Local Infrastructure and the Development of the Private Sector: Evidence from a Randomized Trial*

Daniel Rogger (World Bank) ⓡ
Leonardo Iacovone (World Bank) ⓡ
Luis F. Sánchez-Bayardo (World Bank) ⓡ
Craig McIntosh (UCSD) ⓡ

December 13, 2023

Abstract

We study how local public infrastructure investment affects neighborhood economies. By tracking the impacts of US$68 million of randomized investments in Mexican municipalities, we document how government investment leads to sustained increases in the size and profitability of treated private-sector companies. Initially, wages rise to compensate for higher costs of living, inefficient firms die, and more efficient firms grow faster. Over the subsequent decade treated firms increase their capital stocks and revenues, suggesting durable improvements in the structure of the local economy. Our results provide novel evidence of linkages between government investment, business growth, and the dynamics of local economies.

Keywords: Infrastructure, Firm Productivity, Randomized Trials
JEL Codes: R11, O18, D22, C93

*Ordering of authors is certified random (AEA randomization code: r4-FYDEDt8MZ). We gratefully acknowledge financial support from the World Bank's Development Impact Evaluation i2i and Knowledge for Change Program trust funds, Bureaucracy Lab and Governance Global Practice. We are grateful to John Loeser, Niclas Moneke, Caio Piza, Marta Santamaria, and Hannah Uckat; and, seminar participants at the 11th European Meeting of the Urban Economics Association at the London School of Economics and Political Science, the University of Oxford, and the World Bank for helpful comments. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

1 Introduction

Infrastructure investment is a central driver in the development of an urban economy [Bryan, Glaeser and Tsivanidis, 2020; Jones and Romer, 2010; Lucas, 1988]. Investment in the built environment alters the incentives for firms to start up, select locations, hire, invest, and shut down [Glaeser and Gottlieb, 2009; Glaeser, Luca and Moszkowski, 2020]. While canonical models of urban geography have long recognized productivity as an important channel to spatial differences in property prices [Diamond, 2016; Roback, 1982; Rosen, 1979], the empirical literature examining this channel directly has been limited by the challenges of matching exogenous shocks in infrastructure to firm-level data. In this paper we study this relationship using experimental variation in infrastructure spending and a granular multi-wave census of firms in urban Mexico to track the impacts of that spending.

The recent literature has highlighted two important features of any investigation of the impacts of local infrastructure development. Firstly, spatial general equilibrium models imply that changes in infrastructure can rearrange the patterns of production, commuting, and land values within a city [Almagro and Domınguez-Iino, 2019; Balboni et al., 2021; Franklin et al., 2023; Tsivanidis, 2019] or across cities [Allen and Arkolakis, 2014; Donaldson and Hornbeck, 2016]. A key question is the extent to which local investments generate net positive benefits to the city, versus delivering zero-sum gains that simply relocate business activity across space. Identifying the geographic distribution of impacts requires granular data as well as localized exogeneity in infrastructure investments.[1] To understand these shifts it is particularly important to be able to distinguish intensive-margin impacts on pre-existing firms from extensive margin changes in composition. Secondly, given the potential for agglomeration benefits to firms from shifts in infrastructure [Greenstone, Hornbeck and Moretti, 2010], it is critical to have well-identified evidence on the extent to which infrastructure impacts become self-reinforcing over time.

Our paper contributes to these debates by bringing together three unique features. First, we follow the large-scale randomized implementation of Mexico's Programa Hábitat, in which US$68m of spending on infrastructure was experimentally allocated across 370 neighborhoods in 68 Mexican municipalities over the course of three years. The empirical study of infrastructure has employed randomized experiments at the household- [Cattaneo et al., 2009; Galiani et al., 2017; Harari, Wong et al., 2018] or block-level [Gonzalez-

[1]Impacts across space are likely to be obscured in aggregated output data [Haughwout, 2002].

Navarro and Quintana-Domeque, 2016], or undertaken more macro non-experimental studies that estimate the aggregate multiplier on national infrastructure investment [Ramey et al., 2020]. Hábitat neighborhoods are small economic spheres of their own (on average 40 blocks), spread across Mexico, offering distributed and independently functioning economic environments in which to test the impacts of the program. Our study uses the experimental rollout to provide a detailed causal picture of the impacts of localized infrastructure investments on the local economy.[2] The experimental design randomized the 'saturation' levels of investment across neighborhoods. We use this feature to make precise causal statements about the impacts of the scheme on firms in neighboring areas as investment levels vary.

Second, we study these neighborhood changes using a census of Mexican firms conducted in 2008 (pre-treatment), 2013 (shortly post-treatment), and 2018 (six years after the end of the program). The fact that this data has linking identifiers across waves means that we are able to document impacts on the birth and death of firms over the course of a decade, as well as studying the dynamics of ongoing firms in both the short and medium term. Finally, both the intervention and the firm dataset are finely geolocated, allowing us to analyze the spatial pattern of firm responses in an unusually granular way. Put together, these features permit a unique window into the decade-long evolution of the economic geography of this important infrastructure program.

Our results prove surprising in a number of dimensions. Despite the fact that the program targeted its infrastructure investments towards residential amenities, it had a substantial effect on commercial activity. In the short-term, changes from the program are consistent with a wage shock coming from increased property values. One year after the end of investments, wage expenditures have jumped by 18%, with impacts of a similar magnitude in both manufacturing and trade and services. In parallel however, we also see an increase in the total number of workers, and a modest increase in revenues and capital investment. On the extensive margin, the treatment leads to an immediate shake-out with roughly 3% of the stock of firms being pushed out by the intervention on top of the natural level of turnover, with these firms being replaced by newer, smaller, faster-growing firms. Firms with high initial value added per worker are less like to 'die' in treatment neighborhoods, and increase their capital investments the most under the program. Hence a set of investments that focused on residential amenities led to substantial private-sector benefits in the short term.

[2] Previous work has shown that the program realized substantial improvements in residential amenities and property price increases of 10%, as well as meaningful improvements in safety and decreases in residential churn [McIntosh et al., 2018].

If anything, these impacts appear to accelerate over time. Looking six years after the end of investment, wages remain elevated, the number of workers has converged to the control group average, but the increases in capital stocks and revenue have accelerated. During this six year period, the impacts diverge by sector. By 2018 service sector firms in treatment neighborhoods have revenues 9% higher and capital stocks 22% higher than the control neighborhoods. Firms have re-optimized the capital-intensity of their production, reflecting a more 'mature' operating model. In manufacturing, the story is different. Revenues are only 3% higher in treatment neighborhoods, with no significant change in capital stocks. These longer-term effects are found almost entirely on the intensive margin and suggest that a self-reinforcing dynamic of growth has taken hold in these communities.

In terms of spatial agglomeration and the geography of impacts beyond the bounds of intervention neighborhoods, we recover a tightly estimated zero on any spillovers. We have an unusually clear window on spatial spillovers based on our ability to narrowly geo-locate infrastructure investment and measure firm activity over continuous space. Our experimental structure lets us define buffers around treatment and control neighborhoods in identical ways. Examining distance buffers as small as 100 meters and looking all the way up to a kilometer distant, we find small increases in firm churn in immediately adjacent blocks but no shifts in firm numbers or business outcomes at any distance, over either the short or the medium term. Proximity to highways has some effect of amplifying impacts, but the firm responses to this localized investment are quite invariant to measures of market access, either within study neighborhoods or surrounding them. Similarly, we uncover no larger-scale spillovers by exploiting the two-level design of the wtudy in which the fraction of neighborhoods treated within a municipality was also randomized. In contrast to the complex spatial tradeoffs demonstrated by the literature on transport infrastructure, then, the small-scale gentrification driven by this type of localized government infrastructure investment does not appear to have driven agglomeration or dislocation in surrounding neighborhoods.

These results contribute to our understanding of the mechanisms through which urban investment generates productivity and contributes to economic growth.[3] On the one hand, improvements in urban amenity values operate like a cost shock to the firm, pushing out a set of unproductive firms. At the same time, increases in property values improve consumer spending power and lead to a dynamic improvement in the growth

[3]Given the important role played by local actors and residents in deciding the specific investments to be made in Hábitat, our study also speaks to the large literature on Community-Driven Development (CDD) programs [Labonne and Chase, 2009; Mansuri and Rao, 2004; Paxson and Schady, 2002].

prospects of more productive firms. In many ways, these impacts mimic the effects of exposure to international trade which acts both to cull less productive firms and a vehicle for expansion, generating impacts on both the extensive and intensive margins of firm productivity [Mayer, Melitz and Ottaviano, 2021; Melitz and Redding, 2014]. What is different here is that these benefits are seen almost entirely among service firms that provide the non-tradeable products that can benefit from localized shifts in demand. The implication is that governments have a tool to drive sustainable increases in private-sector efficiency through investments in neighborhood infrastructure.[4]

The paper also contributes to the burgeoning empirical literature on the mechanisms of growth and gentrification in urban economics. Work on gentrification has emphasized increases in the skill of workers [Su, 2022], decreases in retail prices [Borraz et al., 2021], employment shifting from manufacturing to services [Hartley, Lester et al., 2013], and an increase in firm churn induced by exit of low-price firms the and growth of larger, higher-priced firms [Glaeser, Luca and Moszkowski, 2020]. A voluminous literature has tackled the impact of transport infrastructure on cities [Duranton and Turner, 2012], on market integration [Brooks and Donovan, 2020; Casaburi, Glennerster and Suri, 2013; Donaldson, 2018], and how such changes may be capitalized into land prices [Donaldson and Hornbeck, 2016; Tsivanidis, 2019]. By exploiting randomized variation in a program that generates infrastructural gentrification, we are able to nail down linkages between the constructed environment and the endogenous location and growth decisions of private firms. As such, this paper provides one of the most granular and well-identified assessments available of the ability of public infrastructure investment to stimulate private sector development.[5]

The rest of the paper is structured as follows: Section 2 provides the context for the study and describes the experimental design; Section 3 describes the data used, Section 4 provides the simple experimental results, and Section 5 explores the mechanisms through which the impacts are realized. Section 6 examines spillovers and conducts aggregate cost-benefit analysis at the neighborhood level, and Section 7 concludes.

[4]To assess the overall impact on government finances of such investments, we conduct an accounting exercise using changes in value added taxes, social security contributions, and revenue taxes to calculate that the tax take via the private sector rises by almost US$5 million per year in study neighborhoods. This implies that the program would pay for itself in 14 years simply through firm taxation. This suggests that cost-benefit evaluations of infrastructure programs using property prices alone to value benefits may miss an important vehicle for cost recovery via the private sector.

[5]By following a large-scale public investment program we also help to build the relatively sparse experimental literature that studies programs implemented at scale by national governments [Muralidharan and Niehaus, 2017].

2 Context and Experimental Design

2.1 The Private Sector in Urban Mexico

The Mexican economy is dominated by small businesses, with the median firm being a self-employed entrepreneur who is the only employee. Average revenue is only $15,600, and assets average just under $2,500. Micro-enterprises (defined by Mexico's Statistical Agency as having less than 10 employees) represent 95 percent of the total existing businesses. Only 2% of Mexican firms are fully formalized, 11% report having access to finance, and 2% report any online sales. This small average size and high level of informality is widely believed to hamper the overall productivity of the Mexican economy [Levy, 2010].

The vast majority of businesses belong to commerce and non-tradable services (over 90% of total). Within this group, grocery stores make up around 25% of the total, while stationery shops and beauty salons make up approximately 4% each. Manufacturing firms tend to be concentrated in activities that serve the local market, focusing on production of food and beverages, or activities related to construction and housing. Around 3% of firms produce corn tortillas and 1% are bakeries. Ironworks, furnishing and milling activities businesses comprise close to 1% of the total each.[6] Consequently, efforts to raise value-added and increase formality are of central concern.

A second pertinent feature of the Mexican business environment is high turnover of firms. To illustrate this, Figure 1 shows the birth and death rates of all firms observed across the three waves of the firm censuses we use (2008, 2013 and 2018). Green lines indicate firms that we see created, the red lines firms that we see closing down, and the blue line indicates firms that exist in all three rounds. Rates of churn are very high: 29% of firms in both subsequent rounds are newly born in that five-year period, and 18-20% of firms previously observed close down in each subsequent round. Only 17% of all firms observed survive through all three rounds of the data. In general the stock of firms grows over time as the rate of new business formation is 10% higher than the firm death rate in both subsequent rounds. Given this environment of rapid creation and destruction of firms, it will be critical to isolate and examine the extensive margin impacts of infrastructure as well as considering how investments may effect continuing firms. Fortunately, the unique firm-level identifier in the census data makes this straightforward in our case.

[6]Statistics based on the 2019 Economic Census.

2.2 The Hábitat Program

The Hábitat program was created under Mexico's Ministry of Social Development (SEDESOL) in 2003 to provide federal support for improvements in the infrastructure of marginalized neighborhoods in cities across the country (see Figure A1 for where the cities were located). The core purpose of the program is to make a suite of coordinated investments in residential amenities for previously under-served neighborhoods, thereby increasing livability and social cohesion [Campuzano et al., 2007]. The targeting and funding rules for the program are formulaic and centralized, and the program has tightly defined eligibility rules which require matching investments from state and municipal governments.[7]

In terms of project selection, Hábitat pursues decentralized community-driven mechanisms to allocate funding across potential investments. The actual investments made in a neighborhood are determined by the interplay of a set of technical experts from the program who make recommendations based on observed infrastructure deficits, and a locally driven project selection component.[8]

Typical Hábitat investment includes a mixture of physical infrastructure (street paving, sidewalk and median construction, electrification and sewerage connections, etc.) with spending on community centers, sports fields, and trainings. Figures A2 and A3 show 'before and after' photographs from Google Streetview in two intervention neighborhoods in Guadalajara, and illustrate the nature of typical changes in the neighborhood: street paving improved, and sidewalks, crosswalks, and bollards were installed. Importantly, nearly all of the spending under the program is for *residential* amenities. While a previous study analyzing the same experiment as this paper has shown that the program results in dramatic improvements in the walkability of and crime levels in treatment communities [McIntosh et al., 2018], Hábitat funds are typically spent on inputs that are not directly productive for the private sector.

Table A1 provides a breakdown of the money spent through the program, showing that roughly half of the spending went to street paving and almost a quarter to a set of social and community development activities (such as after-school youth activities in community centers and domestic violence prevention training). Even the money spent on

---

[7]These cost-sharing rules require local governments to providing 50% of project costs: municipalities provide 40%, the states 8%, and the beneficiaries 2%. So the study universe consists of municipalities that were willing and able to meet these matching requirements.

[8]The carefully orchestrated role played by local residents in proposing and vetting the use of funds makes this program similar in spirit to the large set of Community Driven Development (CDD) programs implemented across the developing world [Mansuri and Rao, 2004]. Explicit in the decision-making process was that municipal government would assume all maintenance costs of Hábitat infrastructure once the construction phase was completed.

roads and paving is primarily used to improve residential neighborhood roads and is not, for example, building trunk roads to better connect these peripheral neighborhoods with central parts of the city. Hence this study examines how the residential livability of a neighborhood, which we might more typically think of under the rubric of 'gentrification', drives outcomes for the private sector.

The program has clearly specified poverty targeting criteria. In order to be eligible to benefit from Hábitat, a neighborhood must consist of settled households in a marginalized urban areas with concentrations of asset poverty greater than 50%, located in cities of 15,000 inhabitants or more, with a deficit of infrastructure and urban services, and with at least 80% of the lots having no active conflict over property rights. This means that our study areas are typically poor outlying neighborhoods of major cities with high poverty and poor infrastructure, but relatively high levels of home ownership.

Eligibility was established in a very concrete spatial manner, whereby Hábitat defined 'polygons' that were clearly demarcated contiguous blocks that met the requirements for the program and in which the local layers of government were willing to invest. A Hábitat polygon is smaller than a locality and is a designation not used by other layers of government. Figures A4 to A7 illustrate the sizes of the treatment and control polygons relative to the overall sizes of Mexico City, Mérida, Tijuana and León, respectively. The average polygon in our study contains 40 blocks, 98 firms, 3800 inhabitants, and covers an area of 0.4 square kilometers. Because of the presence of simultaneous investment across multiple dimensions of urban infrastructure, undertaken in a highly targeted way, the program provides a unique opportunity to observe the impacts of dramatic improvements in residential amenities.

2.3 The Design of the Hábitat Experiment

We follow an experimental phase of the implementation of Hábitat in 2009-2012, in which a set of 370 'polygons' (or neighborhoods) in 68 municipalities across urban Mexico were randomly assigned to treatment.[9] The study featured a randomized saturation design, which first selected the fraction of study polygons that would be treated within each municipality using a uniform probability between .1 and .9 (so that all municipalities have some treated and some control polygons), and then assigned treatment to polygons according to this fraction. These sites contain 14,276 distinct blocks located in 38 cities, representing most of the large urban areas of Mexico (as exhibited in Figure

[9] Full details of the experimental design are provided in Ordóñez-Barba et al. [2013] and McIntosh et al. [2018]).

A1). Study polygons contained 3% of the population and 1% of the surface area of study municipalities. The randomization was conducted in 2009, the project selection process began immediately thereafter in treatment neighborhoods, and investments in the experimental locations ran from 2010-2012. 176 polygons were assigned to the treatment, and 194 to the control. US$68 million in federal, state, and municipal funding was invested in treatment polygons during the period of the study.

