

Slum Infrastructure Upgrading and Budgeting Spillovers: The Case of Mexico's Hábitat Program

Craig McIntosh*
UC San Diego, CEGA, and JPAL

Tito Alegria
El Colegio de la Frontera Norte

Gerardo Ordoñez
El Colegio de la Frontera Norte

René Zenteno
UT San Antonio and El Colegio de la Frontera Norte

October 2013

Abstract

This paper reports on the results of a large infrastructure investment experiment in which \$67 million in spending was randomly allocated across a set of urban slums in Mexico. We show that the program resulted in meaningful improvements in the access of the average household to numerous forms of infrastructure, such as electric lighting, street lights, sidewalks, medians, and road paving. We exploit the study's randomized saturation design to understand how compensating expenditure changes at the municipal level might work to undermine causal inference when multiple levels of government provide forms of investments that are close substitutes. The program increased the aggregate real estate value in program neighborhoods by two dollars for every dollar invested.

Keywords: Infrastructure impacts, flypaper effects, evaluation

JEL Codes: H54, H73, C93

* Corresponding author, ctmcintosh@ucsd.edu. 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, and to participants in seminars at Claremont-McKenna, CU Boulder, El Colegio de la Frontera Norte, IFPRI, Stanford, UCSD, and UC Berkeley for helpful comments. The authors were paid as consultants to perform this evaluation by the Inter-American Development Bank and the Social Development Secretariat (SEDESOL) of the Mexican government.

1. INTRODUCTION

Despite a long-standing belief in the critical role of infrastructure investment (Sachs 2005), much of the recent push for evidence-based development policy has focused primarily on micro-level interventions such as cash transfers (Skoufias & Parker 2001, Fiszbein et al. 2009), bed nets (Dupas 2009, Tarozzi et al. 2011), and individual financial services (Bannerjee et al., 2013, Ashraf et al., 2006). The compelling evidence of positive impact from many micro-interventions has arisen largely as a result of the use of Randomized Controlled Trials (RCTs) to assess these interventions. In contrast, the evidence base for infrastructure investment is almost entirely observational; examples of the strategies used to identify the effects of macro-investment in infrastructure include studies using staggered rollout (Dercon et al. 2009, Galiani et al. 2008, Galiani & Shargrodsky 2010), matching estimators (Chase 2002, Newman et al. 2002), discontinuity designs (Cellini et al. 2012, Casaburi et al. 2013) and instrumental variables (Paxson & Schady 2002, Duflo & Pande 2007). In this environment, it becomes critical that we develop an equally well-identified evidence base on large-scale investments like infrastructure (Newman et al., 1994), or else we risk seeing evidence-based funding flow towards interventions simply because they are easily evaluable.

This paper joins a recent but rapidly-growing literature using RCTs to examine the impact of improvements in infrastructure.¹ We present the results of a major federal infrastructural spending experiment implemented across Mexico during the years 2009-2011 by the Social Development Secretariat (SEDESOL) through its major urban investment program, Hábitat. The study is very large in absolute scale: \$67 million in infrastructure investment was randomly allocated, making this the largest experiment attempted in Mexico since Progresá. Hábitat primarily builds heavy infrastructure such as roads, water, sewerage, lighting, and sidewalks, but also invests in community centers, parks, and sports facilities. The experiment includes most of the urban parts of the country (60 municipalities across 20 different Mexican states), and is accompanied by detailed household & block-level data collection (9,702 households surveyed in a two-period panel). To measure real estate impacts of the program, professional valuation of the change in price of empty lots was conducted by the Management Institute of National Property Appraisals (INDAABIN). This is an evaluation ‘at scale’ of an actual program administered by the federal government.

¹ Kremer et al. 2011 randomize placement of water infrastructure across 186 springs in Kenya and find meaningful decreases in child diarrhea. Closer to both the location and the spirit of this paper, Gonzales-Navarro & Quintana-Domeque (2011) examine a street-paving experiment in the city of Acayucan (Veracruz), Mexico, and document a corresponding increase in private investment, housing values, and satisfaction with local politicians. Galiani et al. (2013) examine the impact of providing slum households with a new, prefabricated house.

A large public finance literature examines the possibility that changes in federal spending may be crowded out or crowded in (the ‘flypaper effect, as in Hines and Thaler, 1995) by local government spending (Dahlberg et al. 2008, Nesbit and Kreft 2009).² Less well appreciated is the fact the re-budgeting implied by these effects is likely to violate the Stable Unit Treatment Value Assumption (SUTVA), and hence to lead to bias even in an experimental estimator. Facing a pre-announced experiment in federal infrastructure spending, a municipal government might be inclined to redistribute spending towards the control if it faced a concave objective function, or it might be required to tax the control in order to meet federal matching requirements. In either case, the counterfactual is contaminated.³ Foreseeing the potential for bias, the Hábitat experiment implemented a ‘randomized saturation’ design (Baird et al, 2013) intended to reveal the ways in which lower levels of government re-optimize around a federal spending experiment.⁴ We began from a frame of ‘polygons’ (hereafter referred to as neighborhoods) in 60 municipalities, all of which were Hábitat-eligible. We first assigned each municipality a treatment saturation drawn from a uniform distribution between .1 and .9 and then assigned treatment randomly at the neighborhood level so as to make the actual study fraction treated as close as possible to the assigned saturation given the integer problem. Figure 1 presents the empirical distribution of the treatment saturations in the experiment and a visualization of the way in which we can exploit this variation using the control group.

This randomization of the intensity of treatment within a municipality allows us to look for crowdout or flypaper effects in two quite distinct dimensions. The first is to ask whether increasing the fraction of neighborhoods treated with federal infrastructure alters the quality of infrastructure in the remaining control neighborhoods. This form of displacement is particularly important because it represents a form of causal interference that will violate the internal validity of standard experimental designs. The second dimension is to look directly for the existence of a flypaper effect

² A recent literature on federal programs intended to promote long-term employment such as works programs or enterprise zones has discussed general-equilibrium effects when worker migration causes arbitrage; see Busso et al. (2013) or Suárez-Serrato and Wingender (2011). Hábitat makes a relatively short-term investment in durable infrastructure, and thus we expect effects driven by long-term relocation for work to be small. We are therefore primarily concerned with spillovers in expenditures rather than in labor markets.

³ This concern was paramount at the time of the design of the experiment because SEDESOL had recently received a propensity-score evaluation of their previous phase of implementation conducted by Mathematica (Campuzano et al., 2007), and this study had found no significant impacts on the core infrastructural outcomes. SEDESOL’s concern was that municipal governments were taking money away from treated locations and spending it in untreated ones, leading to a downward-biased measure of their real impact.

⁴ The use of a two-level randomized saturation experiment follows a relatively new literature that is seeking to understand spillover effects using experiments designed explicitly for this purpose, such as McConnell et al. 2011, Gine and Mansuri 2011, Crepon et al. 2011, and Callen et al. 2013.

in overall municipal spending on infrastructure as federal investment increases. We show that instrumenting for actual federal spending variation across municipalities with the treatment saturation delivers experimental variation ranging from almost nothing to over two million dollars per municipality. Using this source of variation, we find evidence consistent with flypaper effects in overall municipal infrastructure spending. The paper develops a simple electoral model to motivate the tension between crowd-out and matching requirements as the two levels of government interact, and we draw the analogy between this problem and the large literature on infra-marginal food aid (Moffit 1989, Gentili 2007). We show how to use the novel estimands uncovered by the randomized saturation design to reveal the extent of interference in the study, and how to use a linearity assumption in the response to saturations to correct the estimate of the Intention to Treat (ITT).

The results of the study show that Hábitat investment resulted in very large improvements in the reported quality of road paving, sidewalks, medians, and public lighting, but less so to forms of infrastructure to which baseline access was very high; namely water, sewerage, and electricity. An index of infrastructure quality, to which we committed as the core outcome indicator in a pre-analysis plan, is significantly improved by the intervention.⁵ Spillover effects are found to be muted overall, although generally consistent with a mild crowd-in or flypaper effect. Control neighborhood infrastructure becomes slightly worse as treatment intensity in a municipality increases, so the estimates of the ITT that correct for this indicate somewhat smaller impacts than uncorrected estimates. Private investment in the housing stock is broadly improved by the program, and the value of a square meter of property in treatment neighborhoods increases by \$2 for every \$1 invested by the program, a sign of under-investment in Mexican slum infrastructure. Contrary to Gonzales-Navarro & Quintana-Domeque (2011) we find the program was very neutral politically, with no electoral advantage having been incurred by the incumbent political party in the 2012 elections.

The paper is organized as follows: Section 2 presents a simple theory for the multi-level budgeting game induced by the experiment and the ways in which the randomized saturation design helps to disentangle the causal inference problem. Section 3 introduces the program and the data

⁵ The companion paper to this, Zenteno et al. 2013, presents the impacts of the program on the other outcomes indicated in the pre-analysis plan, and shows that while the program did not improve public health outcomes or transportation times, it resulted in a meaningful improvement in trust between neighbors and large decrease in the rate of violent crime.

collection strategy, Section 4 presents the estimates of the impact of the program, and Section 5 concludes.

2. THE MULTILEVEL BUDGETING PROBLEM

2.1. Multilevel Budgeting and Causal Inference.

The response of local governments when given block grants from the central government has been a major focus of inquiry in public economics for many decades. A basic model of rationality at the local level would suggest when the pattern of federal spending is well understood, such grants should be almost entirely crowded out by changes in local spending behavior. Nonetheless numerous studies have documented the ‘flypaper effect’ in which money “sticks where it hits” and the corresponding changes in total spending arising from block grants are close to 100% of the size of the grant (Hines and Thaler 1995, Nesbit and Kreft 2009). Indeed, some recent empirical studies have found evidence not of crowd-out but of crowd-in, whereby federal grants induce an absolute increase in local spending (Dahlberg et al, 2008). Recognizing the likelihood that federal spending is targeted towards local areas with specific counterfactual spending patterns, researchers have attempted to exploit discontinuities (Gordon 2004) or instrumental variables (Dahlberg et al., 2008) to identify casual effects on municipal spending. Our first empirical analysis in Section 4.1 uses the randomized component of the variation in federal spending per municipality to examine this relationship experimentally.

We then turn to a more standard analysis of impacts at the neighborhood level, using untreated neighborhoods in study municipalities to form counterfactuals. Even with an experiment in federal spending, the causal effect recovered by this exercise is far from clear. At the simplest level, if local governments understand the design of the experiment and fully re-optimize they may completely crowd out the federal expenditure variation. An experiment in this context would recover a zero impact even under conditions where improvements in spending created large benefits. If infrastructure is subject to a budget constraint, then increasing the fraction of a municipality treated can create an increasingly strong budgetary spillover effect on other parts of a municipality; as the matching required by municipal governments goes up then this budgetary effect is more likely to be negative (implying overestimation of treatment effects). Finally, federal spending in infrastructure may crowd local governments in or out of overall infrastructure spending. We now develop a simple theoretical environment in which to consider these competing effects, and tie these concepts to the quantities recovered by the randomization of the intensity of treatment across municipalities.

