# Public Housing Spillovers in a Developing Country*

Benjamin Bradlow
*Brown University*

Stefano Polloni†
*Brown University*

William Violette
*Federal Trade Commission*

This Version: October 4, 2018

**Abstract**

Does subsidized housing improve neighborhood quality? We estimate economic spillovers from a large public housing program in South Africa using geocoded deeds records, building growth, and census data. To identify project impacts, we compare constructed and unconstructed projects at fine geographic levels, in a differences-in-differences design. Within project footprints, housing quality improves and subsidized formal structures successfully crowd-out slums. Yet just outside of project footprints, we detect no change in housing quantity, quality, or price.

*We are grateful to Matthew Turner, Jesse Shapiro, John Friedman, Andrew Foster, Brian Knight, and Daniel Bjorkegren for their feedback, advice and support. We also thank Adelaide Steedley and the Centre for Affordable Housing and Finance in Africa as well as Geo-Terra Image for generous research guidance and for providing excellent data. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the the Federal Trade Commission or its Commissioners.

†Department of Economics, Brown University, Box B, Providence, RI 02912 E-mail: stefano_polloni@brown.edu

# 1. Introduction

In developing countries, 30% of urban populations live in slums where households often suffer from high rates of crime, low access to infrastructure, insecure property rights, and unsanitary conditions (United Nations [2015]). These negative congestion externalities often combine to create poverty traps, preventing long-term economic development (Marx et al. [2013]). Governments have responded by replacing slums with new homes and moving slum dwellers to public housing projects. These policies are intended to provide not only direct health and economic benefits to recipients, but also greater incentives for neighbors to invest in their homes and communities, reducing negative externalities and steering communities away from poverty traps. At the same time, public housing can also attract nearby slum growth by improving access to bulk services like water and sanitation as well as providing undeveloped land within and nearby housing projects. In this way, public housing projects may ultimately exacerbate the same negative externalities that they were designed to remediate.

In this paper, we analyze the impacts of public housing on the development of surrounding neighborhoods. We study a large-scale housing program in South Africa, which has allocated over 4.3 million dwellings and houses over 13% of the total population (Department of Human Settlements [2012,2015]; GHS [2009-2013]). Already one of the largest housing programs in the developing world, this program continues to respond to large backlogs in housing demand with an even mix of upgrading slum areas with new houses as well as constructing stand-alone developments. This program is intended to not only serve as "a key strategy for poverty alleviation" for direct beneficiaries, but also generate community-wide benefits, "leveraging growth in the economy," "combating crime, promoting social cohesion and improving quality of life for the poor," and "utilizing housing as an instrument for the development of sustainable human settlements, in support of spatial restructuring" (Department of Human Settlements [2004]).

We combine administrative records for 68 completed housing projects with data on property transactions, demographics, and house construction to measure the local impacts of these projects.

We find evidence of greater access to services, improved home quality, and a greater share of formal housing stock within project areas but no effects on surrounding areas. Examining nearby housing transactions, we find no statistically significant effects of public housing on prices in the formal market. To interpret these findings, we provide a simple conceptual framework balancing the amenity effects of improved housing quality against the price and quantity effects of a housing supply shock. This framework helps shed light on the absence of spillover effects detected in our analysis.

To identify these effects, we use a differences-in-differences strategy leveraging both the exact timing of housing project construction as well as the precise geographic proximity of surrounding areas. Like prior studies in the US, the substantial uncertainty in project timing due to difficulties coordinating many stakeholders and sources of funding limits the extent to which local housing markets are able to anticipate these projects (Diamond and McQuade [2016]; Tissington [2011]). To address the potential endogenous placement of housing projects, we use 68 planned but unconstructed projects as counterfactuals and detect no impacts of these projects on local housing markets.

Our zero spillover estimates stand in contrast to a large literature in development that has found positive impacts of public housing on direct recipients. Relying on small-scale experimental designs, previous studies have linked public housing to improvements in employment outcomes (Franklin [2016]), self-reported wellbeing (Galiani et al. [2017]; Devoto et al. [2012]), and child health outcomes (Cattaneo et al. [2009]). Data limitations both in finding a large enough sample of housing projects and in identifying outcomes at a precise spatial scale have prevented previous studies from identifying spillover effects. Taken alongside these previous studies, our findings suggest that policymakers may be able to continue emphasizing direct benefits to recipients in designing future housing policy since spillover affects appear to be minimal.

We proceed by first providing background on the South African housing program in Section 2. In Section 3, we develop a conceptual framework of housing supply and housing externalities to help to interpret the results. Section 4 describes the data used to measure outcomes and details our approach to identifying housing projects while Section 5 provides descriptive evidence. We present spillover results for residential home prices in Section 8 and demographic outcomes in Section 6.1. Section 9 includes a discussion of our findings before providing some concluding thoughts.