There was a subsequent, non-randomized expansion of the Hábitat program nationally in the years after the experiment ended, and a large share of the polygons involved in the study received some funding from that subsequent roll-out. However, not only was spending per surface area roughly one tenth of what was experienced during the experiment (roughly a million USD per square km in our experiment versus 100,000 USD per km in subsequent years), but the later program is perfectly balanced on the experiment (see Table A2 demonstrating that later implementation effectively ignored the experiment in targeting). Hence spending under Hábitat after the experiment effectively becomes a part of the regular background flow of infrastructure investment made by multiple levels of government.

3 Data

3.1 Hábitat database

The Hábitat database contains detailed geospatial information of the blocks, called *manzanas*, included in the study. The Hábitat study relies on Mexico's Statistical Agency (INEGI)'s identification system of blocks, which in most part is standardized across the agency's different projects. This makes it relatively simple to intersect data from the program with broader data sources compiled by the Mexican government, such as the population and firm censuses. Each block is located within a polygon and has a corresponding treatment/control status.

The Hábitat data contains substantial richness; it is possible to observe the exact type, amount, and location of each infrastructure upgrade a polygon received and on which year it occurred (2009, 2010 or 2011). Because the actual investments made in a given location were endogenous (both to the decisions of the Hábitat engineering team and to the community-driven selection process) we abstract away from this and analyze the treatment with a simple binary indicator.

3.2 Economic census database

The second data source is the Economic Censuses implemented by INEGI every five years. For this project, we use information of the firm censuses conducted in 2008, 2013 and 2018. The timing of these censuses is remarkably fortuitous for a study of Hábitat, given that the first interval allows us to conduct a before-after analysis of the short-term impacts of the program on the private sector, and the 2018 wave allows us to examine impacts 7 years after the cessation of investment.

The objective of these censuses is to capture information on firms which have a fixed location (i.e. not stands, stalls or other temporary buildings), irrespective of their formality status. Data is collected on income and expenses, labor, capital stock and a range of other variables. Businesses covered by the census are classified into manufacturing, services, and construction sectors. The census has a very high response rate, above 98% of all firms surveyed.

INEGI uses unique identifiers for each business surveyed. Thus, it is possible to follow firms through the censuses and hence to measure firm creation and destruction. The core variables used as outcomes for analysis are: firm revenue, number of paid workers, wage bill and capital stock. All financial variables are adjusted for inflation so as to represent constant 2008 US dollar values.

The INEGI survey also contains detailed information regarding firms' geographical location. Thus, firms can be placed on the block on which they are located within a city. This is crucial, as this geospatial information makes possible to cross this database with the Hábitat database and identify those firms contained within Hábitat polygons. We are able to locate 84,119 firms within Hábitat polygons.[10]

3.3 Summary Statistics and Balance

The universal nature of INEGI's firm census allows us to contextualize the study universe in a very simple way, by comparing the Hábitat control polygons to the broader universe of the cities in which these firms are located. Table 1 provides summary statistics for both firms in Hábitat polygons, and separately for those in the same city but outside of our study polygons. As can be seen, on average Hábitat firms are poorer, smaller and less productive.

However, the differences across the distribution of firms we study are not as sub-

[10]Given that there are slightly more control polygons, the majority of businesses are located in such polygons (roughly 60% of firms).

stantial as one might expect, given the poverty-targeting approach of the Hábitat rules. Figure 2 provides a visual representation of this comparison, showing the densities for our four major study outcomes: log revenue, number of paid workers, log wage bill, and log capital stock (we do not represent paid workers in logs because the majority of firms in the census have one paid employee, meaning that the firm owner is the only employee). The firms in control neighborhoods prove to be surprisingly representative of their cities as a whole. They are indeed slightly smaller in terms of revenues across the distribution, and they are somewhat more likely to have only one paid worker. In terms of wage bill and capital stocks they track the broader distributions quite closely. Overall, the data suggests that our study neighborhoods contain firms that are similar to broader urban Mexico as a whole, albeit using slightly less labor and generating slightly lower revenue.

Table A3 focuses on the firms located within study polygons to examine the balance of the experiment. It uses the pre-treatment data (2008) to present comparative summary statistics for the treatment and control polygons. The penultimate column presents simple comparisons, and the final column regression-based comparisons including municipality fixed effects and clustering standard errors at the polygon level, as in the main impact analysis. The adjusted comparisons are balanced across all sectors for all main outcomes (top panel), and when we disaggregate by manufacturing (middle panel) and services (bottom panel) we see one imbalanced outcome for each, in line with what we would expect by random chance. Overall these results suggest a well-balanced experiment. In our impact analysis we include the baseline polygon-level average level of the outcome variable as an ANCOVA control, which should remove any residual imbalances that do exist.

4 Results

4.1 Firm-Level Impacts

Our analysis uses a post-treatment cross-sectional ANCOVA specification:

$$Y_{ijm1} = \beta_0 + \delta\tau_{jm1} + \rho\bar{Y}_{jm0} + \gamma_{jm} + \epsilon_{ijm1} \tag{1}$$

where Y_{ijm1} is the post-treatment outcome for firm i in polygon j and municipality m, $\bar{Y}_{jm0}$ is the ANCOVA control (baseline mean outcome in that polygon), γ_{jm} is a set of fixed effects for municipality and for baseline polygon size, and ϵ_{ijm1} is a random error which we cluster at the polygon level to account for the design effect. In this specification, the estimand $\hat{\delta}$ on the post-treatment polygon-level dummy τ_{jm1} gives the intention-to-treat effect (ITT) of Hábitat on firms in treatment polygons. We use both 2013 and 2018 as outcome data, but always use 2008 as the year for the ANCOVA control. The variable $\bar{Y}_{jm0}$ is calculated at the polygon-level to solve the problem that would otherwise arise in using a firm-level baseline outcome (whose existence is endogenous if the treatment leads to extensive margin impacts).

Table 2 provides our main analysis of the impact of the program, using all extant firms in each round of the data and so providing an omnibus test that combines the intensive and extensive margin impacts of the program. The first two columns pool all types of firms together, and present impacts in 2013 (three years after the end of treatment) and 2018 (eight years later) in separate columns. Columns 3-4 analyze only manufacturing firms, and Columns 5-6 only trade and services firms.

Looking first at the short-term results that pool sectors, we see that a program has been shown elsewhere to have led to an 18% increase in residential rents and a 10% increase in property prices [McIntosh et al., 2018] has an impact that is consistent with a response to a cost shock: a substantial increase in the wage bill. However, far from cutting back on this now more-expensive labor, we also see an increase in the number paid workers in treatment areas relative to control. The wage bill increases by 18% (US\$.196 thousand over a base of US\$1.06 thousand), and the number of paid workers increases by 3% (.05 workers over a base of 1.35) indicating that both the number of workers and the wage per worker have increased in treatment areas. Both capital stock and revenue rise in the short term to an extent that is quantitatively meaningful (∼6%) but not significant (although both t-statistics above 1). Hence within a year or two of the cessation of the Hábitat investment, costs and employment have risen substantially and revenues have not kept pace.

Over the longer term however, the 2018 data paints a substantially rosier picture. Now 6-7 years after investments ended, revenue has risen by 8% (US\$1.98 thousand over a base of US\$23.4 thousand), capital stock by 17% (US\$1.07 thousand over a base of US\$6.2 thousand), and while the impacts on the number of employees have largely faded the impacts on the wage bill remain largely intact. Taken as a whole, this time path of impacts is suggestive of Hábitat investments acting in the short term as a cost shock to firms without compensation on the revenue side. However, over the longer term as the dynamics of greater residential wealth lead to superior demand, firms grow more quickly while remaining able to cover the higher wage bills necessitated by higher local residential costs. The positive longer-term impacts on revenues indicate that Hábitat induces meaningful medium-term changes to the demand faced by local firms.

The subsequent columns of Table 2 disaggregate these impacts by firm sector. In line with the literature on gentrification [Glaeser and Gottlieb, 2009; Glaeser, Luca and Moszkowski, 2020; Lester and Hartley, 2014] we find the impact of these residential amenity improvements to be entirely confined to service-sector firms. Service sector firms are poised to benefit from localized changes in the demand for (and possibly price of) local non-tradeable goods. In this sector we see revenues jump even in the short term by 4%, and over the longer term service firm revenues in treatment areas are higher by 9%, with capital stocks soaring by 22%. By 2018, manufacturing firms capital and labor stocks have converged with the control group, and revenue is only slightly higher, by approximately 3%. Hence this highly localized program has a very substantial benefit for, and only for, firms operating in the non-tradeable sector.

As discussed previously, in a business environment with such a high degree of churn, impacts on the stock of surviving firms could arise either through intensive margin changes for surviving firms, or through the selective margin by driving firm birth and death. We investigate these two dimensions in turn, beginning by considering firm creation and destruction as outcomes in a standard experimental context. To do this, Table 3 examines firm birth and death as outcomes of the treatment, so as to understand the extent to which the overall treatment impacts of the program on the composition of firms may be arising from entry and exit. The top panel of this figure defines the universe as all firms that existed in 2008, and examines an outcome variable which is a dummy for that firm having exited the market by the time of the post-treatment survey (2013 or 2018). The probit regression results show that the program leads to short-term excess firm death of 1.5 percentage points, or an increase of 3.5% over the control group death rate of 43 percent.[11] This differential actually decreases slightly when we look at 2018, providing preliminary evidence that most of the firm exit generated by the program is experienced immediately. These effects are confined entirely to service sector firms and the program had no impact on exit in the manufacturing sector.

The lower panel of this table looks at firm entry, now taking the universe as the endline sample of firms and defining a dummy variable for whether that firm is newly born since the baseline. Here the story is a mirror image (although less significant); increases in the rate of firm birth of around 1.3 percentage points, entirely concentrated

[11] Note that the percentages in this table are the fraction of baseline firms, while the percentages presented in Figure 1 are the percentage of all firms ever observed.

in the service sector, and mostly experienced in the short term. Taken as a whole then, we can summarize the extensive margin results quite simply by saying that as of 2018 the treatment had resulted in about 1.5 percent of the total distribution of firms being different than the ones that would have existed in the absence of the program, with no effect on the total number of firms.

We can dig deeper into the dynamics of entry and exit by looking at the firm birth and death that occurs between 2013 and 2018; this helps us to understand whether the program continues to exert a dynamic selection effect on the composition of firms. In Table A4 we therefore use the post-treatment 2013 survey as our baseline and examine entry and exit between that year and 2018. Using this (admittedly endogenous) post-treatment yardstick for subsequent growth we see no extensive margin impacts, meaning that the compositional effects of the program were relatively immediate. Hence the program leads to a short-term shake-out on the extensive margin but does not exert subsequent composition effects.

Given these meaningful but not qualitatively massive extensive margin effects, we suspect that the program has led to growth of firms on the intensive margin. While this story is difficult to tell with perfect experimental clarity, a simple way of posing the question is to restrict the sample to the (endogenous) group of firms that survive from baseline, and looking at impacts on these continuing market participants. Table 4 conducts this exercise and finds impacts that are roughly twice as large as the overall impacts found in Table 2. Here we see really large effects; for example service firms in the treatment area that survive from 2008 to 2018 see revenues that are 13% higher and capital stocks that are fully one third higher than comparable firms in the control. Hence the overall treatment effect is a composite of a large increase in the size of surviving firms with a relatively small increase in the turnover of firms on the extensive margin. Because newly entering firms are on average smaller than incumbents, this increase in churn actually dampens the total ITT effect of the treatment on firms size relative to the impact on ongoing firms. We now turn to a more detailed analysis of the ways in which the treatment altered the composition of market participants.

4.2 Heterogeneity in Impacts

We can perform a straightforward test of heterogeneity for firms that were observed at baseline; hence we begin our analysis by looking at the intensive margin heterogeneity of treatment effects for firms present both at baseline and endline. This analysis is presented in Table 5. This table defines dummies using the baseline distribution of

Revenue and Value Added, identifying firms that were in the top 25, 50, or 75 percent of the relevant distribution before the program began. It then interacts treatment with these classifications to ask whether the impact of Hábitat is larger for firms that were more productive to begin with. While the short-term results are more equivocal, it is clear from this table that by 2018 firms that were in the top half of the original distribution of value added have grown more in the treatment neighborhoods. Revenue, capital stock, and wage bill have all grown significantly more for firms that began with higher productivity, and indeed insignificant uninteracted treatment terms meaning that all of the impact of the program arises in the top half of the original productivity distribution. So the new opportunity provided in these neighborhoods is exclusively seized by productive firms.

Similarly, for firms present at baseline it is straightforward to ask whether firms that were initially less productive are those most likely to exit as a result of the program. Recalling that the treatment effect on death of firms appears in 2013 and not 2018, we again find evidence of the strongest firms surviving best. In Table A5, we see the uninteracted Hábitat treatment dummy suggesting an elevation of 3.2 percentage points in the probability of firm death, and the interaction effect on being in the top quartile of baseline productivity is -2.5 percentage points, meaning that these most efficient 25% of firms see little elevation in exit. Therefore, virtually all of the short-term firm death caused by the program is occurring in the unproductive firms.

The analysis of firm entry is less straightforward in that by definition we do not observe pre-treatment heterogeneity. What we can do is to examine whether there are differences between the attributes of newly created firms between the treatment and control; these differences would be a composite of true extensive margin selection effects on entry as well as the intensive impacts of the treatment on firm growth between creation and the time of the survey. This analysis, in Table A6, also lines up with the idea that the treatment is having meaningful impacts on the distribution of firm productivity, with entering service-sector firms being superior on most core outcomes in 2018. So, while we cannot cleanly say that these firms *entered* being more productive, it does appear to be the case that firm growth was fastest and firm death lowest among firms that were originally productive in the treatment, and new treated firms grew faster. Thus heterogeneity in the response to treatment by more productive firms plays an important role in explaining the total effects we observe.

5 Mechanisms for Firm-Level Impacts

The question of what is driving these results is important not just to better understand the nature of change in Hábitat neighborhoods, but also to appreciate whether this was a broader structural change in the local economy. We turn now to each of three core elements of structural transformation - access to credit, firm formality, and the response of residents and consumers.

5.1 Access to Credit and Financial Services

First, we look at the ways in which financial and technological access may have been a driver of business expansion. We find significant evidence of a financial channel behind the transformation and expansion of firms in treated polygons. As shown in Table 6, these businesses have a 20 percent higher probability of having secured a loan in 2013, an effect driven entirely by firms in the service sector. Point estimates, while small in absolute terms, are sizeable when we compare them to the control mean. While this variable includes all sources of credit, answers are very similar if restricted to formal lenders only, implying that new loans are obtained through the formal financial system (i.e. banks or savings cooperatives). We then see a higher probability of surviving firms having a bank account in 2018, which we interpret as the loan directly helping businesses to obtain formal access to the financial system.

The nature of the businesses in these areas is such that their key margin along which they can expand their economic complexity is through purchasing new and improved machinery and fixed assets rather than the adoption of more sophisticated types of technologies such as IT equipment. As shown previously, we find the businesses in the treated polygons expand their capital stock and purchases of machinery and equipment, supported by their expanded access to loans. The right-hand columns of Table 6 illustrate that use of computers is not changed, though service firms see an increase in access to the internet in 2013 which while strongly significant is only a half of a percentage point change in absolute value.

Importantly, this expanded access to credit does not appear to be mechanically driven by an expansion in the value of businesses' collateral as the value of the owned property increases. This is shown in Table A7, which analyzes the uptake and sources of credit, as well as the uses to which it is put, splitting the sample according to whether businesses own the property on which they operate or not. While businesses with land collateral have somewhat better baseline access to credit (13% versus 9% for those without), the treatment effects of Hábitat are virtually identical: a short-term expansion of 2.5 percentage points in 2013 and no longer-term effect. Unsurprisingly firms with land collateral are more likely to be served by formal banks and less likely to rely on savings banks, and non-landed firms put more of their money into land acquisition and inputs. But the reduced-form change in credit access is not being driven by those businesses that own land, removing as a potential explanation for mechanisms the fact that the private sector expands under residential investment strictly through the collateral value channel. Since collateral value does not seem to drive changes on the supply side, it appears that demand-side shifts arising from improved sales opportunities and profitability are the most reasonable explanation for the credit expansion.

5.2 Firm Formalization

The second mechanism at play which we think explains the changes occurring among businesses in the treated polygon is formalization. Our measure of formalization relies on the definition of Busso, Fazio and Levy [2012] and focuses on contributing to social security (i.e. having formal workers).[12] The advantage of using the level of social security contribution as a measure of formality is that this better captures the fact that the business is not only legal from a tax perspective but it is substantially contributing to generate higher quality formal jobs as its employees are covered by social security benefits (i.e. pension, health insurance, etc.). As suggested by Busso, Fazio and Levy [2012] we estimate that businesses should be on average paying the equivalent of 18 percent of total wages in social security contributions to be fully complying with their social security regulations ("strict formality definition"). However, this is an upper bound of their contribution and firms paying social security contributions that are below 18 percent of the wage total could still be fully compliant with social security regulations. We therefore also assess a second measure of formality as those firms that pay any social security contributions ("relaxed formality definition"). Accordingly, the latter is our preferred measure of business formality.

As shown in Table 7 we find that Habitat increases the likelihood of being formal that is driven primarily by businesses in the services sector (although this is a rare case where we see positive impacts of Hábitat on manufacturing firms as well). While point estimates are small in absolute terms we should observe that the prevalence of formalization among these types of businesses is very low (0.3 percent using our stricter

[12] We do not focus on the formal aspect of having a tax ID as it is much more common among businesses in Mexico and less stringent measure of the degree of formalization of a business in Mexico.

definition of formality and 2.4 percent using our more liberal definition) and our result imply an increase in the likelihood of being formal compared to the control mean equal to 17 percent in 2013 and 4 percent in 2018. While the results are mainly driven by the transformation of incumbent businesses that change their status from informal to formal, there are also some effects in the service sector along the extensive margin, with newly entered firms being more likely to be formalized as well.

