (2018) find gains from foreign MNEs in retail, mainly driven by a lower cost of living.

[9] Sharma and Cardenas (2018) is a notable exception: using state-level panel data from Mexico between 2005 and 2015, the paper shows that FDI inflows are not associated with changes in employment in the informal sector. McCaig et al. (2022) find that in Vietnam, a country with high informality, industries most exposed to sudden and large U.S. tariff reductions experienced faster formal employment growth, driven by the entry of foreign MNE affiliates. This paper does not specifically examine the relationship between foreign MNE employment growth and informality, nor does it provide a general equilibrium framework to assess the welfare implications of the foreign MNE employment growth. Rothenberg et al. (2024) evaluate Indonesia's Integrated Economic Development Zone (KAPET) program, which offered tax breaks to firms (not limited to MNEs) operating in designated districts. Using a spatial model with formal and informal sectors, based on Zárate (2022), the authors show that the program's overall welfare effects were small.

distortions (in addition to other neoclassical channels). This framework enables us to highlight the role of allocative efficiency in the welfare gains, compared to those from productivity spillovers.

Second, both pioneering and more recent work highlights how the effects of international trade are influenced by the weak institutions, market failures, and firm-level distortions prevalent in LMICs (for reviews, see Krueger, 1984; Atkin and Khandelwal, 2020; Atkin et al., 2022; Atkin and Donaldson, 2022). Closest to our work are papers that focus on the distortions associated with informality, where informal firms pay little or no taxes and do not comply with regulations, particularly those related to labor.[10] On the one hand, papers such as Dix-Carneiro and Kovak (2019), Blyde et al. (2020), Ulyssea and Ponczek (2022), and Cisneros-Acevedo (2022) show that exposure to foreign competition has increased informality among low-skill workers in Brazil, Mexico, and Peru, with informality acting as a buffer against unemployment. On the other hand, Aleman-Castilla (2006) and McCaig and Pavcnik (2018) find that greater access to foreign markets in Mexico and Vietnam has reallocated workers from informal firms to more productive formal firms.

We differentiate ourselves from this work in three ways. First, we analyze the effects of FDI, specifically those of foreign MNE affiliates. Trade integration and FDI are distinct forms of globalization, each with unique mechanisms, impacts, and policy implications. Second, while we also differentiate domestic establishments based on their decision to operate formally or informally, we go beyond the conventional focus on labor market regulation compliance. Namely, we construct additional direct measures of distortions that unevenly affect formal and informal establishments, including those stemming from crime, corruption, and varying compliance with other non-labor taxes. Third, our paper is related to Dix-Carneiro et al. (2024), whose model features heterogeneous firm dynamics and an endogenous (in)formality choice in response to economic changes, and examines how distortions associated with informality affect the gains from trade. While our model is static, it incorporates foreign MNEs and their distinctive effects on the domestic economy (e.g., through productivity spillovers), along with an endogenous (in)formality choice for heterogeneous domestic establishments. We can then study how MNEs affect the domestic economy through the informality margin and how this, in turn, influences the aggregate welfare impact of MNEs. Moreover, our model adds two important dimensions for a study of the interplay between foreign MNEs, a distorted recipient economy, and allocative efficiency: (i) location-specific distortions within a country, where locations are linked through trade and migration flows, and (ii) uneven output and input distortions across establishments within each location, where establishments are categorized by ownership (domestic or foreign) and formality status.

Third, a vast literature has examined the prevalence of distortions in LMICs and their implications for firm-level and aggregate TFP (e.g., Hsieh and Klenow, 2009; Hopenhayn, 2014; Baqaee

[10] Alternatively, Asturias et al. (2019), Arkolakis et al. (2019), and Dhingra and Morrow (2019) study the gains from trade in settings with variable markups, while Lanteri et al. (2023) examines a setting with capital-reallocation frictions. Berthouy et al. (2021) and Bai et al. (2024) show that in second-best environments, trade can lead to welfare losses because the income losses from distortions may outweigh the gains from trade; they then use model-inferred firm-level distortions to provide empirical examples. Zárate (2022) investigates the effects of another form of market integration—transit improvements—on informality rates in Mexico City, highlighting how infrastructure investments can drive aggregate economic gains by improving the allocative efficiency of a distorted economy.

and Farhi, 2020).[11] This literature has highlighted Mexico as a country subject to widespread distortions (e.g., Hanson, 2010; Kehoe and Ruhl, 2010; Hsieh and Klenow, 2014; Hsieh and Olken, 2014; Levy, 2018; Alvarez and Ruane, 2024; Bloom et al., 2025). The standard approach in this literature is indirect: it assumes a production technology and market structure, then uses data to infer wedges in the first-order conditions that characterize an efficient allocation. These wedges are interpreted as reflecting distortions to efficient allocations. However, this approach carries several risks, such as those arising from measurement error and model misspecification (Restuccia and Rogerson, 2017; Bils et al., 2021; Buera et al., 2023). Instead, we use direct measures of various output and input distortions, which can be linked to policy levers and used to assess the burden of specific distortions or their combinations. Our approach, however, cannot account for all possible distortions in Mexico, but only those for which direct measures are available.[12]

The remainder of the paper is organized as follows. Section 2 describes the data and context on foreign MNEs and distortions in Mexico. Section 3 presents the empirical evidence on the effects of increased exposure to foreign MNEs in a commuting zone on the domestic economy. Section 4 introduces the theoretical framework used to analyze the welfare implications of foreign MNEs in an economy with distortions, along with the steps taken to map the theory to the data. Section 5 outlines the model estimation and calibration procedure. Specifically, we consider the counterfactual welfare change in Mexico between 1994 and 2019 in the absence of foreign MNEs, as well as the welfare effect of foreign MNEs when all distortions, except for the productivity spillovers, are assumed away. Section 7 concludes.

2 Data and context

2.1 Establishments in Mexico

Data on establishments in Mexico The primary data source on establishments is the Economic Census collected by the National Institute of Statistics and Geography (INEGI in Spanish) every five years since 1994. The Economic Census is an almost complete enumeration of fixed establishments in Mexico.[13] Critically for our analysis, the Economic Census includes establishments that are not registered or, even if registered, are not compliant with government regulation. Our baseline definition of an informal establishment labels as informal any establishment that does not comply with the legal requirement of paying social security contributions for its workers (as in

[11] In ongoing work, Atkin et al. (2024) combine rich administrative data from Chile with a general equilibrium model of heterogeneous agents in a distorted economy (building on Baqaee and Farhi, 2020). Their aim is to study the incidence of distortions across different groups of individuals in Chile (e.g., categorized by income, education or gender).

[12] Misch and Saborowski (2020) use establishment-level data from the 2013 Economic Census in Mexico and a model-based approach, as proposed by Hsieh and Klenow (2009), to document significant disparities in resource misallocation at the state level, even after controlling for the composition of industries across states. The authors then run regressions at the industry-state level to explain the large variation in resource misallocation, finding compelling evidence that misallocation increases with observable distortions, such as the prevalence of labor informality, crime, and corruption.

[13] For manufacturing, commerce and non-financial services, the Economic Census has a reduced coverage in more isolated rural locations. Businesses that do not have a permanent location and without fixed installations – such as street vendors – are not covered by the Economic Census (irrespective of their industry).

Levy, 2008; Ulyssea, 2020), irrespective of whether the establishment is registered or not.

The Economic Census microdata allows us to build an unbalanced panel of all establishments in Mexico between 1994 and 2019.[14] This panel tracks typical balance sheet variables such as total sales, number of workers, wage bill, value added, and assets (our proxy for capital). The Economic Census also tracks the establishment-level employment and wage bill by worker type. In particular, salaried workers are split into administrative workers (who we label high-skill or white-collar) and operative workers (who we label low-skill or blue-collar).[15] We also observe the municipality of each establishment, which we then map to a commuting zone – abbreviated CZ hereafter. Mexico will therefore be divided into 455 CZs.[16]

Last and crucially for this paper, we also observe the establishments with foreign ownership. We refer to an establishment as a foreign MNE if it is at least 50% foreign-owned in a Census year. Additionally, we track the country of origin of an establishment's foreign capital (which is a key input for our leading instrumental variable). In the rare instances where we observe more than one country of origin, we keep the country with the largest share of ownership. Appendix A includes further details on the data, sample build, and summary statistics.

Other data We use the publicly available Population Census microdata from INEGI to build 1990-2020 panel data for the population of each municipality. We divided the population proportionally into five categories: in the labor force, not in the labor force, informal workers, employed, and unemployed using shares from the National Survey of Occupation and Employment (ENOE) and its predecessor, the National Survey of Employment (ENE). We also use data from the "Integrated Public Use Microdata Series" (IPUMS) and Comtrade con construct CZ-level controls in our empirical regressions. Appendix A.3 presents additional details.[17]

Differences between foreign MNEs and domestic establishments Table A1 presents summary statistics on the sample of establishments in our analysis, separated into three groups: foreign MNEs, domestic formal establishments, and domestic informal establishments.[18] Table A1 shows the average establishment characteristics by establishment type at the beginning and end of our sample period (1994 and 2019). Specifically, we look into the total number of workers and sales. Across all characteristics, domestic informal establishments perform about an order of magnitude

[14]The Economic Census gathers data about the previous calendar year. We refer to years by the Economic Census year (e.g., 1994) but note that the data corresponds to the prior calendar year (e.g., 1993). The Economic Census also records the industry classification of each establishment. So far, our analysis abstracts from industry classifications.

[15]Workers are classified as either salaried or non-salaried. The salaried group is further divided into operative workers, who operate machinery for manufacturing goods, and administrative workers, who perform general office tasks, accounting functions, and executive roles. Non-salaried workers include owners, relatives, active partners, and workers who do not receive a fixed salary. In addition, this group also includes interns, individuals working for tips, those engaged in social service, participating in training and development programs, and volunteers.

[16]We start with the 781 CZs of Blyde et al. (2023). We then join CZs that have zero high or low-skilled formal or informal workers to the closest neighbor. We end with 455 CZs in our estimation sample.

[17]We also describe our concordance tables for changing commuting zones and industry classifications across time. Between 1994 and 2019, Mexico both created new municipalities and merged some municipalities, making it necessary to create time-consistent CZs. Moreover, the 1994 Economic Census industry classification differs from the NAICS classification used in the subsequent Censuses.

[18]The further separation of foreign MNEs into maquila and non-maquila is done in Table A3.

more poorly than domestic formal establishments, which in turn perform about an order of magnitude more poorly than foreign MNEs. For instance, in 2019, the average domestic informal (formal) establishment employed 2.4 (23.6) workers, while foreign MNEs employed 301.2 workers. While small, there were almost two million informal establishments in 1994, and the national share of employment in informal establishments increased from 31.70% to 36.83% (see Table A2). By 2019, domestic informal establishments account for a non-negligible share of sales (9.61%) and wage bill (9%). Strikingly, they also employ more than 80% of workers in some CZs in the Southern half of Mexico (see Figure 1).

Table A2 also shows the importance of foreign MNEs in the economy-wide totals, again in 1994 and 2019. While in 2019, they represent only 0.2% of the number of establishments, they account for 9.99% of employment, 18.39% of sales, and 25.72% of the wage bill. Moreover, foreign MNEs have also grown in importance at the national level between 1994 and 2019. Figure 2 presents the percentage of CZ-level employment in foreign MNEs at the beginning and the end of our sample. We can also see that while distance to the United States border remains a good predictor of the importance of foreign MNEs in CZ-level employment, between 1994 and 2019, foreign MNEs have also expanded into Central and Southern Mexico.

Spatial clustering of foreign MNEs by country of origin We now turn our interest to the country of origin of foreign MNEs in Mexico. Panel (a) in Figure C1 shows the changing composition of foreign MNE employment by the country of origin of the foreign capital. We focus on Mexico's eight largest investor countries: the United States, Japan, Germany, Netherlands, Great Britain, Canada, France, and Switzerland (plus the rest of the world labeled "other"). While MNEs from the United States remain dominant, Mexico has also been able to increasingly diversify the origin countries of its foreign MNEs. In panel (b) of the same figure, we normalize to 1 the nation-wide employment in 1994 in establishments owned by Mexico's eight largest investor countries. Each line reveals the change since 1994 of foreign MNE employment from these largest investor countries. Most of the growth in employment in foreign MNEs since 1994 has occurred in establishments owned by British, German, Canadian, Japanese, and French investors.

Figure C2 showcases the spatial clustering by country of origin of foreign MNEs, measured as the employment in foreign MNEs from a given origin country in a given CZ as a share of the total employment in foreign MNEs in that CZ. MNEs from Canada are more likely to favor Northern Mexico, with particular preferences for Baja California Sur and Coahuila. While MNEs from Japan rarely targets East, Southeast and Southwest Mexico, MNEs from Western Europe are particularly well represented in those regions. While MNEs from the United States are prevalent across Mexico, it is particularly dominant in areas such as the Yucatán Peninsula and Chiapas. We will leverage this spatial clustering by country of origin in our identification strategy for the effects of foreign MNEs on the domestic economy.

Maquila vs. non-maquila foreign MNEs The maquila status of a foreign MNE is associated with duty-free and tariff-free operations. Maquila MNEs tend to take raw materials and assemble, manufacture, or process them and export the finished products. The separation between maquila

and non-maquila MNEs is meant to acknowledge both the different tax distortions faced by these MNEs and the plausibly different effects they trigger on the domestic economy.

Table A4 presents summary statistics separately for foreign MNEs in Mexico in 2019 with or without income from maquila activities. We learn that while non-maquila MNEs are 81% of the number of foreign MNE establishments, they only employ 55% of all workers in foreign MNEs. Relative to non-maquila foreign MNEs, we find that maquila foreign MNEs are, on average, substantially larger employers (employing, on average, 702 workers compared to 204 in non-maquilas). Moreover, maquila MNEs use significantly less capital per worker and generate less sales per worker. Last, maquila MNEs are more import-reliant and export-oriented than non-maquila MNEs.[19]

2.2 Distortions in Mexico

One of our key contributions is to build a comprehensive dataset that directly estimates the level of a broad set of output and input distortions for the initial year (1994) across establishment types and CZs in Mexico. In particular, we focus on two types of distortions: Those related to taxes and subsidies and those related to weak institutions and the rule of law (Atkin and Khandelwal, 2020). The source data for all these distortions is Mexico-specific and draws either from the Economic Census directly or from national representative surveys.

The first category of distortions we measure stems from taxes and subsidies. For distortions affecting output, we calculate the share of sales accounted for by value-added taxes, government subsidies, and other output taxes, such as tariffs imposed on exporters. For those impacting labor inputs, we assess social security contributions made on behalf of workers and other labor-related costs, including severance payments. The second category of distortions we measure relates to weak institutions and the rule of law. These distortions include the burden of regulation and bureaucracy, which measures the costs of bureaucratic processes. They also cover the burden of crime and corruption, accounting for losses associated with crime and bribes paid to public officials.

An alternative approach to building these various measures of distortions ourselves would be to use indirectly inferred measures of distortions à la Hsieh and Klenow (2009). This indirect approach infers misallocation from measured gaps in average products. This approach has obvious benefits, including the parsimonious set of variables needed to implement it (i.e., typical balance sheet variables) and the comparability of such distortion measures with existing research. There are drawbacks to this approach as well. Bils et al. (2021) acknowledge that "differences in measured average products need not imply differences in true marginal products. First, marginal products are proportional to average products only under Cobb-Douglas production. Second, measured differences in revenue over inputs could simply reflect poor measurement. [...] More-

[19] Verhoogen (2013) shows that despite their high foreign ownership and export share, maquila establishments in Mexico tend to resemble non-exporting plants in the traditional non-maquila sector more than they resemble non-maquila exporters. For instance, maquila establishments have, on average, a lower share of workers with 12 or more years of schooling, a higher share of blue-collar employment, and fewer years of schooling among blue-collar workers than non-maquila establishments (exporters and non-exporters alike).

over, there could be transitory mismeasurement from classifying some revenue or inputs in an adjacent year to when it actually occurred."

An important appeal of our direct measurement approach is that it allows us to unpack the relative importance of a large set of output and input distortions. This means that our measures can be traced back to specific government policies or market failures, which is what Restuccia and Rogerson (2017); Bils et al. (2021) call for in their conclusion.[20] One caveat to our direct measurement approach is that it is exceptionally data-intensive. Moreover, one can always argue that there is an additional distortion that we should have incorporated in our list, despite our intention to be as exhaustive as the Mexican data allows. With these caveats in mind, we move forward with our direct measurement approach.

For those distortions measured directly in the Economic Census, we observe a value for each establishment in the Census in each Economic Census year when that establishment is in operation. For those distortions measured via national surveys covering a representative sample of establishments, we project survey responses on establishment-level observables and use those projections to generate values for all establishments in the Census based on their observables. We assume that the projection coefficients are time-invariant (e.g., that the predictive role of establishment size for a given distortion is the same in 1994 and 2019).[21]

Data on distortions in Mexico Distortions coming from taxes and subsidies are directly measured in the Economic Census. This Census tracks establishment-specific (i) taxes paid for supplying goods and services to producers and final consumers, (ii) subsidies granted by the government to establishments associated with their programs and (iii) contributions to social security on behalf of workers and other labor-related contributions that each establishment made in a given year.

The data source for the distortion associated with the excessive burden of regulation is the "National Survey on Productivity and Competitiveness of Micro, Small and Medium Enterprises" (ENAPROCE in Spanish, collected in 2015 and 2018). ENAPROCE allows us to estimate (i) the value of time spent by all employees to address government requirements and (ii) any additional payments to the federal government due to regulation (other than taxes). For the distortion associated with crime and corruption, we use the National Survey of Business Victimization (ENVE in Spanish, collected in 2012, 2014, 2016, 2018, and 2020) to obtain estimates for the cost of crime and corruption to establishments in Mexico. For crime, we measure actual damages caused by crimes. For corruption, we look into the bribes paid by establishments to public servants or third parties.[22]

[20] This also relates to Rodrik (1987), who argues that "a policy broadly aimed towards addressing observed market wedges will not work because different mechanisms — even if they lead to similar divergence between social and private costs — could behave differently to policies."

[21] Appendix B provides a detailed account of these data sources, variables used and the imputation procedure.

[22] Olken and Pande (2012) review literature pointing to the prevalence and distortive effect of corruption on firm behavior in the developing world (e.g., Sequeira and Djankov, 2014).

2.2.1 Facts on distortions in Mexico

Fact 1: The ranking of establishment types by the burden of distortions they face depends on the specific distortions considered. Table 1 reports weighted averages of distortion values by establishment type (domestic informal, domestic formal, foreign non-maquila MNE and foreign maquila MNE). We learn that distortions vary substantially across establishment types. For instance, domestic informal establishments do not pay labor taxes (by construction), whereas foreign MNEs and domestic formal pay around 19% of their wage bill. Other distortion types also vary significantly across the four types of establishments, reinforcing the importance of keeping track of as many empirically relevant distortions as the data allows. In addition, we note that the ranking of the distortion values across establishment types depends on the distortion. For instance, crime and corruption account for, on average, 7% of informal establishments' sales while the same "tax" amounts to around 3% for other establishments.[23] The table also shows that informal and maquila MNE establishments are less burdened by net taxes (taxes minus subsidies) than domestic formal and non-maquila MNE establishments.

Fact 2: Even within the same type of establishment, there is significant spatial variation in the burden of distortions an establishment faces. The data shows substantial heterogeneity across CZs regarding how distorted a given establishment type is. The maps in Figure 3 plot the values of the average output and input distortions in 1994 by commuting zone (averaging across establishment types and industries).

The spatial variation remains even within each establishment type. This is illustrated in Figure 4, which presents four empirical density plots showing the average distortion values for each CZ in 1994. Each plot contains one histogram per establishment type. Panel (a) depicts the cost of crime and corruption, Panel (b) shows the cost of bureaucracy, Panel (c) illustrates the cost of net taxes (VAT minus government subsidies), and Panel (d) represents labor-related taxes (social security contributions plus other labor taxes). Panels (a)–(c) measure output distortions as the monetary cost divided by total sales, while Panel (d) captures labor input distortions, calculated as the cost of labor taxes divided by the wage bill.

If all CZs experienced the same burden of distortions, the density plots would collapse into a single mass point. However, all panels reveal significant geographic variation in distortion values, both within and across establishment types. Remarkably, these patterns persist even after residualizing each establishment's distortion by 4-digit industry fixed effects and a polynomial of size (number of workers). Figure B1 in Appendix B.4 presents histograms of the residualized distortions, indicating that geographic dispersion cannot be attributed to differences in industry composition or establishment size across regions.

Fact 3: When considering all distortions combined, informal establishments and foreign maquila MNEs are the least distorted on average. Fact 1 highlights that analyzing distortions individually may yield differing conclusions about which types of establishments are more

[23] This supports the view that property rights are protected for formal firms but not for informal ones (De Soto, 1989).

burdened by distortions. However, from the perspective of allocative efficiency, the impact of distortions is determined by the extent to which they create dispersion in the value of marginal revenue products. This dispersion is proportional to a "total distortion," which we will refer as an index that aggregates the total output and input distortions (Hsieh and Klenow, 2009, and subsequent literature). Establishment types burdened by a higher total distortion remain inefficiently small with respect to the planner's solution, and their relative expansion would reduce misallocation. Define the total output distortion (t^Y) as the sum of all wedges influencing output decisions, and let the total input distortion (t^L) be the equivalent from the input side. A key question remains: how do these total distortions relate to the value of marginal revenue products? We leverage the structure of the model introduced in Section 4, which builds on a widely used class of models, to relate the marginal revenue product of labor (VMPL) to the measured distortions. Through the lens of our model, the marginal revenue product of labor (VMPL) satisfies $VMPL \propto \frac{1+t^Y}{1-t^L}$. We call this fraction the total distortion.

Figure 5 includes an empirical density plot of the value of total distortions in 1994, one histogram for each establishment type across all 455 CZs in Mexico. We learn that domestic informal establishments tend to be similarly distorted than foreign maquila MNEs, which in turn tend to be less distorted than domestic formal establishments and foreign non-maquila MNEs. The comparison between the distortions of foreign MNEs relative to the domestic sector (formal and informal) highlights that the mere expansion of foreign MNEs in the Mexican economy has reallocative implications (even without distinguishing between the relative effects of this expansion on the domestic formal vs. informal sectors).

3 The local effects of changes in exposure to foreign MNEs

In this section, we estimate the effect of relative changes in the CZ-level importance of foreign MNEs on the relative performance of the CZ-level domestic economy. The purpose of this analysis is twofold. First, while the effects of foreign MNEs on host domestic economies have been studied before (e.g., Alfaro-Ureña et al., 2021; Setzler and Tintelnot, 2021), the sign and magnitude of the reallocation between the formal and informal segments of the economy have not been explored prior. Second, the reduced-form estimates presented in this section provide critical inputs for calibrating key model parameters in Section 5. The calibrated model will allow us to quantify the effects of MNEs on aggregate welfare and explore counterfactual scenarios.

3.1 Empirical strategy

Measure of exposure to foreign MNEs We denote $\widehat{X}^F_{cz,t}$ as the exposure to foreign MNEs in commuting zone cz between two Economic Census years $(t-5)$ and t. We define this exposure as the change in the number of workers employed by foreign MNEs between $(t-5)$ and t divided by the population of that CZ in $(t-5)$:

$$\widehat{X}^{F}_{cz,t} \equiv \frac{L^{F}_{cz,t} - L^{F}_{cz,t-5}}{Pop_{cz,t-5}} \tag{1}$$

where $L^{F}_{cz,t-5}$ and $L^{D}_{cz,t-5}$ are the total number of workers employed at foreign MNEs and domestic establishments, respectively, in CZ cz in year $(t-5)$ and $Pop_{cz,t-5}$ is the population in in CZ cz in year $(t-5)$.

OLS specification We start with the following regression equation:

$$\frac{y^{F}_{cz,t} - y^{F}_{cz,t-5}}{Pop_{cz,t-5}} = \beta \widehat{X}^{F}_{cz,t} + \theta' K_{cz,t} + \Delta\epsilon_{cz,t} \tag{2}$$

where cz indexes commuting zones in Mexico and t indexes Economic Census years (i.e., 1994, 1999, 2004, 2009, 2014, 2019). $y_{cz,t}$ is the outcome associated with domestic establishments in CZ cz in Economic Census year t; our main outcomes are the total sales, number of workers, wage bill, and capital of all domestic establishments in CZ cz. We follow the standards from the literature studying local labor market effects of economic shocks (e.g, Autor et al., 2013) and include in the vector $K_{cz,t}$: (i) economic region-year indicators (controlling for shocks common across CZs in an economic region); (ii) time-varying CZ-specific controls (such as urban concentration rates, employment rates, the importance of manufacturing in employment, the prevalence of a secondary education or routine occupations, in addition to demographic characteristics such as the share of indigenous or foreign-born individuals); (iii) controls for changes in CZ-level exposure to the "China shock" (both via import competition and competition in Mexico's historical export markets) and to rising imports from the United States into Mexico (following Autor et al., 2013; Blyde et al., 2020); (iv) a control for the "missing share" (Borusyak et al., 2022b); and (v) measures of indirect exposure to the growth in foreign MNE employment in other CZs (Adao et al., 2019; Borusyak et al., 2022a). We present more details about the control variables in Appendix A.3. All observations are weighted by the number of workers in domestic establishments in 1994. We cluster standard errors at the census division level.[24]

Potential endogeneity concerns We want a consistent estimate of β measuring the effect of employment growth at foreign MNEs on the CZ-level outcomes of domestic establishments. However, the OLS estimate presents two main potential concerns. First, there is the classic omitted variables concern. The OLS estimate might overestimate the effects of foreign MNE employment growth if foreign MNEs tend to flow into CZs set to grow (e.g., because foreign MNE parents learn about a new government commitment to invest in productive infrastructure). Conversely, the OLS estimate might underestimate the effects of foreign MNEs if they tend to enter and grow in CZs experiencing a decline (e.g., in reaction to falling wages or local government incentives). Second, there might be measurement error in the growth of the employment share of foreign MNEs in a CZ (for instance, if an establishment misreports its foreign ownership).

[24]Mexico is divided into seven census division regions: Bajío, Center, Gulf, Metropolitan Zone, North, Pacific, and South.

Instrumental variable (IV) approach Before introducing the instrument, it is first helpful to denote $L^{F,o}_{cz,t}$ as the employment of foreign MNEs from origin country o in CZ cz in year t and rewrite the expression of $\widehat{X}^F_{cz,t}$ as follows:

$$\widehat{X}^F_{cz,t} \equiv \frac{L^F_{cz,t} - L^F_{cz,t-5}}{Pop_{cz,t-5}} = \sum_o \Delta L^{F,o}_{cz,t} \times S^o_{cz,t-5}$$

where $\Delta L^{F,o}_{cz,t} \equiv (L^{F,o}_{cz,t} - L^{F,o}_{cz,t-5})/L^{F,o}_{cz,t-5}$ and $S^o_{cz,t-5} = \frac{L^{F,o}_{cz,t-5}}{Pop_{cz,t-5}}$. The expression above spells out the fact that the overall growth in foreign MNE employment in a CZ cz between $(t-5)$ and t is a weighted sum across origin countries o of the growth experienced by foreign MNEs from each origin country (where the weights are the $(t-5)$ share of employment from that origin in the population in that CZ and year). Consider two hypothetical CZs with the same share of foreign MNE employment in the overall population in $(t-5)$, $\left(L^F_{cz,t-5}/Pop_{cz,t-5}\right)$ but with different compositions of that foreign MNE employment across countries of origin $\left(L^{F,o}_{cz,t-5}/L^F_{cz,t-5}\right)$. Whenever Mexico experiences an increase in foreign MNE employment from origin o (for reasons unrelated to cz-specific shocks), cz would be predicted to benefit from that foreign MNE employment growth from o proportional to the original importance of origin country o for that CZ. We propose an IV strategy that takes advantage of this past composition of foreign MNE employment in a CZ to predict its growth.

