American Economic Journal: Applied Economics 2018, 10(3): 263–286
https://doi.org/10.1257/app.20160429

The Neighborhood Impacts of Local Infrastructure Investment: Evidence from Urban Mexico†

By Craig McIntosh, Tito Alegría, Gerardo Ordóñez, and René Zenteno*

This paper reports on the results of a large infrastructure investment experiment in which $68 million in spending was randomly allocated across a set of low-income urban neighborhoods in Mexico. We show that the program resulted in substantial improvements in access to infrastructure and increases in private investment in housing. While a pre-committed index of social capital did not improve, we find an apparent decrease in the incidence of personal assault and teen misbehavior in neighborhoods where investments were made. The program increased the aggregate real estate value in program neighborhoods by two dollars for every dollar invested. (*JEL* H76, O18, R23, R31, R53, Z13)

When governments invest in infrastructure, what is the impact, and who ultimately realizes the benefits from these investments? Despite global infrastructure spending of $4.2 trillion in 2013 (PwC report[1]), these questions are difficult to answer in any straightforward way. Infrastructure investment takes place at a highly aggregated spatial level and therefore is not generally amenable to randomized experimentation in the same way as individually targeted programs (Newman, Rawlings, and Gertler 1994; Field and Kremer 2006; Hansen, Andersen, and White 2012). Even where impacts can be measured, the incidence of these benefits is complex. A long literature stretching back to the Rosen-Roback models of compensating differentials suggests that improvements in amenities will be priced into property and real wages. This implies that renters will not realize welfare

* McIntosh: School of Global Policy and Strategy, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093 (email: ctmcintosh@ucsd.edu); Alegría: El Colegio de la Frontera Norte, A.C. Km 18.5 Carretera Escénica Tijuana–Ensenada San Antonio del Mar Tijuana, Baja California, México C.P. 22560 (email: talegria@colef.mx); Ordóñez: El Colegio de la Frontera Norte, A.C. Km 18.5 Carretera Escénica Tijuana–Ensenada San Antonio del Mar Tijuana, Baja California, México C.P. 22560 (email: ordonez@colef.mx); Zenteno: University of Texas at San Antonio, One UTSA Circle, MB 1.410, San Antonio, TX 78249 (email: rzenteno2012@gmail.com).
Thanks to Beatriz Alfaro, Miguel Ángel Ramírez, Camilo Contreras, Mario Jurado, Silvia López, Gabriela Pinillos, Ruth Rodríguez, and Wilfrido Ruiz for their invaluable work on the survey and index construction, to Thomas Dickinson and Ana Quiroz for GIS work, and to participants in seminars at CIDE, Claremont-McKenna, CU Boulder, El Colegio de la Frontera Norte, Georgetown, IFPRI, Stanford, UCSD, USC, UC Berkeley, and the World Bank for helpful comments. The project was funded by the Mexican Social Development Secretariat (SEDESOL) under Loan 1928/OC-ME for the Inter-American Development Bank. The authors were paid as consultants to conduct the research.
† Go to https://doi.org/10.1257/app.20160429 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.
[1] PwC, "Capital Project and Infrastructure Spending," https://press.pwc.com/News-releases/infrastructure-spending-to-more-than-double-to-9-trillion-annually-by-2025/s/e4ea4334-fdfc-4504-9273-c2e545faeb8e.

gains from improved amenities on average because they will be forced to pay the full compensating differential to enjoy these amenities. In this paper, we present the results of a major federal infrastructural spending experiment implemented in low-income neighborhoods across Mexico during the years 2009–2011 by the Hábitat program of the Social Development Secretariat (SEDESOL). We exploit detailed household and block-level data to measure impacts on specific forms of infrastructure and then use professional property price valuations to compare the overall amenity value of the improvements to the cost of the program (Greenstone and Gallagher 2008). Our work joins a rapidly growing literature using experiments to examine the impact of improvements in infrastructure, such as road paving (Gonzalez-Navarro and Quintana-Domeque 2016), new home construction (Galiani et al. 2013), and improvement of water resources (Kremer et al. 2011).[2]

The analysis of urban housing prices has long been considered in a general equilibrium setting where locations are bundles that comprise rents, wages, and amenities (Rosen 1979). In the presence of costless relocation by workers and firms, differences in cost-adjusted wages across locations provide a measure of the amenities present in each location (Roback 1982, Blomquist et al. 1988). When neighborhood amenities are improved via public investment, property prices and rents will both shift upward, passing the incidence of the benefits to preexisting owners of land. We study a program that targeted low-income neighborhoods with no ongoing conflicts over land tenure. This, combined with the very low mortgage lending rates in Mexico, presents an environment in which 80 percent of study households own their homes outright. There is therefore scope for a successful program to generate a substantial real wealth transfer to low-income households through this type of neighborhood infrastructural investment. From a welfare perspective, urban slums may prove a particularly attractive place to make new infrastructure investments precisely because they have been historically underserved (see Turley et al. 2013 for a systematic review of the evidence on slum infrastructure upgrading).

This study tracks the impact of \$68 million in infrastructure investment, which was randomly allocated to poor urban neighborhoods in 60 municipalities across 20 different Mexican states. This represents the largest-scale experimental evidence to date on the impact of infrastructure investment on critical outcomes, such as private housing investment, social capital, and aggregate property prices. Hábitat investment is primarily used to build heavy infrastructure, such as roads, water, sewerage, lighting, and sidewalks, but is also put towards community centers, parks, and sports facilities. The study is accompanied by a detailed panel household survey tracking outcomes on almost 10,000 different city blocks. The results demonstrate the social and economic benefits of investment in underserved neighborhoods and help to address the paucity of infrastructure evaluation studies that are performed "at scale."

The results of the study show that Hábitat investments resulted in very large improvements in the quality of road paving, sidewalks, medians, and public lighting. An index of infrastructure quality, to which we committed as the core outcome

[2] For recent non-experimental evaluations of infrastructure spending, see Paxson and Schady (2002); Newman et al. (2002); Duflo and Pande (2007); Cattaneo et al. (2009); Dercon et al. (2009); Khandker, Bakht, and Koolwal (2009); and Casaburi, Glennerster, and Miguel (2013).

indicator in a pre-analysis plan, is significantly improved by the intervention. Private investment in the housing stock is crowded in by the program; homeowners in treated neighborhoods are more likely to install cement floors and flush toilets. Using survey data, we show that monthly rents increase by \$18 off a base of \$88, but this estimate likely overestimates the causal effect of public spending because it includes the effects of private investment in the housing stock. We attempt to isolate the pure effect of public investment by using professional assessors to provide estimates of the prices of unbuilt lots at baseline, and then for the same lots again at endline. The value of a square meter of land in treatment neighborhoods increases by more than \$2 for every \$1 invested by the program. This estimate is substantially larger than corresponding numbers from infrastructure investment in the United States (Pereira and Flores de Frutos 1999; Cellini, Ferreira, and Rothstein 2010), suggesting underinvestment in infrastructure in these low-income neighborhoods.

A primary purpose of the designers of Hábitat has been that, through a multidimensional set of investments such as walkability or sports and cultural opportunities for youth, the program would build a sense of community. Despite this emphasis, we find an insignificant positive impact of the program on our pre-committed index of social capital. We exploit the rich household survey data to examine social capital more closely in an exploratory analysis not disciplined by a pre-analysis plan. Domains, such as participation and trust, are not improved, but the treatment group entirely avoids a strong decay in security that is experienced by the control. Consistent with investment in community sports, arts, and music, the program generated an absolute improvement in an index of youth behavior. So, while the program does not improve most of the measures used in our social capital index (as shown in Ordóñez Barba and Ruiz Ochoa 2015), these more tenuous results suggest that it may have led to meaningful improvements in public security.

The paper is organized as follows. Section I introduces the program and the data collection strategy, Section II presents the estimates of the impact of the program on primary outcomes, Section III presents secondary analysis, and Section IV concludes.

I. The Hábitat Program

A. *Program Description*

Urban Mexico has been a tumultuous place to live over the past decades. Rapid urbanization, combined with the dramatic increase in violent crime, has put pressure on the social fabric of Mexican cities (Vilalta 2014, BenYishay and Pearlman 2014). The Mexican federal government created the Hábitat program in 2003 in order to provide infrastructure investments to marginalized urban parts of the country and to provide public resources to improve the quality of life in these communities. It combines centralized targeting and spending rules with decentralized implementation in a way that makes the program comparable to the larger set of community-driven development (CDD) interventions (see Mansuri and Rao 2004 for an overview and Chen, Mu, and Ravallion 2009 or Casey, Glennerster, and Miguel 2012 for more recent examples). The premise of the program is that simultaneous investment across multiple dimensions of neighborhood amenities (public infrastructure, walkability, community centers, sports fields, etc.) can substantively improve social capital and strengthen the social fabric binding neighborhoods together.