The basic evaluation problem is to measure the intention-to-treat (ITT) of an experiment that randomizes an additional injection of federal spending. Consider a municipal government that takes federal spending as given, and then maximizes a concave electoral return function where spending in each neighborhood i within municipality j yields a higher vote total at an ever-decreasing rate. We assume that neighborhoods within a municipality are homogenous other than the treatment status they are assigned in the federal experiment. Further, assume that voters have no ability to attribute spending to the correct entity, so that baseline federal expenditures S_{ij} and municipal expenditures s_{ij} are perfect substitutes. Municipal governments maximize this objective function subject to an overall budget constraint $\sum_i s_{ij} \leq B_j$, leading to an optimized pre-experimental per-neighborhood spending s_{ij}^* , with $s_{ij}^* = s_{kj}^* \forall i \neq k$. Because the two forms of spending are perfect substitutes and the objective function is concave, the municipal government will want to counteract any exogenous change in spending by the federal government one-for-one.

From this initial equilibrium, the federal government initiates a randomized experiment in which fraction σ_j of the locations within each municipality are selected to be a part of the study, a fraction τ_j of study locations treated per municipality is randomized at the municipality, and then conditional on this saturation and the locations within each municipality are randomly assigned a binary treatment indicator T_{ij} . Assume that the federal government spends a fixed amount K per treatment neighborhood. Foreseeing the desire towards crowd-out, H abitat like many similar federal programs requires matching of federal monies by state and municipal government, so that municipal governments are required to contribute spending of mK towards H abitat investments in treatment neighborhoods. Municipalities thus face competing incentives to crowd out windfall federal spending, and yet if the matching requirements are high enough it may force them to crowd in municipal spending to the treatment.

Requiring municipal governments to spend their money in specific places represents a kind of ‘tied’ aid, and creates an analogy between the budgeting spillover question and the substantial literature on food versus cash aid (Moffit 1989). The core intuition of this literature is that as long as the post-transfer desired spending on the mandated category would have been larger than the required spending, the individual will be able to internally re-budget so as to make cash and food aid equivalent. The analogy to our problem is that as long as the re-optimized post-transfer municipal

spending per neighborhood is larger than the required matching amount, the municipal governments will be able completely to unwind the experiment.

We make a slight change to the standard counterfactual notation to consider the impacts recovered by an experiment in this context where SUTVA is violated. Let s_{0ij}^* be the optimized spending per neighborhood in the absence of *the experiment*, and s_{1ij}^* be the optimized spending in the presence of the experiment. By the properties of randomization, $E(s_{0ij}^* | T = 1) = E(s_{0ij}^* | T = 0)$, but the possibility of strategic spillovers means that $E(s_{1ij}^* | T = 0)$ is not necessarily equal to $E(s_{0ij}^* | T = 0)$.

The experiment induces a ‘Net Budget Effect $NBE(\tau)$ ’ equal to $\sigma_j \tau_j K$; this is the amount by which spending would increase in every neighborhood if the additional federal resources were spread evenly within the municipality. After the experiment, municipal governments will *want* a total amount of $s_{0j}^* + \sigma_j \tau_j K$ to be spent in every location, meaning that absent constraints, $s_{1ij}^* = s_{0j}^* - K(1 - \sigma_j \tau_j)$ in treatment locations and $s_{1ij}^* = s_{0j}^* + \sigma_j \tau_j K$ in control locations. The ability of municipal governments to achieve this given the matching requirements will determine what is measured in the experiment.

If the matching constraint does not bind, meaning that $mK < s_{0j}^* - K(1 - \sigma_j \tau_j)$, then every neighborhood sees its spending rise by the NBE, the experiment is completely unwound, and the Average Spillover on the Control $ASC(\tau)$ will be the $NBE(\tau)$. Therefore, if matching constraints do not bind then we have $\frac{dNBE(\tau)}{d\tau} = \frac{dASC(\tau)}{d\tau} = \sigma_j K > 0$, and we will observe a Net Budget Differential (the difference between final optimized spending in the treatment and the control) $NBD(\tau) = ITT(\tau) = 0$ for every τ (even though the $NBE(\tau)$ increases with τ).

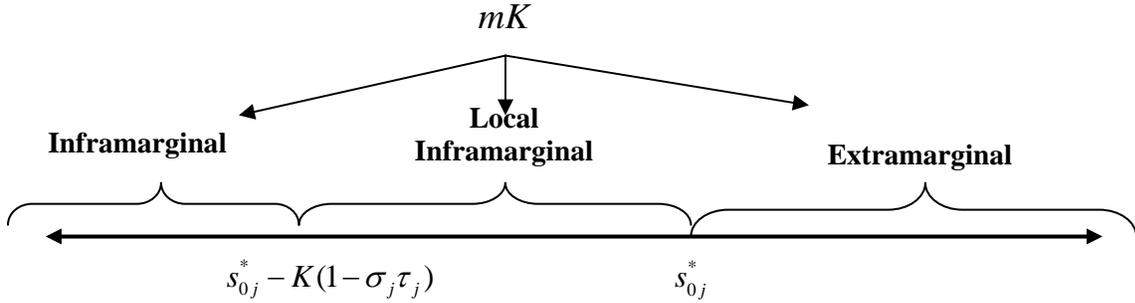
If on the other hand the matching constraint is binding meaning that $mK > s_{0j}^* - K(1 - \sigma_j \tau_j)$, then municipal spending will equal mK in all treatment locations. Decisionmaking is no longer done to balance the marginal conditions; instead the matching dictates the redistribution across treatment and control that can (or must) be done. Within this range if $mK > s_{0ij}^*$ then the control locations must be taxed to pay for the matching money, while if $mK < s_{0ij}^*$ then the additional resources from the experiment will be partially redistributed to the control. When the matching

constraint is binding then $ASC(\tau) = \frac{\sigma_j \tau_j}{1 - \sigma_j \tau_j} (s_{0j}^* - mK)$ and $\frac{dASC(\tau)}{d\tau} = \frac{s_{0j}^* - mK}{(1 - \sigma_j \tau_j)^2}$, both of

which have the same sign as $s_{0j}^* - mK$. Also, $NBD(\tau) = K - \frac{(\sigma_j \tau_j)^2}{1 - \sigma_j \tau_j} (s_{0j}^* - mK)$ and

$$\frac{dNBD(\tau)}{d\tau} = -\frac{(1 - \sigma_j \tau_j) 2\sigma_j^2 \tau_j + \sigma_j^3 \tau_j^2}{(1 - \sigma_j \tau_j)^2} (s_{0j}^* - mK),$$

both of which have the opposite sign to $s_{0j}^* - mK$. The fact that the signs of these derivatives across the saturation distribution flip depending on whether $mK < s_{0j}^*$ allows us to divide the decision space into three empirically relevant regions:



If the matching requirement is higher than the re-optimized spending per neighborhood, then the matching is strictly binding and the municipality will have to tax the control locations in order to be able to balance its budget. This is an ‘extramarginal’ transfer. If pre-treatment spending was anyways greater than the matching requirement, then the experiment can be completely unwound and we are in the ‘inframarginal’ case. In between these two values (the ‘local inframarginal’ case) the matching constraint will be binding but rather than taxing controls to pay for the matching, the experiment will generate a net transfer to controls (although not a complete unwinding of the experiment).

We can now consider how randomizing the treatment saturation τ_j allows us to distinguish these cases from each other when we consider some outcome $Y_{ijt} = f(S_{ijt} + K_{ijt} + s_{ijt}^*)$, where $f(\cdot)$ is a strictly non-decreasing function of total expenditure. Our core outcome will be an index of infrastructural quality calculated at the neighborhood level.

Consider the following regression equations using panel neighborhood-level data:

$$(1) \quad Y_{ijt} = \alpha_{ij} + \delta_t^A + \gamma(T_{ij} * \delta_t) + \varepsilon_{ijt}$$

Here, α_{ij} is a set of neighborhood-level fixed effects, δ_t^A measures the average change in outcomes for the control group, and γ gives:

$$I\hat{T}T(\tau) = E\left(Y(S_{ij} + \bar{K} + s_{1ij}^*) \mid \tau_j, T_{ij} = 1\right) - E\left(Y(S_{ij} + s_{0ij}^*) \mid \tau_j, T_{ij} = 0\right).$$
 We can also write the term recovered by the experiment as $I\hat{T}T(\tau) = E(\Delta Y(NBD(\tau)))$.

The estimated quantity is not necessarily the same as the desired true ITT, which is

$$I\tilde{T}T(\tau) \equiv E\left(Y(S_{ij} + \bar{K} + s_{1ij}^*) \mid \tau_j, T_{ij} = 1\right) - E\left(Y(S_{ij} + s_{0ij}^*) \mid \tau_j, T_{ij} = 0\right)$$

We can linearize the slope terms identified by the Randomized Saturation experiment to achieve a simple parameterization of saturation effects:

$$(2) \quad Y_{ijt} = \alpha_{ij} + \delta_t^B + \beta(T_{ij} * \delta_t) + \mu_1(\tau_j * \delta_t) + \mu_2(\tau_j * T_{ij} * \delta_t) + \varepsilon_{ijt}$$

In this regression, μ_1 linearizes the slope of $E(\Delta Y(\cdot) \mid \tau)$ in the control, and μ_2 measures whether this saturation-driven change is different for treated neighborhoods than control neighborhoods. Experimental identification of spillover effects allows us to posit four observationally distinct possibilities for how treatment and control outcomes will respond to randomized variation in municipal-level saturations.

1. No strategic response.

Here, local governments do not re-optimize, and we recover the standard causal estimands under SUTVA:

- $\hat{\gamma} = \hat{\beta} = ITT(K)$
- $\hat{\mu}_1 = \hat{\mu}_2 = 0$

2a. Inframarginal strategic response. No treatment effect, treatment and control outcomes improve equally with saturations by the Net Budget Effect.

- $\hat{\gamma} = 0$
- $\hat{\mu}_1 = \frac{dASC(\tau)}{d\tau} > 0, \quad \hat{\mu}_2 = \frac{dNBD(\tau)}{d\tau} = 0$

2b. Local Inframarginal strategic response. Partial treatment effect, positive net spillover effect to the control.

- $0 < \hat{\gamma} = ITT(NBD) < ITT(K)$

- $\hat{\mu}_1 = \frac{dASC(\tau)}{d\tau} > 0$, $\hat{\mu}_2 = \frac{dNBD(\tau)}{d\tau} < 0$

2c. Extramarginal strategic response. Estimated treatment effect includes municipal spending, negative net spillover effect to the control.