We use an instrument that follows a shift-share structure and will predict the increase in foreign MNE employment from origin country o in cz $\left(\Delta L^{F,o}_{cz,t}\right)$ by leveraging the insight that new foreign MNE employment in Mexico from origin country o will disproportionately be created in CZs where that country of origin clustered in the past. Specifically, we construct the instrument $\widehat{Z}^F_{cz,t}$ as follows:

$$\widehat{Z}^F_{cz,t} \equiv \sum_o \frac{\sum_{cz' \neq cz} \left(L^{F,o}_{cz',t} - L^{F,o}_{cz',t-5}\right)}{\sum_{cz'} L^{F,o}_{cz',t-5}} \times S^o_{cz,t-5} = \sum_o \Delta L^{F,o}_{-cz,t} \times S^o_{cz,t-5} \tag{3}$$

where $\Delta L^{F,o}_{-cz,t} \equiv \sum_{cz' \neq cz} \left(L^{F,o}_{cz',t} - L^{F,o}_{cz',t-5}\right) / \sum_{cz'} L^{F,o}_{cz',t-5}$. Each shift in $\Delta L^{F,o}_{-cz,t}$ is the counterpart for $\Delta L^{F,o}_{cz,t}$, which uses the prediction for the overall change in employment in Mexico in foreign MNEs from origin o (i.e., the overall change in foreign MNE employment in Mexico from that origin *excluding* the change in the CZ of interest) and apportions it to cz based on the previous clustering of foreign MNEs from o in cz. For instance, if 30% of the employment in Mexico of foreign MNEs from Germany was clustered in $(t-5)$ in a CZ in Puebla, that CZ is predicted to receive 30% of any new increase between $(t-5)$ and t of employment in Mexico in foreign MNEs from Germany.

The rationale of this IV is analogous to that of a classic IV from the migration literature, which instruments the immigrant inflows from a given source country to a particular location by using

the historical settlement patterns of immigrants from that source country and the total number of newly-arriving immigrants from that country (Card, 2001). Several studies have documented herding patterns in MNC entry decisions for parent firms from the same origin country. For example, Head et al. (1995) find that Japanese firms tend to locate affiliates in states across the United States that already host many other Japanese-owned firms, Barry et al. (2003) find co-localization patterns for United States-owned affiliates in Ireland, while Bernard and Thomas (2019) bring similar evidence for MNC entry in Eastern Europe. A variant of this instrument was recently used by Setzler and Tintelnot (2021) to study the effects of foreign MNEs in the United States.

Panels C3a and C3b from Figure C3 present two examples from Mexican newspaper articles that discuss the pattern of spatial clustering for German and Japanese MNEs in Mexico. In those articles, the state of Puebla is heralded as the leader in attracting MNEs from Germany, whereas MNEs from Japan have an "appetite" for Guanajuato. Panels C3c and C3b exemplify that the geographic clusters for MNEs from a given origin country are also clusters for amenities relevant to expatriates from that country. In this example, German schools (relevant to German expatriates) also cluster in Puebla. The intuition of our IV relies on the predictive power of factors such as the presence of German schools to predict where German MNEs will expand in Mexico.

The relevance of this IV requires that the growth of foreign MNE employment from an origin country o in a CZ and its growth in all other CZs in Mexico are correlated. This is the case whenever both growth rates are affected by improvements in the overall appeal of Mexico as an investment destination for parent firms from origin country o or exogenous shocks to the performance of these parent firms that lead to increased investments worldwide.[25] Moreover, the relevance of the instrument requires that the past clustering patterns of foreign MNE employment by country of origin is indeed a good predictor of the spatial split of future foreign MNE employment growth from that country.

In both panels of Figure C4, the X-axis presents the residualized measure of exposure to foreign MNEs ($\widehat{X}^F_{cz,t}$) and the Y-axis presents its residualized instrument ($\widehat{Z}^F_{cz,t}$). Both variables are residualized by controlling for the vector of fixed effects and other controls $K_{cz,t}$. The difference between the plots in Panels C4a and C4b is that in Panel C4b, we trim the top and bottom 1% of both the X and Y axes to ensure that the relationship between X and Y is not driven by outliers. The slope coefficients (standard errors) are 1.09 (0.13) for Panel C4a and 0.94 (0.10) for Panel C4b. This figure is a graphical check on the first stage of our instrument and the importance of outliers for the first stage coefficient.

The exclusion restriction requires that the unexplained determinant of growth of the domestic economy in cz and the instrument are orthogonal (conditional on the potential determinants of growth already controlled for through $K_{cz,t}$). While the exclusion restriction is not testable, it is not unreasonable to expect it to be satisfied. First, $\widehat{Z}^F_{cz,t}$ is not a function of cz-specific changes in input or output market between $(t-5)$ and t. Moreover, prior papers on the spatial clustering of

[25] For instance, after each revision of the "Global Agreement" between Mexico and the European Union (which first came into force in 2000), Mexico became a more attractive destination for MNEs from the European Union.

foreign MNEs by origin country argue that differences across locations in endowments of natural resources, labor and infrastructure do not seem to explain this pattern of localization choice (Head et al., 1995; Barry et al., 2003; Bernard and Thomas, 2019). Instead, this pattern suggests stronger agglomeration and demonstration effects among firms from the same origin country.

The definition of the instrument in equation (3) highlights that this IV has a shift-share structure. Identification in our shift-share IV regressions relies on the assumption that the assignment of the shocks $\left(\Delta L^{F,o}_{-cz,t}\right)$ is exogenous, whereas the exposure shares $\left(S^{o}_{cz,t-5}\right)$ are allowed to be endogenous (as discussed in Borusyak et al., 2022b). In the same paper, Borusyak et al. (2022b) recommend to control for the "missing share". We add this control to all our regressions to avoid comparing CZs with too different levels of importance of foreign MNEs in their overall population.

To address remaining threats to identification, we include in the control vector $K_{cz,t}$ a conservative set of CZ-level controls. First, we include economic region-year indicators to control for shocks common across commuting zones in an economic region. Second, we include time-varying controls for CZ-specific urban concentration rates, employment rates, the importance of manufacturing in employment, the prevalence of secondary education or routine occupations, and demographic characteristics such as the share of indigenous or foreign-born individuals. Third and last, we include controls for CZ-level shocks in import and export markets. Namely, we control for the "China shock" using two variables: one capturing the rising importance of imports from China into Mexico and another capturing the increasing competition from China in the main export markets of Mexico in 1994 (in the spirit of Autor et al., 2013; Blyde et al., 2020). We also include a control for the rising importance of imports from the United States into Mexico. Appendix A.3 provides additional details on these controls.

Specification for the effects of maquila vs. non-maquila foreign MNEs We now modify the specification in equation (2) by separating the change in foreign MNE employment in a CZ between that occurring in maquila (F_M) vs. non-maquila (F_N) foreign MNEs:

$$\frac{y^{F}_{cz,t} - y^{F}_{cz,t-5}}{Pop_{cz,t-5}} = \beta_M \widehat{X}^{F_M}_{cz,t} + \beta_N \widehat{X}^{F_N}_{cz,t} + \theta' K_{cz,t} + \Delta\epsilon_{cz,t}, \tag{4}$$

where $\widehat{X}^{F_M}_{cz,t} \equiv \frac{L^{F_M}_{cz,t} - L^{F_M}_{cz,t-5}}{Pop_{cz,t-5}}$ and $\widehat{X}^{F_N}_{cz,t} \equiv \frac{L^{F_N}_{cz,t} - L^{F_N}_{cz,t-5}}{Pop_{cz,t-5}}$, and $L^{F_M}_{cz,t}$ and $L^{F_N}_{cz,t}$ denote the employment in a commuting zone cz in year t in maquila (F_M) and non-maquila (F_N) foreign MNEs. All other features of this regression remain the same as those in equation (2).

We instrument one type of foreign MNE employment change at a time while controlling for the other. For instance, we instrument the maquila foreign MNE employment change $\left(\widehat{X}^{F_M}_{cz,t}\right)$ while controlling for the non-maquila foreign MNE employment change in the same CZ over the same period $\left(\widehat{X}^{F_N}_{cz,t}\right)$.[26]

[26]The magnitude of the IV estimates for β_M and β_N is not sensitive to whether we instrument both exposures simultaneously or only one at a time (while controlling for the other). The F-statistic is substantially lower when instrumenting for both maquila and non-maquila foreign MNE exposure in one go. Given the stability of the IV estimates, we use as our baseline IV specification the one where we instrument for each foreign exposure at a time.

3.2 Effects of exposure to foreign MNEs on the local domestic economy

Table 2 reports the OLS and IV estimates from running the regression in equation (2) for six outcomes. In the order of the columns, these outcomes are the working-age population, followed by five measures of the size of the local domestic economy (aggregated across all domestic establishments operating in a census year): sales, value added, assets, employment, and wage bill.

Panel (a) in Table 2 presents the OLS estimates, whereas Panel (b) shows the IV estimates and corresponding F-statistics. A first takeaway is that while both OLS and IV estimates are positive and significant, the IV estimates are larger in magnitude. As discussed in Section 3.1, one reason behind this pattern could be measurement error in the exposure to foreign MNEs. Alternatively, this pattern could reflect the selection of foreign MNEs into CZs with declining conditions (e.g., falling wages and/or being attracted by local government incentives).

The F-statistic is 107.8, and the first-stage coefficient is 0.796 (see Table C1). Hence, the IV defined in equation (3) is a strong predictor of foreign MNE employment growth. In the right-hand side panel of Table C1, we also show the reduced-form regression of our outcomes of interest on the instrument. These regressions serve to convince the reader that the IV actually affects the outcome variables. According to Angrist and Pischke (2009), the reduced-form regressions act as a weak-instruments test. It is reassuring that we find effects in the reduced form as well, and the estimates do not vary substantially between the reduced form and the IV.

Focusing on the IV estimates, we find that increased exposure to foreign MNEs leads to expansion of the domestic economy across most size measures. We estimate that an increase of one foreign MNE worker per capita in a CZ increases the employment of domestic establishments by 0.33 workers per capita. Most of this domestic employment increase comes from blue-collar workers (see the decomposition between blue and white-collar workers on Table C3). The same increase in MNE employment leads to an increase of $35,708 (2018 USD) in sales per capita.[27] Value-added, assets, and wage bills also increase by $15,231, $6,227, and $3,532 per capita (2018 USD), respectively.

Effects of exposure to maquila vs. non-maquila foreign MNEs Table 3 presents the IV estimates of the regression described in equation (4). The three panels differ in the set of foreign MNEs whose effects we estimate (and the subsequent controls used). In the upper panel, we keep the baseline sample of all foreign MNEs, which corresponds to the regression in equation (2) and the results presented in Table 2. In the middle panel, we focus only on the exposure to maquila foreign MNEs. In these regressions, we control for the change in exposure to non-maquila foreign MNEs. In the lower panel, we focus only on the exposure to non-maquila foreign MNEs. In these regressions, we control for the change in exposure to maquila foreign MNEs.

By contrasting the estimates in the three panels, one can notice that, on average, maquila

[27] Note that the $35,708 (2018 USD) increase corresponds to an extreme change in $\widehat{X}^F_{cz,t}$ of 1. In practice, the difference between the 10th and 90th percentiles of $\widehat{X}^F_{cz,t}$ during the period is 0.07, implying an increase in sales per capita of around 120 dollars per year. This figure represents around 1.2% of Mexico's 2018 GDP per capita.

foreign MNEs have weaker positive effects on the size of the local domestic economies than the non-maquila foreign MNEs. When focusing on high- vs. low-skill workers (see table C4), we learn that non-maquilas tend to favor the expansion of the high-skill and low-skill workforce.

3.3 Effects of exposure to foreign MNEs on the domestic formal vs. informal sector

To study the reallocation effects between the formal and informal segments of the domestic economy, we use the exact regression specification and IV strategy as in equation (2) but restrict the outcomes to either the formal or informal segment of the domestic economy in CZ cz. For example, when we refer to the domestic informal segment in cz, the value of an outcome (e.g., number of workers) is the summation of that variable across all domestic informal establishments in that CZ cz and year t.

Table 4 reports the IV estimates. We learn that the formal sector is the primary beneficiary of the positive effect of foreign MNEs on the domestic economy. In contrast, capital and workers in the informal sector seem to be unaffected by the changing exposure to foreign MNEs.[28] Table C2 includes the reduced form regressions (outcomes regressed on the instrument).

We also run the specification (4) to study the differential effects on the formal and domestic informal sectors of either change in maquila or non-maquila foreign MNE employment. The two panels in Table 5 summarize these differential effects. The results indicate that most of the positive effect of foreign MNEs on the domestic formal sector comes from its exposure to non-maquila MNEs. In contrast, expansions of maquila MNEs reduce the size of the formal sector.

4 Model

We now describe the theoretical framework for our analysis. The main goal is to understand the welfare and output implications of foreign MNEs in an economy where market failures distort the allocation of resources. We model the foreign MNE shock as a productivity increase in foreign-owned MNEs that changes their labor demand. The shock affects the economy depending on the potential spillovers from foreign MNEs to domestic ones and its effect on the relative supply and demand of factors and how these effects interact with the initial distribution of distortions.

We develop a model combining elements from the informality literature with recent advancements in the spatial economics literature. Domestic firms are heterogeneous, face exogenous wedges and endogenously decide whether to operate in the formal or the informal sector. In the case of informal firms, they also face a distortion that increases with size to capture the probability of getting caught by the government. The model also incorporates trade and migration across regions. Figure 6 shows a diagram of the model and how the different economic agents interact with each other.

[28] This finding aligns with the result in Sharma and Cardenas (2018) that higher FDI inflows are less associated with changes in informal employment in Mexican states between 2005 and 2015. Table C5 shows the decomposition of employment and wage bills by worker type.

4.1 Preferences

There are multiple locations, which correspond to different commuting zones in Mexico and an additional region that represents the rest of the world. These locations will generally be indexed by c when they are consuming and by j when they are producing. In each area, there is a mass of potential domestic producers $\mathcal{M}_j$ that is exogenous. Domestic firms are heterogeneous and are indexed with their productivity value φ that is drawn from a Pareto distribution with p.d.f $\vartheta_j(\varphi)$. Moreover, within each area there is a single foreign MNE that produces output and sources inputs from local (formal) domestic producers. This means that there are three different firm types: domestic formal, domestic informal, and foreign MNEs. The utility function of consumer ω that decides to live in location c and work in firm type s is:

$$U_{c,s,\omega} = a_c \ell_{c,s,\omega} \epsilon_{c,\omega} q_{c,s,\omega},$$

where a_c is an amenity level in location c, $q_{c,s,\omega}$ corresponds to the CES consumption level of traded commodities of worker ω in sector s and location c (which has the same functional form as the one given below in equation 6 for aggregate consumption), and $\ell_{c,s,\omega}$ and $\epsilon_{c,\omega}$ are idiosyncratic shocks drawn from a nested Fréchet distribution. Specifically, ℓ represents an efficiency-units shock that determines the sector a worker participates in and ϵ represents an amenity shock that affects the location a worker decides to settle in. The indirect utility of worker ω is:

$$V_{c,s,\omega} = \frac{a_c \epsilon_{c,\omega} (\ell_{c,s,\omega} w_{c,s} + r_c)}{P_c}, \tag{5}$$

where $w_{c,s}$ corresponds to the wage per efficiency unit in firm type s in c, r_c to the per-capita revenue from distortions, and P_c to the ideal price index. The aggregator of traded commodities is a CES composite good of domestic and foreign varieties:

$$Q_c = \left(\sum_j \int_{\varphi \in \Omega_{j,D}} q_{jc}(\varphi)^{\frac{\sigma-1}{\sigma}} d\varphi + q_{jc,M}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{6}$$

where $q_{jc}(\varphi)$ represents the variety produced by domestic firm φ in region j and traded to c and, $q_{jc,M}$ corresponds to the variety produced by foreign firms in j and traded to c.

4.2 Technology of domestic producers

Domestic firms operate with a technology that is linear in labor. We denote total production as x. The production function for domestic firms is:

$$x_{j,s}(\varphi) = A_{j,s} \varphi l_{j,s}(\varphi),$$

where $A_{j,s}$ is an aggregate TFP of location j and sector s, φ corresponds to the index of the firm, and also to its productivity, and $l_{j,s}(\varphi)$ is the amount hired by firm φ. In the case of formal firms, total production is composed of sales to final consumers and sales to MNEs. If a firm operates in the informal sector, it only sells the good in its own location and does not sell to MNEs.

The problem of formal domestic firm φ in region j consists of maximizing the following profit function:

$$\begin{aligned}\pi_{j,F}(\varphi) &= \underbrace{(1-t^F_{j,Y})}_{\text{Output wedge}} \sum_c \left(\underbrace{P_c^\sigma Q_c}_{\text{Agg demand}} + \underbrace{J_{c,M} P^\sigma_{c,FM}}_{\text{MNEs demand}} \right) \cdot \underbrace{p_{jc,F}(\varphi)^{1-\sigma}}_{\text{Price function}} \\ &- \underbrace{(1+t^F_{j,L})}_{\text{Labor wedge}} \underbrace{w_{j,F} l_{j,F}(\varphi)}_{\text{Labor cost}}, \end{aligned} \tag{7}$$

where X_c represents the aggregate expenditure in location c, and $J_{c,M}$ and $P^{\sigma-1}_{c,FM}$ denote the total demand by MNEs in location c. These terms capture the aggregate demand for formal varieties in each location. All producers are atomistic and take these variables as given, which are determined in general equilibrium. The term $(1-t^F_{j,Y})$ and $(1+t^F_{j,L})$ corresponds to output and labor wedges respectively in the case firms decide to operate in the formal sector. For example, the output wedge includes how much firms have to pay in bribes or to criminal gangs in location j, and the labor taxes capture social contributions and labor regulations for formal establishments in the Mexican context.

Firms choose the price to maximize the sum of profits across locations and we assume that the market structure is monopolistically competitive. Let's define, $\Phi_c \equiv P_c^\sigma Q_c + (1-\beta)\frac{\sigma-1}{\sigma} R_{c,M}(1-t^M_{c,Y}) P^{\sigma-1}_{c,FM}$ as the aggregate demand for domestic formal varieties. The solution to the problem of firm φ in the case in which it operates in the formal sector is:

$$p_{jc,F}(\varphi) = \left(\frac{\sigma}{\sigma-1}\right) \frac{w_{j,F}(1+t^F_{j,L})\tau_{jc}}{A_{j,F}\varphi(1-t^F_{j,Y})}, \tag{8}$$

where TFP is given by the product $A_{j,F} \cdot \varphi$ and $\tau_{jc} \geq 1$ corresponds to an iceberg trade cost of sending output from j to c. We can also solve for the total amount of labor hired by formal firms, which is:

$$l_{j,F}(\varphi) = (A_{j,F}\varphi)^{\sigma-1} \left(\tilde{\sigma} w_{j,F}(1+t_{j,T})\right)^{-\sigma} \sum_c \Phi_c \tau_{jc}^{1-\sigma}, \tag{9}$$

where $\tilde{\sigma} \equiv \sigma/(\sigma-1)$ represents the markup charged by firm φ and $(1+t_{j,T}) \equiv (1+t^F_{j,L})/(1-t^F_{j,Y})$ corresponds to the wedge faced by firm φ in the formal sector.

On the other hand, firm φ maximizes the following profit function if it operates in the infor-

mal sector:

$$\pi_{j,I}(\varphi) = \underbrace{(1-t^I_{j,Y})}_{\text{Output tax}} \underbrace{(Q_j)^{\frac{1}{\sigma}} P_j}_{\text{Agg. demand}} \underbrace{\left(A_{j,I}\varphi l_{j,I}(\varphi)\right)^{\frac{\sigma-1}{\sigma}}}_{\text{Price function}} - \underbrace{(1+l_{j,I}(\varphi)/b_j)}_{\text{Labor wedge}} \underbrace{w_{j,I} l_{j,I}(\varphi)}_{\text{Labor cost}}.$$

As in the case of formal firms, the terms Q_j and P_j capture the aggregate demand of final consumers in region j and are determined in general equilibrium. The terms $(1-t^I_{j,Y})$ and $(1+l_{j,I}(\varphi)/b_j)$ correspond to output and labor wedges if firm φ decides to operate in the informal sector. In this case, $(1-t^I_{j,Y})$ captures similar wedges than in the case of the formal sector. They correspond to taxes and payments to crime, bribes, or bureaucracy as a share of value added that we calibrate for informal firms using the information from the different surveys. Since informal firms do not pay social contributions, we follow Ulyssea (2018) and model the labor wedge as a function that increases with size to capture the probability that informal firms get caught by the government. This function increases with firm size, implying a convex marginal cost. Notice that the marginal cost is constant in the case in which $b_i \to \infty$. The idea of this wedge is to capture the fact that as the size of the firm grows, it is more likely that the government will discover that the firm is not paying social contributions to its workers and impose a sanction.

The solution to firm's φ problem in the informal sector is:

$$l_{j,I}(\varphi) = \left(\frac{\frac{\sigma-1}{\sigma}(1-t^I_{j,Y})\left(A_{j,I}\varphi\right)^{\frac{\sigma-1}{\sigma}} \Xi_j}{w_{j,I}\left(1+\frac{2l_{j,I}(\varphi)}{b_j}\right)} \right)^{\sigma},$$

where $\Xi_j \equiv (Q_j)^{\frac{1}{\sigma}} P_j$ is a constant that captures aggregate demand. Using the implicit function theorem, it is easy to show that the total amount of labor depends positively on productivity, implying that more productive firms, even in the informal sector, hire more labor.

After solving the problem of firm φ in the formal and informal sector, firms decide by backward induction whether to operate in the formal or in the informal sector. Firm φ decides to operate in the formal sector if the following condition holds:

$$\pi_{j,F}(\varphi) - FC_j \geq \pi_{j,I}(\varphi), \tag{10}$$

where FC_j is a fixed cost in terms of the final good in each location that firms need to pay to operate in the formal sector. The idea of this fixed cost is to capture the "red tape" that firms incur to operate formally in Mexico. We calibrate this cost to match the share of formal firms observed in each location j. Moreover, we assume that since the labor wedge increases with size for informal firms, the following condition holds:

$$\frac{\partial[\pi_{j,F}(\varphi) - \pi_{j,I}(\varphi)]}{\partial\varphi} > 0 \quad \forall j, \varphi. \tag{11}$$

This condition implies that as the productivity of a firm increases, the profits it obtains in the

formal sector increase at a higher rate than the ones it would obtain in the informal sector. As a result, there exists a marginal firm $\tilde{\varphi}$ that is indifferent between operating in the formal sector and doing so in the informal sector. All firms with $\varphi \geq \tilde{\varphi}$ decide to operate in the formal sector, and all firms with $\varphi < \tilde{\varphi}$ decide to operate in the informal sector.[29]

4.3 Technology of foreign multinationals

There is a representative foreign multinational in each location that operates as a monopoly and is indexed by (j, M). The problem of this firm consists of maximizing total profits subject to the following production technology:

$$q_{j,M} = A_{j,M} \left(\frac{L_{j,M}}{\beta} \right)^{\beta} \left(\frac{J_{j,M}}{1-\beta} \right)^{1-\beta}, \tag{12}$$

where $L_{j,M}$ corresponds to the labor hired by the multinational and $J_{j,M}$ to the domestic inputs sourced by the MNE in location j. This consists of a CES aggregator that only includes the domestic formal firms across all locations:

$$J_{j,M} = \left(\sum_{h} \int_{\varphi \in \Omega_{h,F}} r_{hj}(\varphi)^{\frac{\sigma-1}{\sigma}} d\varphi \right)^{\frac{\sigma-1}{\sigma}}, \tag{13}$$

where $\Omega_{h,F}$ corresponds to the set of formal domestic producers in location h.

The problem of the single MNE in location j is to maximize the following profit function:

$$\pi_{j,M} = \underbrace{(1 - t^{M}_{j,Y})}_{\text{Output wedge}} \underbrace{\sum_{c} p_{jc,M} q_{jc,M}}_{\text{Total revenue}} - \underbrace{(1 + t^{M}_{j,L})}_{\text{Labor wedge}} \underbrace{w_{j,M} L_{j,M}}_{\text{Labor cost}} - \underbrace{P_{j,FM} J_{j,M}}_{\text{Int. Inputs cost}},$$

subject to:

$$\begin{aligned} q_{j,M} &= \sum_{c} \tau_{jc} q_{jc,M} \\ q_{jc,M} &= \left(\frac{p_{jc,M}}{P_c} \right)^{-\sigma} Q_c, \end{aligned}$$

where $P_{j,FM}$ corresponds to the unit price of the intermediate input bundle used by MNEs. The solution to the maximization problem implies that the price charged by MNEs from location j to consumers in c is:

$$p_{jc,M} = \tilde{\sigma} \frac{(w_{j,M}(1 + t^{M}_{j,L}))^{\beta} P^{1-\beta}_{j,FM} \tau_{jc}}{A_{j,M}(1 - t^{M}_{j,Y})}, \tag{14}$$

and total profits are: $\pi_{j,M} = [(1 - t^{M}_{j,Y}) \sum_{c} p_{jc,M} q_{jc,M}] / \sigma$. We assume that profits from foreign MNEs

[29] In the calibration, we verify that this condition holds for all locations.

are redistributed towards the RoW region. Finally, the total amount of labor hired by MNEs is:

$$L_{j,M} = \frac{\beta}{A_{j,M}} \left(\frac{w_{j,M}(1+t^M_{j,L})}{P_{j,FM}} \right)^{\beta-1} \left(\tilde{\sigma} \frac{(w_{j,M}(1+t^M_{j,L}))^{\beta} P^{1-\beta}_{j,FM}}{A_{j,M}(1-t^M_{j,Y})} \right)^{-\sigma} \sum_c \tau^{1-\sigma}_{jc} \Psi_c, \tag{15}$$

where $\Psi_c = \Xi^{\sigma}_c$.

4.4 Cross-spillovers

Based on the findings from previous literature (e.g., Alfaro-Ureña et al., 2022b), we assume that MNEs generate positive spillovers to the local economy. In particular, we assume the following functional form:

$$A_{j,F} = \tilde{A}_{j,F} L^{\gamma_{M,F}}_{j,M} \tag{16}$$

$$A_{j,I} = \tilde{A}_{j,I} L^{\gamma_{M,I}}_{j,M}, \tag{17}$$

where $\tilde{A}_{j,F}$ and $\tilde{A}_{j,I}$ are parameters that capture exogenous scale effects of total factor productivity at the local level, $L_{j,M}$ corresponds to the labor force of MNEs, and the γ parameters capture the strength of the spillovers from foreign MNEs to the domestic economy. We calibrate these elasticities by matching the reduced-form coefficients from our previous section using the predictions generated by the model.

4.5 Total expenditure and price index of final consumers

To close the model from the labor demand side, we need to compute the final consumption expenditure, the revenue from MNEs, and the price indices, we obtain that:

$$P_c Q_c = \underbrace{w_{c,I} L_{c,I} + w_{c,F} L_{c,F} + w_{c,M} L_{c,M}}_{\text{Labor income}} \tag{18}$$

$$+ \underbrace{\mathcal{M}_c \left(\int_{\varphi \in [\underline{\varphi}_c, \tilde{\varphi}_c]} \pi_{c,I}(\varphi) \vartheta_c(\varphi) d\varphi + \int_{\varphi \in [\tilde{\varphi}_c, \bar{\varphi}_c]} \pi_{c,F}(\varphi) \vartheta_c(\varphi) d\varphi \right)}_{\text{Domestic Profits}} + \underbrace{\sum_j \pi_{jc,M}}_{\text{Foreign MNE Profits}} \tag{19}$$

$$+ \underbrace{t^I_{c,Y} \mathcal{M}_c \int_{\varphi \in [\underline{\varphi}_c, \tilde{\varphi}_c]} p_{c,I}(\varphi) x_{c,I}(\varphi) \vartheta_c(\varphi) d\varphi}_{\text{Informal Output Tax Rev.}} \tag{20}$$

$$+ \underbrace{t^F_{c,Y} \mathcal{M}_c \int_{\varphi \in [\tilde{\varphi}_c, \bar{\varphi}_c]} p_{c,F}(\varphi) x_{c,F}(\varphi) \vartheta_c(\varphi) d\varphi}_{\text{Formal Output Tax Rev.}} + \underbrace{t^M_{c,Y} \sum_r q_{cr,M} p_{cr,M}}_{\text{MNE Output Tax Rev.}} \tag{21}$$

$$+ \underbrace{\mathcal{M}_c \int_{\varphi \in [\underline{\varphi}_c, \tilde{\varphi}_c]} t^I_{c,L}(l_{c,I}(\varphi)) w_{c,I} l_{c,I}(\varphi) \vartheta_c(\varphi) d\varphi + t^F_{c,L} w_{c,F} L_{c,F} + t^M_{c,L} w_{c,M} L_{c,M}}_{\text{Labor Distortion Rev.}}, \tag{22}$$

where the last term captures labor taxes paid by formal firms and MNEs as well as the revenue associated with the non-linear distortions in the informal firm's problem, and $\pi_{jc,M}$ captures the profits from foreign MNE's in location j that are sent to location c. In practice, we will assume that $\pi_{jc,M} = 0$ if $c \neq$ RoW and $\pi_{jc,M} = \pi_{j,M}$ if $c =$ RoW. The price index for final consumers in location c and its sub-components are given:

$$P_c = \left[P_{c,D}^{1-\sigma} + P_{c,M}^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{23}$$

$$P_{c,D} = \left(\sum_j \mathcal{M}_j \left[\int_{\varphi\in[\tilde{\varphi}_j,\bar{\varphi}_j]} p_{jc,F}(\varphi)^{1-\sigma}\vartheta_j(\varphi)d\varphi\right] + \mathcal{M}_c \int_{\varphi\in[\underline{\varphi}_c,\tilde{\varphi}_c]} p_{c,I}(\varphi)^{1-\sigma}\vartheta_c(\varphi)d\varphi\right)^{\frac{1}{1-\sigma}} \tag{24}$$

$$P_{c,M} = \left[\sum_j p_{jc,M}^{1-\sigma}\right]^{\frac{1}{1-\sigma}}. \tag{25}$$

With these ingredients we can solve the model from the labor demand side. In the next section, we explain the labor supply and then the labor market clearing conditions.