Hábitat represents an injection of federal spending into the types of local infrastructure more typically provided by municipal governments. To be able to receive investment from Hábitat, a polygon must be located within a state and municipality whose respective governments are willing to cooperate with Hábitat's cost-sharing rules (which involve local governments providing 50 percent of project costs; in our projects, the municipalities provided 40 percent, the states 8 percent, and the beneficiaries 2 percent).[3] Hábitat defines eligibility in a spatial manner, using a geographic information system (GIS) to define a "polygon," which is an explicit geographic perimeter outlining a neighborhood that meets the criteria for inclusion. A Hábitat polygon is smaller than a locality and is a designation not used by other layers of government. In order to be eligible to benefit from Hábitat, a polygon must consist of settled households in marginalized urban areas with concentrations of asset poverty greater than 50 percent, located in cities of 15,000 inhabitants or more, with a deficit of infrastructure and urban services, and with at least 80 percent of the lots having no active conflict over property rights (this will prove important in interpreting the welfare implications of our results and presents a contrasting ownership environment to the property titling programs studied by Field 2005 and Galiani and Schargrodsky 2010).

Once Hábitat begins to operate in a polygon, two sets of actors worked together to determine the specific investments to be made. Program officials would make recommendations to the community about the observed infrastructure deficits, and a community-driven process (including local government officials and neighborhood leaders) would identify which were most pressing. A "project executant" then submits a proposal, subject to a set of program rules of operation (such as the fact that drainage should be installed or repaired prior to street paving). Spending caps imposed by the federal program governed the amount that could be spent on any specific item (although, there is some variation in the magnitude of spending per capita driven by the scale of the project undertaken in a particular location). It is made explicit at the time of the decision that all maintenance costs would be borne by the municipal governments moving forward.

While the program invested in a diverse set of activities, including community development centers, job training, and health and nutrition training for young mothers, more than two-thirds of Hábitat investment went into improving localized infrastructure. This includes investment in street paving within the polygon (48 percent of total spending), piped water and sewerage (11 percent), and the construction of medians and sidewalks to make the neighborhood more pedestrian-friendly (4 percent). Online Appendix Table A1 provides a detailed breakdown of how the money was spent in the 155 treatment polygons studied in this paper, and Ordóñez et al. 2013 provides a more in-depth description of the program as well as some preliminary infrastructure impact measures.

[3] This is the reason that our study does not include a pure control consisting of municipalities with no treatment. SEDESOL felt that in return for going through the cost-sharing negotiations municipal governments should be guaranteed that they would receive at least one neighborhood assigned to treatment.

The randomized evaluation of the Hábitat program was designed in 2008 following on the heels of a propensity-score matching analysis that was conducted by Mathematica on the prior wave of Hábitat investment from 2003–2004 (Campuzano, Levy, and Zamudio 2007). Despite the fact that 300 million dollars had been spent by Hábitat in those two years, the prior study returned quite disappointing results, finding a 3 percentage point increase in access to sewerage and no improvement in access to drinking water or electricity. Because that study used outcomes available in the Mexican census and *Conteo*, however, it was only able to look for impacts on infrastructure, such as electricity and piped water, that is already nearly universal in the country. In this study, we conduct detailed household surveys and so are able to paint a richer picture of the impacts of local infrastructure. While we also fail to find strong impacts on the headline infrastructure outcomes studied in the Campuzano, Levy, and Zamudio paper, we uncover a wide range of impacts on indicators to which access at baseline was less universal.

B. *Study and Survey Design*

The original study universe consisted of all eligible polygons within the set of 65 municipalities that were recruited into the experimental wave of the study by Hábitat in 2008.[4] The polygons included in the study were required to satisfy two additional eligibility restrictions beyond the standard conditions for the program: we excluded municipalities that had only one polygon and cities that had fewer than four polygons (large Mexican cities may be divided into multiple municipalities). The randomization was conducted at the level of the 370 polygons that satisfied all eligibility criteria for inclusion, with 176 treatment and 194 control polygons. This covers 65 municipalities, with an average of 5.7 polygons per municipality. The experiment featured a two-level randomization (first the saturation of treatment was randomly assigned at the municipality level between 0.1 and 0.9, and then treatment was randomly assigned at the polygon level to match the municipality saturation as closely as possible). In the end, the small eligibility coverage of the program meant that only 1 percent of the surface area and 3 percent of the population of study municipalities are included in study polygons. (Figure 1 shows the map of the municipality of Tijuana, illustrating the small size of the Hábitat polygons relative to the overall surface area of the municipality). This means that the cross-municipality saturation results lack statistical power, and so we do not focus on them here, other than to note that no strong spillover effects were detected. Online Appendix Table A2 presents the distribution of polygons by state and treatment group.

The data collection instruments for the study consisted of panel block- and household-level surveys, as well as an assessment of property values by real estate professionals. To obtain a random sample of blocks and houses for the study, we began by listing all blocks from the 2005 *Conteo* contained within study polygons. We drew into the study all blocks in polygons with 100 or fewer blocks and randomly sampled 100 blocks in larger polygons (only 4.3 percent of the total). Study blocks

[4] Mexico contains 2,438 municipalities in all, but many of the most urban municipalities in the country are contained within this study.

Figure 1. Map of Hábitat Polygons in the Municipality of Tijuana (*polygons are dark-shaded*)

Source: Author's mapping based on shapefiles of Hábitat polygons from SEDESOL

then had a visual block-level survey filled by a trained enumerator to describe the basic condition of roads, sidewalks, and other public facilities. One household on each block was randomly sampled to receive a household survey, which was itself conducted in a long and short form. The short survey included questions on the construction and amenities of the house, the quality of local infrastructure, access to neighborhood facilities, health, and satisfaction with the urban environment. The long survey included all these questions and in addition asked about transport modalities, social capital, crime victimization, and teen behavior. Households were randomly assigned to answer the short versus the long version of the questionnaire resulting in 6,419 long-form and 5,065 short-form questionnaires. These instruments were conducted between the months of March to July 2009 (baseline observation) and January–March 2012 (follow-up observation) in the 370 polygons. The analysis is weighted by the number of households per block (and by the number of blocks in larger polygons) to make the study representative of the population of study polygons.

Accurate panel measurement of property values in a randomized control trial (RCT) context, and in a country without digital transaction records, is a challenge. First, increases in private investment (such as installation of concrete floors or bathrooms with indoor plumbing) confound the measure of increases in property values because the housing stock itself improves because of private, as well as public, investment. Price increases driven by private expenditures are a valid causal effect of the program, but they complicate an accounting of the per-dollar returns to public investment. Secondly, recent empirical work suggests that urban Mexican households typically provide overestimates of the sale value of their own properties (Gonzalez-Navarro and Quintana-Domeque 2009). To overcome the first of these issues, we use price estimates only on empty lots that have no construction on them as the baseline, so our estimate of price per square meter of raw land is not polluted by changes in the nature of the private housing stock. In order to get a high-quality estimate of sales prices in an environment in which there is no regular recording of sales prices, we used professional property assessors from the Instituto de Administración Avaluos de Bienes Nacionales (INDAABIN), the Mexican government's institute of real estate valuation.

These assessors provided estimates of the value of every one of the 464 unbuilt lots that were for sale in the study polygons at baseline and then returned to the same lots at the time of follow-up and provided new estimates of the raw land value of the lot at that time (whether or not a structure had by then been built). In each round, the assessors assembled information from comparable sales and put together estimates according to established INDAABIN methodology. Assessors were blinded to the treatment design (meaning that they did not know whether they were providing estimates in treatment or control communities). While the total number of empty lots for sale at baseline was small, this analysis provides a precise and readily interpretable impact on land values.