- $\hat{\gamma} = ITT(NBD) > ITT(K)$
- $\hat{\mu}_1 = \frac{dASC(\tau)}{d\tau} < 0$, $\hat{\mu}_2 = \frac{dNBD(\tau)}{d\tau} > 0$

In an RCT in which treatment had been blocked at the municipal level, there would be no way to investigate this bias in the ITT empirically because $\tau_i = \bar{\tau}$, and so controls within every municipality would receive identical spillover effects. This two-layer experiment with randomization of both τ_j and T_{ij} allows us to directly test the ways in which spillovers may be biasing causal inference.

2.2. Recovering Internally Valid Estimates in the Face of Interference.

If we find evidence of spillover effects, even an RCT will be subject to bias if it is using within-cluster observations as a counterfactual. The problem comes from the fact that the panel counterfactual outcome $\delta^A = E(\Delta Y(s_{ij}^*) | T_i = 0)$ includes the mean of the $ASC(\tau)$ across the empirical saturations used in the experiment, while the desired unperturbed control outcome $E(\Delta Y(s_{0ij}^*) | T_i = 0)$ would have no spillovers. As shown in Baird et al (2012), by imposing a continuity assumption on the distribution of outcomes as a function of τ we can estimate the control outcome that would have obtained in the absence of the experiment (that is, at zero treatment saturation). With this, we can adjust the ITT by the difference between the desired and observed counterfactual outcome. This is useful because even in an experiment without a pure control (completely untreated municipalities), a functional form assumption in combination with the randomization of saturations allows us to back out the desired unperturbed control outcome $\lim_{\tau \rightarrow 0} E(\Delta Y(s_{ij}^*) | \tau_j, T_{ij} = 0)$. A very simple means of establishing this desired quantity is $\hat{\delta}^B$ from Equation (2) above. This intercept is the estimated control group outcome at a saturation of $\tau = 0$ given a linear (affine) relationship between the saturation and the outcome.

The actual counterfactual change in outcome used in the simple ITT regression, on the other hand, is estimated by $\hat{\delta}^A$ in Equation (1).. This implies that the bias in the simple \hat{ITT} is given by

$\delta_t^B - \delta_t^A$, which is the average spillover to the control observed in the trial, and the ‘corrected’ ITT that is purged of this spillover effect is given by $\gamma + (\delta_t^A - \delta_t^B)$.

Graphical intuition for this correction is provided in Figure 2. The figure shows a scatterplot of the average change in the infrastructure index for the treatment group (black) and the control group (gray), as well as the linear fit of the change in outcomes across saturations for each group. The point represented by a star illustrates the average empirical saturation in the experiment (.462) and the average weighted change in outcomes among control locations. This value ($\delta_t^A = .115$) serves as the actual counterfactual for the experiment. The point represented by a diamond is the projected change in outcomes for the control where the saturation equals zero; this quantity ($\delta_t^B = .230$) gives the estimated pure control outcome that would have obtained in the absence of the experiment, and represents the desired counterfactual outcome. The difference between these two terms is weighted average bias caused by spillovers into the control group, and suggests that the treatment effect is overestimated by the difference, .115. The standard ITT estimate is .220, and so the ‘corrected’ ITT, given by $\gamma + (\delta_t^A - \delta_t^B)$, is .105. The downward slope for both the treatment and control as saturations increase suggests that municipal spending is extramarginal: as the total amount of federal spending they are asked to match increases, improvements in infrastructure within each location become smaller due to binding budget constraints. Without the use of the randomized saturation design we would have been unable to investigate the average spillover to untreated locations, and hence unable to correct for this bias.

3. IMPLEMENTATION AND DATA

3.1. Program Description.

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. The program targets the urban poor and focuses on slum upgrading, pouring money into urban infrastructure investment but also investing heavily in Community Development Centers (CDCs) and skills upgrading such as job training for the unemployed and health and nutrition training for young mothers.

When it intervenes, Hábitat defines a ‘polygon’ which is effectively a shapefile designating a specific slum neighborhood. A Hábitat polygon is smaller than a locality and is a designation not

used by any other layer of government. The target population of the program are settled households in marginalized urban areas in which has a concentration of households in asset poverty of at least 50%, located in cities of 15,000 inhabitants or more, present a deficit of infrastructure, equipment and urban services, and with an occupancy of at least 80% of the lots having no problems of land tenure. In order to be eligible to benefit from Hábitat, a polygon must have a state and municipal government that is willing to cooperate with Hábitat's cost-sharing rules (which involve local governments providing 50% of project costs; in our projects the municipalities provided 40%, the states 8%, and the beneficiaries 2%).⁶ While early iterations of the program focused heavily on the issue of regularizing land tenure, the program was reformed in 2008 and the wave of implementation studied here actually required that residents not be in breach of land ownership laws.

The lion's share of Hábitat investment goes into a set of activities it calls 'Urban Environment Improvement'. These consist of the introduction or improvement of basic urban infrastructure networks (water, drainage and electrification) street lighting, paving, curbs, sidewalks and wheelchair ramps; the construction or improvement of roads, neighborhood gardens and community sports fields, installation or strengthening systems for trash collection, water sanitation and the improvement and equipping of community development. The second most important is 'Social and Community Development', which consists of job trainings and gender violence workshops, the development and updating of community development plans, and social service provision for students secondary and higher education. Table 1 provides the breakdown of how the money was spent in the 155 treatment polygons studied in this paper.

3.2. Sample Selection and Survey.

The polygons included in the study were required to satisfy several eligibility restrictions. The original sample of polygons was provided by Hábitat, and satisfied their standard conditions for inclusion. The sample was formed from a shortlist of 516 polygons (19,427 blocks) that SEDESOL considered viable for the study because they remained untreated as of early 2009. For the selection of the study sample, two additional restrictions were applied: we excluded the municipalities which had only had one polygon, and cities that had fewer than four polygons. With these criteria, the impact study was restricted to a total of 370 polygons. From this universe of polygons and the

⁶ 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.

randomized saturation design we selected 176 neighborhoods to serve as the treatment group and 194 for the comparison. Table 2 presents the distribution of polygons by state and treatment group.

This analysis is based primarily on two data collection instruments. First is a block-level survey completed by the surveyor based on the externally visible condition of the infrastructure on the streets in a block. The second is a household survey with two versions, one short and one long. 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 according to the following strategy. The block-level survey and the household survey was conducted in all the blocks in those polygons with 100 or fewer blocks. For polygons with more than 100 blocks (only 4.3% of total polygons) 100 blocks were randomly selected and each of them got up the same number of cards from apples and household surveys. Households were randomly assigned to answer the short versus the long version of the questionnaire (the questions on the short version are all included on the long), resulting in 6,419 long-form and 5,065 short-form questionnaires.

The study had a pre-analysis plan submitted to the Mexican government at the time the analysis of the baseline survey was being completed. The outcomes included in that pre-analysis plan include those evaluated here (core infrastructure impacts, private investment, and real estate impacts) as well as impacts on social capital, crime, health, transport time, and access to and use of community centers. An overview of the project is given by Ordoñez et al. (2013), and the impacts on pre-committed outcomes not presented here are given in Ordoñez and Ruiz (2013). The pre-analysis plan included the study of spillover effects using the specification given in Equation (2).

3.3. Attrition and Balance.

3.3.1. Attrition.

The analysis sample begins with 22,841 household surveys, of which 11,380 are in the baseline (R1) and 11,461 are in the followup (R2). There are three distinct factors causing the number of observations in the final analysis to differ from this number:

(1) The first factor determining the number of observations in the final analysis data is the population weights. The survey was designed to be weighted to be representative of the population of households in the study, and hence each block-level observation has a population weight attached to it (one each for the short-form and the long-form survey, which were sampled differently). There are 710 blocks sampled into the study that had no residents; surveyors were able to record some

basic data on the quality of infrastructure for these blocks, but they are dropped when we consider impact on residents, rather than blocks.

(2) The unit of analysis for the project is the block, and so a second step is to impose the condition that panel analysis will only be conducted on blocks that appear both in the baseline and the follow-up data. This restriction causes us to lose a further 23 observations in the baseline and 82 in the follow-up.

(3) After the intake sample had been defined and the baseline had been conducted, implementation problems arose with 5 of the original 65 municipalities in the sample (the municipal governments did not meet the matching requirements).⁷ 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 another 793 observations, almost 7% of the sample.

The final dataset used for analysis thus consists of 19,417 panel observations on populated blocks located in municipalities in which Hábitat was able to conduct treatment, and in which COLEF was able to conduct panel surveys. This panel of 9,702 blocks contains one household survey per block in the baseline and a few duplicate surveys within blocks in the follow-up, leaving us with 9,702 R1 observations and 9,715 R2 observations located in 342 polygons and 60 municipalities.

Table 3 tests for whether each potential type of attrition proves to be correlated with the treatment. In columns 1-2, we begin with the 10,670 sample of inhabited blocks, from (b) 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 (8.8%), there appears to be no systematic correlation between original treatment status and the polygons in which Hábitat was able to administer treatment. This is true whether we examine the simple treatment-control attrition differential (1) or the difference controlling for covariates (2). Columns 3-4 then examine the attrition driven by the success of COLEF field teams in conducting a panel block-level survey. Overall attrition at the block level was low (98.5% of the potential panel blocks were successfully tracked) and appears similarly to be balanced by treatment. The dummy on treatment is insignificant, indicating that the final analysis sample should be well representative of the original intake sample.

⁷ These five municipalities are Cuajimpala de Morelos, La Magdalena Contreras, Xochimilco, Almoloya de Juarez, and Ecatepec de Morelos.

This study considers infrastructural outcomes at the level of the structure, not the resident. For this reason we are less concerned with tracking the same households than with tracking the same houses. If at followup 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. When we look at the success of field teams at locating first the same household, and then the same house, 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% lower in the treatment than the control, although this difference is statistically insignificant. When we look at attrition at the household level, we see evidence of a very high rate of ‘churn’ of households during the three years of the study, with almost 40% of the houses having different residents in the control group. Most importantly, this rate is substantially lower in the treatment group, indicating that the Hábitat intervention decreased residential turnover by almost 9 percentage points. This means that the program decreased the rate of churn by about a quarter. In summary, the attrition of *houses* is balanced in the study, but the endline survey is more likely to return to structures with pre-existing *households* as a result of the treatment. We return to a discussion of this issue when analyzing changes in private investment in the housing stock.

3.3.2. Balance at the Neighborhood Level.

Table 4 examines the baseline treatment-control comparison within the final analysis sample. The overall balance of the experiment looks very good. There is some evidence (10% significant) that the treatment went into locations with worse access to lighting, but an overall index of infrastructural quality is higher in the treatment than the control. The other variable here displaying slight imbalance is the index of satisfaction with social conditions, indicating that the treatment fares slightly worse. In a broader analysis of balance (available upon request), baseline balance was checked using the entire battery of 65 variables designated in the pre-analysis plan as central. We find $5/65 = 7.6\%$ of the outcomes are unbalanced at the 5% level, very close to what we should observe by random chance. Overall, there do not appear to be ‘families’ of variables that are systematically different between treatment and control. One implication of this good overall balance is that the results prove very robust to control structure: a post-treatment single-difference,

a simple difference-in-differences, a DID using FE at the municipality level and a DID using FE at the block level give nearly identical results.