5.3 Neighborhood Population

Together, our results imply broad changes in the characteristics of firms serving residents in treatment neighborhoods. How do residents respond in turn? This matters for the interpretation of our findings, with our firm level results implying changes in the nature of local consumption. With greater revenues for dominantly non-tradeable items, it is likely that much of the implied consumption is local. To what extent are our results driven by new and different populations moving into treatment neighborhoods, or by existing populations changing their consumption patterns?

McIntosh et al. [2018] document an increase in private investment in housing, with householders incorporating the higher amenity value of their surroundings into home upgrading. In particular, they observe significant upgrades to flooring and plumbing, with a 12 percent increase in the likelihood of a home containing a flush toilet. They also document the fact that though home ownership rates do not change significantly in treatment areas, property values rise substantially and rental costs rise by almost 20 percent. This finding is consistent with the increase in wage bills that we observe for firms in Table 2.

We assess the issue more broadly by analyzing the Mexican Census of Population and Housing 2020, also provided by INEGI and integrated with the same set of blocks and polygons as our core analysis. Table 8 presents our results. We regress, at the polygon level, measures from the 2020 Population Census on a treatment dummy and values of the variable from the 2010 Population Census. As such, the analysis is in the form of an ANCOVA specification, allowing us to present the most precise assessments our data allow. We split the analysis into variables related to the levels (or corresponding percentage) of the variable (Panel A) and log transformations of those same variables (Panel B) to check for robustness from outliers.

Overall, we do not find significant effects of the Hábitat program on the structure or characteristics of the population. Columns 1-3 assess the size of the population within our study polygons in terms of total, female and male populations. In each case the coefficient is small and insignificant at the usual levels. Similarly, we find no evidence that there is a higher proportion of adults or children in treatment neighborhoods (Column 4). Columns 5-7 indicate that the population of treatment neighborhoods are similar to control along a number of important margins. They are no more educated, no more likely to be employed, nor married.

It does not, therefore, seem that the changes in the private sector we observe are driven by significant changes in the demographic characteristics of populations in Hábitat neighborhoods. This is consistent with the limited change in home ownership rates observed previously. Rather, the results are consistent with the upgrading of neighborhoods changing the consumption patterns of local residents.

A remaining question is whether these changes were driven by wealth effects from the injection of investment capital from Programa Hábitat or other indirect effects of the program. McIntosh et al. [2018] provides detailed measurements of house price changes based on the assessments of professional property assessors from the Instituto de Administración Avaluos de Bienes Nacionales (INDAABIN), the Mexican government's institute of real estate valuation. They find that "the treatment group had almost triple the real rate of appreciation as the control", implying significant increases in the wealth of many of the treatment polygon's residents.

Bringing together the insights from the economic and population censuses, and the results from McIntosh et al. [2018], we see that Programa Hábitat had impacts on the nature of the private sector that do not seem to have been driven by changes in the underlying population being served but rather their core spending power. The program seems to have shifted the structure of the local economy - by which we mean the consumption choices of neighborhood residents and production choices by neighborhood firms - to a different equilibrium. That equilibrium had many characteristics of a more mature service economy - bigger, more capital-intensive firms with a greater likelihood of indicators of formality for example. This interpretation implies that government infrastructure investments can induce structural economic change at a very localized level.

6 Understanding the Urban Geography of Hábitat

6.1 Spillovers to Adjacent Neighborhoods

Spatial spillover effects are critical for interpreting the underlying model of economic geography revealed by the program. If these investments are driving agglomeration effects, we should see increases in investment and TFP in surrounding areas as in Greenstone, Hornbeck and Moretti [2010], and the narrow consideration of the Hábitat polygons would represent an under-estimate of total benefits. Alternatively, if the program has simply driven improvement in neighborhood amenities ala Rosen-Roback, economic activity might be spurred by an increase in demand from greater local housing wealth, but no underlying change in factor productivity would have occurred. Even in this case, if business growth came at the expense of adjacent neighborhoods, it would help to inform our understanding of the urban geography of demand. For these reasons, an understanding of spatial externalities is key both for welfare interpretations as well as for testing the 'productivities' interpretation against the 'amenities' story.

Our study provides an unusually clear window on the existence of spillovers, using a granular analysis using geographic buffers around the study polygons. Because we know the exact physical boundaries of the infrastructure and we have an experimental counterfactual in which no investment took place, we have a straightforward and statistically well-powered way of examining whether Hábitat drove changes in adjacent neighborhoods. Further, the neighborhoods actually treated are relatively small and non-contiguous, meaning that both treatment and control polygons are small islands surrounded by non-study neighborhoods. Therefore, we can look at buffers of up to 1 km without running into problems of overlapping treatment statuses, as can be a problem in more intensive spatial treatments such as the one studied in Franklin et al. [2023].

We begin from the outlines of the study polygons and define buffers as small as 100m and as large as 1 km in the way for the treatment and control alike (the distance buffers are non-inclusive, meaning that the 100-250m buffer does not include the 100m buffer). We then locate INEGI firms within each buffer and examine the differential outcomes for firms at different distances treatment and control polygons. This approach is simple and experimental in an attractive way, and has very similar statistical power to the overall study (especially as we look at larger buffers that contain more firms).[13]

[13]As in the main analysis we cluster at the polygon level and include municipality fixed effects.

This approach reveals some increase in the churn of firms in immediately adjacent neighborhoods, but surprisingly weak spillovers overall. Appendix Table A8 examines the likelihood that a firm dies or is born during the ten years from 2008-2018, finding the probability of both entry and exit rising by 3-4 percentage points with the overall number of firms staying relatively stable. Table 9 finds no consistent evidence of spillovers on firm outcomes at any distance; there are no results significant at the 5% level for all firms at any distance.[14] The simple takeaway is therefore that the impacts of the program are highly localized, generating no consistent spillovers even at 100 meters, and so we have neither any measurable agglomeration effects, nor any observable crowd-out of business activity from adjacent firms.

6.2 Heterogeneity by Market Access and Road Networks

To connect the effects from Hábitat more deeply into the geography of the surrounding cities function, we can examine heterogeneity of treatment and spillover effects in several key dimensions of market access. Neighborhoods with strong market access may themselves benefit more from infrastructure investment if they have more opportunities to exploit dense linkages, and certainly we may expect that the spillover effects of the program to surrounding neighborhoods will be larger when market access is better. We explore four metrics of market access: local population, local non-poor population, polygon size, and distance to highways.[15]

We first look for evidence of impact heterogeneity within study polygons, on the extensive margin in Table A9 and on firm-level outcomes in Table A10. Neither metric of market access generates meaningful heterogeneity. Polygon size does matter, in that larger polygons see lower churn on the extensive margin but larger increases in revenue and wage bills, at least in the short term. Proximity to roads drives a short-term

[14]Given that spillovers may both shift the composition of firms as well as altering outcomes on the intensive margin for pre-existing firms, for each buffer we show the total effect (all endline firms), the impact on 'survivors' (firms that existed at both baseline and the relevant endline), and 'entrants' (new firms since baseline). Here we do see limited evidence of positive spillovers in revenue for existing firms at the nearest distances, but these fade quickly with distance and are not present in any of the other outcomes or for the overall group of firms.

[15]The first calculates an inverse-distance weighted average population density around the centroid of each Hábitat polygon (treatment and control), as a measure of the local total market size. The second measure emphasizes the spending power of the adjacent population by calculating the same metric but using only the population above the poverty line. Our third metric explores the idea that our polygons may simply be too small to drive spillover effects. To investigate this we calculate the size of each polygon in square meters and use geographic size as a dimension of heterogeneity. Finally, we try to tie the study to the commuting map of the city by measuring the minimum distance from each polygon centroid to the nearest major road.

improvement in revenues and capital stock, and a longer-term improvement in paid workers. We can then apply this same approach to our spillover analysis, using buffers and firms defined as before but now studying how the magnitude of spillovers varies with attributes of the market access of the polygon. Table A11 looks at the extensive margin, and Table A12 the intensive margin. All four measures of market access have the effect of increasing churn among immediately adjacent firms, but as above new births are roughly balanced by new deaths, leaving the total number of firms relatively stable. In no case do we see effects that are clearly monotonic in distance, as we would expect in any continuous spatial model of market access over such small areas.[16]

6.3 Saturation-driven Spillovers at the Municipality Level

Spillover effects need not be spatial in the manner analyzed in the previous section; general equilibrium or market-level impacts of the program may alter commercial patterns more broadly. The two-level randomization embedded in Hábitat's rollout offers a unique experimental lens on these city-wide effects, and we have the universe of firms in every city and so can capture them very comprehensively. By exploiting the varying saturation levels of Hábitat investment implemented as part of the experimental design, our analysis has a rather unique lens on how higher levels of investment might lead to greater spillovers on neighboring firms.

This approach suffers from low power in that Hábitat neighborhoods represent a small fraction of the overall city ($\tilde{2}$% of surface area), but nonetheless represents an attractive and design-based way to understand the impact of localized treatment on the city as a whole.[17] Table 10 analyzes outcomes in non-study neighborhoods as a function of the treatment saturation in study neighborhoods in that municipality, clustering at the municipality-level to reflect the design effect from this component of the experiment. Consistent with the overall findings from our analysis of geographic spillovers, the results again suggest a lack of significant spillovers.

[16]Given the large number of coefficients presented in Table A12 it is important to view them through the lens of multiple comparisons; we present 192 interaction effects in this table and find 8 significant at the 1% level, 15 significant at the 5% level, and 14 significant at the 10% level, not far above what would be expected by random chance.

[17]We construct the sample for this analysis by first eliminating the study polygons from the data, then collapsing the remaining blocks at the AGEB level (a geographical unit between a block and locality, typically comprising about 100 firms). We then regress outcomes for all remaining AGEBs in the data on the municipal-level treatment saturation, now clustering standard errors at the municipal level to reflect the design effect from this component of the experiment. Table A13 illustrates that the saturation experiment is balanced on baseline outcomes.

Given these results, what can we conclude about the economic geography of this type of program? Whether one expects to see spillover effects that are positive (agglomeration) or negative (diversion), our design allows us to say is that improvements in local private sector opportunity are not achieved by helping or hurting surrounding areas. This says that the pixel size of our neighborhoods (typically 40 blocks) is not large enough to trigger larger neighborhood dynamics that either contribute to or detract from adjacent parts of the city. Rather, it appears to have generated a relatively homogeneous and remarkably localized set of impacts that are narrowly concentrated in the immediate location of the investments. The conclusion is that at the scale of investment we study, the total welfare effect of the program is likely localized within study neighborhoods.

7 Conclusion

We bring together a large-scale experiment in the construction of infrastructure with a three-round census on firm activity in urban Mexico. Studying a program that spent US$68 million over three years and across 65 municipalities, we find a powerful and durable response of private-sector firms to improvements in neighborhood amenities. Wages, employment, and capital investment all rise. Infrastructure improvements accelerate churn in firms, and appears to lure 'better' entrants, particularly in the service sector. Firms use credit to expand whether or not they have land as collateral, and formalization rates rise. If anything, these impacts accelerate over time. Firms appear to shift onto a higher path of revenue growth that is continuing to expand relative to the control six years after the end of the program. This pattern is consistent with wealth-driven increases from local consumers boosting demand, and suggests that our study captures a kind of small-scale randomized gentrification. Improvements in urban amenity value directly enhance opportunities for local firms in the non-tradeable sector.

In contrast to much of the recent urban literature we find minimal spillover effects, whether using a simple buffer neighborhood approach, looking at how measures of market access mediate treatment and spillover effects, or assessing the impact of the scale of infrastructure investment on its consequences for neighboring firms. Given the type of infrastructure built here this may not be surprising. In contrast to the large recent literature on transport infrastructure that shifts commuting patterns, Hábitat did very little to shift linkages *between* neighborhoods and instead improved community amenities *within* them. Unlike programs that achieve infrastructure improvement through large-scale public employment such as Ethiopia's Urban PSNP [Franklin et al., 2023], Hábitat created few direct local jobs and its impact on urban labor markets appear to arise mostly through stimulating local demand.[18]

The limited spillovers imply that the incidence of the program is confined to study areas, allowing a simple assessment of the ways in which tax receipts could recoup the costs of the program. We undertake a back-of-the envelope calculation that calculates the effects of the intervention on the total value of firm outcomes that are taxed by the government (Table A14), and then applies the appropriate marginal tax rates to these totals to back out the fiscal implications of the intervention (Table A15). These rates are 16% for Value Added, 2% for firm revenue, and then contributions to the pension system *IMSS* are directly asked in the INEGI survey. These calculations imply that the Hábitat investments we study increased annual government receipts of US$4.7 million; which put up against the US$67 million cost of the program, suggests that the program will pay for itself in a little over 14 years just through tax intake from private-sector firms. This estimate is conservative in that it ignores impacts on property taxes (either for firms or households), but is strongly suggestive that private sector taxation represents a meaningful channel through which the costs of infrastructure programs can be recouped by governments even when the private sector was not the target of those investments, and even when the impacted businesses are small.

Our results imply that while property values may provide a sufficient summary statistic in terms of static welfare, the existence of a causal relationship between property values and private sector growth adds a dynamic multiplier on to the benefits of this type of spending. The fact that the private sector responds to shifts in property values, that these changes are dynamic and durable, and do not appear to come at the cost of growth in neighboring areas, all suggest that private sector taxation is an underappreciated channel for recovering the costs of infrastructure spending.

Overall, the distributed, experimental, and independent nature of the Hábitat investments allow us to isolate how infrastructure investments kick-start localized agglomeration effects. The granularity and precision of our estimates showcase the dynamics of this process and its limitations. Local infrastructure investment can be a spur to the development and reshaping of the private sector, even in the most local of economies.

[18] The 40-block size of the neighborhood improvements generated in this experiment may be unnaturally small relative either to standard large infrastructure projects, or to the size of neighborhood that we might typically see 'gentrify'. Seen in this light, our results shed light on the minimum scale of investment required to kick-start localized agglomeration effects. Indeed, the subsequent rollout of Hábitat that took place in the years subsequent to the study did typically operate over substantially larger contiguous geographic areas.

References

Allen, Treb, and Costas Arkolakis. 2014. "Trade and the Topography of the Spatial Economy." *The Quarterly Journal of Economics*, 129(3): 1085–1140.

Almagro, Milena, and Tomás Domınguez-Iino. 2019. "Location sorting and endogenous amenities: Evidence from amsterdam." *NYU, mimeograph.*

Balboni, Clare, Gharad Bryan, Melanie Morten, and Bilal Siddiqi. 2021. "Could Gentrification Stop the Poor from Benefiting from Urban Improvements?" Vol. 111, 532–537, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Borraz, Fernando, Felipe Carozzi, Nicolás González-Pampillón, and Leandro Zipitría. 2021. "Local retail prices, product varieties and neighborhood change." *American Economic Journal: Economic Policy (forthcoming).*

Brooks, Wyatt, and Kevin Donovan. 2020. "Eliminating uncertainty in market access: The impact of new bridges in rural Nicaragua." *Econometrica*, 88(5): 1965–1997.

Bryan, Gharad, Edward Glaeser, and Nick Tsivanidis. 2020. "Cities in the developing world." *Annual Review of Economics*, 12: 273–297.

Busso, Matías, Maria Fazio, and Santiago Levy. 2012. "(In) formal and (un) productive: The productivity costs of excessive informality in Mexico." *IDB Working Paper No. IDB-WP-341.*

Campuzano, Larissa, Dan Levy, Andres Zamudio, et al. 2007. "The Effects of Hábitat on Basic Infrastructure." *Princeton, NJ: Mathematica Policy Research.*

Casaburi, Lorenzo, Rachel Glennerster, and Tavneet Suri. 2013. "Rural roads and intermediated trade: Regression discontinuity evidence from Sierra Leone." *Available at SSRN 2161643.*

Cattaneo, Matias D, Sebastian Galiani, Paul J Gertler, Sebastian Martinez, and Rocio Titiunik. 2009. "Housing, health, and happiness." *American Economic Journal: Economic Policy*, 1(1): 75–105.

Diamond, Rebecca. 2016. "The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000." *American Economic Review*, 106(3): 479–524.

Donaldson, Dave. 2018. "Railroads of the Raj: Estimating the impact of transportation infrastructure." *American Economic Review*, 108(4-5): 899–934.

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American economic growth: A "market access" approach." *The Quarterly Journal of Economics*, 131(2): 799–858.

Duranton, Gilles, and Matthew A Turner. 2012. "Urban growth and transportation." *Review of Economic Studies*, 79(4): 1407–1440.

Franklin, Simon, Clement Imbert, Girum Abebe, and Carolina Mejia-Mantilla. 2023. "Urban public works in spatial equilibrium: Experimental evidence from Ethiopia."

Galiani, Sebastian, Paul J Gertler, Raimundo Undurraga, Ryan Cooper, Sebastián Martínez, and Adam Ross. 2017. "Shelter from the storm: Upgrading housing infrastructure in Latin American slums." *Journal of urban economics*, 98: 187–213.

Glaeser, Edward L, and Joshua D Gottlieb. 2009. "The wealth of cities: Agglomeration economies and spatial equilibrium in the United States." *Journal of economic literature*, 47(4): 983–1028.