The study had a pre-analysis plan submitted to the Mexican government at the time the analysis of the baseline survey was being completed. All of the outcomes included in that pre-analysis plan are presented in this paper. They include an overall infrastructure index, availability of six-specific infrastructure types (electricity, piped water, sewerage, paved streets, streetlights, and sidewalks/medians), the impact on an index of social capital, and the impact on property prices. The social capital index is constructed at the polygon level using principal-components techniques on the baseline data and comprises a weighted average of 14 survey questions on participation in neighborhood groups; trust in social institutions, neighbors, and household members; levels of conflict between neighbors and community members; and the degree of knowledge about community organizations and social problems. A principal-components method was used to identify factors loadings across these various attributes using baseline data, and the loadings were held constant and used on the endline data to provide a panel measure of changes in social capital for each polygon across the two waves of the survey (a more detailed presentation of the construction of the index can be found in Ordóñez Barba and Ruiz Ochoa 2015). That paper also shows that using a panel difference-in-difference strategy there is a small and insignificant increase in the social capital index. However, the index was heavily weighted towards group participation and social cohesion, outcomes that proved not to be affected by the program. The only significant result reported in that paper is on "confidence between neighbors," a question which included willingness to offer mutual protection. In this section, we show that variables describing stranger-driven crime and teen misbehavior (which were not included in the index at all) did in fact see substantial improvements as a result of the program.

We analyze impacts using population-weighted polygon-level averages for the two study rounds, including fixed effects at the polygon level and clustering standard errors at the municipal level to account both for spatial covariation as well as for the component of the municipal-level design effect arising from the randomized saturation design.

C. *Attrition and Balance*

Panel tracking was done at the *house* level (not the household level) to capture changes in the attributes of structures.[5] This study primarily considers infrastructural outcomes at the level of the residence, not the resident, and so naturally attempted to conduct a survey in the same house as the baseline. If at follow-up the survey teams were unable to find the same household, they were instructed first to survey the new residents of the same house so as to provide a panel on changes to that specific structure, and if they were not able to find residents in the same house, then they were to randomly sample a new house and household from the same block. We always capture whether the household answering the survey is the same at endline as baseline and so can address migration both as an outcome of interest and as a potential confound in measuring household-level outcomes such as social capital.

The analysis sample begins with 10,670 baseline household observations. Online Appendix Table A3 provides a Consort diagram of the attrition that occurred during the course of the study. Two distinct factors cause us to lose observations from the study. First, implementation problems arose with 5 of the original 65 municipalities in the sample (the municipal governments did not meet the matching requirements).[6] Since Hábitat could not treat the polygons in these municipalities, we have removed them from the study altogether, treatment and control alike. This causes us to lose 748 observations, almost 7 percent of the sample. Second, in forming the block-level panel, we lose 220 household observations. The dataset used for analysis thus consists of 19,417 panel household surveys that provide two periods of data at the block level. We also see two additional forms of replacement that do not cause attrition from the study, but alter the interpretation of impacts. First, we have 17.6 pecent of houses whereby the surveyors could not locate anyone to survey in the same house as was surveyed at baseline, and hence, they replaced the baseline house with a different house on the same block. In 25.7 percent of cases where the survey teams were able to locate the same house, the household currently residing in the structure had changed since baseline, and hence, the household (but not the house) is replaced.

Table 1 tests for whether these four distinct types of attrition prove to be correlated with the treatment. In columns 1–2, we begin with the sample of inhabited blocks and examine the attrition caused by the dropping of the five municipalities in which treatment was not possible. While this attrition represents a large part of the original sample, there appears to be no systematic correlation between original treatment status and the polygons in which Hábitat was able to administer treatment. Columns 3–4

[5] The data from the evaluation survey are posted at http://www.normateca.sedesol.gob.mx/es/SEDESOL/ProgramaHabitat.

[6] These five municipalities are Cuajimpala de Morelos, La Magdalena Contreras, Xochimilco, Almoloya de Juarez, and Ecatepec de Morelos.

Table 1—Attrition

	Attrition between rounds 1 and 2							
	Attrition at municipal level (municipality selected to be part of study but removed by Hábitat)		Attrition at block level (block sampled at baseline and in study municipalities, but panel dependent variable not observed)		Attrition at house level (baseline sampled house replaced with alternate at follow-up)		Attrition at household level (baseline sampled household replaced with alternate at follow-up)	
Baseline values of:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	−0.014	−0.00996	−0.00412	0.000965	−0.0363	−0.0297	−0.0702	−0.0612
	(0.048)	(0.049)	(0.005)	(0.001)	(0.030)	(0.028)	(0.030)	(0.026)
Index of basic services		−0.00000749		−0.0000254		0.000614		0.000647
		(0.001)		(0.000)		(0.000)		(0.001)
Satisfaction with social infrastructure		0.00166		−0.000922		−0.0029		0.0123
		(0.008)		(0.000)		(0.005)		(0.005)
Observation weight		0.0000488		−1.22e-05		0.000260		0.000352
		(0.000)		(0.000)		(0.000)		(0.000)
Average fraction attrited in control group	0.095		0.016		0.176		0.257	
Observations	10,670	10,436	9,922	9,745	9,702	9,702	8,302	8,302

Notes: Regressions include fixed effects at the municipality level and are weighted to be representative of all residents in the study neighborhoods. Standard errors in parentheses are clustered at the polygon level to account for the design effect.

then examine the attrition driven by the success of El Colegio de la Frontera Norte (COLEF) field teams in conducting a panel block-level survey. Overall attrition at the block level was low (98.5 percent of the potential panel blocks were successfully tracked) and appears similarly to be balanced by treatment. When we look at the success of field teams at locating first the same house (columns 5–6) and then the same household (columns 7–8), we begin to see evidence of significant differences across the treatment and control. The probability of being unable to find any resident in the baseline-surveyed house is 3.6 percent lower in the treatment than the control, but remains insignificant (columns 5–6). Once we examine the turnover rate of the actual households answering the survey, however, we see that the treatment neighborhoods had dramatically lower rates of residential churn. More than a quarter of the houses that were panel tracked saw the resident household change in the three years of the study, indicating a very high rate of residential churn in the study. More importantly, the treatment had a strong effect on this rate of churn; in treatment neighborhoods, this rate of churn dropped by 7 percentage points, or a quarter of the control-group rate. While we are concerned only with the attrition through column 4 for the basic analysis of infrastructure impacts, we return to a discussion of this differential turnover of households when analyzing changes in private investment and social capital.

Table 2 examines the baseline treatment-control comparison within the final analysis sample, using the short-form survey questions on public infrastructure in panel A and the long-form questions on health, transportation, and social capital in panel B. The experiment was stratified on access to water and electricity, and so the balance here is necessarily excellent. Overall, the infrastructure variables appear very well balanced, and indeed, different impact specifications (posttreatment single-difference, simple difference-in-differences, or polygon fixed

Table 2—Balance in Panel Sample

Panel A. Public infrastructure (short-form survey questions)

	Index of basic infrastructure	Piped water	Sewerage service	Electric lighting	Streetlights	Medians	Sidewalks
Simple treatment-control difference	−0.195 (0.161)	−0.0264 (0.020)	−0.0472 (0.043)	−0.00444 (0.003)	−0.009 (0.024)	−0.0682 (0.052)	−0.0779 (0.062)
Baseline control mean	2.740	0.926	0.829	0.989	0.555	0.588	0.589
Observations	342	342	342	342	342	342	342
R^2	0.011	0.006	0.007	0.004	0	0.012	0.014

Panel B. Household outcomes (long-form survey questions)

	Health index	Local facilities index	Social capital index	Social capital: Participation	Social capital: Trust	Social capital: Security	Social capital: Youth
Simple treatment-control difference	0.178 (0.107)	0.0954 (0.337)	−0.0190 (0.009)	−0.24 (0.162)	−0.294 (0.222)	−0.132 (0.178)	−0.0114 (0.085)
Baseline control mean	−0.138	−0.297	0.407	0.330	0.444	0.076	−0.129
Observations	342	342	342	342	342	342	342
R^2	0.012	0.001	0.028	0.014	0.018	0.005	0

Notes: The polygon-level analysis is weighted by a population weight to be representative of all residents in the study neighborhoods. Standard errors clustered at the municipality level are in parentheses.

effects analysis) all arrive at very similar impacts. The only evidence of imbalance is in the social capital index, which is significantly lower in the control than the treatment. We return in Section III to an analysis of this potential imbalance in the social capital variables and discuss how this affects our interpretation.