4. RESULTS.

4.1. The flypaper effect at the municipality level.

A natural starting point for the relationship between federal and local spending is an analysis of aggregate infrastructure budgets at the municipal level. We exploit the randomized saturations as an instrument for total federal spending through Hábitat to examine this relationship. Our study contains only 60 municipalities, and while every Hábitat-eligible slum neighborhood in the study municipalities was included in the study, these neighborhoods cover only an average of 3% of the population and 1% of the surface areas of study municipalities. Because the *poligono* is a definition of a neighborhood used only by Hábitat, municipal governments do not record spending at this level, and the most disaggregated data available to us is total annual spending by the municipality on ‘Public works and social action’, a category which includes public infrastructure, public safety, and economic development. This investment category maps relatively well to the activities in which Hábitat invests, but the randomized saturation experiment was conducted within a very small subset of the overall population of the municipalities. This implies that a 90% treatment saturation, while treating almost all of the Hábitat-eligible slums, may still generate a small budgeting spillover to the municipality as a whole.

The relationship that is most informative of the flypaper effect is to run:

$$\Delta MS_i = \alpha + \eta FS_i + \varepsilon_i,$$

where ΔMS_i is the change in municipal spending on infrastructure once the experiment has begun, and FS_i gives the federal spending in municipality i arising through the experiment. If FS_i were directly randomized, we could simply estimate the municipal response. An estimate of $\hat{\eta} = -1$ would imply perfect crowdout, $\hat{\eta} > -1$ would indicate some flypaper effect, $\hat{\eta} = 0$ indicates no strategic response, and $\hat{\eta} > 0$ indicates crowd-in.

The problem with this regression in practice is that municipal-level spending through Hábitat has an endogenous component arising from the fraction of neighborhoods in a municipality that are eligible for the program. To address this, we instrument for federal spending with the assigned saturations. Figure 3 illustrates the instrumentation strategy using two sets of scatterplots. The first plots the *true* population treatment saturation of a municipality against the *assigned* treatment

saturation that was directly randomized. There is substantial variation in the true saturations not explained by the assigned, arising mostly from the fact that poorer municipalities will have a larger share of their population eligible for Hábitat. The right-hand figure plots the key right-hand side measure of FS_i (namely federal Hábitat spending as a fraction of baseline municipal infrastructure budgets); the fitted line then represents variation that can be exploited by an instrumentation of federal spending with the randomized saturations. The conclusion from this graph is that while Hábitat spending is overall small relative to municipal infrastructure budgets, the saturations alone can explain a properly randomized fraction of spending that varies from 1% to 9% of baseline annual municipal budgets.

Figure 4 provides a visual take on the reduced-form variation in the IV regression, plotting the dependent variable of ‘% change in municipal spending, pre- to post-experiment’ against the treatment saturations. From inspection of this figure, Tijuana emerges as a major outlier in the regression. This municipality saw a huge increase in infrastructural spending related largely to long-planned main road construction, not an activity with which Hábitat is directly involved. Tijuana is also a high-saturation municipality in the study, meaning that its inclusion tends strongly to increase the apparent marginal effect of treatment saturations on spending changes. When we move to presenting the ‘flypaper’ results in Table 5, therefore, we show all results with and without Tijuana in the sample so that the reader can gauge the effect of this data point on the overall relationship.⁸

The first column in Table 5 validates the cross-municipality experiment, showing the variation in saturations to be orthogonal to baseline municipal expenditure levels. Columns 2 and 3 show the reduced-form impacts depicted in Figure 4; namely the percent change in municipal infrastructure spending by the assigned treatment saturation.⁹ Point estimates are positive but insignificant, indicating a mild crowd-in effect in spending, even when Tijuana is excluded. The next two columns present the simplest way in which to measure the magnitude of crowd-in, by regressing the peso value raw change in municipal expenditures on the peso value of Hábitat investment in the municipality. Because the latter is endogenous, we instrument for it using the assigned saturation as depicted in Figure 3. Here again we get positive and insignificant results, but the magnitude of the effect is worthy of comment: the result from column 5 indicates that for every peso spent in a

⁸ Tijuana is not an outlier in the other regressions in this paper that consider treatment versus control changes in infrastructure within slum neighborhoods only; its exclusion does not affect other results.

⁹ Note that here we do not weight regressions since the intent of Table 5 is to be representative for municipality decision-makers rather than for the inhabitants of municipalities. All variables used in this analysis enter directly at the municipal level and do not require the construction of weighted averages within municipalities.

municipality by Hábitat, local infrastructural spending rises by 5.6 pesos!¹⁰ While this regression on 58 observations is certainly not adequately powered to detect the difference between 0 (no strategic response) and -1 (perfect crowdout), we are almost able to reject that this coefficient is negative. The last two columns of Table 5 try to further remove municipal scale effects by regressing the percent change in municipal spending on the instrumented Hábitat investment as a percentage of baseline municipal spending. The column excluding Tijuana indicates that when Hábitat spending as a fraction of municipal spending increases by 1%, municipal spending increases by 2%.

Clearly, this analysis of flypaper effects in municipal infrastructure spending suffers from low power. Nonetheless, we can exploit access to an unusual source of randomized variation in the magnitude of federal investment within municipalities. Given that the average annual municipal budget is \$23 million, our IV-randomized federal expenditure varies from \$230,000 to \$2,140,000 per municipality. Using this variation in spending we are unable to reject perfect crowdout or a lack of strategic response, but we uncover marginal effects consistent with a relatively strong crowd-in effect of federal infrastructure spending on the expenditures of municipalities.

4.2. Neighborhood-level Impacts on Basic Infrastructure.

The most basic question for a program intended to improve infrastructural quality is its impact on the availability of core services at the household level. To measure this, we calculate neighborhood (polygon)-level averages of the 10 infrastructure variables indicated by our pre-analysis plan as the core indicators of the program. These neighborhood averages are then analyzed with a fixed-effects difference-in-differences regression (Equation (1) in Section 2). This regression is weighted by two sets of weights in order to make it representative of the average household in the study. The first is the standard survey weight, the number of households in the neighborhood (and neighborhood-level averages are calculated weighting by the number of households per block). Secondly we use ‘saturation weights’ (Baird et al. 2013) to undo the structural weighting caused by the randomized saturation design: by definition we observe more treated neighborhoods and fewer untreated neighborhoods as saturations increase. To make it so that each potential outcome is given equal weight across the saturation distribution, we must weight treatment observations by the product of the sampling weights and $.5*(1/saturation)$, and each control outcome by the sampling weights times $.5*(1/(1-saturation))$. Standard errors are clustered at the municipal level to reflect the component of the design effect that enters through the randomized saturation experiment.

¹⁰ One US dollar buys 13 Mexican pesos.

Table 6 presents these core infrastructure results. We first discuss 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% for sewerage to 98.9% for electricity). The treatment estimate on water is negative, and sewerage indicates an insignificant increase of almost 2%. Electric lighting improved by just less than a percentage point, significant at the 10% level. When we examine 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% at baseline, rising to 62.5% in the control at followup, but increased to almost 70% in the treatment. An index of basic infrastructure improves in the treatment at triple the rate of the control, significant at the 99% level. Reported satisfaction of residents with the quality of infrastructure improves, but not significantly.

In summary, for this project in which \$67 million was spent building infrastructure for 118,000 households in treatment-eligible slum neighborhoods (\$567 per household), significant improvements in the quality of infrastructure enjoyed by the average resident were achieved.

4.3. Spillover Effects and Causal Inference at the neighborhood level.

Section 2 discussed the problems with causal inference that arise when distinct layers of government provide complementary infrastructure spending and local governments may re-optimize around the experimental design. Before proceeding farther with the analysis, we therefore provide a set of empirical tests that exploit the randomized saturation design to measure the extent to which crowdout and spillover effects appear to be biasing our results. The ability to compare control polygons in heavily treated municipalities to control polygons in lightly treated municipalities gives a direct experimental evidence of saturation effects, and provides the ability to correct for them if they are found to be present.

We analyze these infrastructure construction spillovers using specification (2) from Section 2. Table 7 provides the results of the spillover regression using the core infrastructural outcomes as dependent variables. Table 7 reports the differential saturation slope in the treatment in the first row (μ_2), the saturation slope in the control in the second row (μ_1), the treatment effect at zero saturation in the third row (β), and the estimated intercept in the fourth row (δ^B), which gives the linearized projected outcome at zero saturation. For the purpose of visual comparison, at the

bottom of the table we also reproduce the average change estimated in the control group from the previous impact regression (δ^A) and the simple IIT estimate (γ), as well as the resulting estimates of the spillover to the control and the corrected IIT.

The core concern with spillover effects from the perspective of evaluation is to determine whether our counterfactual may have been polluted. To this end we focus first on the second row of results (μ_2), which give the extent to which control neighborhood outcomes are a function of municipal-level treatment saturations. These results provide little evidence of strong spillover effects of the program, although the majority of coefficients are negative. Consistent with a lack of strong spillover effects, the Treatment on the Uniquely Treated given by the Treatment * R2 interaction (which projects what the treatment effect would be at zero saturation, where no spillovers should occur) in general look similar to the simple IIT effects presented in the previous table. The sole significant control group saturation slope term (on road paving) is negative, and the sign of this term on the overall index and the satisfaction index are both negative. The general picture is therefore one of mild crowd-in, meaning that control outcomes become slightly worse as the intensity of treatment within a municipality rises.

The next step is to estimate the magnitude of the implied bias, and to provide spillover-corrected estimates of program effects following the methodology in Section 2. To this end, we are interested in the comparison between the counterfactual actually used in the impact regressions (the 'Round 2' dummy δ^A from Table 6, which gives the change in the control) and the desired counterfactual implied by linearization of the saturation effects as in Equation 2 (the zero-saturation intercept δ^B from Table 7). The actual minus the desired counterfactual gives a measure of the bias in the actual estimate of the IIT, and so we can then subtract this bias off the estimated IIT to get the 'corrected IIT', presented in the final row of Table 7. When we compare the corrected IIT at the bottom of Table 7 to the uncorrected IIT estimated in Table 6, the results suggest that several very large-footprint activities such as water and sewer installation had positive spillover effects on control neighborhoods. Consequently the control outcome is distorted upwards, and so the corrected IIT for these types of infrastructure is larger than the uncorrected. For several highly localized and easily divertable types of construction such as the installation of medians and sidewalks and the paving of roads, on the other hand, it appears that improvements in the treatment came *at the expense* of control neighborhoods, meaning that the counterfactual has been depressed and hence the naïve IIT is overestimated. When we correct for this, several of the very strong impacts of the

program disappear, and the index of infrastructure itself is barely significant.¹¹ Correction for spillover effects therefore increases the apparent impact of the program on large-footprint infrastructure, and decreases the apparent impact on more localized types of investment such as sidewalks, medians, and road paving. Nonetheless, the absolute magnitude of spillovers in this study is muted.