4.6 Labor supply

Workers make two sequential decisions. First, they decide the location to live in, and second, they choose the firm type to work in. We assume that the idiosyncratic shock of worker ω, denoted as $\epsilon_{\ell c,s,\omega}$ depends on two components: a shock that determines the migration decision and a shock that determines the labor supply decision across firm types. The former shock is to amenities, while the latter is to efficiency units as in Galle et al. (2023).

The distribution of labor supply across firm types depends on workers' decisions about where to work. Let $\boldsymbol{\ell} = (\ell_1, \ell_2, \ldots, \ell_S)$ and let $I_{j,s} \equiv \{\boldsymbol{\ell} \mid w_{js}\ell_s \geq w_{jk}\ell_k \text{ for all } k\}$. A worker with productivity vector $\boldsymbol{\ell}$ in region j will apply to firm type s if and only if $\boldsymbol{\ell} \in I_{j,s}$. Let $F_j(\boldsymbol{\ell})$ be the joint probability distribution of $\boldsymbol{\ell}$ for workers in j that we assume is drawn from a Fréchet or extreme value type II distribution. From Lagakos and Waugh (2013) and Hsieh et al. (2019), the share of workers that decide to work in firm type k in location j is

$$\lambda_{jk|j} = \frac{B_{jk}w_{jk}^{\kappa}}{\sum_s B_{js}w_{js}^{\kappa}}, \tag{26}$$

where κ is the dispersion parameter from the Fréchet distribution which also captures the labor supply elasticity across firm types, B_{jk}, are relative efficiency parameters, $\lambda_{jk|j}$ represents the share of workers in region j that decide to work in firm type k in location j, and $\Phi_j^{\kappa} \equiv \sum_s B_{js}w_{js}^{\kappa}$ is a normalization term that captures the ex-ante average income of workers in location j. The labor supply in terms of efficiency units, L_{js} in location j and sector s is:

$$L_{js} = \zeta\frac{\Phi_j}{w_{js}}\lambda_{js|j}\tilde{L}_j,$$

where $\tilde{L}_j$ is the total amount of workers in j and $\zeta = \Gamma\left(1 - \frac{1}{\kappa}\right)$ is a constant term that arises after integrating the Fréchet distribution. We also assume an amenity idiosyncratic shock that determines the migration decisions. As in Fajgelbaum and Gaubert (2020), spatial mobility is determined by the following equation:

$$\lambda_c \equiv \frac{\tilde{L}_c}{\bar{L}} = \frac{W_c^{\eta}}{\sum_h W_h^{\eta}}, \tag{27}$$

where η corresponds to the migration elasticity, $\bar{L}$ to the total amount of workers, and W_c corresponds to the ex-ante welfare measure of workers in c, which is:

$$W_c = \frac{a_c I_c}{P_c},$$

where a_c captures amenities, I_c corresponds to a nominal income measure that depends on the ex-ante average income and the revenue from wedges, and P_c represents the ideal price index in location c.

4.7 Equilibrium

Labor market clearing conditions imply that in equilibrium the following equations must hold:

$$L_{j,F} = \mathcal{M}_j \int_{\varphi \in [\tilde{\varphi}_j, \bar{\varphi}_j]} l_{j,F}(\varphi) \vartheta_j(\varphi) d\varphi \tag{28}$$

$$L_{j,I} = \mathcal{M}_j \int_{\varphi \in [\underline{\varphi}_j, \tilde{\varphi}_j]} l_{j,I}(\varphi) \vartheta_j(\varphi) d\varphi \tag{29}$$

$$L_{j,M} = \left(\frac{\sigma - 1}{\sigma}\right) \left(\frac{\beta(1 - t_{j,M}^{Y})}{w_{j,M}(1 + t_{j,M}^{L})}\right) Y_{j,M}, \tag{30}$$

where $\tilde{\varphi}_j$ represents the marginal firm that is indifferent between operating in the formal and the informal sector, and $Y_{j,M}$ corresponds to the gross total revenue of MNEs in j. Then, the equilibrium is defined by the following vector of endogenous variables:

$$x = \{w_{j,s}, \Phi_j, P_{ck}, P_c, \tilde{\varphi}_j, L_{js}, L_j, W_j, I_j\},$$

given a set of exogenous parameters:

$$\Delta = \{\tau_{js}, \tilde{A}_{js}, t_{j,s}^{Y}, t_{j,s}^{L}, \mathcal{M}_j, b_j, \vartheta_j(\varphi), FC_j, B_{js}, a_c, \beta, \sigma, \kappa, \eta, \gamma_{M,F}, \gamma_{M,I}\},$$

that solve the following system of equations: labor demand for all formal, informal, and MNEs firms from equations (9), (10), and (15); firm type choice probabilities from equation (26); migration decisions from equation (27) and the labor market clearing conditions from equations (28)-(30).

4.8 Theoretical results

In this section, we examine the impact of multinational enterprises (MNEs) on local economic activity, focusing on three key outcomes: (i) the relative employment changes between formal and informal domestic producers, (ii) changes in real income across different locations, and (iii) the overall welfare effects at the aggregate level. The latter two outcomes depend on the interaction between MNE expansion at the local and aggregate levels and the initial distribution of economic distortions.

Relative employment effects: We aim to understand how the relative demand for formal versus informal employment, $\frac{L_{jF}}{L_{jI}}$, responds to an increase in the productivity of MNEs at the local level. Assume $\tau_{jc} \to \infty$ for all $j \neq c$, $b_j \to \infty$ for all j and that there is no change in the limit $\tilde{\varphi}_j$. Under these conditions, the change in the ratio of formal to informal employment driven by a change in MNEs' productivity is:

$$\frac{d \log \frac{L_{j,F}}{L_{j,I}}}{d \log A_{j,M}} = (\sigma - 1) \left(\underbrace{(1 - \lambda_{j,F})[1 + (1 - \beta)\delta_{j,M}\gamma_{M,F}]}_{\text{Input-output linkages}} + \delta_{j,M} \underbrace{(\lambda_{j,F}\gamma_{M,F} - \gamma_{M,I})}_{\text{Spillovers}} \right), \quad (31)$$

where, $\lambda_{j,F}$ denotes the initial share of sales to final consumers relative to total revenue for formal firms in location j, and $\delta_{j,M} = \frac{d \ln L_{j,M}}{d \ln A_{j,M}}$ represents the elasticity of MNE employment with respect to its own productivity.

Equation (31) highlights that employment reallocation between formal and informal domestic establishments depends on two key components. The first term depends on $1 - \beta$, the expenditure share of MNEs in intermediate inputs, and the spillovers from MNEs to formal domestic firms. This term captures the role of input-output linkages between formal domestic firms and MNEs. The second term, which depends on $\gamma_{M,F}$ and $\gamma_{M,I}$, captures the relative productivity spillovers from MNEs to formal and informal establishments.

For instance, an increase in MNE productivity can lead to higher formal employment relative to informal employment if input-output linkages are strong or if the difference in spillovers between formal and informal varieties is substantial. Conversely, employment may reallocate towards informal firms relative to formal firms if the spillovers from MNEs to informal establishments are very strong, outweighing both the input-output linkages and the spillovers to formal establishments.

In the model calibration, we take β directly from the data, and calibrate the spillovers to match the reduced-form coefficients from Section 3.

Real income effects: The second question that we are interested in is how real income at the local level responds to an increase in the productivity of MNEs. The per capita real income measure in

location c is:

$$W_c = \frac{\Phi_c + r_c}{P_c}, \tag{32}$$

where Φ_c is the ex-ante average labor income and $r_c = \frac{R_c}{L_c}$ is the per-capita income obtained from the distortion revenue, including all wedges and markups. Define

$$\mu^L_{j,s} = \tilde{\sigma}\left(\frac{1+t^L_{j,s}}{1-t^Y_{j,s}}\right), \text{ and } \mu^F_{j,M} = \tilde{\sigma}\left(\frac{1}{1-t^Y_{j,s}}\right)$$

as wedge measures that incorporate distortions. We also define $\Psi_{jc,s}$ as the exposure of consumers in c to the change in the producer price j,s.[30] The change in real income due to a productivity change of MNEs is:

$$d\ln W_c = \underbrace{\sum_s \iota_c \lambda_{c,s|c} d\ln w_{c,s} - \sum_{j,s\in\{F,I\}} \Psi_{jc,s} d\ln w_{j,s} - \sum_j \beta \Psi_{jc,M} d\ln w_{j,M}}_{\text{Change in Factoral ToT}} \tag{33}$$

$$+ \underbrace{\sum_j \Psi_{jc,M} d\ln A_{j,M}}_{\text{Technology Change}} + \underbrace{\sum_{j,s\in\{F,I\}} \Psi_{jc,s}\gamma_{M,s} d\ln L_{j,M}}_{\text{Productivity spillovers}} \tag{34}$$

$$- \underbrace{\sum_{j,s\in\{F,I\}} \Psi_{jc,s} d\ln \mu^L_{j,s} + \sum_j \Psi_{jc,M}(\beta d\ln \mu^L_{j,M} + (1-\beta) d\ln \mu^F_{j,M})}_{\text{Change in the wedges}} \tag{35}$$

$$+ \underbrace{(1-\iota_c) d\ln r_c}_{\text{Change in the distortion revenue}}, \tag{36}$$

where ι_c represents the share of factor income in total income in location c. This expression is similar to one from Atkin and Donaldson (2022) and Baqaee and Farhi (2024). It suggests that the change in real income in each location can be decomposed into the following components:

1. **Change in the factoral terms of trade:** This component captures how factor prices change due to productivity shocks. It depends on how the wage component of exports changes relative to the wage component of imports. The exposure measure is determined by the initial expenditure share of each variety. In all trade models, with or without wedges, this term affects real income in each location.

2. **Technology change:** This component corresponds to the technological change driven by the increase in the productivity of MNEs. It reflects how the price index in each location responds

[30] In the case of formal firms: $\Psi_{jc,F} = \left(\phi_{jc,F} + (1-\beta)\sum_h \phi_{hc,M}\cdot\phi_{jh,FM}\right)$ where $\phi_{jc,F}$ is the initial expenditure share of formal varieties of location c from location j and $\phi_{jh,FM}$ is the initial expenditure share within intermediate inputs of MNEs in h on formal varieties from location j. For informal firms, $\Psi_{jc,I} = 0$ if $j \neq c$, and $\Psi_{cc,I} = \phi_{cc,I}$, which is the expenditure share of location c in informal varieties. And finally, in the case of MNEs, $\Psi_{jc,M} = \phi_{jc,M}$, which corresponds to the initial expenditure share of location c in MNEs' varieties from j.

to perturbations in MNEs' productivity. The magnitude of this effect depends on the initial expenditure shares that each region allocates to MNE varieties, captured by $\Psi_{jc,M}$.

3. **Productivity spillovers:** This component captures the potential spillovers that MNEs generate for the local economy. It is represented by the change in MNE employment, $d \ln L_{j,M}$ and the strength of the spillovers, $\gamma_{M,s}$. This term quantifies how the price index in each location changes due to shifts in the productivity of formal and informal domestic firms. The exposure to these changes depend on the initial expenditure shares and the importance of input-output linkages to MNEs.

4. **Change in the wedges:** This component measures how much the price index changes due to variations in wedges. An increase in wedges, akin to higher markups or taxes, raises consumer prices and reduces welfare. Exposure to this shock depends on the initial expenditure share of each variety and, $\phi_{jc,s}$, and in the case of formal domestic firms, the strength of input-output linkages, captured by β.

5. **Change in distortion revenue:** This component reflects how revenue from distortions, in per-capita terms, changes with productivity variations. It depends on the importance of distortion revenue in total income, captured by the initial share $1 - \iota_c$. In a scenario without wedges, distortion revenue is absent.

In a perfectly efficient economy with no distortions or spillovers, the change in real income depends only on the factoral terms of trade and technological changes.

Aggregate welfare: The final outcome that we study is how the expansion of MNEs affect welfare at the aggregate level. We define the welfare function as $\bar{U}^{\eta} = \sum_c W_c^{\eta}$, where η corresponds to the migration elasticity. Also, assume that:[31] i) $b_j \to \infty$ for all j; ii) population in each location is fixed and exogenous to omit congestion externalities; iii) $\mu_{j,M}^{F} = 1$, implying that there are only value-added and labor wedges; and iv) the rebate is proportional to real income, $r_c = \Phi_c \bar{t}$ to prevent fiscal externalities.[32] Under this scenario, the change in aggregate welfare due to a perturbation in the productivity of MNEs is:

$$d \ln U \quad = \quad \sum_c \lambda_c d \ln W_c, \tag{37}$$

where λ_c corresponds to the initial population weights. Using the results for the change in real income, the market clearing conditions, and summing across locations, the change in aggregate

[31] We make these assumptions to understand the economic impact of wedges in the absence of other market failures. Other papers such as Fajgelbaum and Gaubert (2020) have studied the welfare impact of these other market failures.

[32] The parameter $\bar{t}$ represents the constant share of income rebated to households.

welfare is:

$$d \ln U = \underbrace{\sum_{j,c} \lambda_c \Psi_{jc,M} d \ln A_{j,M}}_{\text{Technology Change}} + \underbrace{\sum_{j,c,s \in \{F,I\}} \lambda_c \Psi_{jc,s} \gamma_{M,s} d \ln L_{j,M}}_{\text{Productivity Spillovers}} \tag{38}$$

$$+ \underbrace{\sum_{s,j} \lambda_j \lambda_{sj|j} \left(\frac{\mu^L_{j,s} - \bar{\mu}}{\bar{\mu}} \right) d \ln L_{j,s}}_{\text{Reallocation Effects}}, \tag{39}$$

where $\bar{\mu} = 1 + (\sum_{j,c} (\mu^L_{j,c} - 1) \lambda_c \lambda_{cj|j}) / (\sum_h \Phi_h \lambda_h)$ is the average wedge in the economy. This equation decomposes the effect of productivity shocks at the aggregate level and illustrates how these shocks interact with the initial distribution of distortions. In the absence of distortions, the economy operates efficiently, and the envelope theorem applies. In this case, the overall change in welfare depends solely on model parameters and is unaffected by changes in the optimizers. However, when the economy is inefficient due to the presence of wedges, the aggregate change in welfare depends on how the shock interacts with the initial distribution of distortions. If workers reallocate toward segments of the economy with higher distortions ($\mu^L_{js} \geq \bar{\mu}$), allocative efficiency improves, leading to an additional positive impact on welfare. Conversely, if workers reallocate toward segments with lower distortions ($\mu^L_{js} \leq \bar{\mu}$), allocative efficiency declines, resulting in a further decrease in welfare. Overall, the effect depends on how the shock affects relative labor demand and supply.

It is important to note that at the aggregate level, other terms related to wage responses and revenues from distortions cancel out. Gains (or losses) in income in one region are offset by losses (or gains) in another region through adjustments in the price index.

5 Model calibration

To quantify the role of MNEs and examine their interaction with distortions at the aggregate level in the Mexican context, we calibrate the model parameters using data and established methods. Table 7 provides an overview of the parameters required for the calibration. These parameters include: (i) scale parameters, such as $A_{j,s}$ (total factor productivity), $B_{j,s}$ (labor supply efficiency), and a_c (amenities); (ii) the expenditure share of MNEs on labor, β; (iii) iceberg trade costs between commuting zones, τ_{jc}; (iv) the production fixed cost for formal firms, FC_j; (v) the mass of firms in each region, $\mathcal{M}_j$; (vi) key elasticities, including demand elasticity (σ), labor supply elasticity across firm types (κ), and migration elasticity (η); (vii) wedges that capture distortions in output and labor markets, such as $t^L_{j,s}$, $t^Y_{j,s}$, and b_j; and (viii) the spillover effects of foreign MNEs on the domestic economy, represented by $\gamma_{M,s}$.

We take the following parameters from the literature: the trade elasticity, $\sigma = 6$ from estimates in trade studies; the labor supply elasticity across firm types, $\kappa = 2.2$ from Zárate (2022); and the migration elasticity, $\eta = 2$ from Monte et al. (2018).

Parameters from the data: We take the following parameters directly from the data:

- *Wedges*: Some of the most important parameters to calibrate are the wedges in output and labor that measure the gaps between prices and marginal costs. To calibrate these wedges we use the information from section 2.2. In the case of the output/value-added distortions, we include: VA taxes, distortions from crime, bureaucracy, corruption, and tariffs, and in the case of labor we include social security contributions and payroll taxes. With this information, we calibrate $t^L_{j,s}$ and $t^Y_{j,s}$ for all types of establishments and locations.

  For the parameter b_j that captures how the marginal cost responds to changes in labor size for informal firms, we use information on audits and calibrate b_j from a regression relating the probability of being audited by the Mexican Government with firm size. We provide more details on this procedure in Appendix B.

- *Iceberg trade costs*: We follow the trade literature and in particular Faber and Gaubert (2019) and calculate iceberg trade costs as a function of the euclidean distance in kilometers (kms) between each pair of commuting zones. We use the following functional form $\tau_{jc}^{1-\sigma} = D_{jc}^{-\zeta}$, where D_{jc} corresponds to the distance from centroid j to centroid c in kms and ζ is the decay distance parameter elasticity. We follow the literature and calibrate a value of $\zeta = 1$.

- *Expenditure share of foreign MNEs on intermediate inputs:* We calibrate this parameter using the empirical average for the value-added share in Mexico conditional on the wedges,

$$\beta = \frac{1}{JSN}\left(\sum_{j,s,n\in\Omega_{j,s}}\left(\frac{1+t^L_{j,s}}{1-t^Y_{j,s}}\right)\frac{w_{j,s}l_{j,s}(n)}{r_{j,s}(n)}\right),$$

  where J, S, N correspond to the number of locations, firm types and total firms in the economy.

- *Initial distribution of wages and the mass of firms:* We calibrate the initial distribution of wages using the average of wages reported by establishments at the commuting zone and firm-type level in 1994. Similarly, we calibrate the mass of firms in each location, $\mathcal{M}_j$, using the number of domestic firms observed in the data.

- *Productivity distribution:* We assume that all locations have the same productivity distribution $\vartheta_j(\varphi) = \vartheta(\varphi)$ for all j. We assume that the productivity draws φ are drawn from a Pareto distribution with shape parameter θ. We use the distribution of total sales to recover θ and construct bounds for the distribution using the values from Zárate (2022) that calibrates these parameters with the same data.

Model inversion to recover scale parameters: With the elasticities at hand and observing the distribution of wages and employment in Mexico, we can invert the model to recover the scale parameters, such as TFP at the firm-type location level, $A_{j,s}$, the amenities vector at the location

level, a_c, and the relative efficiency parameters in the labor supply function, $B_{j,s}$. We define the geometric mean of variable, x_j, as $\bar{x}_j = \exp\left(\frac{1}{S}\sum_{s=1}^{S}\ln(x_{j,s})\right)$ and we normalize the variable as $\tilde{x}_j = x_j/\bar{x}_j$. We do a similar normalization across locations. We invert the model to recover the following parameters:

- *Labor supply efficiency parameters:* We recover, B_{js} with the following formula:

$$B_{js} = \frac{\tilde{\lambda}_{js|j}}{\tilde{w}_{j,s}^{\kappa}},$$

  with this equation we match perfectly the employment shares predicted by the model with the ones observed in the data in the initial period.

- *Total factor productivity at the location firm-type level:* With data on wages and employment, we solve the model for the vector $A_{j,s}$ using the algorithm from section Appendix E.1 to match the equilibrium implied by the model with the observed data.

- *Fixed cost for formal firms:* In each location, we observe the share of firms that are formal. We calibrate the productivity value $\tilde{\varphi}_j$ in the initial equilibrium to match the share of formal domestic firms and solve the fixed cost of formal firms FC_j, such that firm $\tilde{\varphi}_j$ is indifferent between operating in the formal and informal sectors in the initial equilibrium.

Inference matching: Conditional on the demand, supply, and migration elasticities, the relative change between domestic formal and informal employment is determined by the spillovers from foreign MNEs to the local domestic economy. We identify these spillover elasticities $\gamma_{M,F}$, and $\gamma_{M,I}$ by looking at the reduced-form coefficient on the relative effect of MNE expansion on the formal and informal employment at the local level. On the one hand, higher spillovers increase total employment for both types of producers. On the other hand, the relative change in employment between the two types of domestic producers determines which type of domestic producers benefit more from the spillovers. In particular, with the data generated by the model we run the following regression:

$$\frac{\Delta y_c}{L_{c,0}} = \beta_y^M \frac{\Delta L_{c,M}}{L_{c,0}} + \epsilon_c, \tag{40}$$

where Δy_c corresponds to the change implied by the model after running the counterfactual of an MNE expansion in our outcomes of interest: i) formal employment, and ii) informal employment; $\Delta L_{c,M}$ corresponds to the change in foreign employment implied by the model in the counterfactual, $L_{c,0}$ represents the initial population in the commuting zone, ϵ_c is the error term of the regression, and β_y^M represents our coefficient of interest for outcome y. We calculate the following moment that matches the coefficient implied by the model with the point estimates of the reduced-form $\Gamma_y = \beta_y^M - \beta_y$ where β corresponds to the IV coefficient from equation (2) for outcome y. The

indirect inference procedure finds the combination of spillover parameters such that:

$$(\hat{\gamma}_{M,F}, \hat{\gamma}_{M,I}) = argmin \;\; \mathbf{\Gamma}(\gamma_{M,F}, \gamma_{M,I})' \mathbf{W} \mathbf{\Gamma}(\gamma_{M,F}, \gamma_{M,I}), \tag{41}$$

where $\mathbf{W}$ is a weighting matrix for which we use the inverse of the variance of the point estimates of the regression from equation (40). Figure 7 plots the loss function of the previous equation for different values of the spillovers and Figure E1 in the appendix plots the results for each spillover holding the other one constant for its respective employment outcome. From the minimization problem, we find a value of $\gamma_{M,F} = 0.125$ and $\gamma_{M,I} = 0.025$, meaning that a 10% increase in foreign employment increases the productivity of formal (informal) domestic firms by 1.2% (0.25%). The positive spillover effects also suggest that from a social planner perspective, MNEs should face lower wedges since they generate a positive externality to the local economy.

6 Counterfactuals

To shed light on the impact of foreign MNEs on real income and the misallocation channel, we run a simulation similar to the one from AD. In particular, for all the places that experience a positive increase in foreign employment in our period of study, we simulate a 10% productivity increase for foreign MNEs. Around 25% commuting zones in Mexico experience the shock. We combine in a single category the foreign MNEs that are maquila and non-maquila.

Table 6 reports the average effect on real income across commuting zones. Column (1) shows the results for the case in which there are distortions and spillovers, column (2) removes the distortions, column (3) the spillovers, and column (4) both market failures. Similarly, Figure 8 plots the percentage change of real income across commuting zones for the four scenarios.

We find that the productivity shock increases real income by an average of 6.4%. This effect is 5% for CZs that do not directly experience the productivity increase from foreign MNEs and 8% for CZs that do. In a hypothetical scenario without pre-existing distortions or spillovers, the real income gains would average around 4%, indicating that distortions and externalities amplify the gains from foreign MNEs by 55%. Column 2 presents results for a counterfactual scenario that includes spillovers but excludes distortions, isolating the role of distortions in shaping overall gains. This result highlights the contribution of reallocating economic activity from the informal (less distorted) to the formal (more distorted) sector. We estimate that this reallocation generates gains of approximately 1.5% (the difference between columns 1 and 2), accounting for 25% of the total gains. Conversely, column 3 isolates the effect of productivity spillovers by considering a counterfactual that includes distortions but excludes spillovers. The results suggest that productivity externalities generate gains of about 1.2% (the difference between columns 1 and 3), explaining 20% of the total gains attributed to foreign MNEs.

Moreover, the importance of each channel varies across the commuting zones that experienced the shock directly vs. the ones that did not. We find that the impact of distortions is greater

in locations that do not directly experience the productivity shock, suggesting that trade linkages between commuting zones facilitate the reallocation of economic activity from low-distortion to high-distortion establishments, even in areas without foreign MNEs. This result aligns with the literature showing that trade openness can reduce informality (McCaig and Pavcnik, 2018), as formal establishments in untreated locations can supply goods and services to MNEs in treated locations. For instance, Figure 8 illustrates that most locations experience larger gains when distortions are taken into account, even when potential spillovers are muted. Spillovers account for most of the gains in locations that directly experience the shock, explaining about two-thirds of the market failure-driven improvements. Panel (b) of Figure 8 further shows that spillovers are more localized, directly benefiting only the zones that experience the productivity shock since MNEs generate local spillovers.

Overall, this counterfactual exercise suggests that distortions and spillovers can amplify the gains from foreign MNEs by a significant amount in low and middle-income countries – even in locations within those countries without foreign MNE presence (due to trade and labor linkages).

7 Conclusion

Countries worldwide provide favorable tax and regulatory treatment to foreign MNEs over domestic firms, justifying this policy non-neutrality with the expectation of productivity spillovers to the domestic economy. Meanwhile, low- and middle-income countries (LMICs) face a second-best world, where multiple distortions keep (typically more productive) parts of their economy inefficiently small. Thus, a complete study of the welfare gains from MNEs in LMICs must account not only for their potential productivity spillovers but also for their impact on the allocation of resources across differently distorted establishments and locations. Such a study poses challenges, including the need for credible measures of distortions, data on all establishments (including informal and nonmanufacturing), reduced-form evidence of the reallocative effects of MNEs, and a general equilibrium model to assess their aggregate welfare effects. We address these in four steps.

First, we build a rich establishment-level panel tracking between 1994 and 2019, the near-universe of four types of establishments in Mexico: domestically-owned establishments (labeled domestic and split into formal or informal), and foreign MNEs (split into maquila and non-maquila). We also construct a novel database with direct measures of distortions, including crime, subsidies, and labor taxes, which differ across the four establishment types and commuting zones (CZs) in Mexico. We find evidence of widespread distortions that vary within CZs by establishment type and across CZs for a given establishment type.

Second, we examine the local reallocation effects of changes in per capita foreign MNE employment within a CZ. Drawing on the migration literature, we use the historical spatial clustering of foreign MNEs by origin country to construct an instrument for the endogenous component of their local employment changes. We find that increased exposure to foreign MNE employment in a CZ boosts the size of the local domestic economy, primarily driven by the expansion of the

formal sector, while informal establishments lose relative importance. Additionally, we find that only non-maquila foreign MNEs have positive effects on the local domestic economy, consistent with their higher productivity and stronger integration into Mexico's economy.

Third, to study the aggregate welfare effects of foreign MNEs and quantify the importance of the allocative efficiency margin in these effects, we develop a quantitative model with heterogeneous domestic firms and foreign MNEs. In our model, domestic firms endogenously decide whether to operate in the formal or the informal economy, and MNEs interact with the local economy through two channels: (i) they generate productivity externalities, and (ii) they establish demand linkages (notably, with formal firms). We show that, theoretically, the effect of foreign MNEs on formal and informal employment depends on the relative importance of demand linkages and the difference in productivity externalities. We also show that, at the aggregate level, the overall impact of MNEs on allocative efficiency depends on whether they reallocate resources toward segments of the economy that were initially inefficiently small due to distortions, or vice versa.