Glaeser, Edward L, Michael Luca, and Erica Moszkowski. 2020. "Gentrification and Neighborhood Change: Evidence from Yelp." National Bureau of Economic Research.

Gonzalez-Navarro, Marco, and Climent Quintana-Domeque. 2016. "Paving streets for the poor: Experimental analysis of infrastructure effects." *Review of Economics and Statistics*, 98(2): 254–267.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings." *Journal of Political Economy*, 118(3): 536–598.

Harari, Mariaflavia, Maisy Wong, et al. 2018. "Slum upgrading and long-run urban development: Evidence from indonesia."

Hartley, Daniel, T William Lester, et al. 2013. "The long-term employment impacts of gentrification in the 1990s." Federal Reserve Bank of Cleveland.

Haughwout, Andrew F. 2002. "Public infrastructure investments, productivity and welfare in fixed geographic areas." *Journal of public economics*, 83(3): 405–428.

Jones, Charles I, and Paul M Romer. 2010. "The new Kaldor facts: ideas, institutions, population, and human capital." *American Economic Journal: Macroeconomics*, 2(1): 224–45.

Labonne, Julien, and Robert S Chase. 2009. "Who is at the wheel when communities drive development? Evidence from the Philippines." *World Development*, 37(1): 219–231.

Lester, T William, and Daniel A Hartley. 2014. "The long term employment impacts of gentrification in the 1990s." *Regional Science and Urban Economics*, 45: 80–89.

Levy, Santiago. 2010. *Good intentions, bad outcomes: Social policy, informality, and economic growth in Mexico.* Rowman & Littlefield.

Lucas, Robert E. 1988. "On the mechanics of economic development." *Journal of monetary economics*, 22(1): 3–42.

Mansuri, Ghazala, and Vijayendra Rao. 2004. "Community-based and-driven development: A critical review." *The World Bank Research Observer*, 19(1): 1–39.

Mayer, Thierry, Marc J Melitz, and Gianmarco IP Ottaviano. 2021. "Product mix and firm productivity responses to trade competition." *Review of Economics and Statistics*, 103(5): 874–891.

McIntosh, Craig, Tito Alegría, Gerardo Ordóñez, and René Zenteno. 2018. "The neighborhood impacts of local infrastructure investment: Evidence from urban Mexico." *American Economic Journal: Applied Economics*, 10(3): 263–86.

Melitz, Marc J, and Stephen J Redding. 2014. "Heterogeneous firms and trade." *Handbook of international economics*, 4: 1–54.

Muralidharan, Karthik, and Paul Niehaus. 2017. "Experimentation at scale." *Journal of Economic Perspectives*, 31(4): 103–124.

Ordóñez-Barba, Gerardo, Tito Alegría-Olazábal, Craig McIntosh, and René Zenteno-Quintero. 2013. "Alcances e impactos del Programa Hábitat en comunidades pobres urbanas de México." *Papeles de población*, 19(77): 231–267.

Paxson, Christina, and Norbert R Schady. 2002. "The allocation and impact of social funds: spending on school infrastructure in Peru." *The World Bank Economic Review*, 16(2): 297–319.

Ramey, Valerie A, et al. 2020. "The macroeconomic consequences of infrastructure investment." National Bureau of Economic Research.

Roback, Jennifer. 1982. "Wages, rents, and the quality of life." *Journal of political Economy*, 90(6): 1257–1278.

Rosen, Harvey S. 1979. "Housing decisions and the US income tax: An econometric analysis." *Journal of Public Economics*, 11(1): 1–23.

Su, Yichen. 2022. "The rising value of time and the origin of urban gentrification." *American Economic Journal: Economic Policy*, 14(1): 402–39.

Tsivanidis, Nick. 2019. "Evaluating the impact of urban transit infrastructure: Evidence from bogota's transmilenio." *Unpublished manuscript.*

Tables

Table 1: Descriptive statistics: by location

	2008				2013				2018			
	Obs.	Mean	Median	Std. Dev.	Obs.	Mean	Median	Std. Dev.	Obs.	Mean	Median	Std. Dev.
Firms in Habitat polygons												
Value added	36,063	7.22	2.48	13.18	44,156	6.76	2.74	11.40	50,822	9.88	4.19	24.91
Revenue	36,063	20.28	8.55	33.17	44,156	16.70	7.52	27.25	50,822	24.10	11.38	60.97
Capital stock	36,063	7.51	1.56	20.15	44,156	6.66	1.57	17.99	50,822	6.78	1.46	23.92
Investment	36,063	0.14	0.00	0.68	44,156	0.15	0.00	0.62	50,822	0.25	0.00	1.61
Value added per paid worker	36,063	4.94	2.11	8.94	44,156	4.91	2.51	7.95	50,822	6.11	3.44	14.19
Paid workers	36,063	1.46	1.00	1.41	44,156	1.40	1.00	1.33	50,822	1.58	1.00	1.76
Wage bill	36,063	1.41	0.00	4.94	44,156	1.20	0.00	4.49	50,822	1.92	0.00	7.34
Wage	6,623	2.90	2.71	1.69	6,461	2.91	2.82	1.59	11,233	3.04	2.91	1.52
Firms in rest of the city												
Value added	1,049,650	30.01	4.82	87.38	1,176,785	29.32	5.01	85.90	1,446,324	46.13	7.37	144.98
Revenue	1,049,650	80.11	13.64	261.05	1,176,785	73.01	11.27	240.88	1,446,324	105.74	17.29	354.33
Capital stock	1,049,650	27.32	2.14	104.99	1,176,785	25.02	1.88	97.26	1,446,324	26.95	1.92	106.32
Investment	1,049,650	0.78	0.00	4.54	1,176,785	0.81	0.00	4.22	1,446,324	0.79	0.00	4.37
Value added per paid worker	1,049,650	11.38	3.34	44.83	1,176,785	12.11	3.71	46.48	1,446,324	17.60	4.94	82.88
Paid workers	1,049,650	2.95	1.00	6.28	1,176,785	2.80	1.00	5.83	1,446,324	3.21	1.00	6.61
Wage bill	1,049,650	8.37	0.00	31.07	1,176,785	7.14	0.00	26.11	1,446,324	9.22	0.00	31.60
Wage	368,070	3.68	3.27	2.77	357,031	3.57	3.18	2.41	562,419	3.67	3.27	2.07

Table provides summary statistics at the firm level for all firms included in INEGI's census within any of the cities in which the Hábitat program was implemented, by location within city. The 2008 wave forms the baseline for this study, and the the 2013 and 2018 waves are post-treatment. For variables other than paid workers, values are in thousands of 2013 US dollars.

Table 2: Main regression results, all firms
Sample: Existing firms in endline, including those missing in baseline

	All sectors		Manufacturing		Trade and Services	
Dependent variable	2013	2018	2013	2018	2013	2018
Revenue	0.732	1.981**	0.687	0.758	0.684	1.984*
	(0.479)	(0.915)	(0.884)	(1.283)	(0.508)	(1.044)
	16.43	*23.43*	*20.36*	*25.86*	*15.90*	*23.10*
Capital stock	0.529	1.071***	-0.062	-0.022	0.648*	1.278***
	(0.371)	(0.399)	(0.783)	(0.711)	(0.358)	(0.412)
	6.340	*6.177*	*10.91*	*8.778*	*5.722*	*5.832*
Paid workers	0.050**	0.022	0.152***	0.009	0.040**	0.025
	(0.020)	(0.027)	(0.057)	(0.073)	(0.019)	(0.028)
	1.352	*1.540*	*1.870*	*2.184*	*1.282*	*1.455*
Wage bill	0.196***	0.198*	0.469**	-0.028	0.156***	0.223**
	(0.061)	(0.102)	(0.229)	(0.276)	(0.058)	(0.109)
	1.061	*1.771*	*2.958*	*4.124*	*0.805*	*1.458*
Observations	[44,156]	[50,822]	[5,339]	[6,034]	[38,791]	[44,750]

The table shows the impact of Hábitat on firms located within the areas treated by the program. Specifically, the value of coefficient δ in the following specification:
$Y_{ijmT} = \beta_0 + \delta\tau_{jmT} + \rho\bar{Y}_{jm2008} + FE_{Municipality} + FE_{Polygonsize} + \epsilon_{ijmT}$
Each coefficient denotes a different regression. Standard errors clustered by Hábitat polygon shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table 3: Effect of Habitat on probability of exit and entry of firms

	All sectors		Manufacturing		Trade and Services	
	2013	2018	2013	2018	2013	2018
Probability of exit						
Habitat	0.015*	0.013	0.006	0.008	0.017*	0.013
	(0.009)	(0.010)	(0.019)	(0.021)	(0.009)	(0.010)
	0.433	*0.596*	*0.461*	*0.624*	*0.429*	*0.592*
Observations	[36,063]	[36,063]	[4,777]	[4,777]	[31,286]	[31,286]
Probability of entry						
Habitat	0.013	0.007	0.005	-0.000	0.015	0.008
	(0.014)	(0.010)	(0.021)	(0.016)	(0.014)	(0.010)
	0.543	*0.718*	*0.534*	*0.713*	*0.544*	*0.718*
Observations	[44,156]	[50,822]	[5,365]	[6,072]	[38,791]	[44,750]

Table presents coefficients from a linear probability model. Specifically, the coefficient δ in the following specification:
$Pr(Y_{ijmT} = 1|\tau_{jmT}) = \beta_0 + \delta\tau_{jmT} + FE_{Municipality} + FE_{Polygonsize} + \epsilon_{ijmT}$
The probability of exit (top panel) is estimated among all firms present at baseline explaining whether they have exited by the indicated round. The probability of entry (bottom panel) is estimated among all firms present in the post-treatment waves explaining whether the firm is a new entrant in that round. Standard errors clustered by Hábitat polygon shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table 4: Main regression results, surviving firms

Sample: Existing firms in both baseline and endline

	All sectors		Manufacturing		Trade and Services	
Dependent variable						
	2013	2018	2013	2018	2013	2018
Revenue	0.611	3.830*	1.185	1.902	0.346	3.814
	(0.602)	(2.038)	(1.111)	(2.380)	(0.664)	(2.348)
	19.42	*29.59*	*24.45*	*33.82*	*18.72*	*29.02*
Capital stock	0.552	2.156**	-0.837	0.127	0.740*	2.505***
	(0.416)	(0.840)	(0.989)	(1.673)	(0.420)	(0.895)
	7.466	*8.531*	*13.87*	*14.21*	*6.579*	*7.762*
Paid workers	0.047**	0.057	0.215***	0.178	0.019	0.035
	(0.023)	(0.054)	(0.079)	(0.141)	(0.023)	(0.055)
	1.368	*1.600*	*1.951*	*2.431*	*1.287*	*1.487*
Wage bill	0.230***	0.398*	0.778**	0.640	0.124	0.349
	(0.083)	(0.238)	(0.301)	(0.595)	(0.077)	(0.251)
	1.196	*2.142*	*3.454*	*5.386*	*0.884*	*1.703*
Observations	[20,147]	[14,271]	[2,565]	[1,784]	[17,582]	[12,487]

The table shows the impact of Hábitat on surviving firms located within the areas treated by the program. Specifically, the value of coefficient δ in the following specification:

$Y_{ijmT} = \beta_0 + \delta\tau_{jmT} + \rho\bar{Y}_{jm2008} + FE_{Municipality} + FE_{Polygonsize} + \epsilon_{ijmT}$

Each coefficient denotes a different regression. Standard errors clustered by Hábitat polygon shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table 5: Heterogeneity in main results

	2013						2018					
	Initial value added per worker			Initial revenue			Initial value added per worker			Initial revenue		
	P-th>=25th	P-th>=50th	P-th>=75th	P-th>=25th	P-th>=50th	P-th>=75th	P-th>=25th	P-th>=50th	P-th>=75th	P-th>=25th	P-th>=50th	P-th>=75th
Revenue												
Habitat*Percentile	-0.185	1.230	0.601	1.169	1.899**	2.283*	3.754	6.445**	5.094	3.833	3.362	-0.591
	(0.817)	(0.881)	(1.165)	(0.827)	(0.944)	(1.377)	(3.008)	(2.749)	(4.563)	(2.867)	(3.241)	(4.462)
Habitat	0.883	0.186	0.630	-0.195	-0.229	0.113	1.188	0.947	2.875	1.082	2.272	4.070*
	(0.796)	(0.697)	(0.617)	(0.742)	(0.657)	(0.618)	(2.451)	(2.261)	(2.360)	(2.464)	(1.716)	(2.383)
Percentile	7.051***	7.530***	9.030***	9.726***	12.276***	16.495***	6.980***	7.305***	9.083***	10.684***	17.160***	23.047***
	(0.486)	(0.561)	(0.675)	(0.498)	(0.579)	(0.898)	(2.014)	(1.933)	(2.506)	(1.849)	(1.722)	(3.089)
	19.42	*19.42*	*19.42*	*19.42*	*19.42*	*19.42*	*29.59*	*29.59*	*29.59*	*29.59*	*29.59*	*29.59*
Capital stock												
Habitat*Percentile	0.451	0.840	0.472	-0.027	1.163*	1.227	2.938**	3.664***	3.873**	2.523*	1.798	3.090*
	(0.581)	(0.584)	(0.808)	(0.641)	(0.641)	(0.895)	(1.470)	(1.304)	(1.508)	(1.488)	(1.201)	(1.679)
Habitat	0.250	0.193	0.476	0.583	0.008	0.277	0.009	0.420	1.295	0.318	1.312	1.485*
	(0.557)	(0.465)	(0.414)	(0.579)	(0.461)	(0.433)	(1.366)	(1.003)	(0.878)	(1.417)	(0.891)	(0.855)
Percentile	1.460***	2.005***	1.923***	2.157***	3.255***	4.982***	-0.262	0.096	0.292	0.013	3.023***	4.039***
	(0.353)	(0.327)	(0.448)	(0.343)	(0.324)	(0.500)	(1.060)	(0.896)	(0.840)	(1.010)	(0.678)	(1.037)
	7.466	*7.466*	*7.466*	*7.466*	*7.466*	*7.466*	*8.531*	*8.531*	*8.531*	*8.531*	*8.531*	*8.531*
Paid workers												
Habitat*Percentile	0.106	0.228	-0.106	0.230	0.335**	0.332	0.665	0.615	0.482	0.446	0.591	0.596
	(0.138)	(0.158)	(0.205)	(0.151)	(0.156)	(0.232)	(0.514)	(0.383)	(0.419)	(0.479)	(0.366)	(0.510)
Habitat	0.160	0.132	0.261***	0.064	0.076	0.160*	-0.086	0.113	0.289	0.076	0.124	0.282
	(0.124)	(0.102)	(0.095)	(0.127)	(0.096)	(0.087)	(0.474)	(0.307)	(0.272)	(0.461)	(0.279)	(0.267)
Percentile	0.367***	0.491***	0.390***	0.554***	1.009***	1.599***	0.013	0.223	-0.069	0.343	1.128***	1.945***
	(0.069)	(0.077)	(0.116)	(0.085)	(0.090)	(0.140)	(0.332)	(0.276)	(0.260)	(0.309)	(0.178)	(0.275)
	1.196	*1.196*	*1.196*	*1.196*	*1.196*	*1.196*	*2.142*	*2.142*	*2.142*	*2.142*	*2.142*	*2.142*
Wage bill												
Habitat*Percentile	0.026	0.068	-0.031	0.051	0.088*	0.097	0.201*	0.176**	0.142	0.138	0.155*	0.099
	(0.044)	(0.049)	(0.059)	(0.046)	(0.047)	(0.068)	(0.104)	(0.081)	(0.096)	(0.094)	(0.079)	(0.112)
Habitat	0.030	0.018	0.056**	0.011	0.008	0.029	-0.088	-0.023	0.026	-0.042	-0.013	0.043
	(0.039)	(0.030)	(0.026)	(0.038)	(0.028)	(0.025)	(0.097)	(0.067)	(0.060)	(0.094)	(0.061)	(0.059)
Percentile	0.108***	0.133***	0.097***	0.168***	0.292***	0.456***	0.037	0.087	0.009	0.137**	0.326***	0.545***
	(0.021)	(0.023)	(0.033)	(0.025)	(0.026)	(0.040)	(0.070)	(0.057)	(0.058)	(0.067)	(0.043)	(0.060)
	1.368	*1.368*	*1.368*	*1.368*	*1.368*	*1.368*	*1.600*	*1.600*	*1.600*	*1.600*	*1.600*	*1.600*
Observations	[20,147]	[20,147]	[20,147]	[20,147]	[20,147]	[20,147]	[14,271]	[14,271]	[14,271]	[14,271]	[14,271]	[14,271]

Table presents interaction effects between Hábitat and initial value added per worker, or initial revenue. Columns interact treatment with dummies for being in or above the 25th, 50th, and 75th percentile of these covariates, with each triad of interaction, Hábitat and Percentile coefficients representing a different regression. The Table shows coefficients ϕ, τ and γ of the following specification:

$Y_{ijmT} = \beta_0 + \phi\tau_{jmT} * Percentile_{ijm2008} + \delta\tau_{jmT} + \gamma Percentile_{ijm2008} + \rho\bar{Y}_{jm2008} + FE_{Municipality} + FE_{Polygonsize} + \epsilon_{ijmT}$

Standard errors clustered by Hábitat polygon shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table 6: Regressions on mechanisms