We can also relate the results of Table 6 back to the hypotheses related to inframarginality in Section 2. The presence of strong treatment effects rules out the possibility that the entire federal program was inframarginal, because clearly complete crowdout did not occur. To distinguish whether we have local inframarginality or extramarginality in the sample, we should examine the magnitudes of the saturation slope terms μ_1 and μ_2 . Not a single differential slope term μ_2 is significant, suggesting that the sample is not extramarginal: if the control was being taxed to meet the matching requirement in the treatment we should see a gap open up between treatment and control outcomes as the saturation increases, and we do not. Similarly, with local inframarginal constraints the control outcome should be increasing consistently with the treatment saturation by the net budget effect, and instead we see a range of signs and values across different types of infrastructure. The combination of strong positive overall treatment effects with saturation effects that are both muted and varied across types of infrastructure leads us to conclude that the data are most consistent with a lack of strategic response on the part of municipal governments. It does not appear that local infrastructure investment was re-optimized across space in a manner that violates causal inference in a systematic way.

4.4. 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 slum communities under-served by past infrastructural investments, and yet in which property rights are robust. Further, 84.4% of households in the baseline reported owning their own homes, and 74% own their homes outright (mortgage financing is difficult to obtain in Mexican slum neighborhoods even with clear property

¹¹ It is interesting to refer again to Figure 2 and to consider the fact that the Mathematica evaluation of Hábitat conducted in an environment in which they treated 100% of eligible neighborhoods found no treatment effect. The treatment-control differential in the infrastructure index is decreasing as the saturation increases, and it appears they would converge very close to a saturation of one. This implies that we only detected significant average ITTs in this study because the use of the RS design held saturations *down*.

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 8 provides evidence of complementarities between private and public investment.¹² Every measure of the private housing stock has improved, 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 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 4.1; in this case private investment appears to have outstripped the measurable improvements in public infrastructure. Home ownership rates in the treatment rise by 2%, 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 at the polygon level, and becomes significant at the 90% level when we analyze the data at the household level.

Returning to the analysis of attrition in Section 3.3.1, recall that the rate of residential relocation between the two waves of the survey fell from 39% in the control to 30% in the treatment. Could this lower rate of turnover itself be an explanation for the greater willingness to invest in houses, and the slight uptick in home ownership and mortgages? A simple way of testing for this is a mediation analysis, in which we control for the share of the houses in the neighborhood that changed owners between the two periods. Results available on request show that the significance level of the significant coefficients in Table 8 are almost unchanged by the inclusion of this variable, indicating that even *within* those who stayed and those who moved, the Hábitat investment induced a meaningful increase in private investment.

The last column of Table 8 shows the impact on rents for the 16% of households that do not own their own homes, and indicates a substantial 218 peso jump in monthly rents, a nearly 20% jump over the baseline control-group average rent of 1,159 pesos. Hence, Hábitat investment does

¹² Having found only weak spillover effects on the first-order infrastructural impacts, we now return to a more standard experimental analysis using only sampling weights (not saturation weights) and we do not provide full spillover-corrected ITT estimates for the remaining analysis. Unreported analysis confirms that spillover effects on the outcomes analyzed in Table 8 are also very small.

appear to have been successful both at crowding private investment into the housing stock in intervention neighborhoods and at increasing property values.

4.5. Impacts on Real Estate Values.

Significant increases in rents in treatment neighborhoods provide enticing evidence that rising public and private investment are translating into a meaningful capitalization of improvements in property values. Particularly because of the high rates of home ownership in Hábitat neighborhoods, the most natural way to examine this is through raw land 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 net benefits realized by 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 suggests ‘money left on the table’, and points to a friction in the political economy of infrastructure delivery.

Measurement of the improvement in property values, however, presents several empirical challenges. 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, not public investment. Value 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 over-estimates of the sale value of their own properties (Gonzales-Navarro & Quintana-Domeque, 2009). To overcome the first of these issues, we use sales prices only on empty lots that have no construction on them as of 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 un-built 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.

Of the 342 baseline polygons used in this analysis, just over 40% had any empty lots for sale at baseline. The average baseline lot 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. Attrition in the real estate analysis (meaning the selection 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. Hence, while the sample selection in this analysis is quite severe, there are no obvious signs that experimental inference on the sample will be invalid.

When we turn to the difference-in-differences impacts in Table 9, we see substantial improvements in prices being induced by the treatment. Relative to a baseline control value of 889 pesos per square meter and a real control group appreciation of 42.7 pesos between 2009 and 2012, the treatment effect of the program was an additional 70.8 pesos per square meter, meaning that the treatment group had almost triple the real rate of appreciation as the control. Figure 5 shows the 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. Perhaps the most meaningful way to put this number 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 total. If the marginal effect estimated above is applied to all inhabited property in the treatment polygons, the resulting increase in total value is 1.8 billion pesos, almost exactly two times the 888 million invested by all three levels of government in the program. The average residence would have enjoyed 15,191 pesos in appreciation from Hábitat investment during 2009 to 2012, while having had 7,525 pesos spent on it. Thus, every peso of public money invested in infrastructure improvement in a polygon yielded two pesos of improvement in the total privately-held value of land there.

4.6. Impacts on Political Behavior.

4.6.1. Attribution.

We motivated the multi-level budgeting game in a model of electoral competition; we now examine directly whether the flow of resources had an effect on political attribution and on voting. A first issue in the political economy of the program is the extent to which residents who had large investments made in their neighborhood are aware of the Hábitat program, and the extent to which they correctly attribute improvements in their local environment to the program. To examine this, we define a dummy for households that had heard of the program, and a dummy for households that had heard of any off a list of *other* organizations that might be working in the local neighborhood. We then construct the same variable for whether a household reports having benefitted from the program, first examining Hábitat and then examining all other programs. The results in Table 10 show quite clearly that while those in program areas are substantially more likely to have heard of Hábitat (19% in the treatment versus 12% in the control), a vanishingly small fraction of households report having benefitted directly from the program (0.8%) and this fraction is actually slightly lower in the treatment than in the control. This number stands in stark contrast to the beneficiary numbers provided by Hábitat itself, whereby they use GIS maps with ‘buffers’ around the locations of investments to suggest that 30.5% of household benefitted from street paving, 16.7% from CDCs, 8.3% from sewerage, 8.1% from sidewalks and medians, 6.5% from street lighting, and 5.4% from drinking water. In short, while the program is both creating real benefits and effectively ‘spreading the word’ as to its own existence, it appears to be generating no positive attribution effects for Hábitat itself as virtually no-one is aware of having benefitted directly from the program.

4.6.2. Voting in the 2012 Elections.

To test the causal relationship between infrastructure spending and political behavior, we examine voting behavior in the 2012 presidential elections. These elections were conveniently timed to occur exactly as the randomized phase was drawing to a close. Data on municipal elections in Mexico is decentralized, and so it is difficult to test voting impacts in local elections. Since Hábitat is a federal program, we examine the extent to which the incumbent National Action Party (PAN) reaped rewards at the ballot box at the end of the experimental phase of the program. While the 2012 election saw the PAN lose to the once-dominant Institutional Revolutionary Party (PRI), we can use fine-grained electoral data to test whether treatment regions display an elevated vote share

for the PAN. To the extent that voters in treatment polygons were attributing to the incumbent national party some of the benefits seen in terms of infrastructure, safety, and property value improvements, we would expect this to improve the vote share for the PAN party relative to other parties in treatment areas

In order to be able to conduct this analysis, it was necessary to map the Hábitat-defined polygons onto the ‘secciones’, which are the most disaggregated level at which the Federal Electoral Institute (IFE) provides shapefiles of electoral districts. IFE provides voting data all the way down to the precinct level, and so we first aggregated the electoral data to calculate vote shares for PAN and PRI at the seccion level, and we then overlaid GIS shapefiles of the Hábitat polygons with the electoral *secciones*. By calculating the fraction of each polygon that lies within each *seccion*, we can then calculate a weighted average of the vote shares in all of the relevant *secciones* to estimate what the vote share for each candidate was in each Hábitat polygon. Table 11 present the cross-sectional differences between treatment and control polygons in the PAN and PRI vote shares at the presidential (national), senatorial (state), and deputy (district) levels.

Consistent with the complete lack of ability to attribute benefits to Hábitat, there is no evidence that treatment polygons turned out to vote more strongly for the incumbent PAN party at the national level. The vote share for PAN in treatment polygons was 0.14% higher than in the control, but this result is far from significant and the increase in the vote share for the (ultimately victorious) PRI party was almost as large. At the senate and the deputy level the analysis similarly shows no evidence that the program has altered party affiliation at these levels.

5. CONCLUSION

This paper presents the results of a large experiment conducted by the Mexican federal government in the construction of infrastructure for slum neighborhoods. In this examination of the effects of \$65 million of spending spread across 118,000 treatment households, we find evidence that infrastructure investment is at sub-optimal levels in these areas. 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, and the total increase in the value of the property in intervention neighborhoods is twice the cost of the program. On the other hand, 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% of their budget on these items). Residents appear completely unable to attribute the improvements to

investment by H abitat, and do not reward the national party that oversaw H abitat at the polls. In short H abitat appears to have been a highly effective program fiscally and yet to have served absolutely no role as a piece of political patronage. This disconnect is most likely due to the poor ability to attribute infrastructure improvements to the correct government entity.

This study has paid particular attention to the possibility of spillover effects. We presented a framework with which to classify the possible responses of municipal governments to an experiment in federal spending, and exploit the randomized saturation design to analyze these responses in two different dimensions. The simple takeaway from this analysis is that there are no substantial spillover effects, and hence standard causal inference is unbiased. Nonetheless, an adjustment of the simple impact estimates for the bias uncovered by the saturation variation is interesting. The spillovers recovered vary substantially across different types of infrastructure, and ultimately appear to relate more to the unique spatial externality of the intervention than to a single underlying budgeting spillover. Large-footprint investments typically drive up outcomes in the control as saturations increase, suggesting positive spillovers and underestimated treatment effects. Granular and easily diverted investments in things like sidewalks create a negative spillover to the control, meaning that the adjusted treatment effects are smaller than the unadjusted.

A reasonable criticism of the spillover analysis in this paper would be that the restrictive eligibility for H abitat treatment makes the saturation experiment of limited power. Saturation variation that is very large in the study sample is quite muted in the total infrastructure budgets of the municipal governments, and hence we are fundamentally underpowered to detect spillover effects. This point is important to remember in interpreting our results, in that our point estimates are typically in line with an extramarginal story of crowd-in, and yet because we cannot reject zero we accept the null of ‘no strategic response’. A more powerful saturation experiment that delivered the same point estimates might conclude that the matching requirements were binding, and that there is a flypaper effect in municipal infrastructure budgets. On the other hand, the very conclusion that the study is too small to perturb outcomes in the control group indicates that SUTVA is satisfied in this program. From a causal inference perspective, then, the fact that the spillovers will be muted due to the small size of federal transfers relative to municipal budgets implies that the experiment is internally valid.