II. Primary Analysis

A. *Impact on Public Infrastructure*

Table 3 presents these main infrastructure results. The first column shows the impact on the core infrastructure index (the sum of the five specific infrastructure variables) and finds an impact of 0.135, significant at the 99 percent level. This implies an 8 percent increase off a baseline mean of 2.7. Relative to the increase in the index observed in the control group, which was 0.115, this treatment effect suggests a more than doubling in the rate of improvement in the overall index over the three years of the study. We then examine the variables used for the stratification of the randomization: availability of piped water, sewerage, and electric lighting. These variables all feature high baseline control means (from 82.9 percent for sewerage to 98.9 percent for electricity). The treatment estimates on all these outcomes are small and insignificant: access to sewerage improves by just over 2 percent, and the improvements in access to piped water and electric lighting are roughly two-tenths of a percent.

When we turn to examining forms of infrastructure to which baseline access was less universal, strongly significant positive effects are apparent. Streetlights, sidewalks, medians, and road paving all see dramatic improvements; the fraction of houses with sidewalks in front of them was 59 percent at baseline, rising to 62.5 percent in the control at follow-up, but increased to almost 70 percent in

Table 3—Public Infrastructure

	Index of basic infrastructure (1)	Piped water (2)	Sewerage service (3)	Electric lighting (4)	Streetlights (5)	Medians (6)	Sidewalks (7)	Paved roads (8)
Intention to treat	0.135	0.00115	0.0203	0.00239	0.0607	0.0617	0.0481	0.0314
	(0.048)	(0.016)	(0.017)	(0.005)	(0.036)	(0.021)	(0.022)	(0.014)
Dummy for round 2	0.149	0.0113	0.0236	0.00609	−0.00587	0.028	0.0428	0.0669
	(0.049)	(0.006)	(0.012)	(0.003)	(0.033)	(0.019)	(0.015)	(0.013)
Baseline control mean	2.740	0.926	0.829	0.989	0.555	0.588	0.589	0.664
Observations	684	684	684	684	682	684	684	684
R^2	0.161	0.013	0.053	0.029	0.021	0.092	0.11	0.209
Number of polygons	342	342	342	342	342	342	342	342

Notes: The polygon-level analysis with polygon fixed effects and standard errors is clustered at the municipal level. The analysis is weighted by population weights to be representative of all residents in the study neighborhoods. Standard errors are in parentheses.

the treatment. Relative to the control group, the pace of road paving more than doubled, sidewalk building tripled, and the rate of construction of medians increased by almost a factor of five. The fraction of blocks with functioning streetlights stayed unchanged at 55.5 percent in the control group, but increased by 7 percentage points in the treatment. When we weight the analysis by the number of blocks rather than the number of inhabitants per polygon, results are very similar; completely unweighted analysis sees impacts fall by about half. The implication is that impacts were larger in the largest polygons.

The program therefore had a really substantial effect on neighborhood attributes, such as paving and walkability, that were far from universal at baseline and a more modest effect on the deep utility infrastructure to which access at baseline was much higher. Hábitat spent $68 million building the infrastructure measured in Table 3, suggesting that by making an investment averaging $567 per household, the program was able to generate significant improvements in the quality of basic infrastructure.

B. *Impact on Private Investment*

The surge in public investment induced by the Hábitat experiment provides an interesting environment in which to investigate potential complementarities between public and private investment. The program places public resources in communities under-served by past infrastructural investments, and yet in which property rights are robust (see de Janvry et al. 2015 for a discussion of the improvement of property rights in Mexico). Further, 84.4 percent of households in the baseline reported owning their own homes, and 74 percent own their homes outright (mortgage financing is difficult to obtain in poor Mexican neighborhoods even with clear property title). Thus, there appears to be substantial scope for the amenity value created by Hábitat investments to pass into the hands of the residents of these neighborhoods. We investigate this interplay between private and public investment by examining privately financed investments in the housing stock of Hábitat neighborhoods.

Table 4 provides evidence of complementarities between private and public investment. All effects are positive, and significant upgrades to flooring and plumbing

Table 4—Impacts on Residents

	Private investment in housing							Financial			Social capital
	Brick walls (1)	Concrete floors (2)	Separate kitchen (3)	Separate bathroom (4)	Flush toilet (5)	Septic system (6)	Piped water (7)	Home-owner (8)	Private bank mortgage (9)	Monthly rent, US$ (for renters only) (10)	Index of social capital (11)
Treatment × Round 2	0.00337	0.0229	0.0112	0.00416	0.0707	−0.0273	0.0146	0.0208	0.00962	16.68	0.0104
	(0.008)	(0.009)	(0.016)	(0.014)	(0.031)	(0.013)	(0.023)	(0.022)	(0.006)	(9.697)	(0.496)
Round 2	0.00763	0.00743	0.0307	0.0184	−0.0475	0.00162	0.0628	0.00557	−0.0134	−0.667	−0.0380
	(0.006)	(0.005)	(0.012)	(0.008)	(0.025)	(0.012)	(0.015)	(0.009)	(0.005)	(7.176)	(0.006)
Baseline control mean	0.942	0.965	0.876	0.930	0.608	0.113	0.703	0.844	0.019	88.4	0.407
Observations	684	684	684	684	684	684	684	684	683	530	684
R^2	0.012	0.065	0.089	0.033	0.037	0.014	0.105	0.019	0.034	0.047	0.145
Number of polygons	342	342	342	342	342	342	342	342	342	299	342

Notes: The polygon-level analysis with polygon fixed effects and standard errors is clustered at the municipal level. The analysis is weighted by a population weight to be representative of all residents in the study neighborhoods. Standard errors are in parentheses.

are made to treatment houses, suggesting that public investment is crowding in private investment from households. The only negative coefficient is on the use of a septic system, but because this is an inferior substitute to the connection to a sewer line, this indicates increasing use of centralized infrastructure. Households are significantly more likely to have installed concrete floors and to have working flush toilets. The improvement in indoor plumbing is particularly interesting given that we did not see significant impacts on sewerage in Section IIIA; in this case, private investment appears to have outstripped the measurable improvements in public infrastructure. Home ownership rates in the treatment rise by 2 percent, although this difference is not significant. The coefficient on having obtained a mortgage from a private bank is very small in absolute magnitude but is almost significant. The penultimate column of Table 4 shows the impact on rents for the 16 percent of households that do not own their own homes and indicates a substantial $17 jump in monthly rents, a nearly 20 percent increase over the baseline control-group average rent of $88.

Given that the rate of churn between the two waves of the survey fell from 39 percent in the control to 30 percent in the treatment, could this lower rate of churn itself be an explanation for the greater willingness to invest in houses and the slight uptick in home ownership and mortgages? We can understand the extent to which differential churn may be a mechanism for the treatment effect by running two sets of regressions; first, we can examine the extent to which the houses that saw new tenants in Round 2 differed from the houses of those who stayed in the same residence throughout the study. Second, we can partition the regressions and run treatment effects only within the group that stayed and within the group that moved. These subgroup comparisons are endogenous and contain selection bias (Angrist and Pischke 2009), but if both subgroups look similar to the overall treatment effect, then the program impact cannot be arising as a result of endogenous selection from one group to the other. Online Appendix Table A4 shows that while households that move during the study are likely to live in lower quality housing and less likely to own their homes, the treatment effect of the program within the "stayer" and the "mover" groups is quite similar, indicating that the program has an impact on housing quality independent of the effect through increasing the duration of residence. Overall, the public spending flowing through Hábitat appears to have induced a meaningful increase in private investment in the housing stock.

C. *Impact on Social Capital*

Column 11 of Table 4 provides an estimate of the impact of the program on the pre-committed index of social capital. Recalling that there was some evidence of imbalance on this variable, we should proceed cautiously to interpret the impacts. The polygon fixed effects estimate from the pre-analysis plan is presented here and shows an effect that is insignificant and very small in absolute terms (an improvement of roughly 2.5 percent relative to the baseline control mean). Additional analysis (not shown) conducts single-difference and an analysis of covariance (ANCOVA) of this outcome and similarly finds small and insignificant effects. This means that three qualitatively different ways of handling the baseline imbalance (differencing out, ignoring, or controlling for) all arrive at the conclusion that the pre-committed index of social capital did not respond to the treatment. In Section III, we dig into the social capital effects in more detail.

D. *Impacts on Property Values*

We now proceed to an analysis of the impact of the program on property values, using professional assessment of unbuilt lots to calculate an effect purged of improvements in housing stock driven by private investment. Of the 342 baseline polygons used in this analysis, just over 40 percent had any empty lots for sale at baseline. The average baseline polygon had 1.25 lots for sale, with a maximum of 23 lots per polygon. The intervention sample provides us with 437 lots located in 138 polygons. Column 1 of Table 5 shows that these neighborhoods are not representative of the overall study, being both larger and poorer than the average study neighborhood. However, the sample selection in the real estate analysis arising from polygons in which at least one lot was for sale is balanced across treatment and control, and the baseline means of polygon-level average prices per square meter in this attrited sample are comparable between the treatment and the control. Hence, our analysis of real estate prices takes place in an unrepresentative subgroup (the population of empty lots for sale at baseline), but appears to be well balanced and internally valid.