Fourth and lastly, we take the model to the data to quantify and decompose the welfare effects of foreign MNEs. We calibrate the distortions and wedges directly from our measures of distortions and recover the scale model parameters, such as, total factor productivity and amenities by inverting the model and matching the initial equilibrium perfectly with the data. We also calibrate the productivity externalities through an indirect inference approach, finding the combination of productivity spillovers from foreign MNEs to domestic firms that replicate the reduced-form coefficients through the lens of the model. The results suggest higher productivity externalities from foreign MNEs to formal domestic firms than informal ones. Armed with the model and the elasticities at hand, we simulate a 10% productivity increase of foreign MNEs and analyze the overall welfare gains and the importance of market failures. We find that the MNE productivity shocks result in 6% aggregate welfare gains. A quarter of these gains stem from reallocation effects in the presence of distortions, while a fifth is attributed to the productivity spillovers of foreign MNEs to the local economy. In our setting, the gains from foreign MNEs increase when we account for pre-existing distortions, as MNEs reallocate economic activity from less distorted to more distorted establishments that were initially inefficiently small.

References

Girum Abebe, Margaret McMillan, and Michel Serafinelli. Foreign Direct Investment and Knowledge Diffusion in Poor Locations. Journal of Development Economics, 158, 2022.

Rodrigo Adao, Costas Arkolakis, and Federico Esposito. General Equilibrium Effects in Space: Theory and Measurement. Technical report, National Bureau of Economic Research, 2019.

Benjamin Aleman-Castilla. The Effect of Trade Liberalization on Informality and Wages: Evidence from Mexico. CEP-LSE Working Paper, -(763), 2006.

Alonso Alfaro-Ureña, Isabela Manelici, and Jose P. Vasquez. The Effects of Multinationals on Workers: Evidence from Costa Rican Microdata. Working Paper, 2021.

Alonso Alfaro-Ureña, Benjamin Faber, Cecile Gaubert, Isabela Manelici, and Jose P. Vasquez. Responsible Sourcing? Theory and Evidence from Costa Rica. NBER Working Paper 30683, 2022a.

Alonso Alfaro-Ureña, Isabela Manelici, and Jose P Vasquez. The Effects of Joining Multinational Supply Shains: New Evidence from Firm-to-Firm Linkages. Quarterly Journal of Economics, 137(3):1495–1552, 2022b.

Jorge A Alvarez and Cian Ruane. Informality and Aggregate Productivity: The Case of Mexico. European Economic Review, 167:104791, 2024.

Vanessa Alviarez, Laura Boudreau, Evangelina Dardati, Jingting Fan, Farid Farrokhi, Ezequiel Garcia-Lembergman, Stefania Garetto, Grace Gu, Galina Hale, David Hemous, Nicola Limodio, Isabela Manelici, Ralf Martin, Nicolas Morales, Nitya Pandalai-Nayar, Nina Pavcnik, Heitor Pellegrina, Natalia Ramondo, Jose P. Vasquez, and Pierre-Louis Vézina. VoxDevLit on "Foreign Direct Investment and Development". VoxDevLit, 13, 2025.

Joshua D Angrist and Jörn-Steffen Pischke. Mostly Harmless Econometrics. Princeton University Press, 2009.

Costas Arkolakis, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple. Innovation and Production in the Global Economy. American Economic Review, 108(8):2128–2173, 2018.

Costas Arkolakis, Arnaud Costinot, Dave Donaldson, and Andrés Rodríguez-Clare. The Elusive Pro-Competitive Effects of Trade. Review of Economic Studies, 86(1):46–80, 2019.

Jose Asturias, Manuel García-Santana, and Roberto Ramos. 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, 2019.

Jose Asturias, Marco Sanfilippo, and Asha Sundaram. Foreign Investment, Geography, and Welfare. Working Paper, 2022.

David Atkin. Endogenous Skill Acquisition and Export Manufacturing in Mexico. American Economic Review, 106(8):2046–85, 2016.

David Atkin and Dave Donaldson. The Role of Trade in Economic Development. Handbook of International Economics, 5:1–59, 2022.

David Atkin and Amit K Khandelwal. How Distortions Alter the Impacts of International Trade in Developing Countries. Annual Review of Economics, 12:213–238, 2020.

David Atkin, Benjamin Faber, and Marco Gonzalez-Navarro. Retail Globalization and Household Welfare: Evidence from Mexico. Journal of Political Economy, 126(1):1–73, 2018.

David Atkin, Laura Boudreau, Rafael Dix-Carneiro, Amit Khandelwal, Isabela Manelici, Brian McCaig, Pamela Medina, Ameet Morjaria, Luigi Pascali, Bob Rijkers, and Meredith Startz. VoxDevLit on "International Trade". VoxDevLit, 4, 2022.

David Atkin, Baptiste Bernadac, David Donaldson, Tishara Garg, and Federico Huneeus. The Incidence of Distortions. Working Paper, 2024.

David H Autor, David Dorn, and Gordon H Hanson. The China Syndrome: Local Labor Market Effects of Import Competition in the United States. American Economic Review, 103(6):2121–2168, 2013.

Yan Bai, Keyu Jin, and Dan Lu. Misallocation under Trade Liberalization. American Economic Review, 114(7):1949–1985, 2024.

David Baqaee and Emmanuel Farhi. Networks, Barriers, and Trade. Econometrica, 2024.

David Rezza Baqaee and Emmanuel Farhi. Productivity and Misallocation in General Equilibrium. Quarterly Journal of Economics, 135(1):105–163, 2020.

Frank Barry, Holger Görg, and Eric Strobl. Foreign Direct Investment, Agglomerations, and Demonstration Effects: An Empirical Investigation. Review of World Economics, 139:583–600, 2003.

Natalie Bau and Adrien Matray. Misallocation and Capital Market Integration: Evidence from India. Econometrica, 2022.

Andrew B Bernard and Catherine Thomas. Multinational Firms' Market Entry and Expansion, with Evidence from Eastern Europe. Global Goliaths: Multinational Corporations in the 21st Century Economy, 2019.

Antoine Berthouy, John Chung, Kalina Manova, and Charlotte Sandoz. Trade, Productivity and (Mis)allocation. Working Paper, 2021.

Mark Bils, Peter J Klenow, and Cian Ruane. Misallocation or Mismeasurement? Journal of Monetary Economics, 124:S39–S56, 2021.

Nicholas Bloom, Leonardo Iacovone, Mariana Pereira-López, and John Van Reenen. Management and Misallocation in Mexico. American Economic Journal: Macroeconomics, 2025.

Juan Blyde, Matias Busso, Kyunglin Park, and Dario Romero. Short-and Long-run Labor Market Adjustment to Import Competition. Review of International Economics, 2023.

Juan S Blyde, Matias Busso, and Dario Romero Fonseca. Labor Market Adjustment to Import Competition: Long-Run Evidence from Establishment Data. IDB Working Paper Series, 2020.

Kirill Borusyak, Rafael Dix-Carneiro, and Brian Kovak. Understanding Migration Responses to Local Shocks. Working Paper, 2022a.

Kirill Borusyak, Peter Hull, and Xavier Jaravel. Quasi-Experimental Shift-Share Research Designs. Review of Economic Studies, 89(1):181–213, 2022b.

Francisco J Buera, Joseph P Kaboski, and Robert M Townsend. From Micro to Macro Development. Journal of Economic Literature, 61(2):471–503, 2023.

Matías Busso, María Victoria Fazio, and Santiago Levy Algazi. (In) formal and (Un) productive: The Productivity Costs of Excessive Informality in Mexico. Technical report, IDB working paper series, 2012.

David Card. Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. Journal of Labor Economics, 19(1):22–64, 2001.

Camila Cisneros-Acevedo. Unfolding Trade Effect in Two Margins of Informality: The Peruvian Case. World Bank Economic Review, 36(1):141–170, 2022.

Hernando De Soto. The Other Path. Harper & Row New York, 1989.

Swati Dhingra and John Morrow. Monopolistic Competition and Optimum Product Diversity Under Firm Heterogeneity. Journal of Political Economy, 127(1):196–232, 2019.

Rafael Dix-Carneiro and Brian K Kovak. Margins of Labor Market Adjustment to Trade. Journal of International Economics, 117:125–142, 2019.

Rafael Dix-Carneiro, Pinelopi K Goldberg, Costas Meghir, and Gabriel Ulyssea. Trade and Domestic Distortions: The Case of Informality. NBER Working Paper 28391, 2024.

Benjamin Faber and Cecile Gaubert. Tourism and Economic Development: Evidence from Mexico's Coastline. American Economic Review, 109(6):2245–2293, 2019.

Pablo D Fajgelbaum and Cecile Gaubert. Optimal Spatial Policies, Geography, and Sorting*. The Quarterly Journal of Economics, 135(2):959–1036, 01 2020. ISSN 0033-5533. doi: 10.1093/qje/qjaa001. URL https://doi.org/10.1093/qje/qjaa001.

Robert C Feenstra and Gordon H Hanson. Foreign Direct Investment and Relative Wages: Evidence from Mexico's Maquiladoras. Journal of International Economics, 42(3-4):371–393, 1997.

Lucie Gadenne, Roland Rathelot, and Tushar Nandi. Taxation and Supplier Networks: Evidence from India. Working Paper, 2024.

Simon Galle, Andrés Rodríguez-Clare, and Moises Yi. Slicing the pie: Quantifying the aggregate and distributional effects of trade. Review of Economic Studies, 90(1):331–375, 2023.

Gordon H Hanson. Why Isn't Mexico Rich? Journal of Economic Literature, 48(4):987–1004, 2010.

Ann Harrison and Andrés Rodríguez-Clare. Trade, Foreign Investment, and Industrial Policy for Developing Countries. Handbook of Development Economics, 5:4039–4214, 2010.

Keith Head, John Ries, and Deborah Swenson. Agglomeration Benefits and Location Choice: Evidence from Japanese Manufacturing Investments in the United States. Journal of International Economics, 38(3-4):223–247, 1995.

Hugo A Hopenhayn. Firms, Misallocation, and Aggregate Productivity: A Review. Annual Review of Economics, 6(1):735–770, 2014.

Chang-Tai Hsieh and Peter J Klenow. Misallocation and Manufacturing TFP in China and India. Quarterly Journal of Economics, 124(4):1403–1448, 2009.

Chang-Tai Hsieh and Peter J Klenow. The Life Cycle of Plants in India and Mexico. Quarterly Journal of Economics, 129(3):1035–1084, 2014.

Chang-Tai Hsieh and Benjamin A Olken. The Missing "Missing Middle". Journal of Economic Perspectives, 28(3):89–108, 2014.

Chang-Tai Hsieh, Erik Hurst, Charles I. Jones, and Peter J. Klenow. The allocation of talent and u.s. economic growth. Econometrica, 87(5):1439–1474, 2019. doi: 10.3982/ECTA11427. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA11427.

Beata S. Javorcik. Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages. American Economic Review, 94(3):605–627, 2004.

Timothy J Kehoe and Kim J Ruhl. Why Have Economic Reforms in Mexico Not Generated Growth? Journal of Economic Literature, 48(4):1005–1027, 2010.

Anne O Krueger. Trade Policies in Developing Countries. Handbook of international economics, 1:519–569, 1984.

David Lagakos and Michael E. Waugh. Selection, agriculture, and cross-country productivity differences. American Economic Review, 103(2):948–80, April 2013. doi: 10.1257/aer.103.2.948. URL http://www.aeaweb.org/articles?id=10.1257/aer.103.2.948.

Andrea Lanteri, Pamela Medina, and Eugene Tan. Capital-Reallocation Frictions and Trade Shocks. American Economic Journal: Macroeconomics, 15(2):190–228, 2023.

Santiago Levy. Good Intentions, Bad Outcomes: Social Policy, Informality and Economic Growth in Mexico. Washington, D.C. The Brookings Institute Press, 2008.

Santiago Levy. Under-Rewarded Efforts: The Elusive Quest for Prosperity in Mexico. Inter-American Development Bank, 2018.

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

Brian McCaig, Nina Pavcnik, and Woan Foong Wong. Foreign and domestic firms: Long run employment effects of export opportunities. CESifo Working Paper, 2022.

Esteban Méndez and Diana Van Patten. Multinationals, Monopsony, and Local Development: Evidence From the United Fruit Company. Econometrica, 90(6):2685–2721, 2022.

Florian Misch and Christian Saborowski. The Drivers and Consequences of Resource Misallocation. Economía, 20(2):61–96, 2020.

Ferdinando Monte, Stephen J. Redding, and Esteban Rossi-Hansberg. Commuting, Migration, and Local Employment Elasticities. American Economic Review, 108(12):3855–3890, 2018.

Benjamin A Olken and Rohini Pande. Corruption in Developing Countries. Annual Review of Economics, 4(1):479–509, 2012.

Jennifer P Poole. Knowledge Transfers from Multinational to Domestic Firms: Evidence from Worker Mobility. Review of Economics and Statistics, 95(2):393–406, 2013.

Natalia Ramondo and Andrés Rodríguez-Clare. Trade, Multinational Production, and the Gains from Openness. Journal of Political Economy, 121(2):273–322, 2013.

Diego Restuccia and Richard Rogerson. The Causes and Costs of Misallocation. Journal of Economic Perspectives, 31(3):151–74, 2017.

Dani Rodrik. Policy Targeting with Endogenous Distortions: Theory of Optimum Subsidy Revisited. Quarterly Journal of Economics, 102(4):903–911, 1987.

Alexander Rothenberg, Yao Wang, and Amalavoyal Chari. When Regional Policies Fail: An Evaluation of Indonesia's Integrated Economic Development Zones. Working Paper, 2024.

Sandra Sequeira and Simeon Djankov. Corruption and Firm Behavior: Evidence from African Ports. Journal of International Economics, 94(2):277–294, 2014.

Bradley Setzler and Felix Tintelnot. The Effects of Foreign Multinationals on Workers and Firms in the United States. Quarterly Journal of Economics, 136(3):1943–1991, 2021.

Amarendra Sharma and Oscar Cardenas. The Labor Market Effects of FDI: A Panel Data Evidence from Mexico. International Economic Journal, 32(4):572–588, 2018.

Felix Tintelnot. Global Production with Export Platforms. Quarterly Journal of Economics, 132(1): 157–209, 2017.

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

Gabriel Ulyssea. Informality: Causes and Consequences for Development. Annual Review of Economics, 12, 2020.

Gabriel Ulyssea and Vladimir P Ponczek. Enforcement of Labor Regulation and the Labor Market Effects of Trade: Evidence from Brazil. Economic Journal, 132, 2022.

Eric Verhoogen. Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector. Quarterly Journal of Economics, 123(2):489–530, 2008.

Eric Verhoogen. Industrial Structure and Innovation: Notes Toward a New Strategy for Industrial Development in Mexico. Boletin Informativo Techint, Edición Especial, Oct., 2013.

Román David Zárate. Spatial Misallocation, Informality and Transit Improvements: Evidence from Mexico City. World Bank Policy Research Working Paper 9990, 2022.

Figures

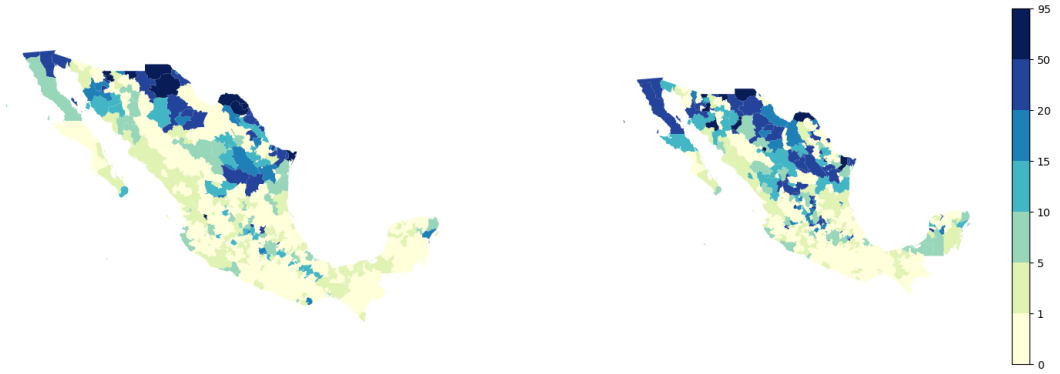
Figure 1: Percentage of CZ-level employment in domestic informal establishments



(a) 1994 (b) 2019

Notes: Figure 1 presents the percentage of CZ-level employment in domestic informal establishments in our initial (1994) and last year of analysis (2019).

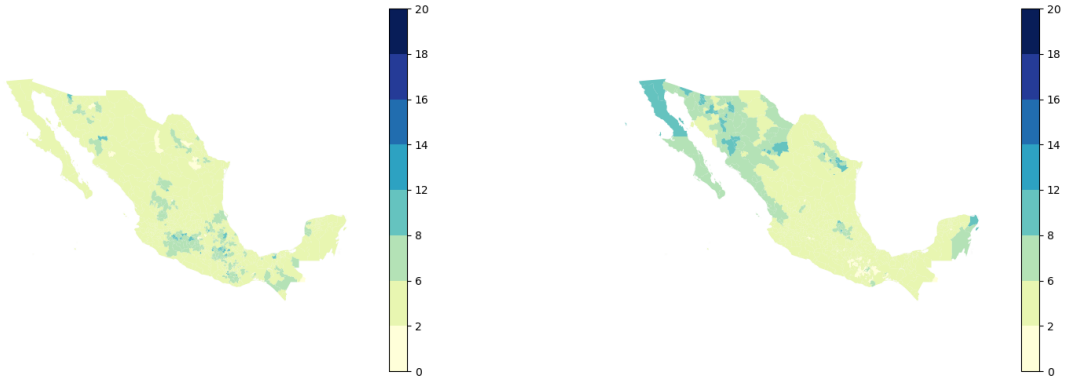
Figure 2: Percentage of CZ-level employment in foreign MNEs



(a) 1994 (b) 2019

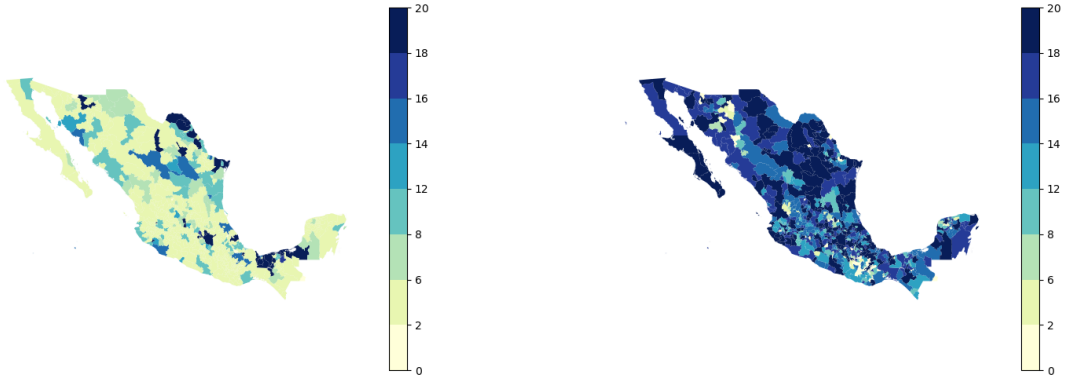
Notes: Figure 2 presents the percentage of CZ-level employment in foreign MNEs in our initial (1994) and last year of analysis (2019).

Figure 3: Spatial distribution of distortions by commuting zone



(a) Crime expenses

(b) Bureaucracy expenses

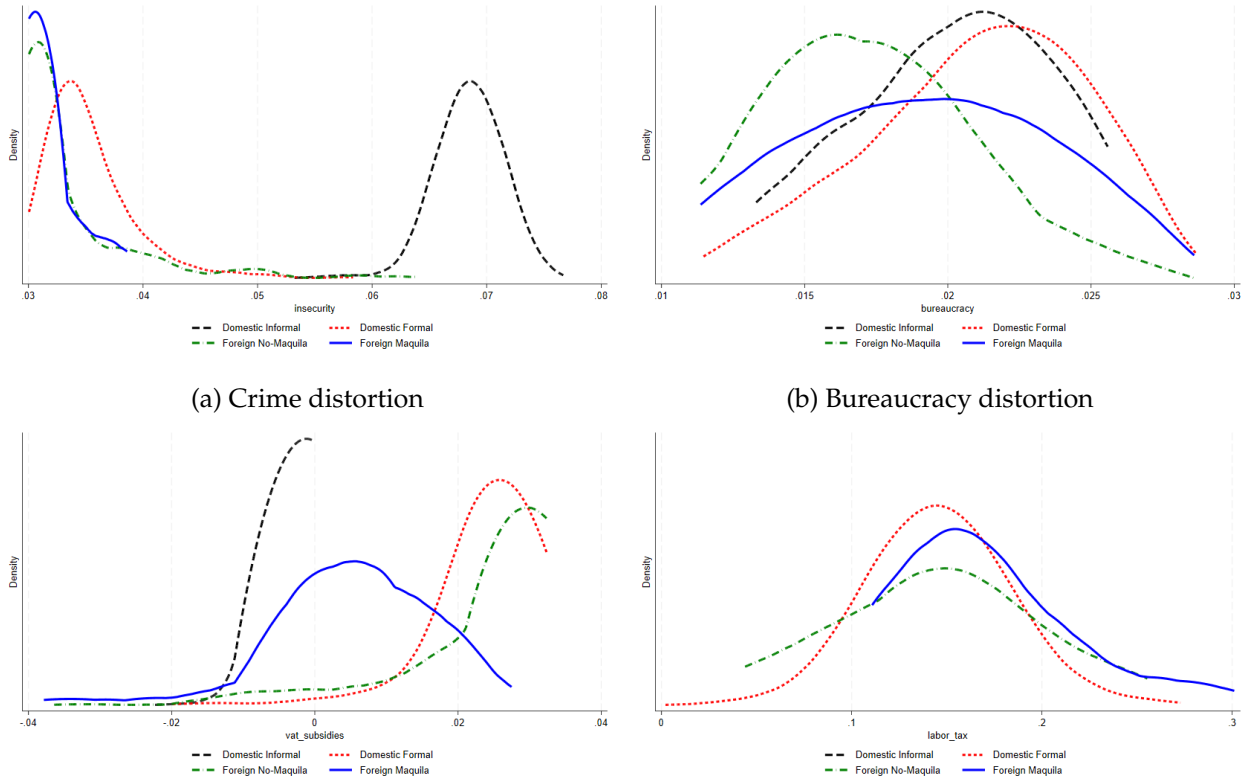


(c) Net sales taxes

(d) Labor tax

Notes: Figure 3 shows the weighted averages of distortions in each CZ as shares of annual sales for our initial year 1994. The weights are the number of workers in each CZ-industry-establishment type. Namely, we plot the expenditure of producers on crime (adding expenses in bribes, crime, and security), bureaucracy (extra payments to governments and time expenses), the net share of taxes (taxes and subsidies), and labor taxes for each CZ in Mexico.

Figure 4: Empirical density of each distortion by establishment-type and CZs



(a) Crime distortion

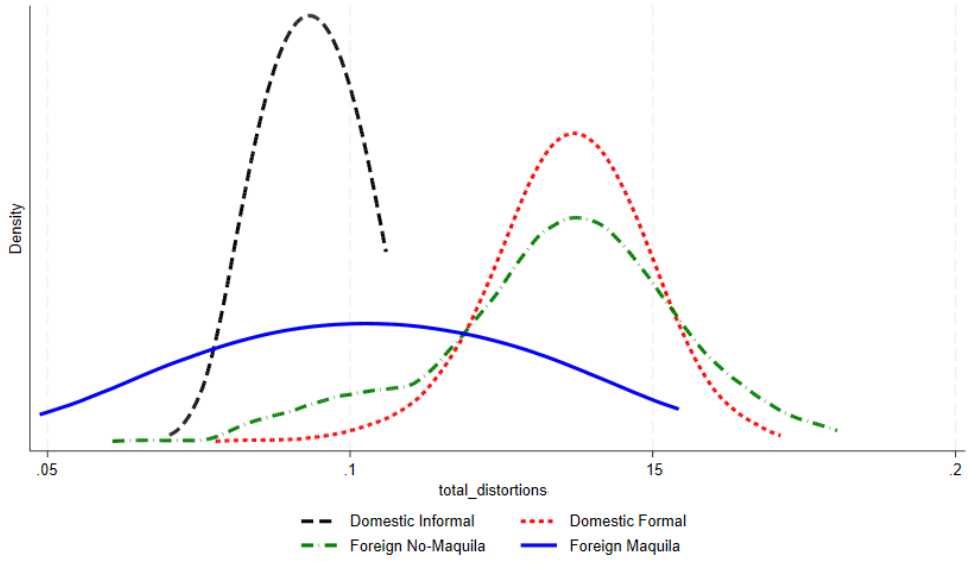
(b) Bureaucracy distortion

(c) VAT & subsidies distortion

(d) Labor tax distortion

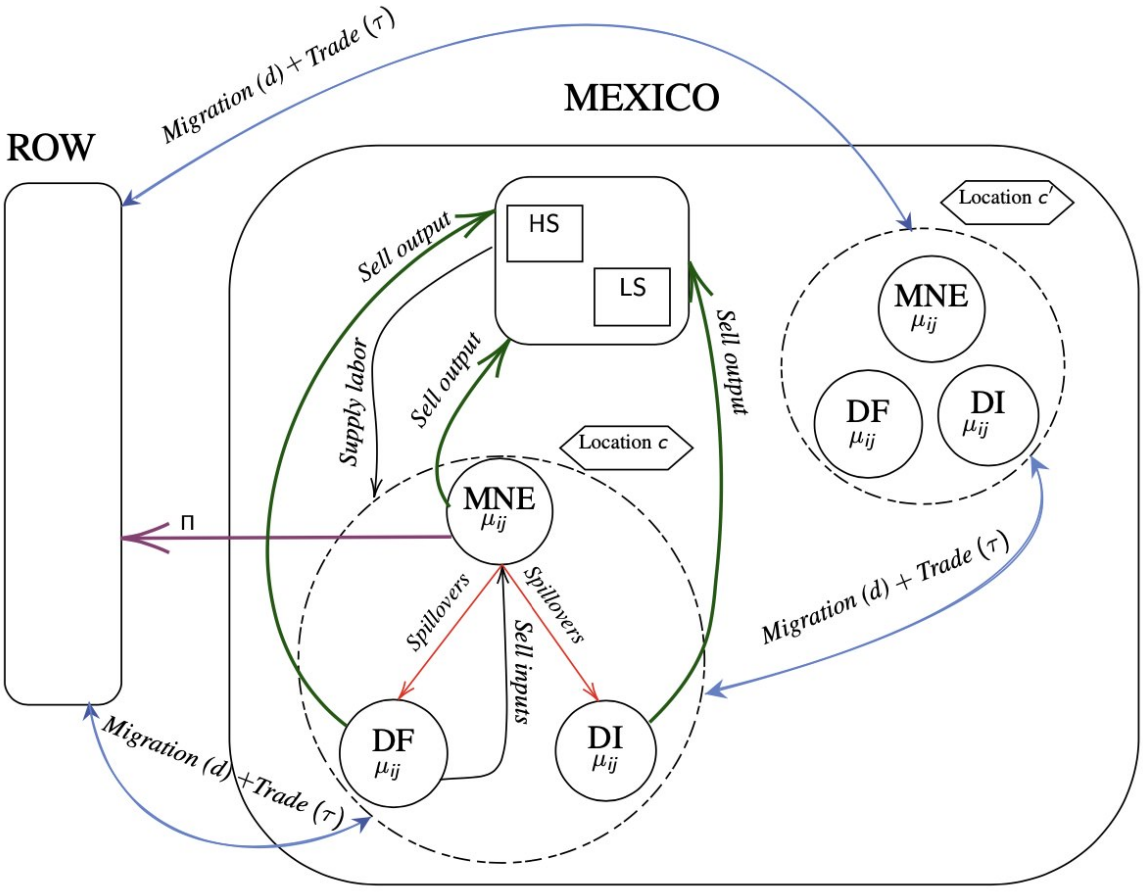
Notes: Figure 4 shows four empirical density plots of each distortion in 1994, including one histogram for each establishment type (domestic formal, domestic informal, foreign maquila MNEs, and foreign non-maquila MNEs) across all 455 CZs in Mexico. Panels (a)-(c) show each output distortion measured as the monetary cost divided by total sales. Panel (d) measures the labor input distortion measured as the cost of labor taxes divided by wage bill. Note that informal establishments do not pay labor taxes by construction, so their histogram is excluded from panel (d). Figure B1 shows the analogous density plots after controlling for industry and establishment size. Section 2.2.1 provides more details.