Then, there is the credibility of the very large increases in the value of private real estate we estimate to have been induced by the program. While this doubling of value appears very large, Cellini et al. (2010) find an increase of \$1.50 in the willingness to pay of homebuyers 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 US, 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 US, a figure of \$2 may not be unreasonable.

We have provided experimental evidence that investments in infrastructure in slum neighborhoods of Mexico have reaped real dividends. These results should bolster the argument that we not overlook large-scale spending on macro programs in the face of micro interventions with demonstrable impact. As a point of comparison, Mexico's cash transfer program *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 as well as a broader set of benefits on social trust and a decrease in crime, as documented in Ordoñez and Ruiz (2013). Improving infrastructure in underserved locations can deliver real social benefits as well as a substantial surge in household wealth.

REFERENCES

- Ashraf, N., D. Karlan, and W. Yin. 2006. "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines." *Quarterly Journal of Economics* 121(2), pp. 635-672.
- Baird, S., A. Bohren, C. McIntosh, and B. Özler. 2013. "Designing Experiments to Measure Spillover and Threshold Effects." *Working Paper*.
- Bannerjee, A., E. Duflo, R. Glennerster, and C. Kinnan. 2013. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." *MIT Department of Economics Working Paper* No. 13-09.
- Busso, Matias, J. Gregory, and P. Kline. 2013. "Assessing the Incidence and Efficiency of a Prominent Place Based Policy", *American Economic Review*, 103(2): 897-947.
- Campuzano, L., D. Levy, and A. Zamudio. 2007. "The Effects of Hábitat on Basic Infrastructure," *Mathematica Working Paper*.
- Casaburi, L., R. Glennerster, and T. Suri. 2013. "Rural Roads and Intermediated Trade: Regression Discontinuity Evidence from Sierra Leone." *Working Paper*.
- Cattaneo, M., S. Galiani, P. Gertler, S. Martinez, and R. Titunik. 2009. "Housing, Health, and Happiness." *American Economic Journal: Economic Policy*. 1:1, pp. 75-105.
- Cellini, S., F. Ferreira, and J. Rothstein. 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *Quarterly Journal of Economics* 125 (1): 215-261.
- Chase, R. 2002. "Supporting Communities in Transition: The Impact of the Armenian Social Fund Investment." *The World Bank Economic Review*, 16(2), pp. 219-240.
- Crepon, B., E. Duflo, M. Gurgand, R. Rathelot, and P. Zamora. 2011. "Do Labor Market Policies have a Displacement Effect? Evidence from a Clustered Randomized Experiment." *Working Paper*.
- Dahlberg, M., E. Mork, J. Rattso, H Agren. 2008. "Using a Discontinuous Grant Rule to Identify the Effect of Grants on Local Taxes and Spending." *Journal of Public Economics*, Vol 92, pp. 2320-2335.
- Dercon, S., D. Gilligan, J. Hoddinott, and T. Wodehanna. 2009. "The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages." *American Journal of Agricultural Economics* 91(4), pp. 1007-1021.
- Duflo, E., and R. Pande. 2007. "Dams." *Quarterly Journal of Economics* 122(2), pp. 601-646.

- Dupas, P. 2009. "What Matters (And What Does Not) in Households' Decision to Invest in Malaria Prevention?" *The American Economic Review Papers and Proceedings*, 99(2), pp. 224-230.
- Fiszbein, A., and N. Schady. 2009. "Conditional Cash Transfers: Reducing Present and Future Poverty." The World Bank, Washington DC.
- Galiani, S., P. Gertler, and E. Schargrodsky. 2010. "Water for Life: The Impact of the Privatization of Water Services on Child Mortality." *Journal of Political Economy*, 113(1), pp. 83-120.
- Galiani, S., M. Gonzales-Rozada, E. Schargrodsky. 2009. "Water Expansions in Shantytowns: Health and Savings." *Economica*, 76(304), pp. 607-622.
- Galiani, S., and E. Schargrodsky. 2010. "Property Rights for the Poor: Effects of Land Titling." *Journal of Public Economics*, 94(9), pp. 700-729.
- Gentili, U. 2007. "Cash and Food Transfers: A Primer." World Food Program, Rome Italy.
- Gine, X., and G. Mansuri. 2011. "Together We Will: Experimental Evidence on Female Voting Behavior in Pakistan." *World Bank Policy research Working Paper No. 5692*.
- Gonzales-Navarro, M., and C. Quintana-Domeque. 2009. "The Reliability of Self-Reported Home Values in a Developing Country Context." *Journal of Housing Economics*, Vol. 18, pp. 311-324.
- Gonzales-Navarro, M., and C. Quintana-Domeque. 2012. "On the Returns to Infrastructure for the Urban Poor and Politicians: Evidence from a Street Pavement Experiment." *Working Paper*.
- Gordon, N. 2004. "Do Federal Grants Boost School Spending? Evidence from Title 1." *Journal of Public Economics* 88, pp. 1771-1792.
- Hines, J., and R. Thaler. 1995. "Anomalies: The Flypaper Effect." *The Journal of Economic Perspectives*, Vol. 9 No. 4, pp. 217-226.
- Khandker, S., Z. Bakht, and G. Koolwal. 2009. "The Poverty Impact of Rural Roads: Evidence from Bangladesh." *Economic Development and Cultural Change*, 57(4), pp. 685-722.
- Kremer, M., J. Leino, E. Miguel, and A. Peterson-Zwane. 2011. "Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions." *Quarterly Journal of Economics*, Vol. 126, pp. 145-205.
- McConnell, M., B. Sinclair, and D. Green. 2010. "Detecting Social Networks: Design and Analysis of Multilevel Experiments." *Working Paper*.
- Meeks, R. 2012. "Water Works: The Economic Impact of Water Infrastructure." *Harvard Environmental Economics Program Discussion Paper* 12-35.
- Moffit, R. 1989. "Estimating the Value of an In-Kind Transfer: The Case of Food Stamps." *Econometrica*, 57(2), pp. 385-409.

- Newman, J., L. Rawlings, and P. Gertler. 1994. "Using Randomized Control Designs in Evaluating Social Sector Programs in Developing Countries." *The World Bank Research Observer*, 9(2), pp. 181-201.
- Newman, J., M. Pradhan, L. Rawlings, G. Ridder, R. Coa and J.L. Evia. 2002. "An Impact Evaluation of Education, Health, and Water Supply Investments by the Bolivian Social Investment Fund." *The World Bank Economic Review*, 16(2), pp. 241-274.
- Nesbit, T., and S. Kreft. 2009. "Federal Grants, Earmarked Revenues, and Budget Crowd-out: State Highway Funding." *Public Budgeting and Finance*, Summer.
- Ordoñez, G., and W. Ruiz (2013). R., T. Alegria, C McIntosh, and G. Ordoñez. 2013. "Formación de Capital Social Comunitario a Partir de Programas Orientados a combatir la Pobreza en México: El Impacto de Hábitat." *Working Paper*.
- Ordoñez, G., T. Alegria, C. McIntosh, and R. Zenteno. 2013. "Alcances e Impactos del Programa Hábitat en Comunidades Pobres Urbanas de México." *Working Paper*.
- Paxson, C., and N. Schady. 2002. "The Allocation and Impact of Social Funds: Spending on School Infrastructure in Peru." *The World Bank Economic Review*, 16(2), pp 297-319.
- Sachs, J. 2005. "Can Extreme Poverty be Eliminated?" *Scientific American*, September, pp. 56-65.
- Skoufias, E., and S. Parker. 2001. "Conditional Cash Transfers and their Impact on Child Work and Schooling: Evidence from the Progresa Program in Mexico." *Economia*, 2(1) 2001.
- Suárez-Serrato, JC, and P. Wingender. 2011. "Estimating the Incidence of Government Spending." *Working Paper*.
- Tarozzi, A., A. Mahajan, B. Blackburn, D. Kopf, L. Krishnan, and J. Yoong. 2011. "Micro-Loans, Insecticide-Treated Bednets and Malaria: Evidence from a Randomized Controlled Trial in Orissa (India). *Working Paper*.

TABLES.

Table 1. Hábitat Expenditures by Activity.

Total Investments in Treatment Polygons, 2009-2011 (Pesos)

Name of Program (Subprogram)	2009-2011				Households Benefitted
	Total Investment	Federal	State	Municipal	
Social and Community Development	182,667,827	92,185,422	4,894,453	85,587,952	256,443
Improvement of Urban Environment:	704,928,229	345,835,448	65,315,350	273,654,669	169,607
Paving	430,993,592	208,677,463	47,058,449	160,669,929	43,054
Sewers	63,996,222	32,345,029	3,954,345	26,867,468	7,672
Drinking water	34,691,248	17,326,549	1,231,531	15,227,487	5,071
Community Development Centers	37,298,216	18,332,709	2,600,280	15,226,006	17,536
Sidewalks and medians	32,274,345	17,062,770	3,414,553	10,894,065	4,447
Public lighting	22,998,836	11,560,567	396,857	10,817,518	5,327
Trash collection	23,636,729	12,014,004	837,954	10,074,326	72,370
Total spending	888,801,056	438,623,370	70,381,053	359,673,871	428,590

Source: SEDESOL

Table 2. Locations of Hábitat Projects.

	Control	Treatment	Total
Baja California	6	14	20
Campeche	3	1	4
Chiapas	2	1	3
Chihuahua	6	5	11
Coahuila	3	3	6
Distrito Federal	16	20	36
Guanajuato	7	13	20
Guerrero	9	7	16
Jalisco	14	10	24
México	46	40	86
Michoacán	6	13	19
Morelos	4	3	7
Nuevo León	4	3	7
Puebla	24	15	39
Quintana Roo	12	1	13
Sinaloa	6	7	13
Sonora	6	1	7
Tamaulipas	8	12	20
Veracruz	9	5	14
Yucatán	3	2	5
Total	194	176	370

Table 3. Attrition.

Attrition Between Rounds 1 and 2:								
	Attrition at municipal level (municipality selected to be part of study but removed by Habitat)		Attrition at block level (block sampled at baseline and in study municipalities, but panel dependant variable not observed)		Attrition at House level (baseline sampled house replaced with alternate at followup)		Attrition at Household level (baseline sampled household replaced with alternate at followup)	
Baseline values of:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	-0.014 (0.048)	-0.00996 (0.049)	-0.00412 (0.005)	0.000965 (0.001)	-0.0363 (0.030)	-0.0297 (0.028)	-0.0874** (0.042)	-0.0737** (0.037)
Index of Basic Services		-0.0000749 (0.001)		-0.0000254 (0.000)		0.000614 (0.000)		0.000984 (0.001)
Satisfaction with Social Infrastructure		0.00166 (0.008)		-0.000922* (0.000)		-0.0029 (0.005)		0.00786 (0.006)
Observation Weight		0.0000488 (0.000)		-1.22e-05*** (0.000)		0.000260* (0.000)		0.000479** (0.000)
Average fraction attrited in control group:	0.095		0.016		0.176		0.388	
# of Obs:	10,670	10,436	9,922	9,745	9,702	9,702	9,702	9,702

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. Stars indicate significance at * 90%, ** 95%, and *** 99%.