Sharp improvements in local infrastructure, lower residential churn, and increases in investment in the housing stock all suggest the possibility that treatment neighborhoods may have seen an improvement in residential amenity values. The significant increases in rents in treatment neighborhoods provide initial evidence that these improvements are being capitalized in property prices. Real estate prices should capitalize the net present value of a flow of amenities from improved infrastructure and thus provide a particularly interesting way of comparing the net costs of an intervention to the

Table 5—Property Price Impacts

	Attrition	Baseline balance		Impact						
		Dependent variable: Changes in real estate price per square meter, real 2012 US dollars								
	Polygon has observation on prices	Baseline price		Change in price per square meter, 2009–2012						
Outcome variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	−0.0986	2.409	3.207	6.523	8.196	5.756	6.991	5.796	8.038	6.477
	(0.07)	(14.94)	(15.80)	(2.59)	(4.19)	(3.41)	(2.57)	(2.94)	(3.09)	(2.95)
Baseline index of services	0.0556									
	(0.076)									
Baseline index of infrastructure	−0.0786									
	(0.036)									
Total number of residences	6.07e-05									
	(0.000)									
Constant	0.474	86.11	44.27	3.017	2.108	−30.75	3.396	−29.06	2.907	−30.40
	(0.19)	(8.67)	(5.69)	(1.55)	(2.89)	(1.23)	(1.31)	(5.12)	(1.77)	(5.07)
Observations	342	138	138	138	138	138	138	138	138	138
R^2	0.216	0.001	0.788	0.055	0.080	0.637	0.062	0.452	0.079	0.547
Weighting	Houses	Houses		None	Houses		Attrition		Attrition × Houses	
Municipality FE	No	No	Yes	No	No	Yes	No	Yes	No	Yes

Notes: All prices are in real 2012 US dollars per square meter. The dependent variable is the price of unbuilt lots as assessed by professionals from INDAABIN. The analysis is conducted at the polygon level; standard errors clustered at the municipal level are in parentheses.

net benefits realized by residents. Because of the high rates of home ownership in Hábitat neighborhoods, increases in land values are likely to translate directly into improvements in the household wealth of the residents. To the extent that a public investment yields total property price increases that are greater than the amount of the investment itself, residents would wish to be taxed to make these investments. The presence of net positive returns would suggest "money left on the table" and points to a friction in the political economy of infrastructure delivery.

When we turn to the difference-in-difference impacts in Table 5, we see substantial improvements in prices being induced by the treatment. Column 6 includes municipality fixed effects and weights the analysis by the number of households represented by each lot at the polygon level. Relative to a baseline control value of \$86.11 per square meter and a real control-group appreciation of \$3.02 between 2009 and 2012, the treatment effect of the program was an additional \$5.76 per square meter, meaning that the treatment group had almost triple the real rate of appreciation as the control. This corresponds to a standardized impact size of roughly 0.5 relative to a control-group price change standard deviation of 11. Figure 2 shows the cumulative dsitribution function (CDF) of the changes in real property prices in treatment and control polygons, demonstrating that improvements in the treatment first-order stochastically dominate the control. These impacts are large and suggest an absolute causal price appreciation of 6.5 percent relative to the counterfactual price (control-group endline). By comparison, Gonzalez-Navarro and Quintana-Domeque (2016) find log impacts of street paving ranging between 9 percent and 16 percent on the professionally appraised value of a house. This suggests that our

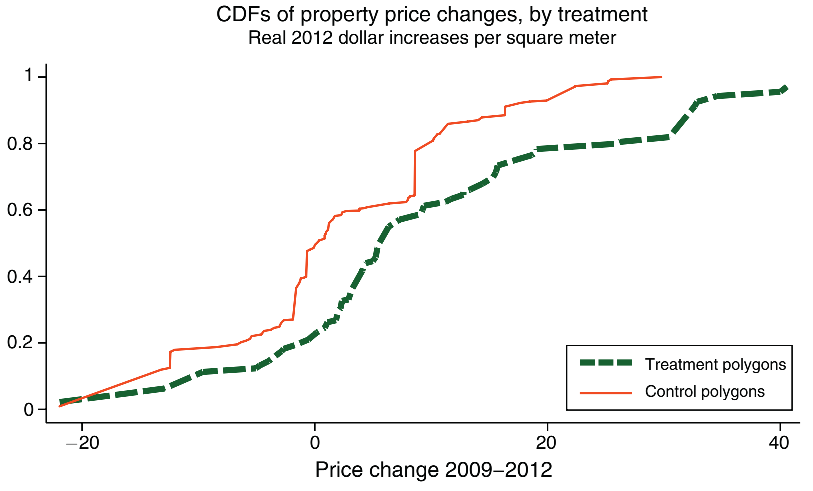

FIGURE 2. CUMULATIVE DISTRIBUTION FUNCTIONS OF PROPERTY PRICE CHANGES

Note: This figure shows polygon-level averages.

estimate may be in a reasonable range, given that their higher estimate provides the Local Average Treatment Effect (LATE) of actually having one's street paved, while ours is an intention-to-treat estimate based on being in a neighborhood where investment was taking place.

Are these estimates robust? Given the explicitly nonrepresentative nature of the valuation sample, one may reasonably ask whether these estimates are externally valid to the entire sample of lots contained in this study. As a way of probing this, we estimate a selection model on the entire sample of polygons, using a battery of baseline covariates to predict the probability that each polygon appears in the real estate sample. With this score, we can then conduct inverse propensity weighting to correct our estimates for observable determinants of selection into the real estate sample (Wooldridge 2002). The results of this exercise are provided in columns 7–10, which show the results first using only the estimated attrition weights and then using the product of the attrition weights and the previous population weights. The results are not sensitive to these corrections, meaning that price change impacts in polygons likely to have valuations are not different from impacts in polygons unlikely to have produced a valuation. It is also worth noting that while the rent variable is selected and suspect in different ways than the real estate analysis, it indicates an appreciation of 19 percent in values. These facts bolster our confidence that there were meaningful price impacts in the overall sample, despite the fact that both of our measures of value come from selected samples.

Perhaps the most meaningful way to put the impact numbers in context is to consider that the treatment polygons contain 118,491 lots with an average of 218 square meters each, for a total of 25.9 million square meters of property. If the marginal effect of $5.76 per square foot estimated above is applied to all inhabited property in the treatment polygons, the resulting increase in total value is $150 million, more than two times the $68 million invested by all three levels of government in the program. We can use the (relatively conservative) impact estimates from column 6 of Table 5 to make a variety of confidence statements: we are 96 percent confident that the benefit is positive, we are 64 percent confident that the benefit/cost ratio is greater than one, and we can be 95 percent confident that the benefit/cost ratio lies between −0.35 and 4.65. The bottom row of Table 5 provides the p-values on the F-test that the benefit-cost ratio is greater than one (meaning that the treatment effect is greater than \$2.62, the amount spent per square meter by the program). These range from 0.36 in the conservative estimate in column 6 to 0.08 (indicating confidence at 90 percent levels) in the specification from column 9 that includes attrition weighting. The average residence would have enjoyed \$1,157 in appreciation from Hábitat investment during 2009 to 2012, while having had \$574 spent on it. Thus, even using our more conservative estimates, every dollar of public money invested in infrastructure improvement in a polygon yielded more than two dollars of improvement in the total privately held value of land there.

III. Secondary Analysis

We now unpack the pre-registered results with two forms of more detailed secondary analysis. First, we examine the implications of our results related to migration. Given the high overall rates of residential churn, who is it that is moving? In which types of neighborhood is the decline in residential churn caused by the program most pronounced? Might impacts on investment or social capital be driven by differential migration rates? Second, we unpack the social capital dimension of the program. While the pre-registered index of social capital was not significantly improved by the intervention, were there dimensions measured by our household survey instrument that *were* improved?

A. *Understanding the Decline in Residential Turnover*

Decades of theory on amenity pricing suggest that improvements in neighborhood amenities will provide net welfare benefits to homeowners, but may have more ambiguous effects on renters. We are therefore interested in understanding whether the overall decline in mobility is masking an acceleration in the turnover of renters. Further, if the departure of renters is occurring disproportionately at the bottom end of the market, this would be evidence of gentrification effects pushing out low-income renters.