Figure 5: Empirical density of an "index" of total distortions by establishment-type and CZs



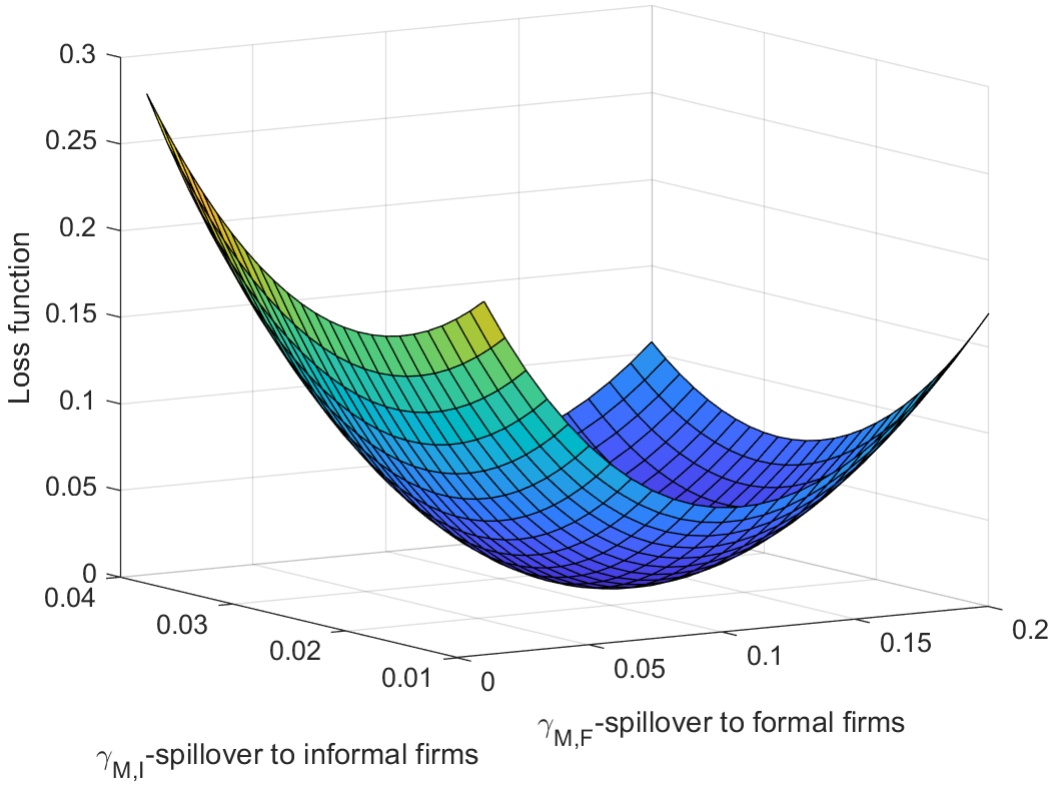
Notes: Figure 5 includes an empirical density plot of the "total distortion" in 1994, one histogram for each establishment type (domestic formal, domestic informal, foreign maquila MNEs, and foreign non-maquila MNEs) across all 455 CZs in Mexico. We build the index of total distortions by computing $\frac{1+t^Y}{1-t^L}$, where the total output distortion (t^Y) is the sum of all wedges influencing output decisions, and the total input distortion (t^L) is the equivalent from the input side. Section 2.2.1 provides more details.

Figure 6: Model sketch



Notes: Figure 6 is a sketch of the main agents and locations in our model environment and the interactions between them. The model features multiple locations (CZs in Mexico and an aggregate rest of the world, RoW) and allows for trade and migration (at costs τ and d), several producer types (foreign MNEs (MNE), domestic formal (DF) and domestic informal (DI)), and two worker types: low- and high-skill (LS and HS). Each producer type in each location potentially faces a different vector of wedges and distortions (μ_{ij}). Foreign MNEs send their profits Π to the rest of the world (RoW). For sketch simplicity, we only depict foreign MNEs as a whole, even if, in the model, we separate them further into foreign MNEs that are in the maquila programs and those that are not.

Figure 7: Loss function to calibrate the spillovers of MNEs to formal domestic ($\gamma_{D_F,F}$) and informal domestic ($\gamma_{D_I,F}$) producers



Notes: This figure plots the loss function to calibrate the spillover elasticities of foreign MNEs to the local economy using the inference matching procedure. We use a grid of $\gamma_{M,F} \in [0, 0.2]$ and $\gamma_{M,I} \in [0, 0.04]$. The coefficients to match correspond to the effect of foreign MNEs on formal and informal domestic employment from table 4. Section 5 provides more details on the indirect inference approach that we use to calibrate these parameters.

Figure 8: Percentage changes in welfare across CZs due to the foreign MNEs productivity shocks



(a) Gains with distortions and spillovers

(b) Gains with no distortions and spillovers



(c) Gains with distortions and no spillovers

(d) Gains with no distortions and no spillovers

Notes: Figure 8 plots the changes in welfare for each CZ in Mexico as a result of the local foreign MNEs productivity shocks. Panel (a) reports the results when we simulate the economy with distortions and spillover effects from increases in foreign MNEs employment. Panel (b) reports the results in the economy with spillovers, but no without distortions. Panel (c) with distortions but no spillovers, and panel (d) without distortions and without spillovers. Section 6 provides more details.

Tables

Table 1: Average distortions (%) in Mexico by establishment type

Establishment Type	Taxes and Subsidies		Weak Institutions and Rule of Law	
	VAT & Subsidies (1)	Labor Contrib. (2)	Bureaucracy & Regulation (3)	Crime & Corruption (4)
Domestic Informal	-0.1	0.0	1.9	6.8
Domestic Formal	2.3	19.5	2.2	3.5
Foreign Non-Maquila MNEs	2.2	19.5	2.1	3.3
Foreign Maquila MNEs	1.3	18.5	1.9	3.2

Notes: Table 1 reports weighted averages of distortion values by establishment type (domestic informal, domestic formal, foreign non-maquila MNEs, and foreign maquila MNEs). The weights in the averages across CZs are the total number of workers in each CZ. The distortion value for each CZ for a given category (e.g., domestic informal establishments) is a weighted average across values for all establishments in that category (where weights are the number of workers in each establishment). VAT & subsidies distortions (column 1) are reported as shares of total net taxes (taxes minus subsidies) in total annual sales. Labor taxes (column 2) are reported as a share among the total annual wage bill. Distortions related to bureaucracy & regulation (column 3) and crime & corruption (column 4) are reported as shares among total annual sales. All values are presented as percentages. Section 2.2.1 provides more details.

Table 2: OLS and IV estimates of the effects of foreign MNEs exposure on the domestic economy

(a) OLS

	Pop	Sales	VA	Capital	Workers	Wage Bill
	(1)	(2)	(3)	(4)	(5)	(6)
Overall	0.088**	9.384***	3.088**	1.846***	0.146***	0.707***
Domestic	(0.036)	(1.999)	(1.222)	(0.488)	(0.027)	(0.141)
Obs	2275	2275	2275	2275	2275	2275

(b) IV

	Pop	Sales	VA	Capital	Workers	Wage Bill
	(1)	(2)	(3)	(4)	(5)	(6)
Overall	1.357***	35.708***	15.231***	6.227***	0.333***	3.532***
Domestic	(0.14)	(3.326)	(0.662)	(0.769)	(0.109)	(1.240)
Obs	2275	2275	2275	2275	2275	2275
F-stat	107.8	107.8	107.8	107.8	107.8	107.8

Notes: Table 2 presents the OLS and IV estimates of the regression described in equation (2). The population variable includes the working-age population only. Sales, value-added, capital (assets), and wage bills are expressed in 2018 CPI-deflated thousands of U.S. dollars per capita. Workers are measured as headcounts. Observations are weighted by the initial CZ domestic employment. We cluster the standard errors at the census division level. We use as controls economic region-year fixed effects, measures of import and export exposure, and various time-variant CZ-level controls. Sections 3.1, 3.2, and Appendix A.4 provide more details about the regression specification and controls.

Table 3: IV estimates of the effects of foreign MNEs exposure on the domestic economy. By foreign MNE type: all, only maquila or only non-maquila

(a) All Foreign

	Pop (1)	Sales (2)	VA (3)	Capital (4)	Workers (5)	Wage Bill (6)
Overall Domestic	1.357*** (0.14)	35.708*** (3.326)	15.231*** (0.662)	6.227*** (0.769)	0.333*** (0.109)	3.532*** (1.240)
Obs	2275	2275	2275	2275	2275	2275
F-stat	107.8	107.8	107.8	107.8	107.8	107.8

(b) Maquila

	Pop (1)	Sales (2)	VA (3)	Capital (4)	Workers (5)	Wage Bill (6)
Overall Domestic	0.970 (0.848)	40.891 (39.550)	7.228 (10.733)	2.363 (7.170)	-0.738* (0.379)	-2.130 (1.742)
Obs	2275	2275	2275	2275	2275	2275
F-stat	157.79	157.79	157.79	157.79	157.79	157.79

(c) Non-Maquila

	Pop (1)	Sales (2)	VA (3)	Capital (4)	Workers (5)	Wage Bill (6)
Overall Domestic	2.664*** (0.351)	69.843*** (8.880)	36.127*** (3.623)	16.636*** (2.720)	1.342*** (0.379)	8.630*** (2.903)
Obs	2275	2275	2275	2275	2275	2275
F-stat	227.27	227.27	227.27	227.27	227.27	227.27

Notes: Table 3 presents the IV estimates of the regression described in equation (4). The three panels differ in the set of foreign MNEs whose effects we estimate (and the subsequent controls used). In the upper panel, we keep the baseline sample of all foreign MNEs, which corresponds to the regression in equation (2) and the results presented in Table 2. In the middle panel, we focus only on the exposure to maquila foreign MNEs. In these regressions, we control for the change in exposure to non-maquila foreign MNEs. In the lower panel, we focus only on the exposure to non-maquila foreign MNEs. In these regressions, we control for the change in exposure to maquila foreign MNEs. The population variable includes the working-age population only. Sales, value-added, capital (assets), and wage bills are expressed in 2018 CPI-deflated thousands of U.S. dollars per capita. Workers are measured as headcounts. Observations are weighted by the initial CZ domestic employment. We cluster the standard errors at the census division level. We use as controls economic region-year fixed effects, measures of import and export exposure, and various time-variant CZ-level controls. Sections 3.1, 3.2, and Appendix A.4 provide more details about the regression specification and controls.

Table 4: IV estimates of the effects of foreign MNEs exposure on the domestic formal vs. informal sectors

	Pop	Sales	VA	Capital	Workers	Wage Bill
	(1)	(2)	(3)	(4)	(5)	(6)
Overall Domestic	1.357***	35.708***	15.231***	6.227***	0.333***	3.532***
	(0.14)	(3.326)	(0.662)	(0.769)	(0.109)	(1.240)
Formal	–	28.142***	12.516***	4.261**	0.295***	2.952***
		(1.959)	(0.459)	(1.652)	(0.083)	(1.088)
Informal	–	7.566***	2.715***	1.966	0.038	0.580***
		(1.544)	(0.520)	(1.495)	(0.036)	(0.197)
Obs	2275	2275	2275	2275	2275	2275
F-stat	107.8	107.8	107.8	107.8	107.8	107.8

Notes: Table 4 presents the IV estimates of the regression described in equation (2). This table includes coefficients for formal and informal domestic establishments. The population variable includes the working-age population only. Sales, value-added, capital (assets), and wage bills are expressed in 2018 CPI-deflated thousands of U.S. dollars per capita. Workers are measured as headcounts. Observations are weighted by the initial CZ domestic overall, formal or informal employment. We cluster the standard errors at the census division level. We use as controls economic region-year fixed effects, measures of import and export exposure, and various time-variant CZ-level controls. Sections 3.1, 3.2, 3.3, and Appendix A.4 provide more details about the regression specification and controls.

Table 5: IV estimates of the effects of foreign MNEs exposure on the domestic formal vs. informal sectors. By foreign MNE type: all, only maquila or only non-maquila

(a) Maquila

	Pop (1)	Sales (2)	VA (3)	Capital (4)	Workers (5)	Wage Bill (6)
Overall Domestic	0.970 (0.848)	40.891 (39.550)	7.228 (10.733)	2.363 (7.170)	-0.738* (0.379)	-2.130 (1.742)
Formal	–	37.139 (23.582)	7.215 (6.917)	8.432*** (3.053)	-0.580*** (0.142)	-0.984 (0.787)
Informal	–	3.752 (16.155)	0.013 (3.851)	-6.069 (4.406)	-0.158 (0.238)	-1.146 (1.251)
Obs	2275	2275	2275	2275	2275	2275
F-stat	157.79	157.79	157.79	157.79	157.79	157.79

(b) Non-Maquila

	Pop (1)	Sales (2)	VA (3)	Capital (4)	Workers (5)	Wage Bill (6)
Overall Domestic	2.664*** (0.351)	69.843*** (8.880)	36.127*** (3.623)	16.636*** (2.720)	1.342*** (0.379)	8.630*** (2.903)
Formal	–	56.679*** (6.394)	28.706*** (2.431)	9.245*** (3.507)	1.074*** (0.267)	7.049** (2.743)
Informal	–	13.163*** (3.989)	7.421*** (2.042)	7.391* (4.395)	0.268** (0.126)	1.581*** (0.284)
Obs	2275	2275	2275	2275	2275	2275
F-stat	227.27	227.27	227.27	227.27	227.27	227.27

Notes: Table 5 presents the IV estimates of the regression described in equation (4). The two panels differ in the set of foreign MNEs whose effects we estimate (and the subsequent controls used). In the upper panel, we focus only on the exposure to maquila foreign MNEs. In these regressions, we control for the change in exposure to non-maquila foreign MNEs. In the lower panel, we focus only on the exposure to non-maquila foreign MNEs. In these regressions, we control for the change in exposure to maquila foreign MNEs. The population variable includes the working-age population only. Sales, value-added, capital (assets), and wage bills are expressed in 2018 CPI-deflated thousands of U.S. dollars per capita. Workers are measured as headcounts. Observations are weighted by the initial CZ domestic overall, formal or informal employment. We cluster the standard errors at the census division level. We use as controls economic region-year fixed effects, measures of import and export exposure, and various time-variant CZ-level controls. Sections 3.1, 3.2, 3.3, and Appendix A.4 provide more details about the regression specification and controls.

Table 6: Welfare effects: averages for three groups of CZs

	Scenario			
	Distortions + spillovers	No distortions + spillovers	Distortions + no spillovers	No distortions + no spillovers
	(1)	(2)	(3)	(4)
All CZs	6.38%	4.83%	5.10%	4.14%
	(0.024)	(0.020)	(0.023)	(0.018)
CZs with foreign MNEs	8.07%	6.57%	5.81%	4.80%
	(0.034)	(0.027)	(0.029)	(0.023)
CZs without foreign MNEs	5.84%	4.28%	4.88%	3.94%
	(0.016)	(0.013)	(0.020)	(0.016)

Notes: This table reports the results of our main counterfactual exercise, which consists of a 10% productivity increase for foreign MNEs in the locations that experience a positive increase of foreign employment between 1994 and 2019. The first column reports the results for the counterfactual that includes wedges and spillovers, column 2 removes the wedges, column 3 the spillovers, and column 4 both sources of market failures. The first row reports the simple average across all the commuting zones, while the second and third rows report the results for the locations that did not experience the productivity shock directly and those that did respectively. The standard deviation is reported in parentheses. Section 6 provides more details.

Table 7: Summary of Parameters for Calibration

Parameter	Description	Value/Source
$A_{j,s}$, $B_{j,s}$, a_c	Scale parameters for productivity, labor supply efficiency, and amenities	Recovered via model inversion
β	Labor share MNEs	We calibrate it using the average value-added share in Mexico
τ_{ij}	Iceberg trade costs	Based on distance, $\zeta = 1$ (Faber and Gaubert (2019))
FC_j	Fixed cost for formal firms	We calibrate it matching the share of formal firms in each location
$\mathcal{M}_j$	Mass of firms in each region	We take it directly from the data
σ	Demand elasticity	6 - different trade studies
κ	Labor supply elasticity across firm types	2.2 (Zárate (2022))
η	Migration elasticity	2 (Monte et al. (2018))
$t^L_{j,s}$, $t^Y_{j,s}$	Wedges in output and labor	Calibrated using data on taxes, crime, corruption, etc.
b_j	Informal firm marginal cost responsiveness to labor size	Estimated from government audits
$\vartheta_j(\varphi)$	Productivity distribution (Pareto)	We calibrate the shape parameter θ using the distribution of sales
$\gamma_{M,F}$, $\gamma_{M,I}$	Spillover elasticities of foreign MNEs to the domestic economy	Indirect inference matching

The Gains from Foreign Multinationals in an Economy with Distortions

Isabela Manelici, Mauricio Ulate, Jose P. Vasquez, and Román D. Zárate

Appendices for online publication

March 2025

These appendices supplement the paper "The Gains from Foreign Multinationals in an Economy with Distortions" with the following material:

- Appendix A provides details on the build of the establishment-level data from the Economic Census and summary statistics and discusses additional data used in the project (e.g., for the CZ-level controls).
- Appendix B describes the construction of the distortions data and presents additional summary statistics.
- Appendix C presents additional empirical evidence referenced in the main text.
- Appendix D presents the details and derivations of the theoretical results.
- Appendix E includes additional quantitative results.

Appendix A Data appendix

Appendix A.1 Establishment-level microdata: Economic Census

Unit of observation The Economic Census is a complete enumeration of establishments in Mexico with the following characteristics: economic units in a single physical location, settled in a permanent place and delimited by constructions or fixed installations, that combine activities and resources under the control of a single owner or controlling entity to carry out activities of production of goods, purchase and sale of merchandise or provision of services; whether for profit or not (INEGI, 2021).

Coverage The activities subject to the Economic Censuses are divided into two large groups: (i) those for which information is collected throughout the entire Mexican territory: fishing and aquaculture; mining; electricity, water and gas; construction; transport, mail and storage, as well as financial and insurance services; and (ii) manufacturing, commerce and non-financial services, whose information is collected through a complete survey in the largest and most economically-important geographical areas in Mexico and, through a sample in rural areas, where the economic activity in these industries is unimportant (INEGI, 2021).

Data We build an unbalanced panel of establishment-level data uniting six consecutive Economic Censuses from 1994, 1999, 2004, 2009, 2014, and 2019. This panel tracks typical balance sheet variables such as total sales, number of workers, wage bill, and assets (our proxy for capital).

Crucially, we also observe the establishments that are at least partially foreign-owned. We refer to an establishment as a foreign MNE if it is at least 50% foreign-owned in one Census year. Additionally, we track the country of origin of an establishment's foreign capital (which is a key input for our leading instrumental variable). In the rare instances where we observe more than one country of origin, we keep the country with the largest share of ownership in the year with the highest foreign ownership share.

We also observe the municipality of each establishment, which we then map to a commuting zone – abbreviated CZ (out of 455 CZs). The Economic Census also records the industry classification of each establishment at the 4-digit level.

Cleaning rules We drop establishment-years in which an establishment reports zero or missing workers, negative sales, negative value-added, or negative production.

Establishment-level identifiers across time INEGI started assigning unique identifiers to each establishment in the 2009 Economic Census, which allows us to track establishments across the 2009, 2014, and 2019 Economic Censuses. In addition, Busso et al. (2018) match the official unique and time-consistent identifiers to the year-specific Census identifiers in 1994, 1999, and 2004. Their iterative procedure finds similar establishments through name matching based on their legal entity name, also considering their geolocation, industry, and birth year. Their algorithm has ten steps to decide on the best match (should it exist) for an establishment comparing two years. If the procedure doesn't find a match, there are two options: the establishment stopped operating or changed locations. Finally, the authors validated their algorithm with the official identifiers, and they found more than 90% of accuracy.

Appendix A.2 Summary statistics

Table A1: Summary statistics by establishment type

	1994			2019		
	Domestic		Foreign	Domestic		Foreign
	Informal	Formal	MNEs	Informal	Formal	MNEs
Number of estab.	1,760,156	436,158	5,228	4,167,031	611,878	9,008
Employment	1.8	13.8	154.2	2.4	23.6	301.2
Sales	20.9	744.2	12,676.2	24.3	1,240.4	21,525.9

Notes: Table A1 presents summary statistics on the sample of establishments in our analysis, separated into three groups (foreign MNEs, domestic formal, and domestic informal). Namely, Table A1 shows the averages across all CZs (weighted by the number of establishments) at the beginning and end of our sample period (1994 and 2019). We measure size as total sales and number of workers. Values for sales are expressed in 2018 CPI-deflated thousands of U.S. dollars.

Table A2: Importance in the Mexican economy of domestic informal and foreign MNEs

	1994		2019	
	Domestic Informal	Foreign Multinationals	Domestic Informal	Foreign Multinationals
Workers	31.70	8.07	36.83	9.99
Sales	8.60	15.50	9.61	18.39
Wage bill	2.00	19.79	9.00	25.72

Notes: Table A2 shows the importance of domestic informal establishments and foreign MNEs as a share of the economy-wide totals of the number of workers, sales, and wage bill in 1994 and 2019. Values are reported as percentages. Domestic formal establishments are the complement to 100%.

Table A3: Summary statistics by establishment type with maquila and non-maquila foreign MNEs

	1994				2019			
	Domestic		Foreign MNEs		Domestic		Foreign MNEs	
	Informal	Formal	No-Maquila	Maquila	Informal	Formal	No-Maquila	Maquila
Number of estab.	1,760,156	436,158	4,802	426	4,167,031	611,878	7,253	1,755
Employment	1.8	13.8	127.6	453.5	2.4	23.6	204.2	702.2
Sales	20.9	744.2	12,726.0	12,114.1	24.3	1,240.4	21,833.7	20,253.6

Notes: Table A3 presents summary statistics on the sample of establishments in our analysis, separated into four groups (non-maquila foreign MNEs, maquila foreign MNEs, domestic formal, and domestic informal). Namely, Table A3 shows the averages across all CZs (weighted by the number of establishments) at the beginning and end of our sample period (1994 and 2019). We measure size as total sales and number of workers. Values for sales are expressed in 2018 CPI-deflated thousands of U.S. dollars.

Table A4: Summary statistics for maquila and non-maquila foreign MNEs in 2019

Any maquila income	Establishments share (1)	Workers share (2)	Sales (3)	Workers (4)	Sales/ worker (5)	Assets/ worker (6)	Exports share (7)
No	0.81	0.55	21,833.7	204.2	106.9	26.5	0.09
Yes	0.19	0.45	20,253.6	702.2	28.8	7.3	0.71

Notes: Table A4 presents summary statistics separately for foreign MNEs in Mexico in 2019 with or without income from maquila activities. The shares in columns (1) and (2) are with respect to the totals in foreign MNEs. Shares in column (7) are with respect to the total sales of the respective type of foreign MNEs (foreign MNEs with or without income from maquila activities). Values for sales, sales per worker and assets per worker are expressed in 2018 CPI-deflated thousands of U.S. dollars.

Appendix A.3 Other data

Time-consistent commuting zones Between 1994 and 2019, Mexico both created new municipalities and merged some municipalities, making it necessary to create time-consistent commuting zones (CZs). We began with an initial mapping of municipalities to CZs for 2019 borrowed from Blyde et al. (2020), which became our reference for the CZ definitions. In 2019, there were 779 CZs in Mexico. Then, to keep track of changes in municipality codes, new municipalities, and municipalities that split or merged, we used the "Catálogo Único de Claves de Áreas Geoestadísticas Estatales, Municipales y Localidades" from INEGI for the 2000-2020 period. For the 1994-2000 period, we used INEGI's census reports of new municipalities.

In 1994, CZs had a mean of 12,196 workers per CZ, a median of 980, a minimum of 4 workers, and a maximum of 2,535,464 workers. In 2019, CZs had a mean of 30,379 workers and a median of 3,081, whereas the minimum and maximum of total workers were 9 and 4,872,287, respectively. CZs have an average area of 2,460.2 km^2 and a median of 666.4 km^2. The smallest CZ has 4.3 km^2, and the biggest has an area of 52,982 km^2.

Additionally, we needed all CZs to have a positive value for all dependent variables of our regressions, so these regressions could be carried out in a balanced CZ-year panel. For example, we don't want to have a CZ where the number of high-skill workers is zero. While this is not a problem for most commuting zones, this is problematic for very small and very rural commuting zones. To solve this, we merged neighboring commuting zones that had at least one of the dependent variables with a zero value. After this, we ended up with 455 CZs.

Migration flows We used various sources of data to construct the 1990 municipality-municipality (MUN-MUN) migration matrix. Firstly, we use the IPUMS 1990 data for Mexico which provide us with the residence of each individual 5 years before the 1990 survey. However, the questionnaire employed in this Census asked for the Mexican state (ENT) or country of residence in 1985. This implied two problems: (i) we had only had ENT-MUN flows, and (ii) we were missing the MUN-Rest of the World (RoW) and RoW-RoW flows. To address the first problem, we made a proportionality assumption to assign the total outflow of each state of origin to each municipality within its boundaries by its share of the population. Once we made this assumption, we were able to construct the MUN-MUN matrix. To address the second problem, the MUN-RoW flows, we used the 1990 ACS US Census data from IPUMS, since the US is the principal destiny of Mexican migration, to find how many people had migrated between 1985 and 1990 and were living in the US. Once we had the aggregated value, we proportionally assign the total outflow of migrants from Mexico to the US to each municipality using the 1993 "Encuesta sobre Migración en la Frontera Norte de México", which collected data on the residence of the surveyed immigrants. Finally, we added the total population in 1990 of all the countries from which people migrating from abroad to Mexico in the 1990 Mexican Census data for the RoW-RoW flow.

Concordance between industry classifications The industry classification used in the 1994 Economic Census differs from the NAICS classification used in the subsequent Economics Censuses (1999-2019). In the 1994 Economic Census, INEGI used the "Mexican Classification of Activities and Products" (CMAP in Spanish). CMAP comprises nine sectors that disaggregate into sub-sectors, branches, and classes. We created a manual mapping between the 30 CMAP sub-sectors (used in 1994) and the 21 NAICS two-digit codes (used between 1999 and 2019) to a set of 16 broad sectors. Table A5 shows the concordance between these 16 broad sectors and each industry classification.

Appendix A.4 Commuting-zone level controls

Data for part of our commuting zone-level controls comes from the "Integrated Public Use Microdata Series" (IPUMS) for Mexico. The Mexico IPUMS data is based on the General Population and Housing Census taking place every ten years. We use beginning-of-decade General Population and Housing Census data for the two years we observe each decade in the Economic Census. Namely, we use the 1990 values for 1994, the 2000 values for 1999 and 2004, the 2010 values for 2009 and 2014, and the 2020 values for 2019. From IPUMS, we use the following commuting-zone level controls: (i) the share of individuals living in a place designated as urban; (ii) the share of individuals working in manufacturing; (iii) the share of individuals with a completed secondary degree; (iv) the share of individuals who are employed; (v) the share of individuals in routine occupations; (vi) the share of indigenous individuals; and (vii) the share of foreign-born individuals.

We also use Comtrade data to build three measures of CZ-level exposure to imports of Mexico from the United States and China and to the growing exports of China to Mexico's main export markets from

Table A5: Concordance table INEGI-NAICS

Sector	CMAP (1994)	NAICS (1999-2019)
Agriculture/Mining	11, 21, 22, 23, 29	11, 21
Energy/Maintenance	41, 42	22, 56
Construction/Rentals	50, 82, 83	23, 53
Foods/Textile	31, 32	31
Chemical/Paper/Non-metallic	33-36	32
Metallic/Machinery/Equipment	37-39	33
Wholesale Trade	61	43
Retail Trade	62	46
Transport	71	48
Postal and Storage	72	49
Media	94	51
Finance and Insurance	81	52
Professional Services	95, 97	54
Healthcare/Education/Social Assistance	92	61-62
Other Services	96	71, 81
Accommodation Services	93	72

1994. These measures adapt to the Mexican setting the exposure measures proposed by Autor et al. (2013). We use two-digit HS import data from Comtrade and map these HS codes to NAICS two-digit industry codes using the R package concordance by Liao et al. (2020). We also use employment data at the CZ level, and at the economic sector (defined by NAICS two-digit industry codes) level in 1994. We used this data to weigh the exposure measures by the CZ-sector employment share of national-sector employment.