## 2. Background: Where are houses built?

Between 2000 and 2010, subsidized housing efforts in South Africa have primarily focused on constructing and allocating single-story, two-room (30 to 40 square meter) dwellings to households in groups of 50 to 500 per project. Housing projects are primarily located on undeveloped state-owned land although in some cases, municipalities will work with private developers to purchase inexpensive, undeveloped private land for these projects. Finding undeveloped land often requires policymakers to locate these projects far from city centers and economic opportunities (Department of Human Settlements [2012,2015]). Often undeveloped land contains preexisting informal settlements, which the government tries to replace with new formal houses and ensure that preexisting residents benefit from the new project houses (Tissington [2011]).

Facing substantial housing demand, the Department of Human Settlements has continued to issue grants to provincial governments to keep the rate of yearly housing allocations roughly constant (Department of Human Settlements [2012,2015]). While the location and types of projects are determined by provincial and municipal governments, construction is subcontracted to private developers who also act as project managers assisting in the allocation of houses to beneficiaries (Durojaye et al. [2013]).

Since housing projects require coordination between many stakeholders, these projects often face unanticipated delays and cancellations due to labor and land procurement issues, difficulties gaining support from local government agencies, environmental impact assessments, and inadequate bulk infrastructure provision (Department of Human Settlements [2012,2015]). In one example, political disagreements with local stakeholders led to the abandonment of a large project near Johannesburg (Dlmini [2017]). In other cases, housing projects can be delayed for upwards of 10 years (DA-GPL [2017]).

### 2.1. Background: Who are the beneficiaries?

The National Department of Human Settlements issues guidelines for eligibility and maintains an official waiting list for eligible households for greenfield developments. Eligibility requires citizenship, no previous property ownership, being married or having financial dependents, and having a monthly household income below R3,500 (Durojaye et al. [2013]).[1] The share of households reporting at least one member on the waiting list has remained stable at over 13% from 2009 to 2013.[2] Before construction, each project is assigned beneficiaries in a first-come, first-served basis according to the waiting list in their province or municipality. For in-situ upgrading projects, previous inhabitants of informal settlements receive renovated houses while any remaining houses are allocated according to the housing waiting list.

In practice, these guidelines are only loosely followed. Recent reports point to cases of corruption in the allocation of houses while in some instances, housing projects are organized with the assistance of local community groups who ultimately select the beneficiaries (Durojaye et al. [2013]; Mathoho [2010]). Research suggests that beneficiaries are often selected over the course of project construction and sometimes even after construction has finished (Durojaye et al. [2013]).

Beneficiaries are expected to pay a small one-time payment in order to receive title for their houses. Guidelines also prevent beneficiaries from reselling their houses within their first 7 years of ownership. Despite these guidelines, only 82% of project houses are reported being still occupied by their original beneficiaries within five years of construction.[3] Anecdotal evidence suggests that project managers are aware of active secondary markets but have difficulty policing these transactions (Matsena [2018]).

[1] The Gauteng Province has implemented their own waiting list since 2008 in order to exert greater control over the allocation process.

[2] This figure is calculated from the General Household Surveys from 2009 to 2013

[3] This figure is calculated from the General Household Surveys from 2009 to 2013.

## 3. Conceptual Framework

We consider housing projects as generating positive shocks to housing quantity as well as quantity. Following a standard model of housing supply and demand, increases in formal housing quantity should (1) lower prices of formal housing nearby and (2) incentivize households to move into project areas, both leading to increases in population. At the same time, gains in housing quality not only benefit direct recipients, but also may have an amenity effect, increasing the desirability of living in neighboring areas (Diamond and McQuade [2016]). This amenity effect is likely to (1) increase house prices nearby and (2) drive increases in local population as well. Taken together, quantity and quality effects have ambiguous predictions for formal housing prices nearby but both predict increases in population both within and around housing project areas.

This simple framework ignores additional features specific to housing markets in developing countries. In the housing production function, there may be complementarities or substitution between project housing and informal housing. By zoning and servicing large swaths of land, housing projects may lower the cost of building new informal housing settlements both within and around these projects. At the same time, clearly defined property rights allocated by these housing projects may increase the cost of squatting, decreasing nearby informal settlements.

The level of informal settlement growth generated by these housing projects may have additional welfare consequences due to general externalities from slum growth (Marx et al. [2013]). Health effects, congested infrastructure, and higher crime combine to form negative amenities, potentially depressing nearby housing prices or lowering population densities in these areas. Although these specific mechanisms are difficult to precisely identify in this setting, these considerations motivate analyzing the differential impacts of housing projects for informal versus formal housing.

## 4. Data

Understanding the spillover impacts of public housing requires (1) outcome measures at high spatial and geographic resolutions as well as (2) a precise measure of the location, timing, and size of housing projects. We locate housing projects using administrative maps compiled by the Gauteng City Regional Observatory. These records include 192 planned housing projects as well as notes on their completion status.[4]

To measure formal housing market impacts, we analyze over 500,000 housing transactions from the South African National Deeds Office covering the universe of transactions for suburbs in the bottom 20% of the housing market between 2003 and 2011 in the Gauteng Province (including the Johannesburg metro area). The bottom 20% suburbs were selected relative to prices in 2003 and followed every year from 2003 to 2011. These data were provided by the Affordable Land and Housing Data Centre, which tracks affordable housing markets. These data include the price, exact location, plot size, buyer name, and seller name for each transaction. To isolate spillover effects, we focus on transactions occurring within 4 kilometers of a housing project. Finally, we exclude the top 1% of prices as well as prices below 2,500 Rand, which are likely composed of measurement error or exchanging of titles between family members.