To examine these questions, we split the baseline sample of structures according to whether the residents in 2007 were renters (11 percent of baseline households are renters, 72 percent own their homes outright, and the remainder are either still paying mortgages or other arrangements). Columns 1 and 5 of Table 6 show the treatment effect of the program on the subsequent rate of churn in these two samples. While the rate of turnover in the control is much higher for renters (58 percent versus 37 percent for non-renters), the impact of the treatment on decreasing churn is extremely similar, a decrease of just over 9 percentage points in both cases. This is comforting initial evidence that Hábitat is not leading renters to be pushed out en masse; in fact, the program stabilizes residential patterns for homeowners and renters alike.

Table 6—Analysis of Renters versus Non-renters

	Outcome variable: Residential churn (*baseline household has moved by follow-up*)							
	Non-renters at baseline				Renters at baseline			
		Baseline covariate interacted with treatment				Baseline covariate interacted with treatment		
	No interaction	Infrastructure index	Social capital index	Average polygon-level sales price at baseline	No interaction	Infrastructure index	Social capital index	Rent paid at baseline
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	−0.0901	−0.0762	−0.0937	−0.0696	−0.0975	−0.0735	−0.103	−0.109
	(0.045)	(0.040)	(0.041)	(0.037)	(0.048)	(0.048)	(0.049)	(0.048)
Treatment × Covariate		−0.0387	−1.553	−0.000130		−0.0474	0.118	0.00115
		(0.016)	(0.533)	(0.000)		(0.031)	(0.819)	(0.001)
Baseline covariate		0.0312	1.041	0.000146		0.0243	−0.658	−0.000741
		(0.014)	(0.536)	(0.000)		(0.025)	(0.617)	(0.000)
Constant	0.366	0.354	0.363	0.343	0.579	0.562	0.573	0.590
	(0.044)	(0.039)	(0.040)	(0.037)	(0.031)	(0.033)	(0.032)	(0.033)
Observations	8,626	8,626	8,626	8,626	1,065	1,065	1,065	1,004
R^2	0.009	0.014	0.019	0.024	0.009	0.014	0.014	0.018

Notes: The polygon-level analysis with polygon fixed effects and standard errors is clustered at the municipal level. The analysis is weighted by a population weight to be representative of all residents in the study neighborhoods. Standard errors are in parentheses.

Concerns about gentrification can also be addressed by examining whether the treatment leads to a disproportionate turnover of residents at the bottom end of the housing scale. To examine this further, we conduct an interaction analysis to understand the types of polygons and households in which the rate of churn is most strongly decreased by the program. To capture heterogeneity in baseline levels of wealth and social capital, we use the block-level infrastructure index and the polygon-level social capital index. We also separately construct an index of baseline property values for homeowners and renters. For homeowners, we use the average professional INDAABIN assessment of value per square meter in the closest available polygon. For renters, we use the reported amount being paid to rent each property at baseline. We then run linear probability regressions explaining a binary indicator of residential turnover between 2009 and 2012 and use the treatment dummy, the baseline values of these covariates, and their interactions as explanatory variables. This interaction analysis reveals patterns that are quite divergent for homeowners and renters. The uninteracted values of the covariates are positive for homeowners, and the interaction terms are negative. This implies that in the absence of the program there is more turnover of homeowners in richer neighborhoods, but that the program reverses this relationship and is most effective at getting homeowners in richer neighborhoods to stay put. For renters, all of the patterns are reversed; in general, there is more turnover of renters in poor neighborhoods, but the program is particularly effective at getting renters of low-value properties to stay put. These results are comforting in multiple dimensions; not only does the program improve residential stability, it improves it even among renters, and it improves it most strongly among renters in the lowest-priced units. We find no evidence of disruptive gentrification within the three-year time span of the study.

B. *Digging Further into Impacts on Social Capital and Crime*

As described in the introduction, the potential to improve the social capital of disadvantaged communities was an explicit objective in the design of the program; with the idea being that while municipalities primarily build infrastructure, Hábitat's coordinated investments across a wide range of activities would be successful at building community. We showed in Table 4 that the core social capital indicator to which we pre-committed was not significantly moved by the program, but the study featured a rich social capital survey, which we now exploit to conduct a more fine-grained analysis of social capital outcomes. The pre-committed index was made up of components measuring group participation, access to information, and levels of neighborhood trust. It did not include dimensions like crime and teen behavior, but these were nonetheless captured in the survey. To aggregate the survey domains, we follow Kling et al. (2007) by constructing four subindexes on social capital; the first two (participation and trust) are comprised of dimensions included in the overall social capital index, and the second two (security and teen behavior) are comprised of questions not included in the original index.[7]

Panel A of Table 7 shows the simple DID impacts of the program on these four indexes of social capital. A first striking feature of these results is the deterioration in the overall social capital observable across the two rounds of the survey. Any impacts must be considered in the context of the problems facing urban Mexico in the period between 2009 and 2012. The first two indexes, which comprise most of the variables making up the pre-committed index, show no significant changes. The index of participation actually decreases slightly, and while the trust index shows an improvement of 0.13 standard deviations, it's not significant. When we move to the security and youth indexes, however, the story changes. Against a backdrop of a 0.28 standard deviation decrease in the average value of the security index, the program achieves an impact of 0.247, significant at the 95 percent level, suggesting that the treatment allowed these communities to hold steady while the control deteriorated. The youth index does not show the same overall decline between 2009 and 2012, but the impacts here are also relatively large (0.164 standard deviation) and significant. It therefore appears that, while the program did not move the participation and trust outcomes on which the pre-committed index was based, it did lead to a meaningful improvement in security and prospects for youth.

Given the large impact on residential churn discussed above, it is important to investigate whether these social capital impacts are arising because of changes

[7] The participation index is comprised of underlying questions on participation in nine different types of community organizations and questions on the existence of community groups, knowledge of community-level problems, and the degree of information exchange between neighbors. The trust index is built from two questions about the level of confidence in neighbors and local institutions and two questions about the degree of trust within the household and trust of neighbors. The security index is based on having not been a victim of any crime in the past 12 months, having not been subject to assault, and a variable counting the number of activities that have not been abandoned in the past 12 months because of physical insecurity. The youth behavior index is based on four variables; an indicator that youth have somewhere other than the streets to convene, the number of "good" activities teens engage in, the additive inverse of the number of "bad" activities teens engage in, and an indicator variable for the absence of gang/crime risk for youth in the neighborhoods. All indexes are constructed as described in the previous section.

Table 7—Detailed Social Capital Impacts

	Social capital index domain							
	Participation		Trust		Security		Youth	
Panel A. Social capital components								
Treatment × Round 2	−0.0168		0.129		0.247		0.164	
	(0.103)		(0.117)		(0.117)		(0.085)	
Round 2	−0.142		−0.379		−0.282		−0.0252	
	(0.088)		(0.102)		(0.081)		(0.069)	
Round 1 mean in control	0.000		0.000		0.000		0.000	
Observations	11,143		11,143		11,143		11,143	
Panel B. Stayers versus movers								
	Stayers	Movers	Stayers	Movers	Stayers	Movers	Stayers	Movers
Treatment × Round 2	−0.0621	0.0805	0.0852	0.174	0.333	0.074	0.253	0.019
	(0.123)	(0.168)	(0.156)	(0.185)	(0.100)	(0.254)	(0.110)	(0.144)
Round 2	−0.144	−0.140	−0.271	−0.561	−0.303	−0.247	−0.132	0.159
	(0.121)	(0.081)	(0.134)	(0.153)	(0.082)	(0.173)	(0.075)	(0.090)
Round 1 mean in control	0.10	−0.16	−0.01	0.01	0.00	−0.01	0.04	−0.05
Observations	7,559	3,584	7,559	3,584	7,559	3,584	7,588	3,597

Notes: The household-level analysis with polygon fixed effects and standard errors is clustered at the municipal level. The analysis is weighted by a population weight to be representative of all residents in the study neighborhoods. Standard errors are in parentheses.

in mobility, or because the outcomes for a specific set of people were improved. To provide insight into this question, panel B of Table 7 segregates the sample into the structures for which the household changed between 2009 and 2012 (the movers) and those for which it did not (the stayers). Beginning again with the time trends, we see that overall the changes in the social capital indexes are similar for both movers and stayers, with the exception of the youth index, which improved for movers and deteriorated for stayers. It is also clear from the summary statistics at the bottom of panel B that the youth index and particularly the participation index are in general superior for stayers than for movers. Looking at impacts, we see that the significant impact on the security and the youth dimensions is confined to the stayers, and no impacts are found for those households that moved. This group is endogenously selected, meaning that mean outcomes among stayers are subject to both an extensive and intensive margin impact from the program. These results suggest that there are three channels open for the impacts on security and youth: a selection effect improving the outcome by decreasing the fraction of movers, a causal impact within the group that did not move, and a second-order effect from increasing the share of people on whom the stronger treatment effect is found. The results in panels B should not be interpreted as straightforward subgroup causal effects, but they help to provide some context as to the mechanisms through which the treatment effects move.