Table 4. Balance.

Balance Tests.

	Average in Control Group	Treatment/ Control Differential	Standard Error of Difference	# Households at Baseline
Piped Water	0.926	-0.0163	(0.016)	9,702
Sewerage Service	0.829	-0.011	(0.027)	9,702
Electric Lighting	0.989	-0.00884*	(0.005)	9,702
Use Water to Bathe	0.287	0.00801	(0.028)	8,649
Flush Toilet	0.613	-0.0325	(0.023)	9,563
Diarrea in Past 12 Months	0.178	-0.0169	(0.016)	9,702
Street Lighting Always Works	0.555	-0.0255	(0.021)	9,702
Street is Paved	0.664	0.0172	(0.025)	9,702
Index of Basic Services	91.492	-1.204	(1.303)	9,702
Index of Basic Infrastructure	68.512	1.097	(2.057)	9,702
Availability of Services + Infrastructure	78.361	0.111	(1.506)	9,702
Satisfaction with Physical Environment	2.844	-0.106	(0.084)	9,702
Satisfaction with Social Environment	83.308	-2.320*	(1.215)	9,702
Knowledge of Public Programs	39.358	-1.322	(0.965)	9,702

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. Stars indicate significance at * 90%, ** 95%, and *** 99%.

Table 5. The Flypaper Effect in Total Municipal Infrastructure Spending

	Municipal Spending on Public Works & Social Action, 2007-09 versus 2010-11						
	Baseline Municipal Spending, millions of pesos	Percent change in annual Municipal Spending		Raw change in Annual Municipal Spending, millions of pesos		Percent change in annual Municipal Spending	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Assigned Municipal Treatment Saturation (0-1)	3.302 (126.6)	56.27 (36.8)	18.74 (29.7)				
Federal Spending, millions of pesos (instrumented with saturation)				28.76 (18.7)	5.615 (8.6)		
Federal Spending as % of baseline municipal spending (instrumented with saturation)						6.438 (4.3)	2.002 (3.0)
Constant (spending at zero saturation)	249.1*** (69.0)	5.161 (20.1)	17.61 (16.0)	-128.6 (131.1)	-3.993 (59.4)	3.752 (21.3)	17.58 (15.1)
# of Obs:	60 0	59 0	58 0	59	58 0	59	58 0
Instrumental Variables Regression	N	N	N	Y	Y	Y	Y
Includes Tijuana?	Y	Y	N	Y	N	Y	N

Municipal-level analysis, not weighted by population. Standard Errors in Parentheses, Stars indicate significance at * 90%, ** 95%, and *** 99%.

Table 6. Polygon-level Infrastructure Impacts.

		Piped Water	Sewerage Service	Electric Lighting	Street Lights	Medians	Sidewalks	Paved Roads	Trash Collection	Index of Basic Infrastructure	Satisfaction with Physical Infrastructure
Intention to Treat	(γ)	-0.0193 (0.020)	0.018 (0.030)	0.00900* (0.005)	0.0684** (0.030)	0.0997*** (0.033)	0.0702** (0.027)	0.0601*** (0.020)	-0.0000452 (0.017)	0.220*** (0.067)	0.694 (0.541)
Dummy for R2	(δ^A)	0.0225** (0.011)	0.0269* (0.015)	0.00404 (0.003)	-0.00714 (0.029)	0.0214 (0.019)	0.0362* (0.020)	0.0502*** (0.014)	0.0188 (0.013)	0.115** (0.056)	-0.16 (0.488)
Baseline control mean:		0.926	0.829	0.989	0.555	0.588	0.589	0.664	0.971	2.740	8.825
Observations		684	684	684	682	684	684	684	684	684	684
R-squared		0.024	0.051	0.036	0.029	0.14	0.13	0.209	0.049	0.18	0.02
Number of polygons		342	342	342	342	342	342	342	342	342	342

Polygon-level analysis with polygon fixed effects and standard errors clustered at the municipal level. Regressions weighted by saturation weights and by polygon populations to make them representative of all inhabitants of study areas. Standard Errors in Parentheses, Stars indicate significance at * 90%, ** 95%, and *** 99%.

Table 7. Spillover Effect Estimates.

		Piped Water	Sewerage Service	Electric Lighting	Street Lights	Medians	Sidewalks	Paved Roads	Trash Collection	Index of Basic Infrastructure	Satisfaction with Physical Infrastructure
Municipal Saturation * Treatment * R2 (differential saturation slope term in treatment)	(μ_2)	-0.0172 (0.063)	-0.0303 (0.136)	-0.0164 (0.020)	-0.0353 (0.126)	-0.168 (0.155)	-0.0585 (0.122)	0.0384 (0.105)	-0.126 (0.076)	-0.107 (0.328)	1.158 (2.255)
Municipal Saturation * R2 (saturation slope term in control)	(μ_1)	0.081 (0.050)	0.0236 (0.059)	-0.0149 (0.013)	-0.00921 (0.125)	-0.0479 (0.084)	-0.0475 (0.092)	-0.121** (0.055)	0.0856 (0.072)	-0.245 (0.232)	-0.304 (1.724)
Treatment * R2: (treatment effect at 0 saturation)	(β)	-0.00989 (0.034)	0.0322 (0.081)	0.0161 (0.013)	0.0842 (0.070)	0.174** (0.086)	0.0956 (0.058)	0.0403 (0.058)	0.0587* (0.031)	0.264* (0.155)	0.166 (0.991)
R2 (trend in control)	(δ^B)	-0.0157 (0.017)	0.0157 (0.022)	0.0110* (0.006)	-0.0028 (0.064)	0.0439 (0.038)	0.0586* (0.034)	0.107*** (0.028)	-0.0215 (0.023)	0.230** (0.099)	-0.0165 (0.627)
# of Obs:		684	684	684	682	684	684	684	684	684	684
Baseline control mean:		0.926	0.829	0.989	0.555	0.588	0.589	0.664	0.971	2.740	8.825
Simple ITT	(γ)	-0.0193 (0.020)	0.018 (0.030)	0.00900* (0.005)	0.0684** (0.030)	0.0997*** (0.033)	0.0702** (0.027)	0.0601*** (0.020)	-0.0000452 (0.017)	0.220*** (0.067)	0.694 (0.541)
SE of Simple ITT											
Average Control Change:	(δ^A)	0.0225**	0.0269*	0.00404	-0.00714	0.0214	0.0362*	0.0502***	0.0188	0.115**	-0.16
Average Estimated Spillover in Control	$(\delta^A - \delta^B)$	0.0382	0.0112	-0.00696	-0.00434	-0.0225	-0.0224	-0.0568	0.0403	-0.115	-0.1435
Corrected ITT	$\gamma + (\delta^A - \delta^B)$	0.019	0.029	0.002	0.064**	0.077**	0.0478*	0.003	0.0402548**	0.105*	0.551

Polygon-level analysis with polygon fixed effects and standard errors clustered at the municipal level. Regressions weighted by saturation weights and by polygon populations to make them representative of all inhabitants of study areas. Standard Errors in Parentheses, Stars indicate significance at * 90%, ** 95%, and *** 99%.

Table 8. Private Housing Investment.
Private Investment, analysis at Polygon level.

	Brick Walls	Concrete Floors	Separate Kitchen	Separate Bathroom	Flush Toilet	Septic System	Piped Water	Home Owner	Private Bank Mortgage	Monthly Rent (for renters only)
treat_r2	0.00337 (0.008)	0.0229** (0.009)	0.0112 (0.016)	0.00416 (0.014)	0.0707** (0.031)	-0.0273** (0.013)	0.0146 (0.023)	0.0208 (0.022)	0.00962 (0.006)	218.8* (127.200)
r2	0.00763 (0.006)	0.00743 (0.005)	0.0307** (0.012)	0.0184** (0.008)	-0.0475* (0.025)	0.00162 (0.012)	0.0628*** (0.015)	0.00557 (0.009)	-0.0134** (0.005)	-8.751 (94.150)
Baseline control mean:	0.942	0.965	0.876	0.930	0.608	0.113	0.703	0.844	0.019	1159.8
Observations	684	684	684	684	684	684	684	684	683	530
R-squared	0.012	0.065	0.089	0.033	0.037	0.014	0.105	0.019	0.034	0.047
Number of polygons	342	342	342	342	342	342	342	342	342	299

Polygon-level analysis with polygon fixed effects and standard errors clustered at the municipal level. Regressions weighted by polygon populations to make them representative of all inhabitants of study areas. Standard Errors in Parentheses, Stars indicate significance at * 90%, ** 95%, and *** 99%.

Table 9. Real Estate Value Impacts.

	<u>Attrition:</u>	<u>Baseline Balance:</u>		<u>Impact:</u>		
	Polygon has Observation on Prices	Baseline Price, Simple Weighted Difference	Baseline Price, Weighted Difference + Municipality FE	Unweighted DID	DID Weighting by number of viviendas per polygon	DID with Weighting + Municipality Fixed Effects
Treatment	-0.0986 (0.06)	31.61 (194.60)	42.07 (140.60)	85.59*** (30.30)	107.5** (49.71)	75.52** (37.68)
Baseline Index of Services	0.0556 (0.081)					
Baseline Index of Infrastructure	-0.0786** (0.035)					
Total # of Residences	6.07e-05*** (0.000)					
Constant	0.474** (0.20)	1,130*** (110.70)	580.8*** (52.80)	39.59* (20.14)	27.66 (31.10)	-403.5*** (70.54)
Observations	342	138	138	138	138	138
R-squared	0.216	0.001	0.788	0.055	0.080	0.637

Analysis of unbuilt lots as polygon averages; dependent variable is price as assessed by professionals from IDAABIN. Standard Errors in parentheses; stars indicate significance at * 90%, ** 95%, and *** 99%.

Table 10. Attribution.

	Heard of Habitat	Heard of non- Habitat programs	Benefited from Habitat	Benefitted from non-Habitat programs
Treatment * R2	0.0765*** (0.027)	0.244 (0.497)	-0.000558 (0.004)	-0.00892 (0.071)
Treat	-0.0267 (0.017)	-0.281 (0.316)	0.000881 (0.002)	-0.0567 (0.053)
R2	-0.0994*** (0.022)	-1.641*** (0.452)	0.00309 (0.003)	-0.0281 (0.052)
R2 Mean in Control:	0.115	7.602	0.008	0.753
# of Obs:	19,417	19,417	19,417	14,394

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. Stars indicate significance at * 90%, ** 95%, and *** 99%.

Table 11. Impacts on Voting Behavior.