	Firm had access to a loan (=0 no, =1 yes)			Firm had a bank account (=0 no, =1 yes)			Firm used computers (=0 no, =1 yes)			Firm had access to internet (=0 no, =1 yes)		
Dependent variables	All sectors	Manufacturing	Trade and Services	All sectors	Manufacturing	Trade and Services	All sectors	Manufacturing	Trade and Services	All sectors	Manufacturing	Trade and Services
All existing firms in 2013												
Habitat	0.022***	-0.013	0.027***	0.003	-0.003	0.004	0.002	0.005	0.002	0.004***	-0.002	0.004***
	(0.008)	(0.012)	(0.009)	(0.010)	(0.012)	(0.010)	(0.005)	(0.010)	(0.005)	(0.001)	(0.003)	(0.001)
	0.111	*0.107*	*0.112*	*0.0610*	*0.0906*	*0.0570*	*0.0897*	*0.0843*	*0.0904*	*0.00738*	*0.0110*	*0.00689*
Observations	[44,156]	[5,339]	[38,791]	[44,156]	[5,339]	[38,791]	[44,156]	[5,339]	[38,791]	[44,156]	[5,339]	[38,791]
All existing firms in 2018												
Habitat	0.001	-0.003	0.001	0.011**	0.006	0.011**	-0.006	-0.017*	-0.004	-0.003	-0.022**	-0.001
	(0.006)	(0.010)	(0.007)	(0.005)	(0.011)	(0.005)	(0.006)	(0.009)	(0.006)	(0.006)	(0.009)	(0.006)
	0.103	*0.0894*	*0.105*	*0.0547*	*0.0856*	*0.0506*	*0.0871*	*0.0853*	*0.0873*	*0.0741*	*0.0756*	*0.0739*
Observations	[50,822]	[6,034]	[44,750]	[50,822]	[6,034]	[44,750]	[50,822]	[6,034]	[44,750]	[50,822]	[6,034]	[44,750]
Surviving firms from 2008 to 2013												
Habitat	0.022**	-0.011	0.027***	0.003	-0.008	0.004	-0.000	-0.000	-0.001	0.003**	-0.008*	0.005***
	(0.010)	(0.015)	(0.010)	(0.008)	(0.015)	(0.008)	(0.005)	(0.012)	(0.005)	(0.001)	(0.004)	(0.002)
	0.111	*0.104*	*0.112*	*0.0611*	*0.102*	*0.0554*	*0.0847*	*0.0917*	*0.0838*	*0.00861*	*0.0189*	*0.00718*
Observations	[20,147]	[2,565]	[17,582]	[20,147]	[2,565]	[17,582]	[20,147]	[2,565]	[17,582]	[20,147]	[2,565]	[17,582]
Surviving firms from 2008 to 2018												
Habitat	0.001	-0.014	0.004	0.017**	0.013	0.017**	0.003	-0.011	0.004	0.004	-0.020	0.007
	(0.008)	(0.015)	(0.008)	(0.007)	(0.021)	(0.007)	(0.008)	(0.013)	(0.009)	(0.009)	(0.014)	(0.009)
	0.0909	*0.0898*	*0.0911*	*0.0596*	*0.116*	*0.0520*	*0.0824*	*0.0956*	*0.0806*	*0.0682*	*0.0849*	*0.0660*
Observations	[14,271]	[1,784]	[12,487]	[14,271]	[1,784]	[12,487]	[14,271]	[1,784]	[12,487]	[14,271]	[1,784]	[12,487]
Entrants from 2008 to 2013												
Habitat	0.022**	-0.013	0.028***	0.002	-0.008	0.005	0.006	0.004	0.008	0.004***	0.001	0.005***
	(0.009)	(0.015)	(0.009)	(0.011)	(0.014)	(0.012)	(0.006)	(0.013)	(0.007)	(0.001)	(0.003)	(0.002)
	0.112	*0.110*	*0.112*	*0.0610*	*0.0808*	*0.0584*	*0.0939*	*0.0778*	*0.0960*	*0.00635*	*0.00413*	*0.00664*
Observations	[24,009]	[2,800]	[21,209]	[24,009]	[2,800]	[21,209]	[24,009]	[2,800]	[21,209]	[24,009]	[2,800]	[21,209]
Entrants from 2008 to 2018												
Habitat	0.000	0.004	-0.001	0.006	0.002	0.006	-0.005	-0.023**	-0.002	-0.006	-0.024**	-0.004
	(0.007)	(0.011)	(0.007)	(0.006)	(0.011)	(0.006)	(0.006)	(0.011)	(0.006)	(0.005)	(0.010)	(0.006)
	0.108	*0.0893*	*0.111*	*0.0527*	*0.0734*	*0.0500*	*0.0889*	*0.0812*	*0.0900*	*0.0764*	*0.0718*	*0.0771*
Observations	[36,551]	[4,288]	[32,263]	[36,551]	[4,288]	[32,263]	[36,551]	[4,288]	[32,263]	[36,551]	[4,288]	[32,263]

Table presents coefficients from a linear probability model. Specifically, the coefficient δ in the following specification:
$Pr(Y_{ijmT} = 1|\tau_{jmT}) = \beta_0 + \delta\tau_{jmT} + \rho\bar{Y}_{jm2008} + FE_{Municipality} + FE_{Polygonsize} + \epsilon_{ijmT}$
Dependent variables are the post-treatment status of having access to a loan, having a bank account, using computers and having access to internet. 'All existing firms' panels include all firms existing at the respective endline. 'Surviving firms' panels only include firms that existed in both baseline and endline. 'Entrants' panels groups firms that did not exist at baseline. Standard errors clustered by Hábitat polygon shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets. Every coefficient denotes a different regression.

Table 7: Effects on Formality

	All sectors				Manufacturing				Trade and Services			
	2013		2018		2013		2018		2013		2018	
	Relaxed	Strict	Relaxed	Strict	Relaxed	Strict	Relaxed	Strict	Relaxed	Strict	Relaxed	Strict
All firms												
Habitat	0.004**	0.001**	0.001	0.001*	0.001	0.004*	-0.015**	-0.001	0.005**	0.001	0.003	0.002**
	(0.002)	(0.001)	(0.002)	(0.001)	(0.008)	(0.002)	(0.007)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)
	0.0237	*0.00262*	*0.0238*	*0.00292*	*0.0513*	*0.00252*	*0.0526*	*0.00498*	*0.0200*	*0.00264*	*0.0200*	*0.00265*
Observations	[44,156]	[44,156]	[50,822]	[50,822]	[5,339]	[5,339]	[6,034]	[6,034]	[38,791]	[38,791]	[44,750]	[44,750]
Survivors												
Habitat	0.005*	0.003***	0.007*	0.003*	0.012	0.008*	-0.026**	0.002	0.003	0.002**	0.010***	0.003*
	(0.003)	(0.001)	(0.004)	(0.001)	(0.012)	(0.005)	(0.013)	(0.007)	(0.002)	(0.001)	(0.004)	(0.001)
	0.0268	*0.00229*	*0.0283*	*0.00288*	*0.0654*	*0.00202*	*0.0792*	*0.00772*	*0.0215*	*0.00233*	*0.0214*	*0.00222*
Observations	[20,147]	[20,147]	[14,271]	[14,271]	[2,565]	[2,565]	[1,784]	[1,784]	[17,582]	[17,582]	[12,487]	[12,487]
Entrants												
Habitat	0.005	0.000	-0.000	0.001*	-0.004	0.001	-0.011	-0.002	0.006*	0.000	0.001	0.002**
	(0.003)	(0.001)	(0.003)	(0.001)	(0.008)	(0.001)	(0.007)	(0.002)	(0.003)	(0.001)	(0.003)	(0.001)
	0.0211	*0.00290*	*0.0221*	*0.00294*	*0.0389*	*0.00295*	*0.0419*	*0.00388*	*0.0188*	*0.00289*	*0.0195*	*0.00282*
Observations	[24,009]	[24,009]	[36,551]	[36,551]	[2,800]	[2,800]	[4,288]	[4,288]	[21,209]	[21,209]	[32,263]	[32,263]

Table presents coefficients from linear probability models explaining whether a firm is 'formal' according to a Relaxed or Strict definition. Namely, the table shows coefficient δ in the following specification:
$Pr(Y_{ijmT} = 1|\tau_{jmT}) = \beta_0 + \delta\tau_{jmT} + \rho\bar{Y}_{jm2008} + FE_{Municipality} + FE_{Polygonsize} + \epsilon_{ijmT}$
'All firms' panel includes all firms existing at endline. 'Survivors' panel only includes firms that existed in both baseline and endline. 'Entrants' panels groups firms that did not exist at baseline.Every coefficient denotes a different regression. Standard errors clustered by Hábitat polygon shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table 8: Impacts of Habitat on Residential Outcomes

	Structure of population				Characteristics of population		
	Total population	Total population: female	Total population: male	Population 18+ yrs old (% of tot. pop.)	Avg. years of schooling	Population 12+ yrs old employed (% of workforce)	Population 12+ yrs married (% of pop. 12+)
Panel A: Levels							
Habitat	-196.396	-102.319	-104.807	-0.004	-0.064	0.000	-0.002
	(175.372)	(91.515)	(87.927)	(0.003)	(0.039)	(0.001)	(0.003)
	18,268	*9,354*	*8,904*	*0.695*	*9.245*	*0.980*	*0.530*
Panel B: Logs							
Habitat	-0.029	-0.027	-0.031	-0.005	0.001	0.001	-0.005
	(0.030)	(0.029)	(0.030)	(0.005)	(0.007)	(0.001)	(0.005)
	8.034	*7.357*	*7.321*	*-0.369*	*2.207*	*-0.0197*	*-0.630*
Observations	[367]	[367]	[367]	[367]	[367]	[367]	[367]

The table presents the impact of Hábitat on population census variables. Population censuses were held in 2010 and 2020, the former is used as baseline year. The table shows coefficient δ in the following specification:

$Y_{jm2020} = \beta_0 + \delta\tau_{jmT} + \rho Y_{jm2010} + FE_{Municipality} + FE_{Polygonsize} + \epsilon_{jm2020}$

Panel A presents variables either in the units they were originally in the censuses (Columns 1, 2, 3 and 5) or were transformed relative to total population or a subgroup of that variable where relevant. Panel B shows those same variables transformed to logarithms. Throughout, all regressions are weighted by the polygon total population in 2010 and conditioned on the value of the variable in 2010. Standard errors clustered by Hábitat polygon shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table 9: Spillover effects

	Period: 2008-2018															
	Revenue				Capital stock				Wage bill				Paid workers			
	0m-100m	100m-250m	250m-500m	500m-1km	0m-100m	100m-250m	250m-500m	500m-1km	0m-100m	100m-250m	250m-500m	500m-1km	0m-100m	100m-250m	250m-500m	500m-1km
All firms																
Treatment buffer	3.711	-2.387	-4.282	2.047	-0.058	-0.380	-0.706	1.569	0.226	0.083	0.100	-0.062	0.043	0.022	0.001	0.004
	(3.225)	(4.753)	(5.538)	(4.196)	(1.248)	(1.197)	(1.835)	(1.762)	(0.371)	(0.346)	(0.328)	(0.212)	(0.072)	(0.071)	(0.068)	(0.040)
	64	*74.30*	*92.85*	*89.72*	*18.19*	*19.88*	*23.63*	*22.53*	*5.333*	*5.948*	*6.630*	*7.241*	*2.335*	*2.477*	*2.644*	*2.766*
Observations	[53,882]	[84,063]	[125,766]	[219,752]	[53,882]	[84,063]	[125,766]	[219,752]	[53,882]	[84,063]	[125,766]	[219,752]	[53,882]	[84,063]	[125,766]	[219,752]
Survivors																
Treatment buffer	1.731	5.954	0.455	2.649	1.428	1.262	0.675	0.952	0.476	0.089	0.010	-0.214	0.099	0.060	0.007	-0.044
	(4.871)	(4.729)	(6.543)	(3.175)	(2.336)	(1.924)	(2.720)	(1.459)	(0.651)	(0.501)	(0.571)	(0.295)	(0.115)	(0.088)	(0.102)	(0.053)
	75.85	*82.92*	*100.3*	*104.8*	*27.02*	*26.97*	*30.96*	*31.18*	*7.016*	*7.538*	*8.270*	*9.143*	*2.652*	*2.768*	*2.942*	*3.123*
Observations	[15,731]	[25,993]	[40,167]	[74,109]	[15,731]	[25,993]	[40,167]	[74,109]	[15,731]	[25,993]	[40,167]	[74,109]	[15,731]	[25,993]	[40,167]	[74,109]
Entrants																
Treatment buffer	16.197*	-0.910	-14.491	-1.254	0.861	-1.662	-4.248	1.025	0.604	0.013	-0.116	-0.307	0.128	-0.007	-0.070	-0.036
	(9.618)	(12.642)	(15.390)	(8.221)	(1.541)	(1.825)	(2.756)	(2.546)	(0.565)	(0.598)	(0.606)	(0.609)	(0.122)	(0.133)	(0.137)	(0.133)
	59.06	*70.47*	*89.49*	*82.11*	*14.51*	*16.74*	*20.32*	*18.16*	*4.631*	*5.244*	*5.889*	*6.279*	*2.203*	*2.347*	*2.509*	*2.586*
Observations	[38,720]	[58,772]	[86,470]	[146,872]	[38,720]	[58,772]	[86,470]	[146,872]	[38,720]	[58,772]	[86,470]	[146,872]	[38,720]	[58,772]	[86,470]	[146,872]

Table estimates the impact of Hábitat only using untreated firms just outside of study polygons. As described in each column, distance from polygon perimeter increases across columns, and the outer bands are non-inclusive of nearer bands. The table shows the esimatation of coefficient δ in the following specification:

$Y_{ikm2018} = \beta_0 + \delta\tau_{mT} + \rho\bar{Y}_{ikm2008} + FE_{Municipality} + \epsilon_{ikm2018}$

where τ_{mT} now denotes exposure to a buffer area belonging to a treatment Hábitat polygon. Note that the ANCOVA control is at AGEB level, denoted by subsscript k.

'All firms' row includes every endline firm in that band, 'Survivors' row includes firms present at both baseline and endline. 'Entrants' row shows firms existing only at the endline. Standard errors are clustered by the closest Hábitat polygon associated to each firm, and are shown in parenthesis. Control group means are in italics, number of observations in hard brackets. Every coefficient denotes a different regression.

Table 10: Saturation effect in areas outside Habitat

	All sectors		Manufacturing		Trade and Services	
	2013	2018	2013	2018	2013	2018
Revenue	-14.947**	-23.750*	-0.439	-4.299	-16.543**	-23.875*
	(7.403)	(13.092)	(4.287)	(8.280)	(7.451)	(12.480)
	79.75	*110.4*	*48.11*	*64.81*	*83.15*	*116*
	[9,653]	[9,650]	[7,988]	[7,963]	[9,603]	[9,600]
Capital stock	-2.227	-5.588*	0.169	1.789	-4.070	-6.660*
	(2.579)	(3.228)	(3.866)	(3.836)	(2.756)	(3.581)
	31.36	*31.90*	*52.37*	*40.96*	*26.90*	*29.17*
	[9,653]	[9,650]	[7,988]	[7,963]	[9,603]	[9,600]
Paid workers	-0.240	-0.247	0.406	0.602	-0.316*	-0.322
	(0.184)	(0.300)	(0.296)	(0.539)	(0.170)	(0.277)
	3.007	*3.381*	*4.619*	*5.116*	*2.710*	*3.076*
	[9,653]	[9,650]	[7,988]	[7,963]	[9,603]	[9,600]
Wage bill	-1.128	-0.939	2.147	3.725	-1.615*	-1.532
	(0.985)	(1.448)	(1.498)	(2.927)	(0.872)	(1.308)
	8.181	*10.30*	*16.13*	*18.74*	*6.731*	*8.761*
	[9,653]	[9,650]	[7,988]	[7,963]	[9,603]	[9,600]
Value added	-5.173**	-9.410*	-0.871	-4.147	-5.366**	-8.491*
	(2.526)	(5.021)	(1.717)	(3.501)	(2.471)	(4.953)
	31.13	*47.20*	*20.25*	*29.48*	*32.25*	*49.35*
	[9,653]	[9,650]	[7,988]	[7,963]	[9,603]	[9,600]
Payments to social security	-0.142	-0.151	0.282	0.455	-0.250	-0.264
	(0.169)	(0.200)	(0.313)	(0.433)	(0.176)	(0.183)
	1.219	*1.176*	*2.325*	*2.182*	*0.998*	*0.975*
	[9,653]	[9,664]	[7,989]	[7,975]	[9,603]	[9,614]

Table estimates the effect of treatment saturation on areas outside Hábitat polygons within cities treated by the program. Namely, it shows the coefficients δ in the following specification: $Y_{kmT} = \beta_0 + \delta\tau_{kmT} + \rho\bar{Y}_{km2008} + \epsilon_{kmT}$ where τ_{kmT} now represents intensity of treatment (i.e. saturation) in the k AGEB. Firm data is aggregated into AGEB level. Standard errors clustered by municipality shown in parenthesis. Control group means are in italics, number of observations in hard brackets. Each coefficient denotes a different regression.