We subject these social capital results to two types of robustness checks; the first examining the problem of multiple inference and the second the imbalance in the social capital index found in Table 2. As a way of looking into the multiple inference problem, online Appendix Table A5 shows results for some of the key underlying social capital variables that make up the indexes used (Trust in Neighbors appears in the trust index, columns 2–4 appear in the security index, and 5–8 appear in the youth index). Our marginal effect of −0.152 on the number of assaults in the past year would indicate that the 118,000 households in treatment polygons suffered a total of almost 18,000 fewer physical assaults as a result of the program. While highly significant, these outcomes need to be corrected for multiple inference because they have been selected from the much broader potential set of social capital variables (Casey, Glennerster, and Miguel 2012; Humphreys, Sanchez de la Sierra, and van der Windt 2013). At the bottom of online Appendix Table A5, we therefore illustrate several different ways of penalizing ourselves for this fishing using Anderson's False Discovery Rate sharpened q-values (Anderson 2008). In the first row of the bottom panel, we present the q-values (comparable to p-values, which are reported in the top panel of the table) when we correct for multiple hypothesis testing across the nine variables presented here. In the second row, we correct using the nine presented here plus all of the variables used in the construction of the social capital index. In the third row, we correct using the entirety of the 70 possible social capital variables. This result suggests that when we examine impacts on the sub-variables that make up the four domains of social capital presented in Table 7 and fully correct for multiple inference, only those describing the improvement in teen participation in music and sports (an area in which Hábitat invested directly) remain significant.

Another robustness check is to ask whether the dramatic deterioration in the control group and steady outcomes in the treatment group may arise from imbalance rather than from a treatment effect that prevented things from getting worse. To examine this, we first calculate the round 1–round 2 change observed for our key index metrics and then fit a regression in the control group explaining these changes using a battery of baseline covariates. The coefficients from this regression are then used to predict changes for treatment and control alike, and we can examine the balance of these predicted changes. Online Appendix Table A6 shows that all of the core outcomes used in the paper appear balanced in terms of predicted changes, with the exception of the two significant social capital indexes: security and trust. For each of these outcomes, we see that the baseline attributes of treatment communities would anyways have predisposed us to expect improvements relative to the control, even in the absence of a treatment effect. However, when we compare the magnitude of this predicted imbalance to the impact coefficients in Table 7, we see that the impact is substantially larger than the predicted imbalance; 0.247 versus 0.093 for security and 0.164 versus 0.072 for youth. Nonetheless, neither of our measured effects would be significant relative to the predicted imbalance, and hence, both the multiple inference corrections and the imbalance adjustment introduce some skepticism into our interpretation of the impacts on crime and teen behavior.

Where does this leave us in terms of policy interpretation? While these results are admittedly equivocal, we do find impacts on a coherent set of public security outcomes that have a clear logical tie to Hábitat investment in streetlights and sidewalks, as well as in community centers intended to give adolescents constructive outlets. Given the tremendous focus on the rise in violent crime in Mexico over the past decade (Dube, Dube, and García-Ponce 2013; Molzahn, Ríos, and Shirk 2012), its linkages to the economic welfare of citizens (BenYishay and Pearlman 2014; Enamorado, López-Calva, and Rodríguez-Castelán 2014), and the lack of credible policy options on the table to deal with it (Phillips 2015), even suggestive evidence of new mechanisms to improve public security are welcome. We therefore present these results in the spirit of intriguing but speculative findings that merit further research. Taken at face value, this study suggests that investing in neighborhood walkability and community infrastructure that provide constructive outlets for youth can have substantial impacts on improvements in public security.

IV. Conclusion

This paper presents the results of a large experiment in which the Mexican federal government improved the quality of infrastructure in low-income urban neighborhoods. We examine the effects of $68 million in spending spread across 118,000 treatment households and find evidence that infrastructure investment in these neighborhoods is suboptimal. Treatment induces a large improvement in the access to well-functioning public lighting, paved roads, and sidewalks; private investment in the housing stock increases, neighborhood churn in real estate decreases by a quarter, crime falls, and the total increase in the value of the property in intervention neighborhoods is more than twice the cost of the program. This high rate of return suggests substantial money "left on the table" via underinvestment. However, a program that spent an average of $550 per beneficiary household did not improve access to water or sewerage (despite having spent more than 10 percent of their budget on these items), and an index of social capital was not substantially improved.

Are the returns to the program using real estate prices of a credible magnitude? Our estimates of two-to-one returns are larger than those available from the developed world. Cellini, Ferreira, and Rothstein (2010) find an increase of $1.50 in the willingness to pay of home buyers for every $1 invested in public schools in California and lay out a simple political theory that says while marginal returns on public investment should be zero, they may be positive in equilibrium because individuals within the community who don't value those things (or already have them) will be unable to support additional spending on the margin. Pereira and Flores de Frutos (1999) use a vector auto-regression model on public spending in the United States, finding that every dollar invested returns 65 cents in private investments. Given that we may expect infrastructure spending in poor Mexican neighborhoods to be farther below efficient levels than in the United States, a figure of $2 may not be unreasonable.

Our results should bolster the argument that we not overlook large-scale spending on macro programs in the face of proven micro interventions, such as conditional cash transfer (CCT) programs (Fiszbein et al. 2009). Given that the flagship CCT was also experimentally tested in Mexico, the *Oportunidades* program provides an interesting point of comparison. *Oportunidades* pays an average of $71 per month to beneficiary households. This means that the Hábitat investment of $550 per household would represent fewer than eight months of cash transfers and has resulted in an increase in the asset wealth of the household of twice this sum and the broader set of amenity benefits flowing from lower crime and more stable neighborhoods. CCT programs are designed to create a temporary flow of consumption benefits that leave behind an improved stock of human capital; here, we see infrastructure spending generating flow improvements in the quality of life and leaving behind a substantially improved stock of property value. Improving infrastructure in underserved locations can deliver real social benefits as well as a substantial surge in household wealth.

REFERENCES

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–95.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Baird, Sarah, Aislinn Bohren, Craig McIntosh, and Berk Özler. 2014. "Designing Experiments to Measure Spillover Effects." Institute for International Economic Policy (IIEP) Working Paper 2014–11.

BenYishay, Ariel, and Sarah Pearlman. 2014. "Crime and Microenterprise Growth: Evidence from Mexico." *World Development* 56: 139–52.

Blomquist, Glenn C., Mark C. Berger, and John P. Hoehn. 1988. "New Estimates of Quality of Life in Urban Areas." *American Economic Review* 78 (1): 89–107.

Busso, Matias, Jesse Gregory, and Patrick Kline. 2013. "Assessing the Incidence and Efficiency of a Prominent Place Based Policy." *American Economic Review* 103 (2): 897–947.

Campuzano, Larissa, Dan Levy, and Andres Zamudio. 2007. "The Effects of Habitat on Basic Infrastructure." https://www.mathematica-mpr.com/-/media/publications/pdfs/habitatbasic.pdf.

Casaburi, Lorenzo, Rachel Glennerster, and Tavneet Suri. 2013. "Rural Roads and Intermediated Trade: Regression Discontinuity Evidence from Sierra Leone." https://scholar.harvard.edu/files/lorenzocasaburi/files/casaburi_glennester_suri_roads.pdf.

Casey, Katherine, Rachel Glennerster, and Edward Miguel. 2012. "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan." *Quarterly Journal of Economics* 127 (4): 1755–1812.

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.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein. 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *Quarterly Journal of Economics* 125 (1): 215–61.

Chen, Shaohua, Ren Mu, and Martin Ravallion. 2009. "Are there lasting impacts of aid to poor areas?" *Journal of Public Economics* 93 (3–4): 512–28.