The first two measures aim to control for changes in the import patterns of Mexico from two important global trade partners, China and the United States. China's becoming a member of the WTO in 2001 is an event with plausible repercussions for Mexican labor markets. The United States gained preferential access to the Mexican market after NAFTA came into force in 1994. Instead of directly using the change in imports of Mexico from these two countries (which might reflect endogenous responses to demand or supply shocks in Mexico), we leverage the changes in the imports of countries Ω from the United States and China, respectively. Namely, Ω includes Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Guyana, Jamaica, Nicaragua, Panama, Peru, Paraguay, Uruguay, and Venezuela (as in Blyde et al., 2020). Changes in imports of countries Ω from China and the United States capture the supply-driven component in Mexico's imports from China and the United States, in addition to changes in demand from China from countries at comparable levels of development as Mexico and distance to China. In addition, for the United States-specific measure, we add Canada to the set Ω, as NAFTA has plausibly affected Canada's imports from the United States similarly to how it has affected Mexico's imports.

To build CZ-level measures of exposure to the imports from the United States and China, we apportion the "shock" (the change in imports of countries Ω from the United States and China) to CZs based on their 1994 employment in the relevant industry j and CZ cz. The measure is then translated into a per-worker measure by dividing it by the CZ cz overall employment in 1994. Namely, we build the following CZ-level measure of exposure to changes in imports from the United States and China as follows:

$$\Delta IPW_{o,cz,t} = \frac{1}{L_{cz,1994}} \sum_{j} \frac{L_{cz,j,1994}}{L_{j,1994}} \Delta M_{jt}^{H \to \Omega} \tag{A1}$$

where cz indexes CZs in Mexico, j is a broad sector category, $H \in \{UnitedStates, China\}$, and L is employment. $\Delta M_{jt}^{H \to RoW}$ represents the change in exports from either the United States or China to the set of countries Ω between $(t-5)$ and t.

Second, we build a CZ-level measure of Mexico's competition with China in the markets to which Mexico traditionally exported. Namely, we build the changes in the exports of China to the main importing partners of Mexico in 1994, i.e., the exports of China to the United States, Canada, Japan, Spain, and France (who jointly account for about 90% of Mexico's export that year).

The measure of exposure to competition with China is given by equation (A2), where $k \in \{$Japan, Spain, France, United States, Canada$\}$. We aggregate across countries k the changes in exports of China to each country k, weighted by the importance of imports from Mexico for that specific country. For a CZ cz and industry j, we weigh the export competition shock by the importance of industry j in the employment of that CZ cz in 1994.

$$Exp_{cz,t} = \frac{1}{L_{cz,1994}} \sum_j \frac{L_{cz,j,1994}}{L_{j,1994}} \left[\sum_k \frac{M_{j,1994}^{MX \to k}}{M_{j,1994}^{k}} \Delta M_{jt}^{China \to k} \right] \tag{A2}$$

where cz indexes CZs in Mexico, j is a broad sector category, L is employment, $\frac{M_{j,1994}^{MX \to k}}{M_{j,1994}^{k}}$ is the share of imports of country k from Mexico out the total imports of country k in 1994, and $\Delta M_{jt}^{China \to k}$ denotes the change of imports of country k from China between $(t-5)$ and t.

Appendix B Establishment-level distortions

Appendix B.1 Output distortions: Taxes and subsidies

Appendix B.1.1 Taxes

Data Economic Census 1994, 1999, 2004, 2009, 2014, 2019. Panel available for all years, should the establishment exist.

Variables Total sales and VAT.

Sources for distortion Taxes.

$$\text{Taxes share} = \frac{\text{Sales VAT}^{\text{annual}}}{\text{Sales}^{\text{annual}}}$$

Extrapolating to other establishments Extrapolation to other establishments is not needed.

Appendix B.1.2 Subsidies

Data Economic Census 1994, 1999, 2004, 2009, 2014, 2019. Panel available for all years, should the establishment exist.

Variables Total sales and subsidies.

Sources for distortion Subsidies.

$$\text{Subsidies share} = -\frac{\text{Subsidies}^{\text{annual}}}{\text{Sales}^{\text{annual}}}$$

Extrapolating to other establishments Extrapolation to other establishments is not needed.

Appendix B.2 Output distortions: Weak institutions and the rule of law

Appendix B.2.1 Regulation

Data ENAPROCE 2015, 2018. This survey reports data from 2014, 2016 and 2017. Panel available: ID to link 2015 with 2018, should the establishment exist. We match observations with the Economic Census.

Variables Total sales (even if the Economic Census was not collected that year), wages of senior executives, average number of senior executives, hours dedicated to administrative requirements (bureaucracy), additional payments, factor weights, ID, sector, CZ, size (number of workers), establishment type (domestic formal, domestic informal, maquila foreign MNEs, and non-maquila foreign MNEs), and census division.

Sources of distortion

1. Time spent (labor factor) used to address government requirements.
2. Additional payments to the government **due to regulation** (other than taxes).

1. Time spent (labor factor) used to address government requirements

$$\frac{\overbrace{\frac{\text{Wage bill}^{\text{annual}}_{\text{senior}}}{N_{\text{senior}}}}^{\text{Average annual senior wage}} \times \overbrace{\frac{1}{5 \times 8 \times 50}}^{\text{Expected worked hours}} \times \text{Hours spent in government requirements}^{\text{annual}}}{\text{Sales}^{\text{annual}}}$$

Superscripts denote the frequency of the variable (weekly, monthly or annual). We assume that senior employees perform these tasks.

2. Additional payments to the government due to regulation (other than taxes)

$$\text{Additional payments markup} = \frac{\text{Additional payments}^{\text{annual}}}{\text{Sales}^{\text{annual}}}$$

Extrapolating to other establishments, matching to Economic Census and obtaining the sales denominator This is a standalone survey; therefore, we take the sales reported in this survey. We match these observations with the Economic Census using different years and project the distortion share to other establishments. The distortions reported in 2014 are matched with those in the Census 2014, where the reported year is 2013. The distortions reported in 2017 are matched with those in the Census 2019, where the reported year is 2018.

The linear regression approach is more transparent than performing the mean imputation; also, the unobserved heterogeneity captured by the fixed effects will improve the estimation. Let the distortion share μ_{fiste} of the establishment f in the census division i of the sector s (manufacturing or non-manufacturing) and establishment type e in year t be explained by:

$$\mu_{fiste} = \alpha + \beta_1 \log(\text{Size}_{ft}) + \beta_2 (\log(\text{Size}_{ft}))^2 + \gamma_i + \delta_s + \eta_e + \varepsilon_{fiste} \tag{B3}$$

where γ_i, δ_s, η_e are the fixed effects of the location, sector, and establishment type, respectively, and the observations are weighted by the expansion factor.

Appendix B.2.2 Crime

Data ENVE 2012, 2014, 2016, 2018, 2020. Panel available: ID to link 2016, 2018, and 2020, should the establishment exist. We drop 2020 due to the COVID-19 pandemic. We match observations with the Economic Census.

Variables Total sales (from the Economic Census), prevention expenditures, cost of crime (damages), factor weights, ID, sector, CZ, size (number of workers), establishment type (domestic formal, domestic informal, maquila foreign MNEs, and non-maquila foreign MNEs), and census division.

Sources of distortion

1. Prevention measures.
2. Actual damages caused by crimes.

1. Prevention measures

$$\text{Prevention share} = \frac{\text{Prevention}^{\text{annual}}}{\text{Sales}^{\text{annual}}}$$

2. Actual damages caused by crimes **Source:** Additional module of ENVE with events by crime type and establishment. Let c be the type of crime.

$$\text{Damages share} = \frac{\sum_c \text{Damages}_c^{\text{annual}}}{\text{Sales}^{\text{annual}}}$$

Extrapolating to other establishments and matching to Economic Census We match these observations with the Economic Census using different years and project the distortion share to other establishments. The distortions reported in 2016 are matched with those in the Census 2014, where the reported year is 2013. The distortions reported in 2018 are matched with those in the Census 2019, where the reported year is 2018.

The linear regression approach is more transparent than performing the mean imputation; also, the unobserved heterogeneity captured by the fixed effects will improve the estimation. Let the distortion share μ_{fiste} of the establishment f in the census division i of the sector s (manufacturing or non-manufacturing) and establishment type e in year t be explained by:

$$\mu_{fiste} = \alpha + \beta_1 \log(\text{Size}_{ft}) + \beta_2 (\log(\text{Size}_{ft}))^2 + \gamma_i + \delta_s + \theta_t + \eta_e + \varepsilon_{fiste} \tag{B4}$$

where γ_i, δ_s, θ_t, η_e are the fixed effects of the location, sector, time, and establishment type, respectively, and the observations are weighted by the expansion factor.

Appendix B.2.3 Corruption

Data ENVE 2012, 2014, 2016, 2018, 2020. Panel available: ID to link 2016, 2018, and 2020, should the establishment exist. We drop 2020 due to the COVID-19 pandemic. We match observations with the Economic Census.

Variables Total sales (from the Economic Census), bribes, factor weights, ID, sector, CZ, size (number of workers), establishment type (domestic formal, domestic informal, maquila foreign MNEs, and non-maquila foreign MNEs), and census division.

Sources for distortion Bribes.

$$\text{Bribes share} = \frac{\text{Bribes}^{\text{annual}}}{\text{Sales}^{\text{annual}}}$$

Extrapolating to other establishments and matching to Economic Census We match these observations with the Economic Census using different years and project the distortion share to other establishments. The distortions reported in 2016 are matched with those in the Census 2014, where the reported year is 2013. The distortions reported in 2018 are matched with those in the Census 2019, where the reported year is 2018.

The linear regression approach is more transparent than performing the mean imputation; also, the unobserved heterogeneity captured by the fixed effects will improve the estimation. Let the distortion share μ_{fiste} of the establishment f in the census division i of the sector s (manufacturing or non-manufacturing) and establishment type e in year t be explained by:

$$\mu_{fiste} = \alpha + \beta_1 \log(\text{Size}_{ft}) + \beta_2 (\log(\text{Size}_{ft}))^2 + \gamma_i + \delta_s + \theta_t + \eta_e + \varepsilon_{fiste} \tag{B5}$$

where γ_i, δ_s, θ_t, η_e are the fixed effects of the location, sector, time, and establishment type, respectively, and the observations are weighted by the expansion factor.

Appendix B.3 Labor distortions

Appendix B.3.1 Labor input distortion

The distortions related to labor input are captured through contributions to social security. By construction, informal establishments don't pay social security.

Data Economic Census 1994, 1999, 2004, 2009, 2014, 2019. Panel available for all years, should the establishment exist.

Variables Social security contributions and wage bill.

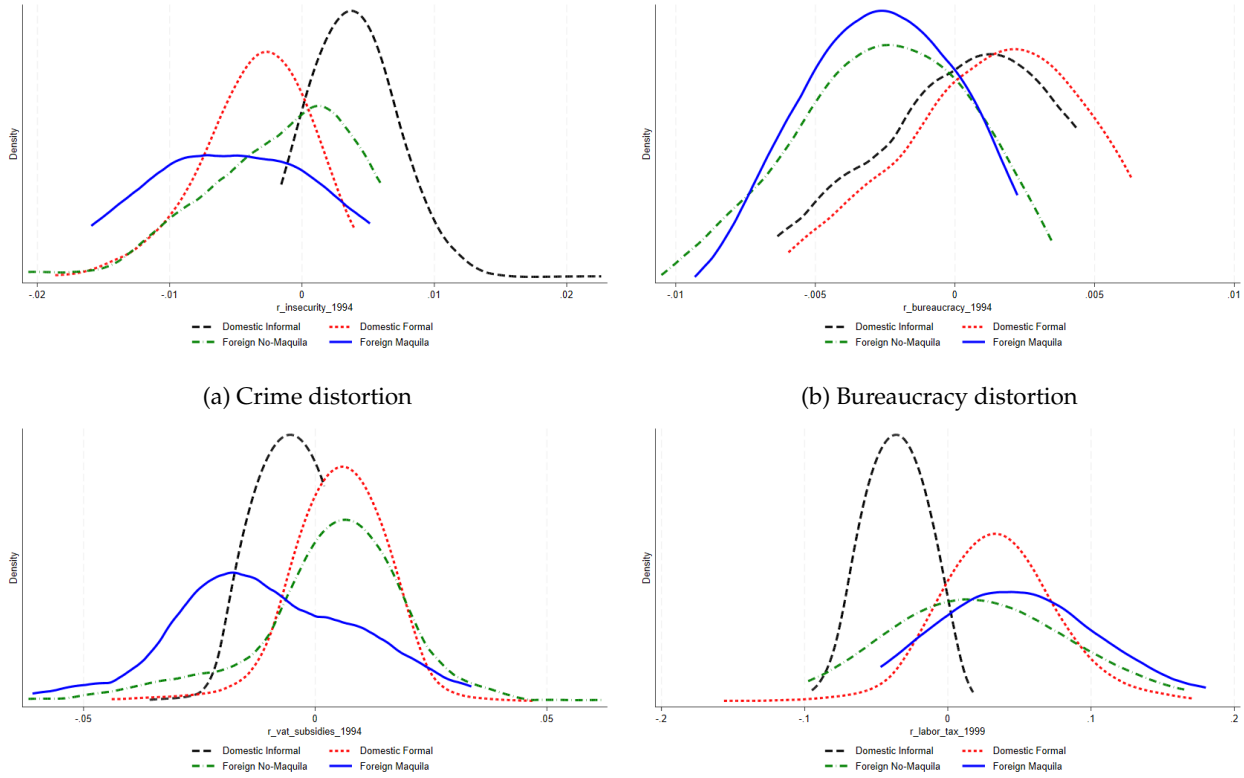
Sources for distortion Social security.

$$\text{Labor distortion share} = \frac{\text{Social security}^{\text{annual}}}{\text{Wage bill}^{\text{annual}}}$$

Extrapolating to other establishments Extrapolation to other establishments is not needed.

Appendix B.4 Histograms about distortions in Mexico

Figure B1: Empirical density of each distortion by establishment type and CZs after controlling for size and sector



(a) Crime distortion

(b) Bureaucracy distortion

(c) VAT & subsidies distortion

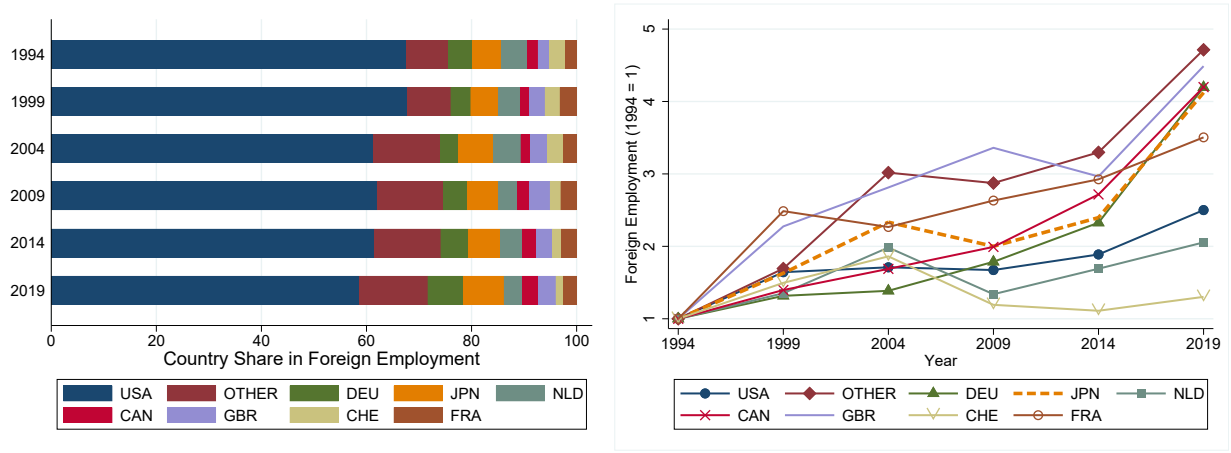
(d) Labor tax distortion

Notes: Figure B1 includes four empirical density plots of each distortion in 1994 after controlling for size and sector, one histogram for each establishment type (domestic formal, domestic informal, foreign maquila MNEs, and foreign non-maquila MNEs) across all 455 CZs in Mexico.

Appendix C Additional empirical results

Appendix C.1 Supportive evidence for the IV

Figure C1: Evolution of employment in foreign MNEs by country of origin

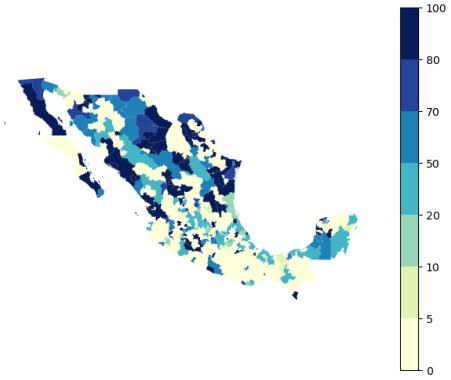


(a) Share of foreign MNEs by country of origin

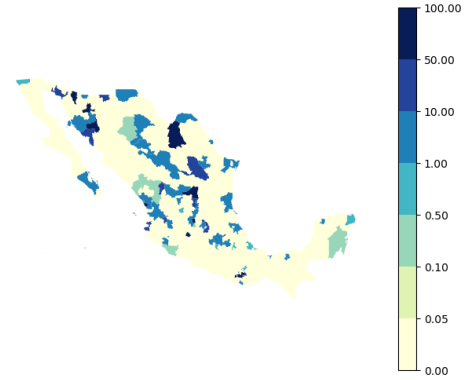
(b) Foreign MNE employment growth (1994=1)

Notes: Panel (a) in Figure C1 shows the composition of employment in foreign MNEs in Mexico by the country of origin of the foreign capital (from 1994 to 2019). Panel (b) normalizes to 1 the Mexico-wide employment in 1994 in establishments owned by Mexico's eight largest investor countries. Each line tracks the evolution since 1994 of foreign MNE employment in Mexico from each of those eight investor countries (plus from any other origin country labelled "other").

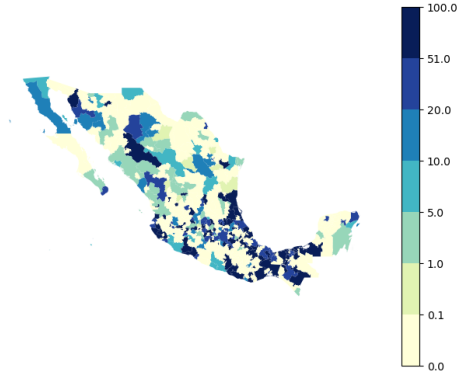
Figure C2: Spatial concentration of foreign MNEs into Mexico by country (region) of ownership (2019)



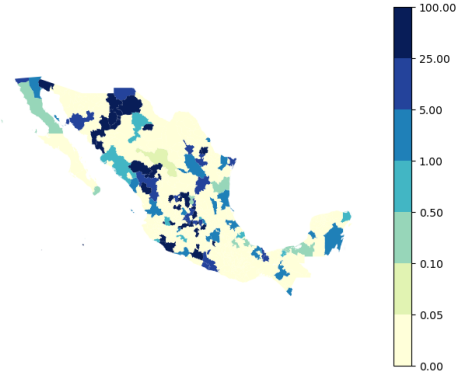
(a) United States



(b) Canada



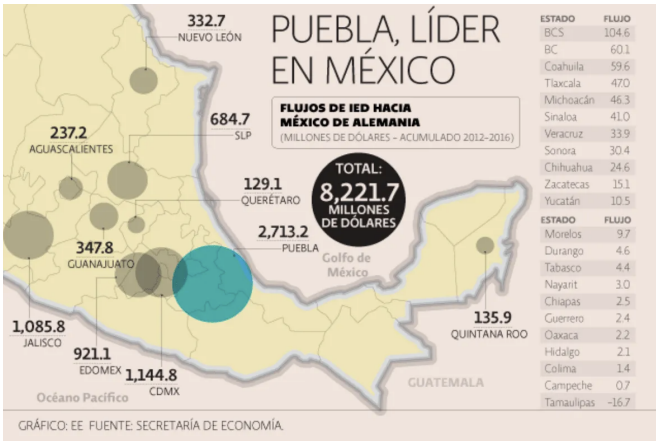
(c) Western Europe



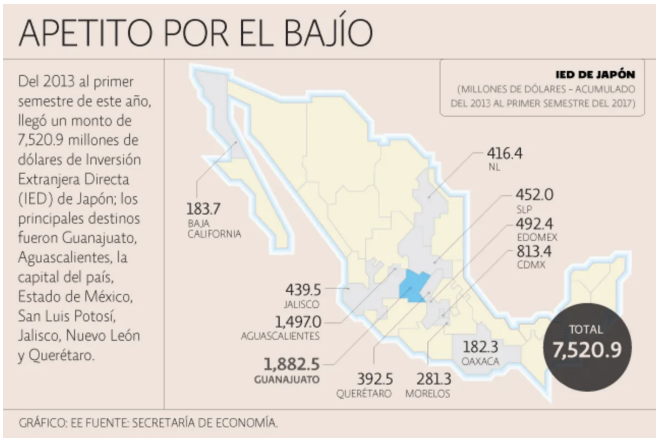
(d) Japan

Notes: Figure C2 displays the spatial variation in the importance of employment at foreign MNEs across commuting zones in Mexico in our end-year 2019. Namely, we plot the share of employment in foreign MNEs from a given country (region) of origin as a share of the total employment in foreign MNEs in a commuting zone. Panel A focuses on foreign MNEs into Mexico from the United States, Panel B on foreign MNEs from Canada, Panel C on foreign MNEs from Western Europe and Panel D on foreign MNEs from Japan. Most foreign MNEs from Western Europe comes from Germany, Netherlands, Great Britain, France, and Switzerland (sorted in descending order based on overall employment in 2019).

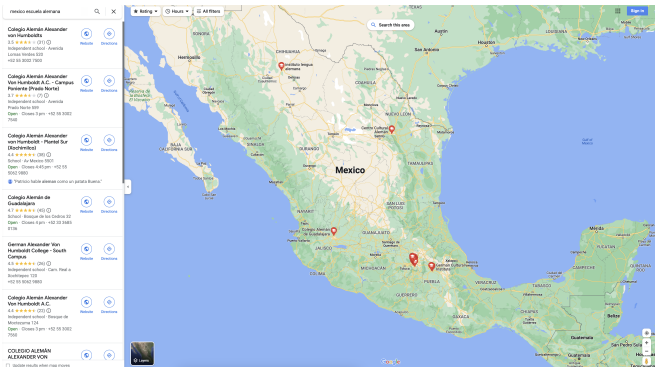
Figure C3: Spatial concentration in Mexico of German and Japanese MNEs and relevant amenities



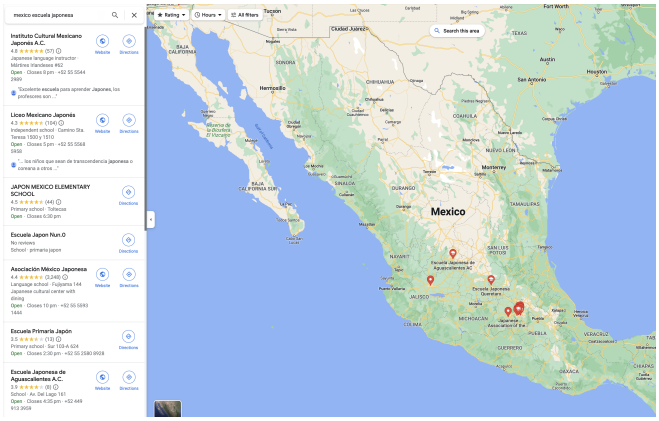
(a) German MNEs cluster in Puebla (2012-2016)



(b) Japanese MNEs cluster in El Bajio (2017)



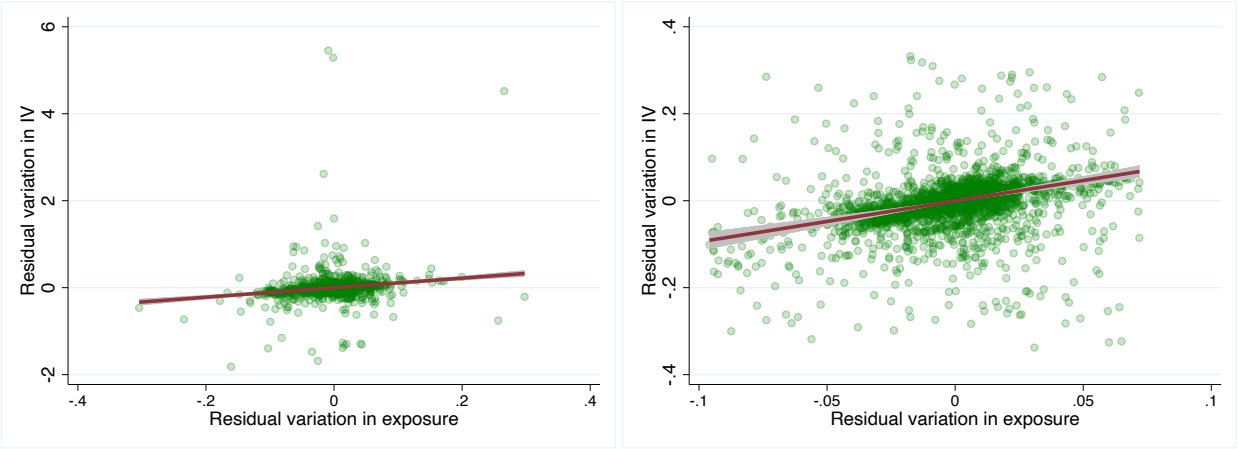
(c) German schools in Mexico (Google Maps)



(d) Japanese schools in Mexico (Google Maps)

Notes: The maps in Panels C3a and C3b come from two articles in the Mexican newspaper *El Economista* (see here and here). The message of these maps and the articles more broadly is twofold: (i) foreign MNEs from specific origin countries (Germany and Japan in these particular examples) tends to be spatially concentrated (e.g., Germany keeps investing in Puebla, whereas Japan keeps investing in Bajío) and (ii) different origin countries have different spatial clustering patterns (even when the countries are similar in GDP per capita and invest in similar industries). The maps in Panels C3c and C3d are print screens of Google Maps searches for German and Japanese schools in Mexico, respectively. The exact searches used were "mexico escuela alemana" and "mexico escuela japonesa." Two patterns stand out. First, as expected, Mexico City concentrates both foreign MNEs and amenities (including amenities relevant to the varied countries of foreign MNE ownership). Second, beyond Mexico City, the geographic clusters of foreign MNEs from a given origin country are also clusters for the amenities relevant to that origin country. For instance, foreign MNEs from Germany clusters in Puebla, and so do German schools.

Figure C4: First stage of the instrument



(a) $\widehat{Z}^F_{cz,t}$ vs. $\widehat{X}^F_{cz,t}$. Full sample

(b) $\widehat{Z}^F_{cz,t}$ vs. $\widehat{X}^F_{cz,t}$. Trimming top/bottom 1%

Notes: In both panels, the X-axis presents the residualized measure of exposure to foreign MNEs ($\widehat{X}^F_{cz,t}$) and the Y-axis presents its residualized instrument ($\widehat{Z}^F_{cz,t}$). Both variables are residualized by controlling for the vector of FE and other controls $K_{cz,t}$. Clustering happens at the census division level.

Table C1: First stage and reduced form estimates of the effects of foreign MNEs exposure on the domestic economy

	First Stage	Pop	Sales	VA	Capital	Workers	Wage Bill
		(1)	(2)	(3)	(4)	(5)	(6)
Overall Domestic	0.796***	1.080***	28.411***	12.119***	4.955***	0.266***	2.811***
	(0.024)	(0.120)	(2.185)	(0.648)	(0.616)	(0.094)	(0.916)
Obs	2275	2275	2275	2275	2275	2275	2275

Notes: Table C1 presents the first stage (unnumbered column) and reduced form estimates (columns 1-6) of the regression described in equation (2). The population variable includes the working-age population only. Sales, value-added, capital (assets), and wage bills are expressed in 2018 CPI-deflated thousands of U.S. dollars per capita. Workers are measured as headcounts. Observations are weighted by the initial CZ domestic employment. We cluster the standard errors at the census division level. We use as controls economic region-year fixed effects, measures of import and export exposure, and various time-variant CZ-level controls. Sections 3.1, 3.2, and Appendix A.4 provide more details about the regression specification and controls.