Figures

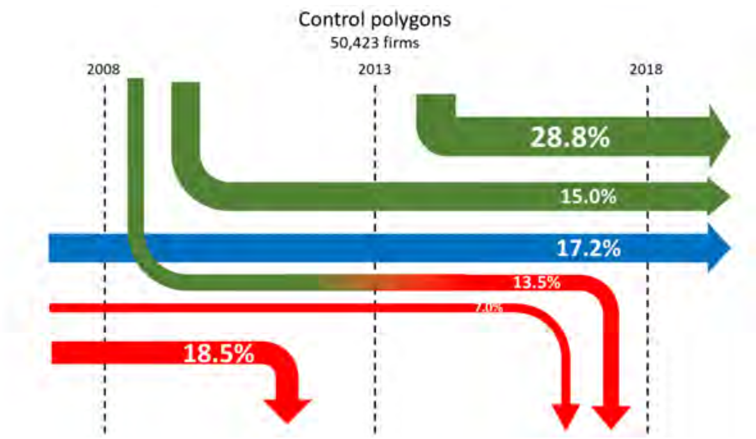

Figure 1: Dynamics of Firm Birth and Death in the Control. Firms across all three census rounds are divided into six strata based on the rounds in which they existed. Green firms are observed to be born, red firms are observed to die, and blue firms persist through all three rounds.

Figure 2: Descriptive densities of outcomes for firms in Hábitat study areas compared to City-wide averages.

For Online Publication: Appendix

Figures

Figure A1: Hábitat cities across the country

Figure A2: Before(above)-and-After(below) Pictures of Hábitat Treatment Neighborhood 1 in Guadalajara

Figure A3: Before(above)-and-After(below) Pictures of Hábitat Treatment Neighborhood 2 in Guadalajara

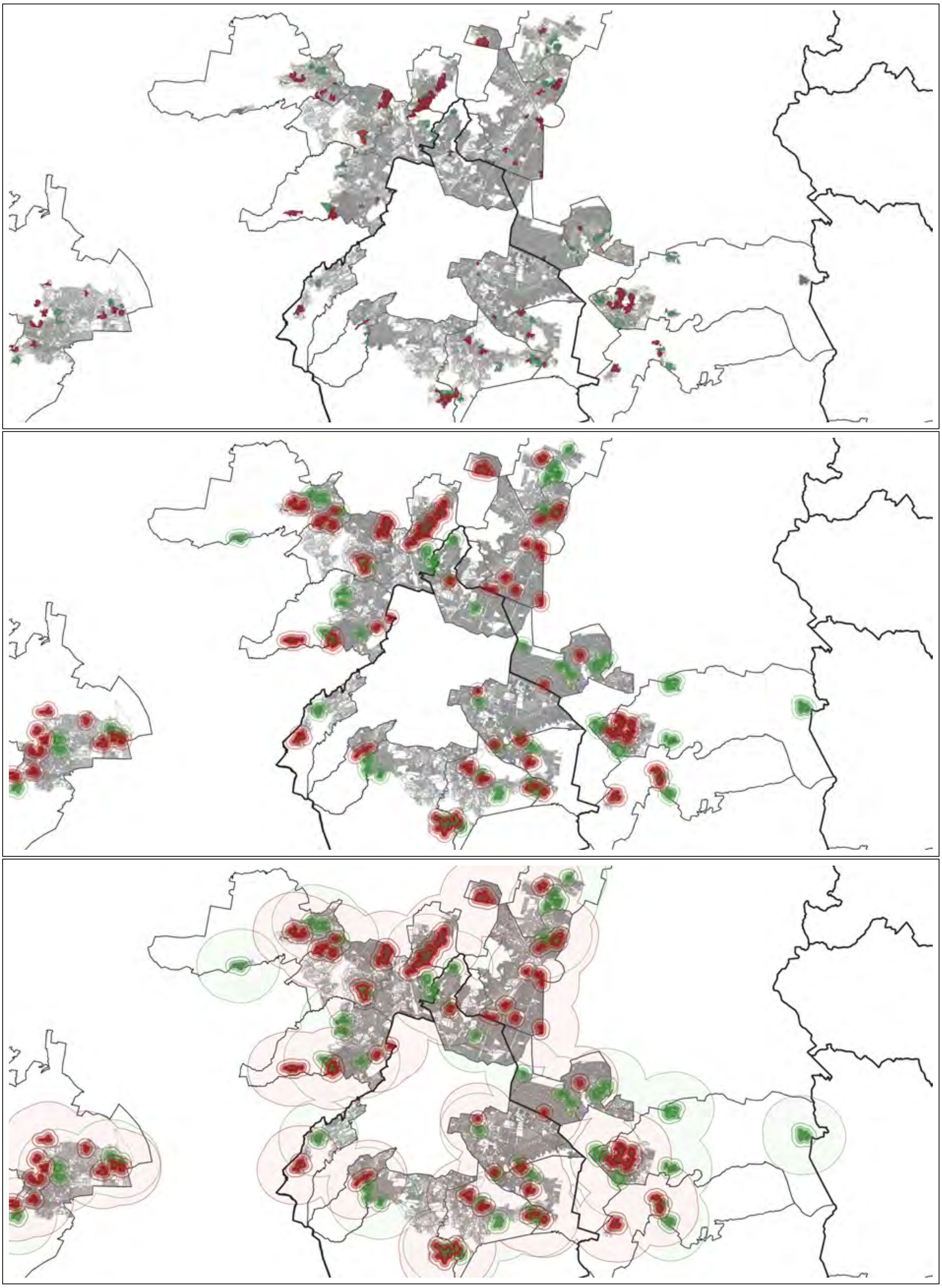

Figure A4: Hábitat program in Mexico City.Top: Hábitat polygons. Center: Polygons with buffer areas up to 1km radius. Bottom: Polygons with buffer areas up to 5km. Green=treatment, Red=control

Figure A5: Hábitat program in the City of Mérida. Top: Hábitat polygons. Center: Polygons with buffer areas up to 1km radius. Bottom: Polygons with buffer areas up to 5km. Green=treatment, Red=control

Figure A6: Hábitat program in the City of Tijuana. Top: Hábitat polygons. Center: Polygons with buffer areas up to 1km radius. Bottom: Polygons with buffer areas up to 5km. Green=treatment, Red=control

Figure A7: Hábitat program in the City of León. Top: Hábitat polygons. Center: Polygons with buffer areas up to 1km radius. Bottom: Polygons with buffer areas up to 5km. Green=treatment, Red=control

Additional Tables

Table A1: Habitat Spending

Total Investments in Treatment Polygons, 2009-2011 (US Dollars)

Name of Program (Subprogram)	2009-2011				
	Total Investment	Federal	State	Municipal	Households Benefitted
Social and Community Development	**13,922,853**	**7,026,328**	**373,053**	**6,523,472**	**256,443**
Improvement of Urban Environment:	**53,729,286**	**26,359,409**	**4,978,304**	**20,857,825**	**169,607**
Paving	**32,850,121**	**15,905,294**	**3,586,772**	**12,246,184**	43,054
Sewers	**4,877,761**	**2,465,322**	**301,398**	**2,047,825**	7,672
Drinking water	**2,644,150**	**1,320,621**	**93,867**	**1,160,632**	5,071
Community Development Centers	**2,842,852**	**1,397,310**	**198,192**	**1,160,519**	17,536
Sidewalks and medians	**2,459,935**	**1,300,516**	**260,256**	**830,340**	4,447
Public lighting	**1,752,960**	**881,141**	**30,248**	**824,506**	5,327
Trash collection	**1,801,580**	**915,702**	**63,868**	**767,860**	72,370
Total spending	**67,743,983**	**33,431,659**	**5,364,410**	**27,414,167**	**428,590**

Source: SEDESOL

Summary of the Hábitat program spending by category and level of government. Source: McIntosh et al. [2018]

Table A2: Balance of Subsequent Habitat Treatment in Original Polygons

Status in subsequent rollout (=0 not treated, =1 treated)	Control			Treatment			t-test (xControl-xTreatment)			
							Unconditional		Conditional	
	Mean	Std. Error	Obs.	Mean	Std. Error	Obs.	Difference	p-value	Difference	p-value
2013	0.985	0.009	194	0.983	0.01	176	0.002	0.905	0.000	0.980
2014	0.418	0.035	194	0.398	0.037	176	0.020	0.700	0.052	0.246
2015	0.381	0.035	194	0.375	0.037	176	0.006	0.899	0.053	0.221
2016	0.33	0.034	194	0.358	0.036	176	-0.028	0.572	0.025	0.560
2017	0.351	0.034	194	0.369	0.036	176	-0.019	0.708	0.027	0.536

Table presents tests of balance between original treatment and control polygons of subsequent Habitat waves from 2013 to 2017. With the exception of 2013, about 35% to 40% of both original control and treatment polygons were treated yearly in subsequent waves of the Hábitat program.

Table A3: Balance of the Experiment at the Firm levels

	Control			Treatment			t-test (xControl-xTreatment)	
	Mean	Std. Error	Obs.	Mean	Std. Error	Obs.	Unconditional	Conditional
All sectors								
Value added	7.122	0.255	21,521	7.376	0.316	14,542	-0.255	-0.146
Revenue	19.983	0.656	21,521	20.724	0.719	14,542	-0.742	-0.061
Capital stock	6.964	0.470	21,521	8.313	0.431	14,542	-1.350**	-0.047
Investment	0.136	0.019	21,521	0.140	0.022	14,542	-0.004	-0.021
Value added per paid worker	4.937	0.136	21,521	4.951	0.188	14,542	-0.014	0.197
Workers: paid	1.447	0.028	21,521	1.483	0.028	14,542	-0.036	-0.054
Wage bill	1.407	0.104	21,521	1.420	0.089	14,542	-0.014	-0.069
Wage	2.950	0.043	3,827	2.840	0.059	2,796	0.110	0.094
Manufacturing								
Value added	11.136	0.71	2,752	9.960	0.84	2,025	1.176	1.324*
Revenue	24.831	1.407	2,752	23.321	1.67	2,025	1.510	2.441
Capital stock	13.484	1.234	2,752	13.222	1.077	2,025	0.262	1.708
Investment	0.231	0.03	2,752	0.203	0.029	2,025	0.028	0.001
Value added per paid worker	4.736	0.224	2,752	4.421	0.3	2,025	0.315	0.422
Workers: paid	2.213	0.093	2,752	2.180	0.112	2,025	0.033	0.064
Wage bill	4.008	0.345	2,752	3.634	0.375	2,025	0.375	0.428
Wage	3.192	0.062	1,058	3.058	0.083	805	0.135	0.105
Trade and Services								
Value added	6.533	0.217	18,769	6.959	0.324	12,517	-0.380	-0.425
Revenue	19.272	0.621	18,769	20.304	0.721	12,517	-0.464	-1.032
Capital stock	6.008	0.395	18,769	7.519	0.426	12,517	-0.309	-1.511***
Investment	0.122	0.019	18,769	0.13	0.023	12,517	-0.023	-0.007
Value added per paid worker	4.967	0.136	18,769	5.037	0.191	12,517	0.147	-0.070
Workers: paid	1.335	0.02	18,769	1.371	0.028	12,517	-0.068**	-0.036
Wage bill	1.025	0.074	18,769	1.062	0.081	12,517	-0.134	-0.370
Wage	2.857	0.051	2,769	2.752	0.07	1,991	0.068	0.105

Table presents summary statistic by treatment arm, and tests of balance between the treatment and control. The first column of balance tests is the simple clustered comparison of means, and the second column uses the municipal FE that are implied by the research design.

Table A4: Entry and exit of firms after 2013

	All sectors		Manufacturing		Trade and Services	
	exit	entry	exit	entry	exit	entry
Habitat	-0.004	-0.005	-0.008	-0.019	-0.003	-0.003
	(0.013)	(0.010)	(0.021)	(0.018)	(0.013)	(0.010)
	0.391	*0.472*	*0.396*	*0.468*	*0.390*	*0.472*
Observations	[44,156]	[50,822]	[5,365]	[6,072]	[38,791]	[44,750]

Table presents LPM estimates on whether firms exit (odd columns) or enter (even columns) between 2013 and 2018. Because the 2013 outcome is endogenous to treatment this analysis is used simply to describe the time path of relative treatment effects. Standard errors clustered by Hábitat polygon shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table A5: Heterogeneity of probability of exit of firms

	2013						2018					
	Initial value added per worker			Initial revenue			Initial value added per worker			Initial revenue		
	P-th>= 25th	P-th>= 50th	P-th>= 75th	P-th>= 25th	P-th>= 50th	P-th>= 75th	P-th>= 25th	P-th>= 50th	P-th>= 75th	P-th>= 25th	P-th>= 50th	P-th>= 75th
All firms												
Habitat*Percentile	-0.025*	-0.010	-0.023*	-0.007	-0.016	-0.012	-0.009	0.005	-0.006	0.003	-0.001	0.008
	(0.014)	(0.011)	(0.014)	(0.015)	(0.012)	(0.014)	(0.012)	(0.011)	(0.012)	(0.012)	(0.011)	(0.013)
Habitat	0.032**	0.017*	0.019**	0.019	0.021**	0.017*	0.018	0.007	0.012	0.009	0.011	0.010
	(0.014)	(0.010)	(0.009)	(0.014)	(0.011)	(0.009)	(0.013)	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)
Percentile	-0.094***	-0.101***	-0.092***	-0.148***	-0.128***	-0.127***	-0.098***	-0.104***	-0.098***	-0.151***	-0.134***	-0.141***
	(0.009)	(0.008)	(0.010)	(0.010)	(0.008)	(0.011)	(0.009)	(0.007)	(0.007)	(0.008)	(0.007)	(0.009)
	0.433	*0.433*	*0.433*	*0.433*	*0.433*	*0.433*	*0.596*	*0.596*	*0.596*	*0.596*	*0.596*	*0.596*
Observations	[36,063]	[36,063]	[36,063]	[36,063]	[36,063]	[36,063]	[36,063]	[36,063]	[36,063]	[36,063]	[36,063]	[36,063]
Manufacturing												
Habitat*Percentile	-0.100***	-0.064**	-0.054	-0.002	-0.017	-0.003	-0.075**	-0.031	-0.039	0.013	-0.014	-0.002
	(0.038)	(0.032)	(0.039)	(0.040)	(0.034)	(0.039)	(0.034)	(0.030)	(0.035)	(0.034)	(0.031)	(0.037)
Habitat	0.078**	0.036	0.017	0.007	0.015	0.003	0.062*	0.022	0.015	-0.002	0.015	0.006
	(0.033)	(0.024)	(0.021)	(0.034)	(0.025)	(0.021)	(0.032)	(0.026)	(0.023)	(0.031)	(0.025)	(0.022)
Percentile	-0.033	-0.049**	-0.040*	-0.130***	-0.113***	-0.138***	-0.040*	-0.051**	-0.056**	-0.129***	-0.102***	-0.120***
	(0.024)	(0.022)	(0.024)	(0.022)	(0.017)	(0.022)	(0.023)	(0.023)	(0.023)	(0.022)	(0.019)	(0.024)
	0.461	*0.461*	*0.461*	*0.461*	*0.461*	*0.461*	*0.624*	*0.624*	*0.624*	*0.624*	*0.624*	*0.624*
Observations	[4,777]	[4,777]	[4,777]	[4,777]	[4,777]	[4,777]	[4,777]	[4,777]	[4,777]	[4,777]	[4,777]	[4,777]
Trade and Services												
Habitat*Percentile	-0.012	-0.001	-0.018	-0.006	-0.016	-0.013	0.002	0.011	-0.001	0.003	0.001	0.010
	(0.015)	(0.012)	(0.015)	(0.015)	(0.013)	(0.015)	(0.013)	(0.012)	(0.013)	(0.012)	(0.011)	(0.014)
Habitat	0.024*	0.014	0.019**	0.020	0.022**	0.019**	0.010	0.005	0.012	0.010	0.010	0.010
	(0.014)	(0.011)	(0.010)	(0.014)	(0.011)	(0.010)	(0.013)	(0.011)	(0.010)	(0.012)	(0.011)	(0.010)
Percentile	-0.102***	-0.108***	-0.101***	-0.151***	-0.131***	-0.126***	-0.107***	-0.112***	-0.105***	-0.155***	-0.140***	-0.144***
	(0.011)	(0.009)	(0.011)	(0.011)	(0.010)	(0.012)	(0.010)	(0.008)	(0.008)	(0.008)	(0.007)	(0.009)
	0.429	*0.429*	*0.429*	*0.429*	*0.429*	*0.429*	*0.592*	*0.592*	*0.592*	*0.592*	*0.592*	*0.592*
Observations	[31,286]	[31,286]	[31,286]	[31,286]	[31,286]	[31,286]	[31,286]	[31,286]	[31,286]	[31,286]	[31,286]	[31,286]

Table presents LPM estimates on whether firms present in the baseline round had exited by each endline round. Heterogeneity is tested through the interaction of treatment and a dummy for being in the Nth percentile of baseline value added and revenue, percentile varied across columns. Standard errors clustered at the polygon level are in parentheses. Mean values of control groups in italics. Number of observations in square brackets.