	Presidential (National) Election		Senatorial (State) Election		Diputado (District) Election	
	Share voting for PAN	Share voting for PRI	Share voting for PAN	Share voting for PRI	Share voting for PAN	Share voting for PRI
Treatment Effect	0.00135 (0.005)	0.000976 (0.005)	-0.00106 (0.005)	0.000322 (0.005)	-0.0016 (0.005)	0.00344 (0.006)
Mean in Control group	0.217	0.280	0.229	0.294	0.228	0.294
Observations	341	341	341	341	341	341
R-squared	0.876	0.82	0.886	0.832	0.878	0.832

Analysis at the polygon level, using weighted averages of the vote outcomes from the polling precincts that overlap with treatment polygons. Regression is a simple difference in the 2012 vote shares. Standard Errors in Parentheses, stars indicate significance at * 90%, ** 95%, and *** 99%.

FIGURES.

Figure 1. The Randomized Saturation Research Design.

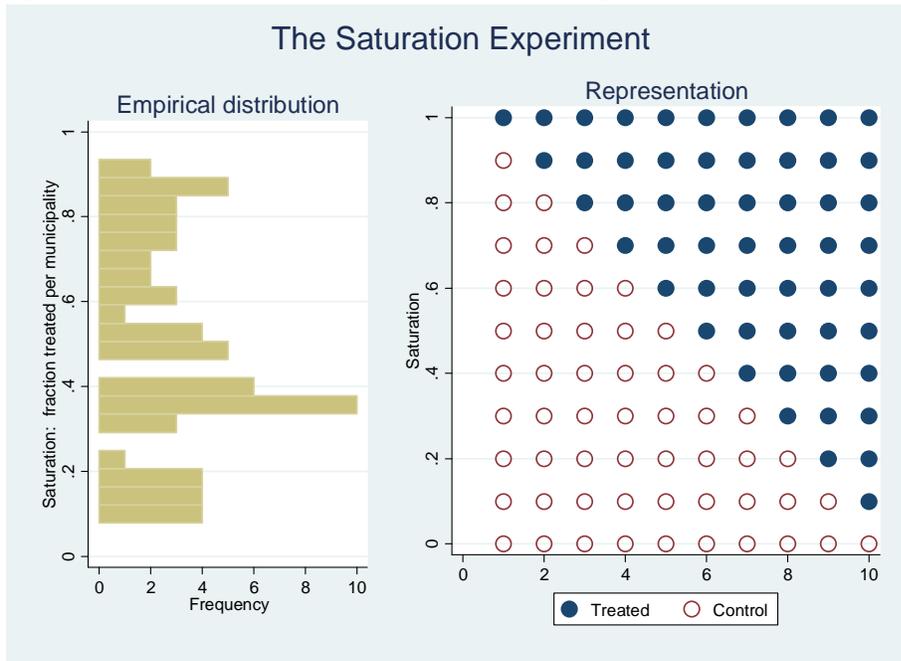


Figure 2. Correction of the ITT for Spillover Effects.

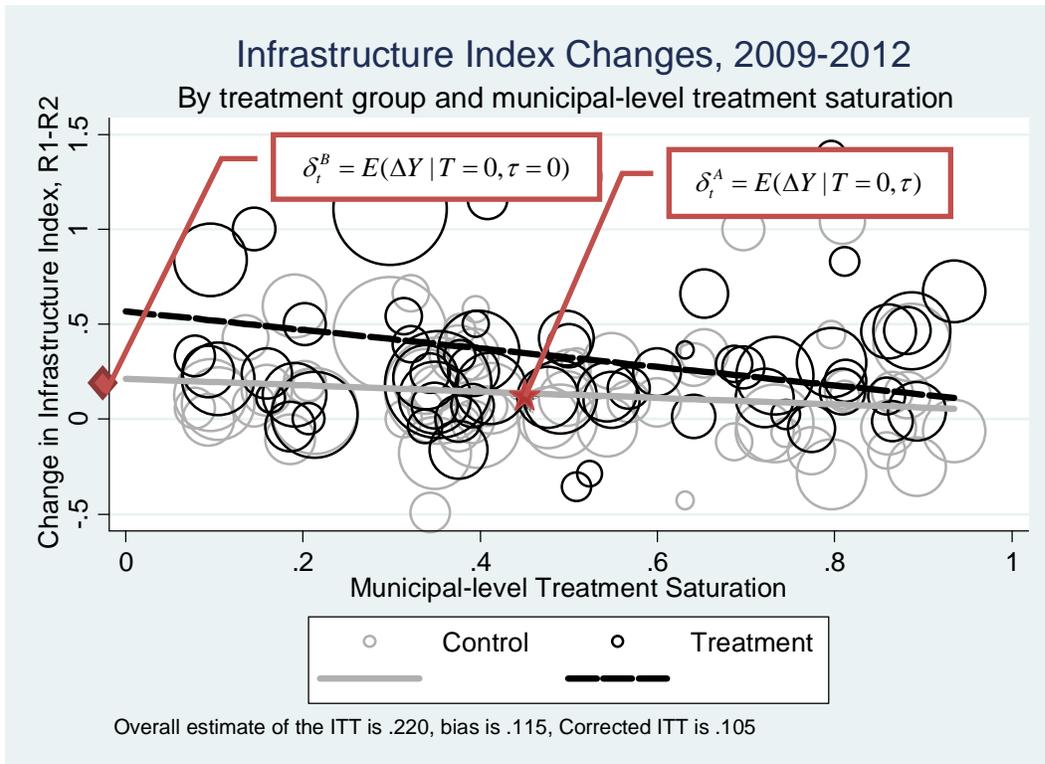


Figure 3. Does Aggregate Municipal Infrastructure Spending Change with Federal Spending?

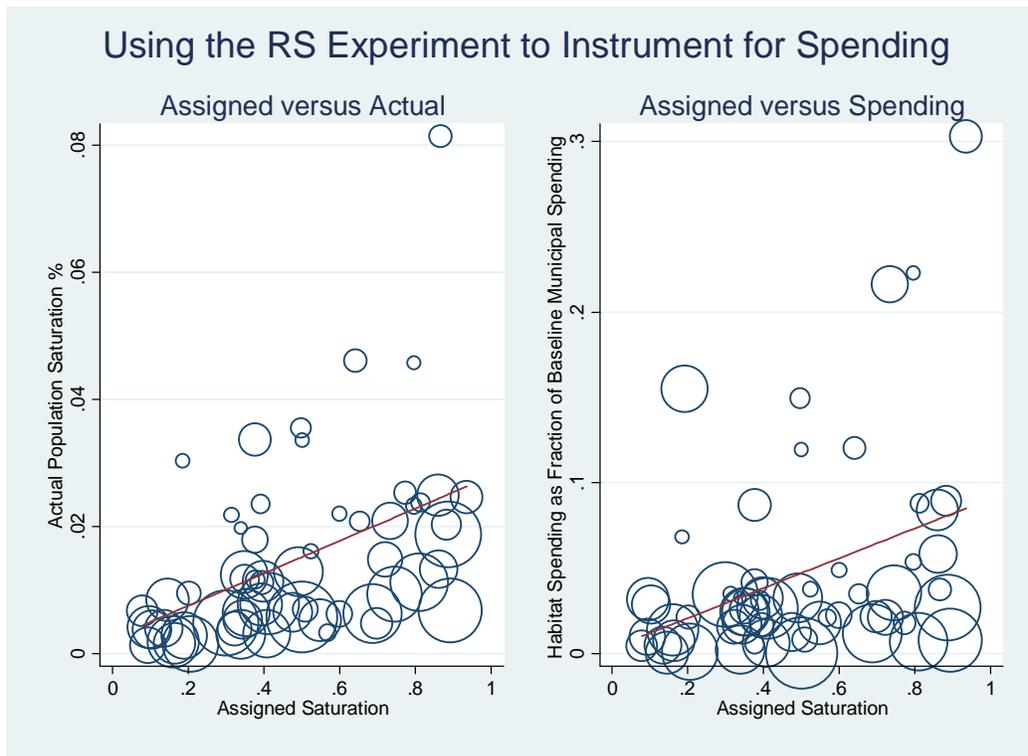


Figure 4. Does Aggregate Municipal Infrastructure Spending Change with Federal Spending?

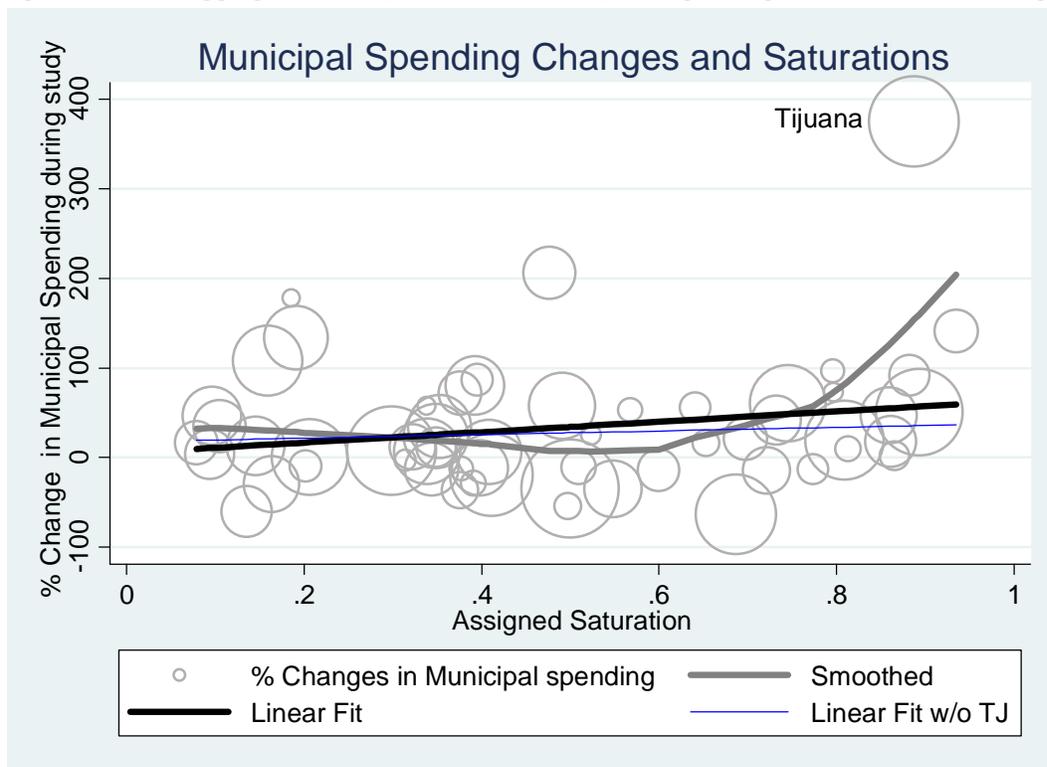


Figure 5. Cumulative Distribution Functions of Property Price Changes.