Cohen, Jeffrey P., and Catherine J. Morrison Paul. 2004. "Public Infrastructure Investment, Interstate Spatial Spillovers, and Manufacturing Costs." *Review of Economics and Statistics* 86 (2): 551–60.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora. 2013. "Do Labor Market Policies have a Displacement Effect? Evidence from a Clustered Randomized Experiment." *Quarterly Journal of Economics* 128 (2): 531–80.

Dahlberg, Matz, Eva Mörk, Jørn Rattsø, and Hanna Ågren. 2008. "Using a discontinuous grant rule to identify the effect of grants on local taxes and spending." *Journal of Public Economics* 92 (12): 2320–35.

de Janvry, Alain, Kyle Emerick, Marco Gonzalez-Navarro, and Elisabeth Sadoulet. 2015. "Delinking Land Rights from Land Use: Certification and Migration in Mexico." *American Economic Review* 105 (10): 3125–49.

Dercon, Stefan, Daniel O. Gilligan, John Hoddinott, and Tassew Woldehanna. 2009. "The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages." *American Journal of Agricultural Economics* 91 (4): 1007–21.

Dube, Arindrajit, Oeindrila Dube, and Omar García-Ponce. 2013. "Cross-Border Spillover: U.S. Gun Laws and Violence in Mexico." *American Political Science Review* 107 (3): 397–417.

Duflo, Esther, and Rohini Pande. 2007. "Dams." *Quarterly Journal of Economics* 122 (2): 601–46.

Enamorado, Ted, Luis F. López-Calva, and Carlos Rodríguez-Castelán. 2014. "Crime and growth convergence: Evidence from Mexico." *Economics Letters* 125 (1): 9–13.

Field, Erica. 2005. "Property Rights and Investment in Urban Slums." *Journal of the European Economic Association* 3 (2–3): 279–90.

Field, Erica, and Michael Kremer. 2006. *Impact Evaluation for Slum Upgrading Interventions*. World Bank. Washington, DC, October.

Fiszbein, Ariel, Norbert Schady, Francisco H. G. Ferreira, Margaret Grosh, Nial Kelleher, Pedro Olinto, and Emmanuel Skoufias. 2009. *Conditional Cash Transfers: Reducing Present and Future Poverty*. Washington, DC: World Bank.

Galiani, Sebastian, Paul Gertler, Ryan Cooper, Sebastian Martinez, Adam Ross, and Raimundo Undurranga. 2013. "Shelter from the Storm: Upgrading Housing Infrastructure in Latin American Slums." National Bureau of Economic Research (NBER) Working Paper 19322.

Galiani, Sebastian, and Ernesto Schargrodsky. 2010. "Property rights for the poor: Effects of land titling." *Journal of Public Economics* 94 (9–10): 700–729.

Gentilini, Ugo. 2007. *Cash and Food Transfers: A Primer*. Rome, Italy: World Food Program.

Giné, Xavier, and Ghazala Mansuri. 2011. "Together We Will: Experimental Evidence on Female Voting Behavior in Pakistan." World Bank Policy Research Working Paper 5692.

Gonzalez-Navarro, Marco, and Climent Quintana-Domeque. 2009. "The reliability of self-reported home values in a developing country context." *Journal of Housing Economics* 18 (4): 311–24.

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): 254–67.

Gordon, Nora. 2004. "Do federal grants boost school spending? Evidence from Title 1." *Journal of Public Economics* 88 (9–10): 1771–92.

Greenstone, Michael, and Justin Gallagher. 2008. "Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program." *Quarterly Journal of Economics* 123 (3): 951–1003.

Hansen, Henrik, Ole Winckler Andersen, and Howard White, eds. 2012. *Impact Evaluation of Infrastructure Interventions*. New York: Routledge.

Hines, James R., Jr., and Richard H. Thaler. 1995. "Anomalies: The Flypaper Effect." *Journal of Economic Perspectives* 9 (4): 217–26.

Humphreys, Macartan, Raul Sanchez de la Sierra, and Peter van der Windt. 2013. "Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration." *Political Analysis* 21 (1): 1–20.

Khandker, Shahidur R., Zaid Bakht, and Gayatri B. Koolwal. 2009. "The Poverty Impact of Rural Roads: Evidence from Bangladesh." *Economic Development and Cultural Change* 57 (4): 685–722.

Kling, Jeffery R., Jeffery B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

Kremer, Michael, Jessica Leino, Edward Miguel, and Alix Peterson Zwane. 2011. "Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions." *Quarterly Journal of Economics* 126 (1): 145–205.

Mansuri, Ghazala, and Vijayendra Rao. 2004. "Community-Based and -Driven Development: A Critical Review." *World Bank Research Observer* 19 (1): 1–39.

McIntosh, Craig, Tito Alegría, Gerardo Ordóñez, and René Zenteno. 2018. "The Neighborhood Impacts of Local Infrastructure Investment: Evidence from Urban Mexico: Dataset." *American Economic Journal: Applied Economics*. https://doi.org/10.1257/app.20160429.

Moffit, Robert. 1989. "Estimating the Value of an In-Kind Transfer: The Case of Food Stamps." *Econometrica* 57 (2): 385–409.

Molzahn, Cory, Viridiana Ríos, and David A. Shirk. 2012. *Drug Violence in Mexico: Data and Analysis Through 2011*. University of San Diego Joan B. Kroc School of Peace Studies Trans-Border Institute. San Diego, March.

Nesbit, Todd M., and Steven F. Kreft. 2009. "Federal Grants, Earmarked Revenues, and Budget Crowd-Out: State Highway Funding." *Public Budgeting and Finance* 29 (2): 94–110.

Newman, John, Menno Pradhan, Laura B. Rawlings, Geert Ridder, Ramiro Coa, and Jose Luis Evia. 2002. "An Impact Evaluation of Education, Health, and Water Supply Investments by the Bolivian Social Investment Fund." *World Bank Economic Review* 16 (2): 241–74.

Newman, John, Laura Rawlings, and Paul Gertler. 1994. "Using Randomized Control Designs in Evaluating Social Sector Programs in Developing Countries." *World Bank Research Observer* 9 (2): 181–201.

Ordóñez-Barba, Gerardo, Tito Alegría-Olazábal, Craig McIntosh, and René Zenteno-Quintero. 2013. "Alcances e impactos del program a hábitat en comunidades pobres urbanas de México." *Papeles de Población* 19 (77): 231–67.

Ordóñez Barba, Gerardo, and Wilfrido Ruiz Ochoa. 2015. "Formación de capital social comunitario a partir de programas orientados a combatir la pobreza en México: El impacto de Hábitat." *Gestión y Política Pública* 24 (1): 3–49.

Paxson, Christina, and Norbert R. Schady. 2002. "The Allocation and Impact of Social Funds: Spending on School Infrastructure in Peru." *World Bank Economic Review* 16 (2): 297–319.

Pereira, Alfredo M., and Rafael Flores de Frutos. 1999. "Public Capital Accumulation and Private Sector Performance." *Journal of Urban Economics* 46 (2): 300–322.

Phillips, Brian J. 2015. "How Does Leadership Decapitation Affect Violence? The Case of Drug Trafficking Organizations in Mexico." *Journal of Politics* 77 (2): 324–36.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90 (6): 1257–78.

Rosen, Sherwin. 1979. "Wage-based Indexes of Urban Quality of Life." In *Current issues in urban economics*, edited by Peter M. Mieszkowski and Mahlon R. Straszheim, 74–104. Baltimore: Johns Hopkins University Press.

Sinclair, Betsy, Margaret McConnell, and Donald P. Green. 2012. "Detecting Spillover Effects: Design and Analysis of Multilevel Experiments." *American Journal of Political Science* 56 (4): 1055–69.

Turley, Ruth, Ruhi Saith, Nandita Bhan, Eva Rehfuess, and Ben Carter. 2013. "Slum upgrading strategies involving physical environment and infrastructure interventions and their effects on health and socio-economic outcomes." *Cochrane Database of Systematic Reviews* 2013 (1): CD010067.

Vilalta, Carlos. 2014. "How Did Things Get So Bad So Quickly? An Assessment of the Initial Conditions of the War Against Organized Crime in Mexico." *European Journal on Criminal Policy and Research* 20 (1): 137–61.

Wooldridge, Jeffrey M. 2002. "Inverse probability weighted M-estimators for sample selection, attrition, and stratification." *Portuguese Economic Journal* 1 (2): 117–39.