Table A6: Characteristics of entrants in Habitat polygons

	All sectors		Manufacturing		Trade and Services	
	2013	2018	2013	2018	2013	2018
Revenue	0.975*	1.637	-0.164	-0.135	1.165*	1.834
	(0.580)	(1.013)	(1.138)	(1.238)	(0.593)	(1.139)
	13.92	*21*	*16.78*	*22.65*	*13.54*	*20.79*
Capital stock	0.530	0.785**	0.874	0.312	0.544	0.885**
	(0.417)	(0.372)	(0.917)	(0.547)	(0.407)	(0.393)
	5.392	*5.252*	*8.322*	*6.592*	*5.004*	*5.075*
Paid workers	0.085***	0.048	0.135*	-0.026	0.080***	0.057*
	(0.030)	(0.031)	(0.082)	(0.082)	(0.029)	(0.032)
	1.338	*1.517*	*1.798*	*2.085*	*1.277*	*1.442*
Wage bill	0.217**	0.203*	0.328	-0.131	0.209**	0.250**
	(0.097)	(0.117)	(0.293)	(0.310)	(0.092)	(0.118)
	0.947	*1.625*	*2.525*	*3.617*	*0.738*	*1.363*
Observations	[24,009]	[36,551]	[2,800]	[4,288]	[21,209]	[32,263]

Table is estimated only among firms that newly entered in each of the endline rounds, using a treatment dummy to examine differences between attributes of entrants. Every coefficient is from a different regression. Standard errors clustered by Hábitat polygon shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table A7: Credit Impacts for Businesses that Do and Do Not Own Land

	Did you get credit, loan or financing for the firm's operation?		What is the source of the credit, loan or financing?				Uses of credit, loan or financing received					
			Bank		Savings bank		Equipment or expansion		Acquisition of building or vehicle		Input acquisition (raw materials)	
	Not owner	Owner	Not owner	Owner	Not owner	Owner	Not owner	Owner	Not owner	Owner	Not owner	Owner
All firms in 2013												
All sectors												
Habitat	0.020**	0.017	0.007*	0.008	0.012***	0.003	0.014**	0.017**	0.001	-0.002**	0.008	0.003
	(0.009)	(0.011)	(0.004)	(0.006)	(0.004)	(0.006)	(0.006)	(0.007)	(0.001)	(0.001)	(0.005)	(0.007)
	0.0983	*0.132*	*0.0333*	*0.0464*	*0.0203*	*0.0308*	*0.0319*	*0.0439*	*0.00124*	*0.00282*	*0.0319*	*0.0515*
Observations	[26,037]	[18,254]	[26,037]	[18,254]	[26,037]	[18,254]	[26,037]	[18,254]	[26,037]	[18,254]	[26,037]	[18,254]
Manufacturing												
Habitat	-0.008	-0.026	0.001	-0.031***	0.001	0.002	0.011	-0.017	-0.002	-0.003	-0.007	-0.006
	(0.013)	(0.019)	(0.007)	(0.012)	(0.006)	(0.008)	(0.007)	(0.013)	(0.003)	(0.004)	(0.008)	(0.012)
	0.0915	*0.142*	*0.0342*	*0.0604*	*0.0188*	*0.0236*	*0.0277*	*0.0491*	*0.00328*	*0.00472*	*0.0324*	*0.0500*
Observations	[3,460]	[1,918]	[3,460]	[1,918]	[3,460]	[1,918]	[3,460]	[1,918]	[3,460]	[1,918]	[3,460]	[1,918]
Trade and services												
Habitat	0.025***	0.023**	0.009*	0.013**	0.013***	0.003	0.014**	0.022***	0.002**	-0.001**	0.011**	0.005
	(0.009)	(0.011)	(0.005)	(0.006)	(0.004)	(0.006)	(0.006)	(0.008)	(0.001)	(0.001)	(0.005)	(0.007)
	0.0993	*0.131*	*0.0332*	*0.0449*	*0.0205*	*0.0316*	*0.0326*	*0.0433*	*0.000928*	*0.00261*	*0.0318*	*0.0517*
Observations	[22,577]	[16,336]	[22,577]	[16,336]	[22,577]	[16,336]	[22,577]	[16,336]	[22,577]	[16,336]	[22,577]	[16,336]
All firms in 2018												
All sectors												
Habitat	-0.001	0.003	0.002	0.003	-0.002	-0.005	-0.005	-0.010**	-0.000	-0.000	-0.002	0.002
	(0.007)	(0.009)	(0.004)	(0.005)	(0.002)	(0.005)	(0.003)	(0.004)	(0.001)	(0.001)	(0.004)	(0.007)
	0.0955	*0.119*	*0.0348*	*0.0487*	*0.0199*	*0.0319*	*0.0260*	*0.0351*	*0.00154*	*0.00251*	*0.0500*	*0.0641*
Observations	[32,507]	[18,496]	[32,507]	[18,496]	[32,507]	[18,496]	[32,507]	[18,496]	[32,507]	[18,496]	[32,507]	[18,496]
Manufacturing												
Habitat	0.003	0.005	0.004	0.007	-0.004	-0.010	-0.007	0.003	-0.002	-0.004*	0.004	0.012
	(0.011)	(0.018)	(0.007)	(0.012)	(0.004)	(0.009)	(0.006)	(0.011)	(0.001)	(0.002)	(0.008)	(0.013)
	0.0771	*0.117*	*0.0342*	*0.0589*	*0.0147*	*0.0263*	*0.0274*	*0.0371*	*0.00159*	*0.00634*	*0.0362*	*0.0571*
Observations	[4,043]	[2,057]	[4,043]	[2,057]	[4,043]	[2,057]	[4,043]	[2,057]	[4,043]	[2,057]	[4,043]	[2,057]
Trade and services												
Habitat	-0.002	0.004	0.001	0.002	-0.001	-0.005	-0.004	-0.011**	-0.000	0.000	-0.004	0.001
	(0.007)	(0.009)	(0.004)	(0.005)	(0.003)	(0.005)	(0.003)	(0.005)	(0.001)	(0.001)	(0.004)	(0.008)
	0.0981	*0.119*	*0.0349*	*0.0476*	*0.0206*	*0.0325*	*0.0258*	*0.0348*	*0.00153*	*0.00207*	*0.0520*	*0.0649*
Observations	[28,464]	[16,439]	[28,464]	[16,439]	[28,464]	[16,439]	[28,464]	[16,439]	[28,464]	[16,439]	[28,464]	[16,439]

Table analyzes experimental impacts on financial impacts at the firm level, splitting the sample according to whether firms did or did not own the land on which the business operated at baseline. Every coefficient is from a different regression. Standard errors clustered by Hábitat polygon shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table A8: Spillover effects on probability of entry and exit of firms

	Period: 2008-2018			
	0m-100m	100m-250m	250m-500m	500m-1km
Exit				
Treatment buffer	0.040***	0.020*	-0.008	-0.004
	(0.010)	(0.011)	(0.015)	(0.010)
	0.600	*0.593*	*0.599*	*0.591*
Observations	[40,452]	[64,231]	[98,063]	[177,610]
Entry				
Treatment buffer	0.032***	0.018	0.001	0.019**
	(0.010)	(0.012)	(0.015)	(0.010)
	0.706	*0.693*	*0.689*	*0.664*
Observations	[54,451]	[84,765]	[126,637]	[220,981]

Table presents coefficients from a linear probability model. The probability of exit (top panel) is estimated among all firms present at baseline explaining whether they have exited by the indicated round. The probability of entry (bottom panel) is estimated among all firms present in the post-treatment waves explaining whether the firm is a new entrant in that round. Standard errors are clustered by the closest Hábitat polygon associated to each firm, and are shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table A9: Heterogeneity in Probability of Exit and Entry of Firms by Market Access

	Period: 2008-2018			
	Population density	Wealth index	Polygon size	Driving distance to main road
Exit				
Habitat*Mkt access	0.012	0.014	-0.018**	0.010
	(0.009)	(0.009)	(0.009)	(0.008)
Habitat	0.011	0.012	0.250**	0.014
	(0.010)	(0.010)	(0.120)	(0.010)
Market access	-0.004	-0.004	0.024**	-0.005
	(0.020)	(0.019)	(0.011)	(0.008)
	0.596	*0.596*	*0.596*	*0.596*
Observations	[36,063]	[36,063]	[36,063]	[36,063]
Entry				
Habitat*Mkt access	0.003	0.003	-0.031***	0.009
	(0.008)	(0.009)	(0.010)	(0.010)
Habitat	0.006	0.008	0.399***	0.008
	(0.010)	(0.010)	(0.123)	(0.010)
Market access	-0.048**	-0.035	0.069***	-0.015*
	(0.022)	(0.023)	(0.012)	(0.008)
	0.718	*0.718*	*0.718*	*0.718*
Observations	[50,822]	[50,822]	[50,822]	[50,822]

Table examines heterogeneity in entry and exit of firms by four different measures of market access: 1) Inverse distance-weighted population at the centroid of each neighborhood, 2) Population above the poverty line instead of total population, 3) Size in square meters of polygon, 4) Distance from polygon to a main road. Standard errors clustered by Hábitat polygon shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table A10: Heterogeneity by Market Access

	All firms								Intensive margin							
	Revenue		Capital stock		Wage bill		Workers: paid		Revenue		Capital stock		Wage bill		Workers: paid	
	2013	2018	2013	2018	2013	2018	2013	2018	2013	2018	2013	2018	2013	2018	2013	2018
Market access proxy: population density																
Habitat*PopDens	0.061	0.634	-0.280	0.219	0.005	0.187*	-0.006	0.027	-0.081	1.307	-0.186	0.762	0.040	0.272	-0.006	0.035
	(0.365)	(0.774)	(0.411)	(0.342)	(0.059)	(0.102)	(0.022)	(0.027)	(0.480)	(2.364)	(0.421)	(0.710)	(0.070)	(0.242)	(0.023)	(0.052)
Habitat	0.726	1.954**	0.558	1.055***	0.199***	0.192*	0.053***	0.022	0.610	3.641**	0.577	2.054**	0.229***	0.371*	0.049**	0.054
	(0.478)	(0.892)	(0.351)	(0.398)	(0.061)	(0.099)	(0.020)	(0.027)	(0.598)	(1.774)	(0.403)	(0.801)	(0.084)	(0.222)	(0.023)	(0.052)
PopDens	-0.074	0.422	0.382	-0.452	0.186*	0.261	0.088**	0.102**	-0.677	-4.087	0.761	-1.769	0.198	-0.097	0.087**	0.019
	(0.714)	(1.340)	(0.965)	(0.693)	(0.109)	(0.178)	(0.036)	(0.045)	(1.022)	(4.000)	(1.112)	(1.324)	(0.129)	(0.452)	(0.037)	(0.096)
	16.43	*23.43*	*6.340*	*6.177*	*1.061*	*1.771*	*1.352*	*1.540*	*19.42*	*29.59*	*7.466*	*8.531*	*1.196*	*2.142*	*1.368*	*1.600*
Observations	[44,156]	[50,822]	[44,156]	[50,822]	[44,156]	[50,822]	[44,156]	[50,822]	[20,147]	[14,271]	[20,147]	[14,271]	[20,147]	[14,271]	[20,147]	[14,271]
Market access proxy: wealth index																
Habitat*Wealth	0.046	-0.023	0.213	0.567	0.057	0.244**	0.003	0.027	-0.173	-0.762	0.473	0.982	0.084	0.243	-0.008	0.035
	(0.436)	(0.768)	(0.397)	(0.361)	(0.064)	(0.105)	(0.023)	(0.029)	(0.556)	(1.954)	(0.432)	(0.764)	(0.076)	(0.246)	(0.024)	(0.056)
Habitat	0.726	1.956**	0.528	1.111***	0.195***	0.206**	0.049**	0.022	0.621	3.854*	0.540	2.203***	0.228***	0.403*	0.045*	0.057
	(0.479)	(0.943)	(0.376)	(0.401)	(0.061)	(0.103)	(0.020)	(0.028)	(0.601)	(2.092)	(0.416)	(0.834)	(0.085)	(0.240)	(0.024)	(0.055)
Wealth	0.324	0.999	0.192	-0.403	0.112	0.287	0.079**	0.115**	-0.371	-1.973	0.530	-1.281	0.094	0.069	0.073*	0.066
	(0.663)	(1.389)	(0.857)	(0.778)	(0.103)	(0.178)	(0.033)	(0.046)	(1.025)	(4.193)	(1.010)	(1.397)	(0.139)	(0.472)	(0.039)	(0.096)
	16.43	*23.43*	*6.340*	*6.177*	*1.061*	*1.771*	*1.352*	*1.540*	*19.42*	*29.59*	*7.466*	*8.531*	*1.196*	*2.142*	*1.368*	*1.600*
Observations	[44,156]	[50,822]	[44,156]	[50,822]	[44,156]	[50,822]	[44,156]	[50,822]	[20,147]	[14,271]	[20,147]	[14,271]	[20,147]	[14,271]	[20,147]	[14,271]
Market access proxy: Polygon size (sq. m)																
Habitat*Size	0.475**	0.193	0.006	0.337	0.067**	-0.064	0.017	-0.018	0.466	-0.516	-0.066	0.471	-0.051	-0.234	-0.018	-0.046
	(0.217)	(0.387)	(0.242)	(0.235)	(0.034)	(0.049)	(0.011)	(0.013)	(0.295)	(1.035)	(0.299)	(0.494)	(0.058)	(0.189)	(0.016)	(0.036)
Habitat	0.518	2.021*	0.482	0.752*	0.173**	0.208*	0.045**	0.023	0.430	3.482	0.501	1.582*	0.290***	0.362	0.065**	0.042
	(0.523)	(1.039)	(0.423)	(0.447)	(0.067)	(0.113)	(0.022)	(0.030)	(0.686)	(2.224)	(0.494)	(0.899)	(0.085)	(0.249)	(0.025)	(0.055)
Polygon size	-0.530*	-0.476	0.086	-0.072	-0.090**	0.113*	-0.022	0.035**	-0.532	1.617	0.213	0.221	-0.015	0.499	-0.000	0.112**
	(0.314)	(0.472)	(0.279)	(0.316)	(0.045)	(0.068)	(0.015)	(0.016)	(0.442)	(1.503)	(0.357)	(0.688)	(0.092)	(0.309)	(0.025)	(0.056)
	16.43	*23.43*	*6.340*	*6.177*	*1.061*	*1.771*	*1.352*	*1.540*	*19.42*	*29.59*	*7.466*	*8.531*	*1.196*	*2.142*	*1.368*	*1.600*
Observations	[44,156]	[50,822]	[44,156]	[50,822]	[44,156]	[50,822]	[44,156]	[50,822]	[20,147]	[14,271]	[20,147]	[14,271]	[20,147]	[14,271]	[20,147]	[14,271]
Market access proxy: Driving distance to main roads (meters)																
Habitat*Distance	-1.165***	-0.725	-0.701*	0.542	-0.066	-0.142	-0.028	-0.071***	-0.471	1.250	-0.456	1.753	0.057	0.155	0.013	0.010
	(0.443)	(1.014)	(0.422)	(0.565)	(0.078)	(0.116)	(0.025)	(0.027)	(0.578)	(2.417)	(0.461)	(1.156)	(0.076)	(0.305)	(0.023)	(0.065)
Habitat	0.557	1.790**	0.436	1.129***	0.187***	0.175*	0.046**	0.011	0.525	3.714*	0.497	2.394***	0.241***	0.422*	0.050**	0.060
	(0.466)	(0.889)	(0.379)	(0.423)	(0.063)	(0.099)	(0.020)	(0.026)	(0.593)	(1.937)	(0.418)	(0.870)	(0.083)	(0.235)	(0.024)	(0.053)
Distance to road	0.511	-1.223	0.591	-0.860*	0.042	0.041	0.015	0.039	0.119	-5.569**	0.452	-1.799*	-0.000	-0.115	-0.000	0.011
	(0.386)	(0.939)	(0.410)	(0.475)	(0.072)	(0.108)	(0.023)	(0.025)	(0.543)	(2.659)	(0.457)	(1.056)	(0.066)	(0.279)	(0.019)	(0.057)
	16.43	*23.43*	*6.340*	*6.177*	*1.061*	*1.771*	*1.352*	*1.540*	*19.42*	*29.59*	*7.466*	*8.531*	*1.196*	*2.142*	*1.368*	*1.600*
Observations	[44,156]	[50,822]	[44,156]	[50,822]	[44,156]	[50,822]	[44,156]	[50,822]	[20,147]	[14,271]	[20,147]	[14,271]	[20,147]	[14,271]	[20,147]	[14,271]

Table examines heterogeneity in overall program impacts by four different measures of market access: 1) Inverse distance-weighted population at the centroid of each neighborhood, 2) Population above the poverty line instead of total population, 3) Size in square meters of polygon, 4) Distance from polygon to a main road. Standard errors clustered by Hábitat polygon shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table A11: Heterogeneity in Spillover Probability of Exit and Entry of Firms by Market Access