Table C2: First stage and reduced form estimates of the effects of foreign MNEs exposure on the domestic formal vs. informal sectors

	First Stage	Pop	Sales	VA	Capital	Workers	Wage Bill
		(1)	(2)	(3)	(4)	(5)	(6)
Overall Domestic	0.796***	1.080***	28.411***	12.119***	4.955***	0.266***	2.811***
	(0.024)	(0.120)	(2.185)	(0.648)	(0.616)	(0.094)	(0.916)
Formal	–	–	22.391***	9.958***	3.390***	0.235***	2.349***
			(1.183)	(0.279)	(1.234)	(0.072)	(0.807)
Informal	–	–	6.020***	2.161***	1.565	0.031	0.462***
			(1.147)	(0.457)	(1.230)	(0.030)	(0.148)
Obs	2275	2275	2275	2275	2275	2275	2275

Notes: Table C2 presents the first stage (unnumbered column) and reduced form estimates (columns 1-6) of the regression described in equation (2). This table includes coefficients for formal and informal domestic establishments. The population variable includes the working-age population only. Sales, value-added, capital (assets), and wage bills are expressed in 2018 CPI-deflated thousands of U.S. dollars per capita. Workers are measured as headcounts. Observations are weighted by the initial CZ domestic overall, formal or informal employment. We cluster the standard errors at the census division level. We use as controls economic region-year fixed effects, measures of import and export exposure, and various time-variant CZ-level controls. Sections 3.1, 3.2, 3.3, and Appendix A.4 provide more details about the regression specification and controls.

Appendix C.2 Additional regression results

Table C3: OLS and IV estimates of the effects of foreign MNEs exposure on the domestic economy blue collar vs white collar

(a) OLS

	Number of Workers			Wage Bill		
	Total	Blue Collar	White Collar	Total	Blue Collar	White Collar
	(1)	(2)	(3)	(4)	(5)	(6)
Overall Domestic	0.146***	0.129***	0.017**	0.707***	0.743***	-0.037
	(0.027)	(0.030)	(0.008)	(0.141)	(0.106)	(0.135)
Obs	2275	2275	2275	2275	2275	2275

(b) IV

	Number of Workers			Wage Bill		
	Total	Blue Collar	White Collar	Total	Blue Collar	White Collar
	(1)	(2)	(3)	(4)	(5)	(6)
Overall Domestic	0.333***	0.211	0.123**	3.532***	1.890***	1.642
	(0.109)	(0.155)	(0.050)	(1.240)	(0.297)	(1.183)
Obs	2275	2275	2275	2275	2275	2275
F-stat	107.8	107.8	107.8	107.8	107.8	107.8

Notes: Table C3 presents the OLS and IV estimates of the regression described in equation (2). This table shows coefficients for total, blue-collar, and white-collar workers and wage bill. Workers are measured as headcounts. Wage bills are expressed in 2018 CPI-deflated thousands of U.S. dollars per capita. Observations are weighted by the initial CZ domestic employment. We cluster the standard errors at the census division level. We use as controls economic region-year fixed effects, measures of import and export exposure, and various time-variant CZ-level controls. Sections 3.1, 3.2, and Appendix A.4 provide more details about the regression specification and controls.

Table C4: IV estimates of the effects of foreign MNEs exposure on the domestic economy blue collar vs white collar. By foreign MNE type: all, only maquila or only non-maquila

(a) All Foreign

	Number of Workers			Wage Bill		
	Total	Blue Collar	White Collar	Total	Blue Collar	White Collar
	(1)	(2)	(3)	(4)	(5)	(6)
Overall	0.333***	0.211	0.123**	3.532***	1.890***	1.642
Domestic	(0.109)	(0.155)	(0.050)	(1.240)	(0.297)	(1.183)
Obs	2275	2275	2275	2275	2275	2275
F-stat	107.8	107.8	107.8	107.8	107.8	107.8

(b) Maquila

	Number of Workers			Wage Bill		
	Total	Blue Collar	White Collar	Total	Blue Collar	White Collar
	(1)	(2)	(3)	(4)	(5)	(6)
Overall	-0.738*	-0.703*	-0.035	-2.130	-3.078	0.948
Domestic	(0.379)	(0.356)	(0.046)	(1.742)	(2.086)	(1.056)
Obs	2275	2275	2275	2275	2275	2275
F-stat	157.79	157.79	157.79	157.79	157.79	157.79

(c) Non-Maquila

	Number of Workers			Wage Bill		
	Total	Blue Collar	White Collar	Total	Blue Collar	White Collar
	(1)	(2)	(3)	(4)	(5)	(6)
Overall	1.342***	1.013**	0.329**	8.630***	5.349***	3.282
Domestic	(0.379)	(0.504)	(0.132)	(2.903)	(0.836)	(3.070)
Obs	2275	2275	2275	2275	2275	2275
F-stat	227.27	227.27	227.27	227.27	227.27	227.27

Notes: Table C4 presents the IV estimates of the regression described in equation (4). The three panels differ in the set of foreign MNEs whose effects we estimate (and the subsequent controls used). In the upper panel, we keep the baseline sample of all foreign MNEs, which corresponds to the regression in equation (2) and the results presented in Table C3. In the middle panel, we focus only on the exposure to maquila foreign MNEs. In these regressions, we control for the change in exposure to non-maquila foreign MNEs. In the lower panel, we focus only on the exposure to non-maquila foreign MNEs. In these regressions, we control for the change in exposure to maquila foreign MNEs. Workers are measured as headcounts. Wage bills are expressed in 2018 CPI-deflated thousands of U.S. dollars per capita. Observations are weighted by the initial CZ domestic employment. We cluster the standard errors at the census division level. We use as controls economic region-year fixed effects, measures of import and export exposure, and various time-variant CZ-level controls. Sections 3.1, 3.2, and Appendix A.4 provide more details about the regression specification and controls.

Table C5: IV estimates of the effects of foreign MNEs exposure on the domestic formal vs. informal sectors. By worker type: total, blue collar, or white collar

	Number of Workers			Wage Bill		
	Total	Blue Collar	White Collar	Total	Blue Collar	White Collar
	(1)	(2)	(3)	(4)	(5)	(6)
Overall Domestic	0.333***	0.211	0.123**	3.532***	1.890***	1.642
	(0.109)	(0.155)	(0.050)	(1.240)	(0.297)	(1.183)
Formal	0.295***	0.189	0.106**	2.952***	1.347***	1.606
	(0.083)	(0.130)	(0.049)	(1.088)	(0.206)	(1.164)
Informal	0.038	0.022	0.016***	0.580***	0.544***	0.035
	(0.036)	(0.036)	(0.003)	(0.197)	(0.183)	(0.022)
Obs	2275	2275	2275	2275	2275	2275
F-stat	107.8	107.8	107.8	107.8	107.8	107.8

Notes: Table C5 presents the IV estimates of the regression described in equation (2). This table shows coefficients for total, blue-collar, and white-collar workers and wage bill. Additionally, it includes the formal and informal decomposition. Workers are measured as headcounts. Wage bills are expressed in 2018 CPI-deflated thousands of U.S. dollars per capita. Observations are weighted by the initial CZ domestic overall, formal or informal employment. We cluster the standard errors at the census division level. We use as controls economic region-year fixed effects, measures of import and export exposure, and various time-variant CZ-level controls. Sections 3.1, 3.2, 3.3, and Appendix A.4 provide more details about the regression specification and controls.

Appendix D Theoretical Results

In this section, we provide proofs for the three theoretical results from section 4.8.

Appendix D.1 Labor supply

Before providing the proofs for the theoretical results, we derive the labor supply functions. A worker chooses to work in firm-type k and live in j if:

$$\frac{\epsilon_j(\ell_{j,k}w_{j,k} + r_j)}{P_j} \geq \frac{\epsilon_c(\ell_{c,s}w_{c,s} + r_c)}{P_c}$$

Workers choose first the location and then, conditional on the residential location, they choose the firm type to work in.

Appendix D.1.1 Labor supply-firm types

Conditional on living in j, the joint distribution of vector $\boldsymbol{\ell}_j$ is:

$$F(\boldsymbol{\ell}_j) = \prod_{k=1}^{K} \exp(-B_{jk}\ell_{jk}^{-\kappa}), \tag{D6}$$

where κ represents the dispersion parameter and B_{js} correspond to the scale parameters from the Fréchet distribution. In this section, we provide two main results. First, the probability that an agent in j chooses firm type k is:

$$\lambda_{jk} = \frac{B_{jk}w_{jk}^{\kappa}}{\sum_s B_{js}w_{js}^{\kappa}}. \tag{D7}$$

Second, that the ex-ante expected average income is:

$$\mathbb{E}[\max_{\{s\}_{s=1}^{K}}\{w_{js}\ell_{js}\}] = \gamma_\kappa \left(\sum_s B_{js}w_{js}^{\kappa}\right)^{\frac{1}{\kappa}} \tag{D8}$$

where γ_κ is a constant equal to $\Gamma\left(\frac{\kappa-1}{\kappa}\right)$ and $\Gamma(\cdot)$ is the gamma function.

To show the first result, note that an agent in location j chooses to work in firm-type k if and only if the following expression holds for all s:

$$w_{jk}\ell_{jk} \geq w_{js}\ell_{js},$$

which is equivalent to $\ell_{js} \leq \frac{w_{jk}}{w_{js}}\ell_{jk}$. Denoting $b_{js} = \frac{w_{jk}}{w_{js}}$, and

$$\Phi_j = \mathbb{E}[\max_{\{s\}_{s=1}^{K}}\{w_{js}\ell_{js}\}].$$

We know that by Bayes rule conditional that the maximum is $w_{jk}\epsilon_{jk}$:

$$\Phi_j = \frac{w_{jk}}{\lambda_{jk}} \int_0^\infty \ell_{jk} \cdot F_{jk}(\ell_{jk}, b_{js}\ell_{jk}, ..., b_{jK}\ell_{jk}) d\ell_{jk},$$

where F_{jk} is the partial derivative of F with respect to the kth element and λ_{jk} is the probability of choosing k. Given the $F(\cdot)$ function above, we get that:

$$\frac{\partial F(\boldsymbol{\ell}_j)}{\partial \ell_{jk}} = \kappa B_{jk}\ell_{jk}^{-\kappa-1} \exp(-\sum_{s=1}^{K} B_{js}\ell_{js}^{-\kappa}) \tag{D9}$$

We then have that the probability of choosing k in location j is:

$$\lambda_{jk} = \int_0^\infty F_{jk}(\ell_{jk}, b_{js}\ell_{jk}, ..., b_{jK}\ell_{jk}) d\ell_{jk} \tag{D10}$$

$$= \kappa \int_0^\infty B_{jk}\ell_{jk}^{-\kappa-1} \exp(-\sum_{s=1}^{\kappa} B_{js}b_{js}^{-\kappa}\ell_{jk}^{-\kappa}) d\ell_{jk} \tag{D11}$$

$$= \frac{1}{\sum_{s=1}^{K} B_{js}b_{js}^{-\kappa}} \int_0^\infty B_{jk} \exp(-u) du \tag{D12}$$

$$= \frac{B_{jk} w_{jk}^{\kappa}}{\sum_{s=1}^{K} B_{js} w_{js}^{\kappa}}, \tag{D13}$$

where $u = \ell_{jk}^{\kappa}$. This equation corresponds to the first result. Computing

$$\Phi_j = \frac{w_{jk}}{\lambda_{jk}} \int_0^{\infty} \ell_{jk} \cdot F_{jk}(\ell_{jk}, b_{js}\ell_{jk}, ..., b_{jK}\ell_{jk}) d\ell_{jk} \tag{D14}$$

$$= \frac{\kappa w_{jk}}{\lambda_{jk}} \int_0^{\infty} B_{jk} \ell_{jk}^{-\kappa} \exp(-\sum_{s=1}^{\kappa} B_{js} b_{js}^{-\kappa} \ell_{jk}^{-\kappa}) d\ell_{jk} \tag{D15}$$

$$= w_{jk} \frac{\sum_{s=1}^{K} B_{js} b_{js}^{-\kappa}}{B_{jk}} \int_0^{\infty} \frac{B_{jk}}{\sum_{s=1}^{K} B_{js} b_{js}^{-\kappa}} \left(\sum_{s=1}^{K} B_{js} b_{js}^{-\kappa} \right)^{\frac{1}{\kappa}} u^{-1/\kappa} \exp(-u) du \tag{D16}$$

$$= \left(\sum_{s=1}^{K} B_{js} w_{js}^{\kappa} \right)^{\frac{1}{\kappa}} \underbrace{\int_0^{\infty} u^{-1/\kappa} \exp(-u) du}_{\Gamma(1-\frac{1}{\kappa})} \tag{D17}$$

This corresponds to the expression that we have in the main text.

Appendix D.1.2 Migration decision

Workers make migration decisions based on idiosyncratic amenity shocks ϵ_c and the expected value of the ex-ante average labor income in each location. A worker decide to lives in c is and only if the following condition holds for all i:

$$\frac{\epsilon_c(\mathbb{E}_{\ell}[\max_{\{s\}_{s=1}^{K}}\{w_{cs}\ell_{cs}\}] + r_c)}{P_c} \geq \frac{\epsilon_i(\mathbb{E}_{\ell}[\max_{\{s\}_{s=1}^{K}}\{w_{is}\ell_{is}\}] + r_i)}{P_i},$$

$$\frac{\epsilon_c(\Phi_c + r_c)}{P_c} \geq \frac{\epsilon_i(\Phi_i + r_i)}{P_i}.$$

Let's define $W_i = \frac{\Phi_i + r_i}{P_i}$ as the real income measure in location i. Then, rewriting the previous condition, a worker lives in c if and only if for all i:

$$\epsilon_c \geq \epsilon_i \frac{W_i}{W_c}.$$

We define $\omega_i = \frac{W_i}{W_c}$. The joint distribution of vector $\boldsymbol{\epsilon}$ is:

$$F(\boldsymbol{\epsilon}) = \prod_{i=1}^{C} \exp(-a_i \epsilon_c^{-\eta}) \tag{D18}$$

This means that the probability of choosing location c is:

$$\lambda_c = \int_0^{\infty} a_c \epsilon_c^{-\eta-1} \exp(-\sum_{i=1}^{C} a_i \omega_i^{-\eta} \epsilon_c^{-\eta}) d\epsilon_c \tag{D19}$$

$$= \frac{a_c}{\sum_{i=1}^{C} a_i \omega_i^{-\eta}} \int_0^{\infty} \exp(-u) d \tag{D20}$$

$$= \frac{a_c W_c^{\eta}}{\sum_{i=1}^{C} a_i W_i^{\eta}} \tag{D21}$$

Finally, we need to prove what is the expected value for the welfare measure. Using a similar analysis to the previous section, we have that the distribution of the maximum from a joint Frechet distribution is:

$$G(x) = \exp(-\sum_{i=1}^{C} a_i W_i^{\eta} x^{-\eta})$$

Then, the expected value of the maximum is:

$$\mathbb{E}(x) = \int_0^{\infty} \eta \sum_{i=1}^{C} a_i W_i^{\eta} x^{-\eta} \exp(-\sum_{i=1}^{C} a_i W_i^{\eta} x^{-\eta}) dx \tag{D22}$$

$$= \int_0^{\infty} \left(\sum_{i=1}^{C} a_i W_i^{\eta} \right)^{\frac{1}{\eta}} y^{-1/\eta} \exp(-y) dy \tag{D23}$$

$$= \left(\sum_{i=1}^{C} a_i W_i^{\eta} \right)^{\frac{1}{\eta}} \int_0^{\infty} y^{-1/\eta} \exp(-y) dy \tag{D24}$$

$$= \left(\sum_{i=1}^{C} a_i W_i^{\eta} \right)^{\frac{1}{\eta}} \cdot \Gamma\left(1 - \frac{1}{\eta}\right), \tag{D25}$$

where $y = x^{-\eta} \sum_{i=1}^{C} a_i W_i^{\eta}$ and $\Gamma(\cdot)$ is a constant term, which corresponds to the Gamma function evaluated at $\frac{\eta-1}{\eta}$. This is the expression that we have in the main text.

Appendix D.2 The Effect of MNEs on Formal vs. Informal Employment

In this section, we try to understand how does the expansion of MNEs affect formal domestic vs. informal domestic employment. Since we want to understand this effect at the local level, we omit trade and migration linkages. Because of this, we also omit the location subindex c since the economy is closed. We are interested in the effect on the ratio $\frac{L_F}{L_I}$, then, we start analyzing the effect on formal varieties and then we study the effect on informal employment and finally we take the difference for the ratio.

Appendix D.2.1 Formal Domestic Employment

From the solution to formal employment, we know that:

$$l_F(\varphi) = \frac{x_F(\varphi)}{A_F \varphi} = \frac{\Phi p_F(\varphi)^{-\sigma}}{A_F \varphi} = \frac{(\alpha P^{\sigma} Q + (1-\beta)\frac{\sigma-1}{\sigma} R_M (1 - t_Y^M) P_{FM}^{\sigma-1}) p_F(\varphi)^{1-\sigma}}{p_F A_F \varphi} \tag{D26}$$

$$= \frac{\sigma - 1}{\sigma} \frac{(1 - t_Y^F)}{w_F (1 + t_L^F)} \left(\alpha P^{\sigma} Q + (1-\beta)\frac{\sigma-1}{\sigma} R_M (1 - t_Y^M) P_{FM}^{\sigma-1} \right) p_F(\varphi)^{1-\sigma} \tag{D27}$$

We can rewrite this as:

$$l_F(\varphi) = \left(\frac{\sigma - 1}{\sigma} \right) \frac{(1 - t_Y^F)}{w_F (1 + t_L^F)} \left(X \cdot \phi_F(\varphi) + (1-\beta)\frac{\sigma-1}{\sigma} R_M (1 - t_Y^M) \cdot \chi(\varphi) \right), \tag{D28}$$

where $X \equiv PQ$ is total expenditure by households (inclusive of fixed cost payments) and $\phi_F(\varphi) = \alpha \frac{p_F(\varphi)^{1-\sigma}}{P^{1-\sigma}}$ and $\chi(\varphi) = \frac{p_F(\varphi)^{1-\sigma}}{P_{FM}^{1-\sigma}}$ are the expenditure shares of final consumers and MNEs in the variety produced by firm φ. Integrating over all firms that decide to operate in the formal sector, we obtain that total formal

employment is:

$$L_F = \left(\frac{\sigma-1}{\sigma}\right)\frac{(1-t_Y^F)}{w_F(1+t_L^F)}\left(X\int_{[\tilde{\varphi},\bar{\varphi}]}\phi_F(\varphi)\vartheta(\varphi)d\varphi + (1-\beta)\frac{\sigma-1}{\sigma}R_M(1-t_Y^M)\right), \tag{D29}$$

where we are using the fact that MNEs spend all their expenditure in intermediate inputs in formal domestic varieties, that is:

$$\int_{[\tilde{\varphi},\bar{\varphi}]}\chi(\varphi)\vartheta(\varphi)d\varphi = 1.$$

Let's define $\phi_F = \int_{[\tilde{\varphi},\bar{\varphi}]}\phi_F(\varphi)\vartheta(\varphi)d\varphi$. This term corresponds to the total expenditure share of final consumers in formal domestic varieties. Then, the total amount of labor in the formal sector is given by:

$$L_F = \left(\frac{\sigma-1}{\sigma}\right)\frac{(1-t_Y^F)}{w_F(1+t_L^F)}\left(X\phi_F + (1-\beta)\frac{\sigma-1}{\sigma}R_M(1-t_Y^M)\right). \tag{D30}$$

The total number of workers in the formal domestic sector depends negatively on the wage in that sector and positively in the expenditure share of consumers in formal domestic varieties and in the revenue obtained by MNEs. The total revenue of MNEs is given by the following expression:

$$R_M = p_M Q_M = \phi_M PQ, \tag{D31}$$

where ϕ_M is the expenditure share of final consumers in the MNE, given by:

$$\phi_M = (1-\alpha)\frac{p_M^{1-\sigma}}{P^{1-\sigma}}. \tag{D32}$$

Let's suppose that there is an increase in the productivity of foreign MNEs, A_M. We are interested in the effect on formal domestic employment. Starting from equation (D30), and abstracting from any income effects (meaning any changes in wages) and choosing X as the numeraire (so that it is not affected by the shock), we obtain that:

$$\begin{aligned}\hat{L}_F &= \frac{\overline{X}\overline{\phi}_F}{\overline{X}\overline{\phi}_F + (1-\beta)\frac{\sigma-1}{\sigma}\overline{R}_M(1-t_Y^M)}\widehat{(X\phi_F)} \\ &+ \frac{(1-\beta)\frac{\sigma-1}{\sigma}\overline{R}_M(1-t_Y^M)}{\overline{X}\overline{\phi}_F + (1-\beta)\frac{\sigma-1}{\sigma}\overline{R}_M(1-t_Y^M)}\left((1-\beta)\widehat{\frac{\sigma-1}{\sigma}R_M}(1-t_Y^M)\right),\end{aligned}$$

where $\hat{x} = d\ln x$ for any variable x. We can write this as:

$$\hat{L}_F = \lambda\hat{\phi}_F + (1-\lambda)\hat{R}_M, \tag{D33}$$

where λ represents the initial share of employment of domestic formal varieties that goes to producing output for final consumers, and $1-\lambda$ is the share that goes to producing output for MNEs. We know that:

$$\hat{\phi}_F = d\ln\phi_F = \frac{d\phi_F}{\phi_F}, \tag{D34}$$

we can apply Leibniz rule to $\phi_F = \int_{[\tilde{\varphi},\bar{\varphi}]} \phi_F(\varphi)\vartheta(\varphi)d\varphi$ and obtain that:

$$d\phi_F = -\phi_F(\tilde{\varphi})d\tilde{\varphi} + \int_{[\tilde{\varphi},\bar{\varphi}]} d\phi_F(\varphi)\vartheta(\varphi)d\varphi. \tag{D35}$$

If we assume that (to first order) the shock does not affect the threshold ($d\tilde{\varphi} = 0$), then:

$$\hat{\phi}_F = \int_{[\tilde{\varphi},\bar{\varphi}]} \frac{d\phi_F(\varphi)}{\phi_F(\varphi)} \frac{\phi_F(\varphi)}{\phi_F}\vartheta(\varphi)d\varphi = \int_{[\tilde{\varphi},\bar{\varphi}]} \hat{\phi}_F(\varphi)\frac{\phi_F(\varphi)}{\phi_F}\vartheta(\varphi)d\varphi. \tag{D36}$$

Recall that $\phi_F(\varphi) = \alpha\frac{p_F(\varphi)^{1-\sigma}}{P^{1-\sigma}}$, so we get:

$$\hat{\phi}_F(\varphi) \quad = \quad (1-\sigma)\hat{p}_F(\varphi) + (\sigma-1)\hat{P}$$

We will use that: $\hat{p}_F(\varphi) = -\hat{A}_F$. We also know that:

$$p_M \quad = \quad \frac{\sigma}{\sigma-1}\frac{(w_M(1+t_L^M))^\beta P_{FM}^{1-\beta}}{A_M(1-t_Y^M)}$$

Since P_{FM} is the aggregate price index at which formal firms sell to the multinational, and all the formal prices satisfy $\hat{p}_F(\varphi) = -\hat{A}_F$, then we also know $\hat{P}_{FM} = -\hat{A}_F$. Then (recall that we are keeping wages, including w_M, unchanged for now), we conclude that:

$$\hat{p}_M \quad = \quad -(1-\beta)\hat{A}_F - \hat{A}_M.$$

Recall that, since:

$$\begin{aligned} A_F &= \tilde{A}_F L_M^{\gamma_{M,F}} \\ A_I &= \tilde{A}_I L_M^{\gamma_{M,I}}, \end{aligned}$$

then:

$$\hat{A}_F = \gamma_{M,F}\hat{L}_M \tag{D37}$$

$$\hat{A}_I = \gamma_{M,I}\hat{L}_M. \tag{D38}$$

Denote with $\delta_M \equiv \hat{L}_M/\hat{A}_M > 0$ the elasticity of MNE employment w.r.t. MNE productivity. We now focus on the revenue side of MNEs, from (D31) we have that

$$R_M = \phi_M X,$$

hence (recall that X was chosen as numeraire and is not changing):

$$\hat{R}_M = \hat{\phi}_M.$$

Use the definition of ϕ_M:

$$\phi_M \quad = \quad (1-\alpha)\frac{p_M^{1-\sigma}}{P^{1-\sigma}}$$

$$\hat{\phi}_M = (1-\sigma)\hat{p}_M + (\sigma-1)\hat{P}. \tag{D39}$$

Putting all the things we have derived so far, we get:

$$\begin{aligned}
\hat{L}_F &= \lambda\hat{\phi}_F + (1-\lambda)\hat{R}_M \\
&= \lambda((1-\sigma)\hat{p}_F + (\sigma-1)\hat{P} - (1-\sigma)\hat{p}_M - (\sigma-1)\hat{P}) + (1-\sigma)\hat{p}_M + (\sigma-1)\hat{P} \\
&= \lambda(1-\sigma)(\hat{p}_F - \hat{p}_M) + (1-\sigma)\hat{p}_M + (\sigma-1)\hat{P} \\
\frac{\hat{L}_F}{\hat{A}_M} &= \lambda(\sigma-1)\left((1-(1-\beta))\frac{\hat{A}_F}{\hat{A}_M} - 1\right) - (1-\sigma)\left((1-\beta)\frac{\hat{A}_F}{\hat{A}_M} + 1\right) + (\sigma-1)\frac{\hat{P}}{\hat{A}_M} \\
&= \lambda(\sigma-1)\left(\beta\gamma_{M,F}\delta_M - 1\right) + (\sigma-1)\left((1-\beta)\gamma_{M,F}\delta_M + 1\right) + (\sigma-1)\frac{\hat{P}}{\hat{A}_M} \\
&= (1-\lambda)(\sigma-1)\left(1-\beta\gamma_{M,F}\delta_M\right) + (\sigma-1)\gamma_{M,F}\delta_M + (\sigma-1)\frac{\hat{P}}{\hat{A}_M} \qquad \text{(D40)} \\
&= (1-\lambda)(\sigma-1)\left(1-\beta\gamma_{M,F}\delta_M\right) + (\sigma-1)\left(\frac{\hat{P}}{\hat{A}_M} - \hat{p}_F\right) \qquad \text{(D41)}
\end{aligned}$$

Appendix D.2.2 Informal Domestic Employment

Recall that, from equations in previous sections (adapted to the notation of a single location), we have that labor for informal firms is given by:

$$l_I(\varphi)^{\frac{1}{\sigma}} w_I \left(1 + a + \frac{2l_I(\varphi)}{b}\right) = \frac{\sigma-1}{\sigma}(1-t_Y^I)\left(A_I\varphi\right)^{\frac{\sigma-1}{\sigma}}\left(\alpha X P^{\sigma-1}\right)^{\frac{1}{\sigma}}$$

Let's log-linearize this (assuming that wages do not change):

$$\frac{\hat{l}_I(\varphi)}{\sigma} + \frac{\left(\frac{2\bar{l}_I(\varphi)}{b}\right)}{1 + a + \frac{2\bar{l}_I(\varphi)}{b}}\hat{l}_I(\varphi) = \frac{\sigma-1}{\sigma}\hat{A}_I + \frac{\sigma-1}{\sigma}\hat{P}$$

So, we obtain:

$$\hat{l}_I(\varphi) = (\sigma-1)\frac{1 + a + \frac{2\bar{l}_I(\varphi)}{b}}{1 + a + (1+\sigma)\frac{2\bar{l}_I(\varphi)}{b}}(\hat{A}_I + \hat{P})$$

Recall that the price of the informal firm is:

$$p_I(\varphi) = \frac{\sigma}{\sigma-1}\frac{w_I\left(1 + a + \frac{2l_I(\varphi)}{b}\right)}{(1-t_Y^I)A_I\varphi} \tag{D42}$$

We log-linearize this to obtain:

$$\hat{p}_I(\varphi) = \frac{\left(\frac{2\bar{l}_I(\varphi)}{b}\right)}{1 + a + \frac{2\bar{l}_I(\varphi)}{b}}\hat{l}_I(\varphi) - \hat{A}_I \tag{D43}$$