	Period: 2008-2018							
	Exit				Entry			
	100m	250m	500m	1km	100m	250m	500m	1km
Population density								
Treatment*Pop. Density	0.018***	0.011	-0.026*	-0.010	0.018*	0.016*	-0.017	0.013
	(0.007)	(0.008)	(0.015)	(0.011)	(0.009)	(0.009)	(0.015)	(0.008)
Treatment buffer	0.033***	0.013	0.005	0.006	0.025**	0.009	0.005	0.009
	(0.010)	(0.010)	(0.010)	(0.009)	(0.010)	(0.011)	(0.011)	(0.010)
Population density	-0.033**	-0.045***	-0.018	0.026	-0.060***	-0.066***	-0.040**	-0.030***
	(0.014)	(0.014)	(0.015)	(0.023)	(0.013)	(0.015)	(0.017)	(0.009)
Wealth index								
Treatment*Wealth	0.017*	0.015	-0.029	-0.015	0.012	0.016	-0.018	0.010
	(0.009)	(0.010)	(0.018)	(0.011)	(0.011)	(0.012)	(0.018)	(0.010)
Treatment buffer	0.038***	0.017	-0.001	0.002	0.031***	0.015	0.003	0.016
	(0.010)	(0.010)	(0.012)	(0.009)	(0.010)	(0.011)	(0.013)	(0.010)
Wealth index	-0.026	-0.043***	0.004	0.040	-0.061***	-0.064***	-0.021	-0.026
	(0.019)	(0.016)	(0.022)	(0.026)	(0.016)	(0.019)	(0.029)	(0.016)
Polygon size								
Treatment*Size	0.008	0.009*	0.010	0.007	0.010**	0.007	0.012*	0.003
	(0.005)	(0.005)	(0.008)	(0.005)	(0.004)	(0.005)	(0.007)	(0.006)
Treatment buffer	0.038***	0.019*	-0.009	-0.004	0.031***	0.018	-0.000	0.020**
	(0.011)	(0.011)	(0.015)	(0.010)	(0.010)	(0.012)	(0.015)	(0.010)
Polygon size	-0.003	-0.001	-0.006	0.001	0.002	0.004	-0.001	0.009*
	(0.004)	(0.003)	(0.005)	(0.005)	(0.003)	(0.004)	(0.005)	(0.005)
Driving distance from main roads								
Treatment*Distance	0.007	-0.001	0.019*	0.004	0.019*	0.019**	0.037***	0.028***
	(0.009)	(0.009)	(0.010)	(0.009)	(0.010)	(0.009)	(0.010)	(0.010)
Treatment buffer	0.039***	0.019*	-0.009	-0.005	0.032***	0.017	-0.001	0.018*
	(0.010)	(0.010)	(0.015)	(0.011)	(0.010)	(0.012)	(0.014)	(0.010)
Distance to road	-0.012*	-0.010	-0.017*	-0.009	-0.022**	-0.020**	-0.027***	-0.021**
	(0.007)	(0.007)	(0.009)	(0.009)	(0.010)	(0.009)	(0.010)	(0.008)
Control mean	*0.600*	*0.593*	*0.599*	*0.591*	*0.706*	*0.693*	*0.689*	*0.664*
Observations	[40,452]	[64,231]	[98,063]	[177,610]	[54,451]	[84,765]	[126,637]	[220,981]

Table examines heterogeneity in spillover probability of exit and entry of firms by four different measures of market access: 1) Inverse distance-weighted population at the centroid of each neighborhood, 2) Population above the poverty line instead of total population, 3) Size in square meters of polygon, 4) Distance from polygon to a main road. Standard errors are clustered by the closest Hábitat polygon associated to each firm, and are shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table A12: Heterogeneity in Spillover Effects by Market Access

	Period: 2008-2018															
	Revenue				Capital stock				Wage bill				Paid workers			
	0m-100m	100m-250m	250m-500m	500m-1km	0m-100m	100m-250m	250m-500m	500m-1km	0m-100m	100m-250m	250m-500m	500m-1km	0m-100m	100m-250m	250m-500m	500m-1km
All firms																
Population density																
Treatment*Pop. Density	-1.582	-9.052	-10.882*	2.024	0.394	-0.583	-3.686*	1.417	0.317	-0.244	-0.211	-0.224	0.007	-0.084	-0.065	-0.018
	(3.148)	(6.746)	(5.866)	(4.808)	(1.265)	(1.432)	(2.111)	(2.502)	(0.407)	(0.465)	(0.330)	(0.153)	(0.080)	(0.095)	(0.068)	(0.032)
Treatment buffer	4.201	1.270	1.197	0.987	-0.145	-0.193	1.324	0.879	0.166	0.199	0.244	0.094	0.051	0.061	0.047	0.024
	(3.133)	(3.280)	(3.608)	(3.522)	(1.102)	(1.049)	(1.004)	(0.953)	(0.327)	(0.270)	(0.266)	(0.221)	(0.068)	(0.057)	(0.055)	(0.042)
Population density	1.082	7.053	0.236	-0.337	0.630	-0.624	1.996	0.151	0.723	0.584	0.416	0.317	0.217*	0.180	0.150	0.083
	(4.094)	(6.724)	(6.211)	(4.227)	(1.759)	(1.975)	(1.863)	(1.298)	(0.591)	(0.605)	(0.480)	(0.251)	(0.121)	(0.127)	(0.092)	(0.051)
Wealth index																
Treatment*Wealth	-0.425	-6.270	-8.193	4.223	1.086	0.238	-3.626	1.091	0.515	0.008	-0.052	-0.079	0.037	-0.036	-0.047	0.007
	(3.674)	(7.125)	(7.096)	(4.863)	(1.218)	(1.529)	(2.462)	(2.603)	(0.370)	(0.484)	(0.385)	(0.204)	(0.075)	(0.104)	(0.081)	(0.039)
Treatment buffer	3.767	-1.684	-2.666	1.356	-0.065	-0.424	0.112	1.431	0.232	0.094	0.134	-0.022	0.048	0.030	0.019	0.010
	(3.210)	(4.201)	(4.229)	(3.606)	(1.212)	(1.125)	(1.260)	(1.281)	(0.367)	(0.317)	(0.283)	(0.205)	(0.073)	(0.065)	(0.057)	(0.039)
Wealth index	2.596	5.880	3.213	1.029	0.998	-0.911	3.299	0.858	0.950	0.435	0.477	0.369	0.303**	0.183	0.187*	0.107
	(5.011)	(8.200)	(7.276)	(5.212)	(2.035)	(2.312)	(2.307)	(1.663)	(0.609)	(0.699)	(0.523)	(0.346)	(0.128)	(0.145)	(0.100)	(0.071)
Polygon size																
Treatment*Size	-0.546	0.381	-1.385	-1.873	0.367	-0.242	1.295	1.474	0.387***	0.286*	0.109	0.388**	0.085***	0.069**	0.026	0.078**
	(3.345)	(3.019)	(2.919)	(2.495)	(0.715)	(0.595)	(0.863)	(0.972)	(0.145)	(0.154)	(0.162)	(0.171)	(0.031)	(0.032)	(0.033)	(0.032)
Treatment buffer	4.130	-2.064	-3.996	1.924	-0.192	-0.235	-0.792	1.587	0.129	0.079	0.135	-0.052	0.023	0.019	0.006	0.005
	(3.344)	(4.940)	(5.768)	(4.173)	(1.330)	(1.226)	(1.863)	(1.798)	(0.388)	(0.353)	(0.331)	(0.213)	(0.075)	(0.073)	(0.068)	(0.041)
Polygon size	1.339	1.963	1.484	-0.741	-0.250	0.578	-0.258	-0.493	-0.065	0.182	0.243*	-0.061	-0.007	0.032	0.037	-0.018
	(1.360)	(2.174)	(2.395)	(2.672)	(0.440)	(0.590)	(0.654)	(0.697)	(0.125)	(0.148)	(0.141)	(0.149)	(0.026)	(0.031)	(0.028)	(0.030)
Driving distance from main roads																
Treatment*Distance	-7.761**	-2.267	-1.066	-4.481	-1.298	-1.422	-2.085	-1.990	-0.040	0.107	-0.170	-0.467	-0.005	0.021	-0.022	-0.116*
	(3.399)	(3.749)	(4.624)	(3.892)	(1.313)	(1.268)	(1.513)	(1.568)	(0.408)	(0.355)	(0.395)	(0.306)	(0.089)	(0.079)	(0.085)	(0.065)
Treatment buffer	3.593	-2.465	-4.285	2.367	-0.093	-0.402	-0.668	1.622	0.235	0.070	0.097	-0.054	0.045	0.020	-0.000	0.005
	(3.154)	(4.761)	(5.644)	(4.071)	(1.217)	(1.162)	(1.860)	(1.746)	(0.371)	(0.339)	(0.324)	(0.209)	(0.072)	(0.070)	(0.067)	(0.039)
Distance to road	4.304	-1.894	-0.555	5.444*	0.460	-0.436	0.382	1.048	0.200	-0.383	-0.241	0.177	0.031	-0.073	-0.064	0.040
	(2.924)	(4.171)	(5.083)	(3.170)	(1.226)	(1.150)	(1.250)	(0.946)	(0.376)	(0.363)	(0.356)	(0.246)	(0.080)	(0.082)	(0.080)	(0.051)
Control mean	*64*	*74.30*	*92.85*	*89.72*	*18.19*	*19.88*	*23.63*	*22.53*	*5.333*	*5.948*	*6.630*	*7.241*	*2.335*	*2.477*	*2.644*	*2.766*
Observations	[53,882]	[84,063]	[125,766]	[219,752]	[53,882]	[84,063]	[125,766]	[219,752]	[53,882]	[84,063]	[125,766]	[219,752]	[53,882]	[84,063]	[125,766]	[219,752]

Table examines heterogeneity in spillover effects by four different measures of market access: 1) Inverse distance-weighted population at the centroid of each neighborhood, 2) Population above the poverty line instead of total population, 3) Size in square meters of polygon, 4) Distance from polygon to a main road. Standard errors are clustered by the closest Hábitat polygon associated to each firm, and are shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table A13: Saturation and variables at baseline

	All sectors	Manufacturing	Trade and Services
Revenue	-24.265	-2.011	-22.994
	(24.012)	(12.503)	(26.098)
	89.20	*64.42*	*91.03*
Capital stock	-5.629	11.289	-10.542
	(9.497)	(14.042)	(8.744)
	35.84	*53.77*	*30.78*
Paid workers	-0.077	1.510	-0.513
	(0.592)	(1.248)	(0.517)
	3.258	*5.103*	*2.860*
Wage bill	-0.435	6.913	-2.549
	(3.123)	(6.802)	(2.686)
	10.27	*18.84*	*8.306*
Value added	-10.874	-0.497	-10.856
	(9.741)	(6.098)	(10.241)
	33.23	*27.95*	*33.31*
Payments to social security	-0.115	0.916	-0.429
	(0.455)	(1.027)	(0.384)
	1.319	*2.498*	*1.026*
Observations	[9,758]	[8,278]	[9,707]

Table examines the relation between the main variables at baseline and the saturation of the program at municipality level. Standard errors clustered by municipality shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table A14: Aggregate effect of Habitat in Polygons

Dependent variable	All sectors				Manufacturing				Trade and Services			
	2013		2018		2013		2018		2013		2018	
	unweighted	weighted	unweighted	weighted	unweighted	weighted	unweighted	weighted	unweighted	weighted	unweighted	weighted
Revenue	19.596	-58.829	175.739	254.076	25.490	18.630	28.572	78.715	-1.426	-69.785	158.056	223.040
	(86.901)	(227.991)	(145.224)	(362.478)	(18.416)	(40.758)	(29.645)	(80.825)	(76.825)	(212.425)	(137.994)	(352.752)
	2,266.0	*8,648.0*	*3,728.0*	*13,858.0*	*357.1*	*1,121.0*	*514.0*	*1,657.0*	*1,942.0*	*7,555.0*	*3,260.0*	*12,265.0*
Capital stock	84.571*	113.738	135.457**	384.730***	15.104	-50.751	6.980	-2.279	68.399*	114.975	141.090**	390.414***
	(46.060)	(132.202)	(67.358)	(142.038)	(14.419)	(32.355)	(14.158)	(29.168)	(39.029)	(104.949)	(64.558)	(136.503)
	877.2	*2,963.0*	*989.4*	*3,140.0*	*192.0*	*552.8*	*174.6*	*514.3*	*709.2*	*2,441.0*	*836.0*	*2,644.0*
Paid workers	3.760	-8.769	-3.151	-17.150	3.245**	2.322	0.387	-2.225	0.957	-9.955	-2.957	-13.268
	(5.365)	(15.074)	(8.479)	(24.565)	(1.579)	(3.724)	(2.339)	(6.273)	(4.359)	(12.889)	(6.844)	(20.847)
	186.5	*722.2*	*245.5*	*958.1*	*32.8*	*101.9*	*43.4*	*137.3*	*156.7*	*625.7*	*205.6*	*825.0*
Wage bill	26.857***	26.910	26.827	33.028	10.784**	11.090	3.302	-1.072	15.603**	7.418	21.258	15.151
	(9.430)	(24.693)	(18.530)	(51.278)	(4.397)	(8.670)	(6.765)	(13.946)	(6.980)	(18.599)	(14.898)	(41.019)
	147.3	*487.1*	*283.3*	*996.3*	*52.1*	*156.2*	*81.9*	*260.2*	*100.2*	*331.2*	*208.1*	*733.7*
Value added	-24.583	-104.644	26.425	42.614	7.635	9.074	19.910	57.989	-26.121	-99.048	20.494	22.432
	(42.894)	(99.763)	(68.878)	(153.804)	(8.863)	(20.256)	(15.513)	(40.070)	(38.473)	(92.615)	(65.207)	(145.931)
	931.5	*3,624.0*	*1,541.0*	*5,813.0*	*168.6*	*530.9*	*258.7*	*841.3*	*781.5*	*3,105.0*	*1,307.0*	*4,987.0*
Payments to social security	2.779***	7.984***	6.398***	14.998***	0.388	0.357	1.048**	0.634	2.296***	7.801***	5.090***	14.464***
	(0.936)	(2.812)	(1.681)	(4.673)	(0.235)	(0.362)	(0.459)	(0.720)	(0.880)	(2.706)	(1.567)	(4.947)
	5.1	*14.6*	*8.2*	*21.2*	*1.6*	*4.7*	*2.3*	*5.9*	*3.7*	*10.1*	*6.0*	*15.6*
Observations	[370]	[370]	[370]	[370]	[354]	[354]	[354]	[354]	[370]	[370]	[370]	[370]

Outcomes in this table are the polygon-level sum of firm-level variables, and so provide impacts on total value of outcomes at the polygon level. Every coefficient is from a different regression. Robust standard errors shown in parenthesis. Mean values of control groups in italics. Number of observations in square brackets.

Table A15: Fiscal Impacts of Program

	Impact per polygon	Total impact across all treatment polygons	Nominal annual taxation rate	Total Annual Tax Treatment Effect
Habitat Treatment Effects:				
Value Added	$42,614	$7,500,064	0.16	$1,200,010
Payments to IMSS	$14,998	$2,639,648	1	$2,639,648
Total Revenue	$254,067	$44,715,792	0.02	$894,316
			Total Annual Tax Recovery:	**$4,733,974**
			Cost of program:	$67,000,000
			Number of years required to recoup cost from business taxes alone:	**14.15**

Table presents the results of a simple accounting exercise that takes the polygon-level impacts from the prior table, scales them to represent the total impact in all treatment polygons, and then uses marginal rates on the three core taxable business outcomes to calculate fiscal recovery by the government. This number is then compared against the total cost of all treatment to provide a number of years required to pay off those costs.

Technical Appendix

This section describes the steps taken in the gathering, cleaning and matching of the Hábitat and Economic Census data sets. The process of matching required several adjustments to make data compatible across the three censuses as well as with the geospatial data used in both sources.

Hábitat database

As broadly described in the data section, the Hábitat database contains detailed geospatial information of the blocks, called *manzanas*, included in the study. The Hábitat study relies on INEGI's identification system of blocks, which in most part is standardized across the Agency's different projects. This makes fairly simple to cross data of different projects. At the same time, each block is identified to the polygon it belongs within the project and its corresponding treatment/control status.

The Habitat data contains substantial richness; it is possible to observe the exact type, amount, and location of each infrastructure upgrade a polygon received and on which year it occurred (2009, 2010 or 2011). Because the actual investments made in a given location were endogenous (both to the decisions of the Hábitat engineering team and to the community-driven selection process) we largely abstract away from this and analyze the treatment with a simple binary indicator.

Economic census database

The economic census microdata provided by INEGI comprises the events held in 2008, 2013 and 2018. The census has a very high response rate, above 98% of all firms surveyed. The timing of these censuses is remarkably fortuitous for a study of Hábitat, given that the first interval allows us to conduct a before-after analysis of the short-term impacts of the program on the private sector, and the 2018 wave allows us to examine impacts 7 years after the cessation of investment.

The census covers all businesses in Mexico that have a fixed location (including informal businesses) and belong to the manufacturing, services or construction sectors. The censuses have a very high response rate (more than 98%), given that firms are required by law to respond, and INEGI has the mandate to make individual firm data confidential.

INEGI uses unique identifiers for each business surveyed. Thus, if a firm appears in two or more censuses, it is possible to link the data collected and create a panel. That is, it is possible to follow firms through the censuses and also identify firm created and destruction.

Another characteristic of the census database is that firms also contain detailed information regarding their geographical location. Thus, firms can be tracked to the block they are located within a city. This is crucial, as this geospatial information makes possible to cross this database with the Hábitat database and identify those firms contained within Hábitat polygons. The variables used for analysis are: firm revenue, capital stock, paid workers and wage bill. Additional variables are used to try to understand mechanisms under the Hábitat program relates to firms' performance.

We are able to locate 84,119 firms within Hábitat polygons. Given that there are slightly more control polygons, the majority of businesses are located in such polygons (roughly 60% of firms). In terms of sectors, the vast majority of businesses belong to trade and services (over 90% of total). This is consistent with the sectoral composition of firms across the country. Within this group, most are grocery stores (around 25% of total), and stationer's shops and beauty salons (approx. 4% each). Manufacturing firms tend to be concentrated in activities related to the production of food and beverages and varied activities related to construction and housing. Around 3% of firms produce corn tortillas and 1% are bakeries. Ironworks, furnishing and milling activities businesses comprise close to 1% of the total each.

In terms of size, most of firms located in Hábitat polygons are microbusinesses. The median of paid workers is 1, which means that the typical firm only "employs" the owner of the firm. However, there are some firms that employ up to 50 workers. In line with the nature of microbusinesses, most firms have rather small yearly revenues (a typical firm makes US$ 15,600) and limited assets (less than US$ 2,500 for the typical firm).