Introduce the previous expression for $\hat{l}_I(\varphi)$ into this to obtain:

$$\hat{p}_I(\varphi) = (\sigma-1)\frac{\left(\frac{2\bar{l}_I(\varphi)}{b}\right)}{1+a+(1+\sigma)\frac{2\bar{l}_I(\varphi)}{b}}(\hat{A}_I+\hat{P})-\hat{A}_I \tag{D44}$$

$$= \frac{(\sigma-1)\left(\frac{2\bar{l}_I(\varphi)}{b}\right)}{1+a+(1+\sigma)\frac{2\bar{l}_I(\varphi)}{b}}\hat{P}+\left(\frac{(\sigma-1)\left(\frac{2\bar{l}_I(\varphi)}{b}\right)}{1+a+(1+\sigma)\frac{2\bar{l}_I(\varphi)}{b}}-1\right)\hat{A}_I \tag{D45}$$

$$= \frac{(\sigma-1)\left(\frac{2\bar{l}_I(\varphi)}{b}\right)}{1+a+(1+\sigma)\frac{2\bar{l}_I(\varphi)}{b}}\hat{P}+\left(\frac{(\sigma-1)\frac{2\bar{l}_I(\varphi)}{b}-1-a-(1+\sigma)\frac{2\bar{l}_I(\varphi)}{b}}{1+a+(1+\sigma)\frac{2\bar{l}_I(\varphi)}{b}}\right)\hat{A}_I \tag{D46}$$

$$= \frac{(\sigma-1)\left(\frac{2\bar{l}_I(\varphi)}{b}\right)}{1+a+(1+\sigma)\frac{2\bar{l}_I(\varphi)}{b}}\hat{P}-\frac{1+a+\frac{4\bar{l}_I(\varphi)}{b}}{1+a+(1+\sigma)\frac{2\bar{l}_I(\varphi)}{b}}\hat{A}_I \tag{D47}$$

For example, if $\sigma=1$, or if $b\to\infty$, this boils down to $\hat{p}_I(\varphi)=-\hat{A}_I$. But, for the general case, introduce the previous equation into the expression for the aggregate price index:

$$\begin{aligned}
P^{1-\sigma} &= \alpha\int_{\underline{\varphi}}^{\tilde{\varphi}} p_I(\varphi)^{1-\sigma}\vartheta(\varphi)d\varphi+\alpha\int_{\tilde{\varphi}}^{\overline{\varphi}} p_F(\varphi)^{1-\sigma}\vartheta(\varphi)d\varphi+(1-\alpha)p_M^{1-\sigma}\\
(1-\sigma)\hat{P} &= \int_{\underline{\varphi}}^{\tilde{\varphi}}\alpha\frac{p_I(\varphi)^{1-\sigma}}{P^{1-\sigma}}(1-\sigma)\hat{p}_I(\varphi)\vartheta(\varphi)d\varphi+\int_{\tilde{\varphi}}^{\overline{\varphi}}\alpha\frac{p_F(\varphi)^{1-\sigma}}{P^{1-\sigma}}(1-\sigma)\hat{p}_F(\varphi)\vartheta(\varphi)d\varphi\\
&+ (1-\alpha)\frac{p_M^{1-\sigma}}{P^{1-\sigma}}(1-\sigma)\hat{p}_M\\
\hat{P} &= \int_{\underline{\varphi}}^{\tilde{\varphi}}\phi_I(\varphi)\hat{p}_I(\varphi)\vartheta(\varphi)d\varphi+\int_{\tilde{\varphi}}^{\overline{\varphi}}\phi_F(\varphi)\hat{p}_F(\varphi)\vartheta(\varphi)d\varphi+\phi_M\hat{p}_M\\
&= \int_{\underline{\varphi}}^{\tilde{\varphi}}\phi_I(\varphi)\hat{p}_I(\varphi)\vartheta(\varphi)d\varphi-\phi_F\hat{A}_F-\phi_M((1-\beta)\hat{A}_F+\hat{A}_M)
\end{aligned}$$

Introduce into this our previous expression for $\hat{p}_I(\varphi)$ to obtain:

$$\hat{P} = \omega\hat{P}-\zeta_I\hat{A}_I-\phi_F\hat{A}_F-\phi_M((1-\beta)\hat{A}_F+\hat{A}_M),$$

where:

$$\omega = \int_{\underline{\varphi}}^{\tilde{\varphi}}\phi_I(\varphi)\frac{(\sigma-1)\left(\frac{2\bar{l}_I(\varphi)}{b}\right)}{1+a+(1+\sigma)\frac{2\bar{l}_I(\varphi)}{b}}d\varphi \tag{D48}$$

$$\zeta_I = \int_{\underline{\varphi}}^{\tilde{\varphi}}\phi_I(\varphi)\frac{1+a+\frac{4\bar{l}_I(\varphi)}{b}}{1+a+(1+\sigma)\frac{2\bar{l}_I(\varphi)}{b}}d\varphi. \tag{D49}$$

While these are complicated objects, they are determined in the initial steady state, so we do not need to worry to much about them (notice, also, that $\phi_I=\zeta_I+\omega$, this might come in handy later on, but perhaps not). So, we can solve for $\hat{P}$, as:

$$(1-\omega)\hat{P} = -\zeta_I\hat{A}_I-\phi_F\hat{A}_F-\phi_M((1-\beta)\hat{A}_F+\hat{A}_M) \tag{D50}$$

$$\frac{\hat{P}}{\hat{A}_M} = -\tilde{\zeta}_I \gamma_{M,I} \delta_M - \tilde{\phi}_F \gamma_{M,F} \delta_M - \tilde{\phi}_M((1-\beta)\gamma_{M,F}\delta_M + 1) \quad \text{(D51)}$$

$$= -\tilde{\zeta}_I \gamma_{M,I} \delta_M - (\tilde{\phi}_F + (1-\beta)\tilde{\phi}_M)\gamma_{M,F}\delta_M - \tilde{\phi}_M, \quad \text{(D52)}$$

where a variable with a tilde represents the corresponding variable without tilde divided by $1-\omega$. Recall that:

$$L_I = \int_{\underline{\varphi}}^{\tilde{\varphi}} l_I(\varphi)\vartheta(\varphi)d\varphi$$

$$\hat{L}_I = \int_{\underline{\varphi}}^{\tilde{\varphi}} \frac{\bar{l}_I(\varphi)}{\bar{L}_I}\hat{l}_I(\varphi)\vartheta(\varphi)d\varphi$$

Introducing our expression for $\hat{l}_I(\varphi)$ into this, we obtain:

$$\hat{L}_I = (\sigma-1)(\hat{A}_I + \hat{P}) \int_{\underline{\varphi}}^{\tilde{\varphi}} \frac{1 + a + \frac{2\bar{l}_I(\varphi)}{b}}{1 + a + (1+\sigma)\frac{2\bar{l}_I(\varphi)}{b}} \frac{\bar{l}_I(\varphi)}{\bar{L}_I}\vartheta(\varphi)d\varphi$$

Lets call the integral (which, once again, is a steady state object) Γ. If $b \to \infty$, then Γ tends to one. But generally, $\Gamma \leq 1$. Then, we finally get:

$$\hat{L}_I = \Gamma(\sigma-1)(\hat{A}_I + \hat{P})$$

$$\frac{\hat{L}_I}{\hat{A}_M} = \Gamma(\sigma-1)\left(\gamma_{M,I}\delta_M + \frac{\hat{P}}{\hat{A}_M}\right)$$

Now, say we are interested in the ratio of formal to informal employment:

$$r \equiv \frac{L_F}{L_I}$$

Then we know:

$$\hat{r} = \hat{L}_F - \hat{L}_I$$

$$\frac{\hat{r}}{\hat{A}_M} = \frac{\hat{L}_F}{\hat{A}_M} - \frac{\hat{L}_I}{\hat{A}_M} \quad \text{(D53)}$$

Introducing our previous expressions we conclude:

$$\frac{\hat{r}}{\hat{A}_M} = (1-\lambda)(\sigma-1)\left(1-\beta\gamma_{M,F}\delta_M\right) + (\sigma-1)\gamma_{M,F}\delta_M + (\sigma-1)\frac{\hat{P}}{\hat{A}_M}$$
$$- \Gamma(\sigma-1)\left(\gamma_{M,I}\delta_M + \frac{\hat{P}}{\hat{A}_M}\right) \quad \text{(D54)}$$

Hence, we get:

$$\frac{\hat{r}}{\hat{A}_M} = (1-\lambda)(\sigma-1)\left(1-\beta\gamma_{M,F}\delta_M\right) + (\sigma-1)\gamma_{M,F}\delta_M$$
$$+ (\sigma-1)\frac{\hat{P}}{\hat{A}_M}(1-\Gamma) - \Gamma(\sigma-1)\gamma_{M,I}\delta_M$$

So, for example, if $\Gamma = 1$ (which happens, e.g., if there is no quadratic component), then we get:

$$\frac{\hat{r}}{\hat{A}_M} = (1-\lambda)(\sigma-1)\left(1-\beta\gamma_{M,F}\delta_M\right) + (\sigma-1)(\gamma_{M,F}-\gamma_{M,I})\delta_M, \tag{D55}$$

which is the expression that we have derived in section 4.8.

Appendix D.3 Change in real income

Let's assume that $b_i \to \infty$ for each region j. Define the following variable for the aggregate expenditure across domestic firms:

$$\phi_{cj,s\in\{F,I\}} = \int_{\varphi\in\varphi_{j,s}} \phi_{cj,s}(\varphi)\vartheta_j(\varphi)d\varphi.$$

We can rewrite the ex-ante real income measure as:

$$W_c = \frac{a_c I_c}{P_c} = \frac{a_c(\Phi_c + r_c)}{P_c} \tag{D56}$$

where a_c is a time-invariant amenity level in location c and r_c corresponds to the per-capita revenue from distortions that accrues to location c. In particular, $r_c = R_c/L_c$, where $R_c = \sum_{h,s\in\Omega_c}(\mu^L_{h,s}-1)\lambda_h\lambda_{hs|h}\bar{L}$ and Ω_c represents the set of transactions that is rebated to c. The change in real income is:

$$d\ln W_c = \iota_c d\ln\Phi_c + (1-\iota_c)d\ln r_c - d\ln P_c, \tag{D57}$$

We proceed to understand the effect of a change in the productivity of MNES in each one of these terms. The change in the ex-ante factoral average income is:

$$d\ln\Phi_c = \sum_s \lambda_{cs|c} d\ln w_{c,s} \tag{D58}$$

On the other hand, the change in the price index is:

$$\begin{aligned}
d\ln P_c &= \sum_{j,s\in\{F,I\}} \phi_{jc,s}(d\ln w_{j,s} + d\ln\mu^L_{j,s} - d\ln A_{j,s}) && \text{(D59)}\\
&+ \sum_j \phi_{jc,M}\beta d\ln w_{j,M} + d\ln\mu^L_{j,M} && \text{(D60)}\\
&+ \sum_j \phi_{jc,M}(1-\beta)d\ln P_{j,FM} + + d\ln\mu^F_{j,M} && \text{(D61)}\\
&- \sum_j \phi_{jc,M} d\ln A_{j,M}, && \text{(D62)}
\end{aligned}$$

where $P_{j,FM}$ represents the cost of MNEs in j of an intermediate input bundle of formal varieties. The change in this variable is:

$$d\ln P_{j,FM} = \sum_h \phi_{hj,FM}(d\ln w_{h,F} + d\ln\mu^L_{h,F} - d\ln A_{h,F}). \tag{D63}$$

Plugging in equation (D63), in the change in the price index we get that:

$$d\ln P_c = \sum_{j,s\in\{F,I\}} \phi_{jc,s}(d\ln w_{j,s} + d\ln \mu^L_{j,s} - d\ln A_{j,s}) \tag{D64}$$

$$+ \sum_j \beta\phi_{jc,M}(d\ln w_{j,M} + d\ln \mu^F_{j,M} - d\ln A_{j,M}) \tag{D65}$$

$$+ \sum_{j,h}(1-\beta)\phi_{jc,M}\phi_{hj,FM}(d\ln w_{h,F} + d\ln \mu^L_{h,F} - d\ln A_{h,F}), \tag{D66}$$

where $\phi_{hj,FM}$ captures the expenditure share of MNEs in location j of formal varieties in h. Defining the following exposure measures for consumers in c to goods from j:

$$\Psi_{jc,F} \equiv \phi_{jc,F} + \sum_h (1-\beta)\phi_{hc,M}\phi_{jh,FM} \tag{D67}$$

$$\Psi_{jc,I} \equiv \phi_{jc,I} \tag{D68}$$

$$\Psi_{jc,M} \equiv \phi_{jc,M}, \tag{D69}$$

the change in the price index is:

$$d\ln P_c = \sum_{j,s\in\{F,I\}} \Psi_{jc,s}(d\ln w_{j,s} + d\ln \mu^L_{j,s} - d\ln A_{j,s}) \tag{D70}$$

$$+ \sum_j \Psi_{jc,M}\beta(d\ln w_{j,M} + d\ln \mu^L_{j,M}) \tag{D71}$$

$$+ \sum_j \Psi_{jc,M}[(1-\beta)d\ln \mu^F_{j,s} - d\ln A_{j,M}]. \tag{D72}$$

Plugging in this expression in the change in real income, using the fact that $\gamma_{M,s} = \frac{d\ln A_{j,s}}{d\ln L_{j,M}}$, and with some algebra manipulation, we get that:

$$d\ln W_c = \sum_s \iota_c \lambda_{c,s|c} d\ln w_{c,s} - \sum_{j,s\in\{F,I\}} \Psi_{jc,s} d\ln w_{j,s} - \sum_j \beta\Psi_{jc,M} d\ln w_{j,M} \tag{D73}$$

$$+ \sum_j \Psi_{jc,M} d\ln A_{j,M} + \sum_{j,s\in\{F,I\}} \Psi_{jc,s}\gamma_{M,s} d\ln L_{j,M} \tag{D74}$$

$$- \sum_{j,s\in\{F,I\}} \Psi_{jc,s} d\ln \mu^L_{j,s} + \sum_j \Psi_{jc,M}(\beta d\ln \mu^L_{j,M} + (1-\beta)d\ln \mu^F_{j,M}) \tag{D75}$$

$$+ (1-\iota_c)d\ln r_c, \tag{D76}$$

which corresponds to the expression in section 4.8 for the change in real income in each location.

Appendix D.4 Change in Aggregate Welfare

Under the assumptions that i) $b_j \to \infty$ for all j; ii) the population shares, λ_j are constant; iii) the rebate in per-capita terms is proportional to income $r_c = \bar{t}\Phi_c$; and iv) the wedges represent taxes to value-added and labor, such that, $\mu^F_{j,M}$=1. The change in aggregate welfare only depends on: i) technological effects; ii) spillovers; and iii) reallocation effects.

Under these assumptions, the market clearing conditions are:

$$\lambda_j \lambda_{js|j} = \frac{1+\bar{t}}{\mu^L_{j,s}} \left(\sum_c \phi_{jc,s} \Phi_c \lambda_c + \{1\}_{s=F} \cdot \phi_{jc,M}(1-\beta) \sum_h \phi_{jh,sM} \Phi_h \lambda_h \right), \tag{D77}$$

where $\{1\}_{s=F}$ is an indicator variable that takes the value of 1 if s corresponds to the formal segment of the economy. The change in aggregate welfare is:

$$d \ln U = \sum_c \lambda_c d \ln W_c \tag{D78}$$

Adding up the change in real income from the previous expression and since the rebate is constant, we obtain:

$$\begin{aligned}
d \ln U &= \sum_{c,s} \lambda_c \lambda_{c,s|c} d \ln w_{c,s} - \sum_{c,j,s \in \{F,I\}} \lambda_c \Psi_{jc,s} d \ln w_{j,s} - \sum_{c,j} \beta \lambda_c \Psi_{jc,M} d \ln w_{j,M} && \text{(D79)} \\
&+ \sum_{c,j} \lambda_c \Psi_{jc,M} d \ln A_{j,M} + \sum_{c,j,s \in \{F,I\}} \lambda_c \Psi_{jc,s} \gamma_{M,s} d \ln L_{j,M} && \text{(D80)} \\
&- \sum_{c,j,s \in \{F,I\}} \lambda_c \Psi_{jc,s} d \ln \mu^L_{j,s} + \sum_{c,j} \Psi_{jc,M} (\beta d \ln \mu^L_{j,M} + (1-\beta) d \ln \mu^F_{j,M}) && \text{(D81)} \\
&+ d \ln(1+\bar{t}). && \text{(D82)}
\end{aligned}$$

From the market clearing conditions and these expressions with some algebra manipulation we get that:

$$\lambda_c \lambda_{c,s|c} - \Psi_{cj,s} = \lambda_c \lambda_{c,s|c} \left(\frac{\mu^L_{j,s} - 1 - \bar{t}}{1+\bar{t}} \right)$$

. Plugging in this expression in the change in welfare, we get:

$$\begin{aligned}
d \ln U &= -\sum_{c,s} \lambda_c \lambda_{c,s|c} \left(\frac{\mu^L_{j,s} - 1 - \bar{t}}{1+\bar{t}} \right) d \ln w_{c,s} && \text{(D83)} \\
&+ \sum_{c,j} \lambda_c \Psi_{jc,M} d \ln A_{j,M} + \sum_{c,j,s \in \{F,I\}} \lambda_c \Psi_{jc,s} \gamma_{M,s} d \ln L_{j,M} && \text{(D84)} \\
&- \sum_{c,j,s \in \{F,I\}} \lambda_c \Psi_{jc,s} d \ln \mu^L_{j,s} + \sum_{c,j} \Psi_{jc,M} (\beta d \ln \mu^L_{j,M} + (1-\beta) d \ln \mu^F_{j,M}) && \text{(D85)} \\
&+ d \ln(1+\bar{t}). && \text{(D86)}
\end{aligned}$$

Calculating the change in the proportional rebate:

$$\begin{aligned}
d \ln(1+\bar{t}) &= \left(\frac{1}{1+\bar{t}} \right) d\bar{t} && \text{(D87)} \\
&= \left(\frac{1}{1+\bar{t}} \right) \sum_{j,s} (\mu_{j,s} - 1) \lambda_j \lambda_{js|j} (d \ln w_{j,s} + d \ln L_{j,s}) + \mu_{j,s} \lambda_j \lambda_{js|j} d \ln \mu_{j,s}. && \text{(D88)}
\end{aligned}$$

Using the previous expression; normalizing GDP to 1, such that, $\bar{t} \sum_{j,s} \lambda_j \lambda_{js|j} = 1$, and with some algebra

manipulations, we get that the change in aggregate welfare is:

$$d \ln U = \sum_{c,s} \lambda_c \lambda_{c,s|c} \left(\frac{\mu_{j,s}^L - 1 - \bar{t}}{1 + \bar{t}} \right) d \ln L_{c,s} \tag{D89}$$

$$+ \sum_{c,j} \lambda_c \Psi_{jc,M} d \ln A_{j,M} + \sum_{c,j,s \in \{F,I\}} \lambda_c \Psi_{jc,s} \gamma_{M,s} d \ln L_{j,M}. \tag{D90}$$

Defining $\bar{\mu} = 1 + \bar{t}$ and replacing $\bar{t}$ in terms of $\bar{\mu}$, we get the expression for the change in aggregate welfare in the main text.

Appendix E Additional quantitative results

Appendix E.1 Algorithm

In this section, we explain the 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), but it incorporates firm heterogeneity. The algorithm works 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 θ. In each location we use a grid with the number of firms $\mathcal{M}_j$.
2. In each location define in the grid the set of firms that are formal or informal by determining the marginal firm $\tilde{\varphi}_j$ that is indifferent between operating in the formal or informal economy.
3. In the first inner loop, guess an initial vector of wages $\vec{w}^0$, the number of residents in each location $\vec{L}^0$, and total expenditure $\vec{P}^0 \cdot \vec{Q}^0$.
4. Given a vector $\vec{w}^t$, $\vec{L}^t$, and Q^t compute the following equations:

   - Labor supply equations:

   $$\lambda_{js|j} = \frac{B_{js} w_{js}^{\kappa}}{\sum_k B_{jk} w_{jk}^{\kappa}}, \tag{E91}$$

   the average income in location j

   $$\Phi_j^{\kappa} = \sum_k B_{jk} w_{jk}^{\kappa}, \tag{E92}$$

   and the total amount of labor units

   $$L_{js} = w_{js}^{-1} \Phi_j \lambda_{js} \lambda_j \bar{L} \tag{E93}$$

   - In a second inner loop, use a contraction mapping to solve the problem for informal firms in each location:
     - Guess an initial solution for the labor of informal firms $l_{j,I}^0(\varphi)$.
     - Update the solution for informal firms with the following equation:

     $$l_{j,I}(\varphi)^{t+1} = \left(\frac{\frac{\sigma-1}{\sigma}(1 - t_{j,Y}^I) \left(A_{j,I}\varphi\right)^{\frac{\sigma-1}{\sigma}} \Xi_j}{w_{j,I}\left(1 + \frac{2 l_{j,I}(\varphi)^t}{b_j}\right)} \right)^{\sigma} \tag{E94}$$

     - For each firm φ, compute the absolute difference $|l_{j,I}(\varphi)^{t+1} - l_{j,I}(\varphi)^t|$.
     - Stop the second inner loop when the difference is lower than some tolerance factor *tol* for all firms.
   - Solve the problem for formal firms:

   $$l_{j,F}(\varphi) = \left(A_{j,F}\varphi\right)^{\sigma-1} \left(\tilde{\sigma} w_{j,F}(1 + t_{j,T})\right)^{-\sigma} \sum_c \Phi_c \tau_{jc}^{1-\sigma}, \tag{E95}$$

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

$$\pi_{j,F}(\varphi) = (1 - t^{Y}_{j,F}) \sum_{c} p_{jc,F}(\varphi) q_{hc,F}(\varphi) - (1 + t^{L}_{j,F}) w_{j,F} l_{j,F}(\varphi)$$

$$\pi_{j,I}(\varphi) = (1 - t^{Y}_{j,I}) p_{j,I}(\varphi) q_{j,I}(\varphi) - (1 + t^{L}_{j,I}) w_{j,I} l_{j,I}(\varphi)$$

- Compute the revenue from distortions that accrues to each location:

$$R_c = \sum_{h,s \in \Omega_c} (\mu^{L}_{h,s} - 1) w_{h,s} L_{h,s}$$

- Compute total expenditure for each final consumers and location:

$$X_c = \left(\Phi_c L_c + R_c\right), \tag{E96}$$

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

$$L_{j,F} = \mathcal{M}_j \int_{\varphi \in [\bar{\varphi}_j, \bar{\varphi}]} l_{j,F}(\varphi) \vartheta(\varphi) d\varphi, \tag{E97}$$

$$L_{j,I} = \mathcal{M}_j \int_{\varphi \in [\underline{\varphi}, \bar{\varphi}_j]} l_{iI}(\varphi) \vartheta(\varphi) d\varphi, \tag{E98}$$

6. Compute total employment for MNEs in each location:

$$L_{j,M} = \frac{\beta}{A_{j,M}} \left(\frac{w_{j,M}(1 + t^{M}_{j,L})}{P_{j,FM}} \right)^{\beta - 1} \left(\tilde{\sigma} \frac{(w_{j,M}(1 + t^{M}_{j,L}))^{\beta} P^{1-\beta}_{j,FM}}{A_{j,M}(1 - t^{M}_{j,Y})} \right)^{-\sigma} \sum_{c} \tau^{1-\sigma}_{jc} \Psi_c, \tag{E99}$$

7. Compute the new vector of total revenue $\vec{P}^{t+1} \cdot \vec{Q}^{t+1}$ using equations 22 and new price indices using equations 25.

8. Calculate the difference between labor demand and labor supply, the number of residents, and the vector of total revenue in each location:

$$z_w = \frac{L^{D}_{j,s} - L^{S}_{j,s}}{L_{j,s}} \tag{E100}$$

$$\tilde{L'}^{t+1}_{c} = \bar{L} \cdot \frac{W^{\eta}_{c}}{\sum_h W^{\eta}_{h}} \tag{E101}$$

$$z_Q = P^{t+1}_{j} \cdot Q^{t+1}_{j} - P^{t}_{j} \cdot Q^{t}_{j} \tag{E102}$$

9. If $||(z_w, \tilde{L}^{t+1}_{c}, z_\chi, z_Q) - (0, \tilde{L}^{t}_{c}, 0, 0)|| < \epsilon^{\text{tol}}$, then the algorithm of the inner loop stops. Otherwise, update:

$$w^{t+1}_{j,s} = w^{t}_{j,s}(1 + \nu_w z_w) \tag{E103a}$$

$$\tilde{L}^{t+1}_{c} = \nu_L \tilde{L'}^{t+1}_{c} + (1 - \nu_L) \tilde{L}^{t}_{c}, \tag{E103b}$$

$$P^{t+1}_{c} \cdot Q^{t+1}_{c} = \nu_Q \tilde{P}^{t+1}_{c} \tilde{Q}^{t+1}_{c} + (1 - \nu_Q) P^{t}_{c} Q^{t}_{c} \tag{E103c}$$

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

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

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

$$z_\varphi = \tilde{\tilde{\varphi}}_j^{t+1} - \tilde{\varphi}_j^t \tag{E104}$$

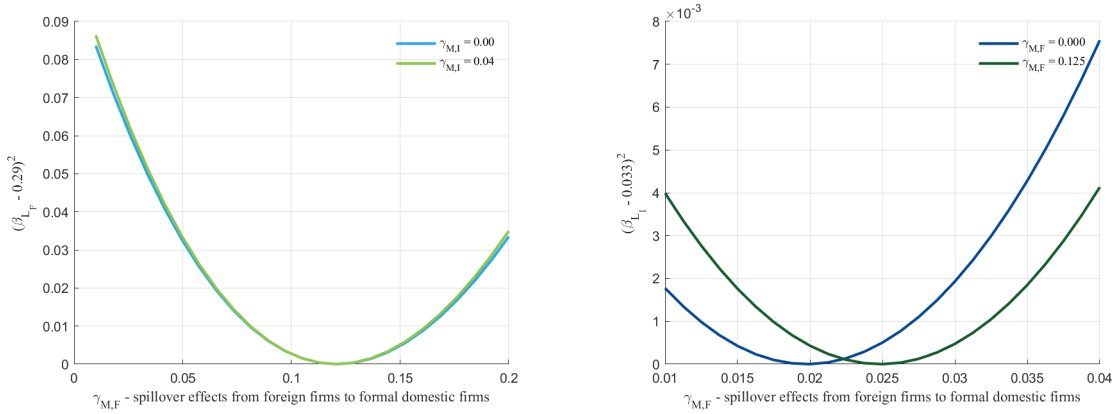
12. Update the new marginal firm:

$$\tilde{\varphi}_j^{t+1} = \nu_\varphi \tilde{\tilde{\varphi}}_j^{t+1} + (1 - \nu_\varphi)\tilde{\varphi}_j^t \tag{E105}$$

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

Appendix E.2 Additional Figures-Model Calibration

Figure E1: Effects on employment for different values of the spillover elasticities to the formal domestic ($\gamma_{M,I}$) and informal domestic ($\gamma_{M,I}$) producers



(a) Effect on formal employment

(b) Effect on informal employment

Notes: This figure plots the coefficients β from equation 40 for domestic formal and domestic informal employment relative to the point estimates from table 4 for several simulations changing the spillover effects from foreign MNEs to domestic establishments. Panel a plots the results for the coefficient of formal domestic employment and the spillover $\gamma_{M,F}$ holding constant the spillover of MNEs to the productivity of domestic informal firms, $\gamma_{M,I}$. Panel b plots the same figure for the coefficient of informal domestic employment for different values of the spillover $\gamma_{M,I}$ holding constant the spillover of MNEs to formal domestic producers, $\gamma_{M,F}$. Section 6 provides more details about the inference matching algorithm.

References in appendices

Fernando Alvarez and Robert Jr. Lucas. General equilibrium analysis of the Eaton-Kortum model of international trade. Journal of Monetary Economics, 54(6):1726–1768, September 2007.

David H Autor, David Dorn, and Gordon H Hanson. The China Syndrome: Local Labor Market Effects of Import Competition in the United States. American Economic Review, 103(6):2121–2168, 2013.

Juan S Blyde, Matias Busso, and Dario Romero Fonseca. Labor Market Adjustment to Import Competition: Long-Run Evidence from Establishment Data. IDB Working Paper Series, 2020.

Matías Busso, Óscar Fentanes, and Santiago Levy. The Longitudinal Linkage of Mexico's Economic Census 1999-2014. Inter-American Development Bank Technical Note, –(IDB-TN-1477), 9 2018.

INEGI. Censos Económicos 2019. Características y Manejo del Negocio. INEGI, 2021.

Steven Liao, In Song Kim, Sayumi Miyano, and Hao Zhang. Concordance: Product Concordance, 2020. R package version 2.0.0.

Stephen J. Redding and Esteban Rossi-Hansberg. Quantitative Spatial Economics. Annual Review of Economics, 9(1):21–58, September 2017.