Since it is unlikely that government deeds records capture informal housing markets, we also include a building census of all structures in the Gauteng Province in 2001 and 2011. Using a combination of high-resolution satellite imagery and local field teams, these data record the precise location of each structure, identifying structures within over 30 categories including formal and informal residential dwellings. Out of 3,817,840 structures, this building census includes 1,628,073 formal structures and 1,560,345 informal structures. These data serve as both outcome measures of informal housing development as well as additional measures of public housing construction.

For demographic and economic outcomes, we have access to the full census of population and housing for 2001 and 2011, which provides information on housing characteristics, population

[4]Since this data comes in the form of overlapping shapefiles, we use the intersection of the shapes excluding tiny shapes below 0.50 square kilometers.

Figure 1: Top-Five Sellers

| Seller Name | Observations |
|---|---|
| City Of Johannesburg Metropolitan Municipality | 29,087 |
| City Of Johannesburg | 27,672 |
| City Of Tshwane Metropolitan Municipality | 24,780 |
| Ekurhuleni Metropolitan Municipality | 21,758 |
| Gauteng Provincial Housing Advisory Board | 13,058 |
| Total Observations | 549,704 |

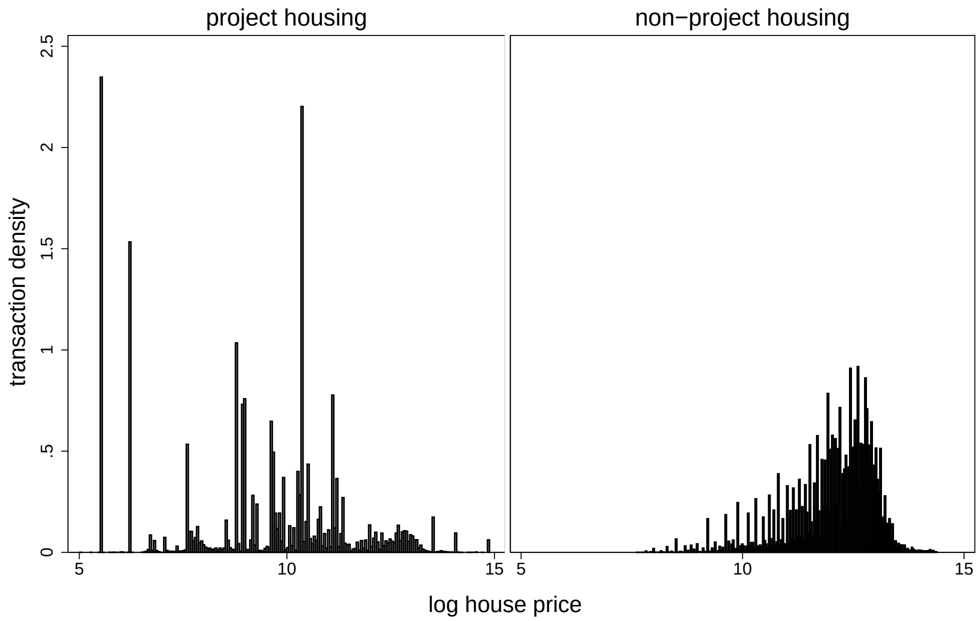
density, as well as income and employment measures. To spatially link these households in both samples to their corresponding housing projects, we conduct the analysis at the census block level, which is the smallest geography available with 28,437 total blocks in Gauteng Province and 20,457 with centroid within 4 kilometers of a housing project.

### 4.1. Constructed Housing Projects

We define projects as successfully constructed when we observe at least some recipients receiving deeds to their new government houses. Deed transfers are assumed to belong to a housing project if the seller name includes a government, municipality, or large developer. As recommended by the data provider, we exclude deeds flagged as large buildings used for commercial purposes ($<$2% of transactions) as well as purchase prices that are more than R50,000 above the yearly subsidy value ($<$4% of remaining transactions). Using this criteria, we are able to identify over 68 constructed housing projects out of 192 planned projects.

This approach may introduces measurement error insofar as we are mis-attributing deeds to housing projects. We use a limited definition of verified seller names to minimize the chances that unconstructed projects are misclassified as constructed projects. However, this approach increases the possibility that this definition mistakenly excludes constructed projects. Descriptive evidence supports our project definition in several ways. First, Figure 1 shows that the top five most frequently occurring seller names clearly correspond to government housing programs in the region. Second, transaction prices for project houses are strongly clustered around subsidy values. Fig-

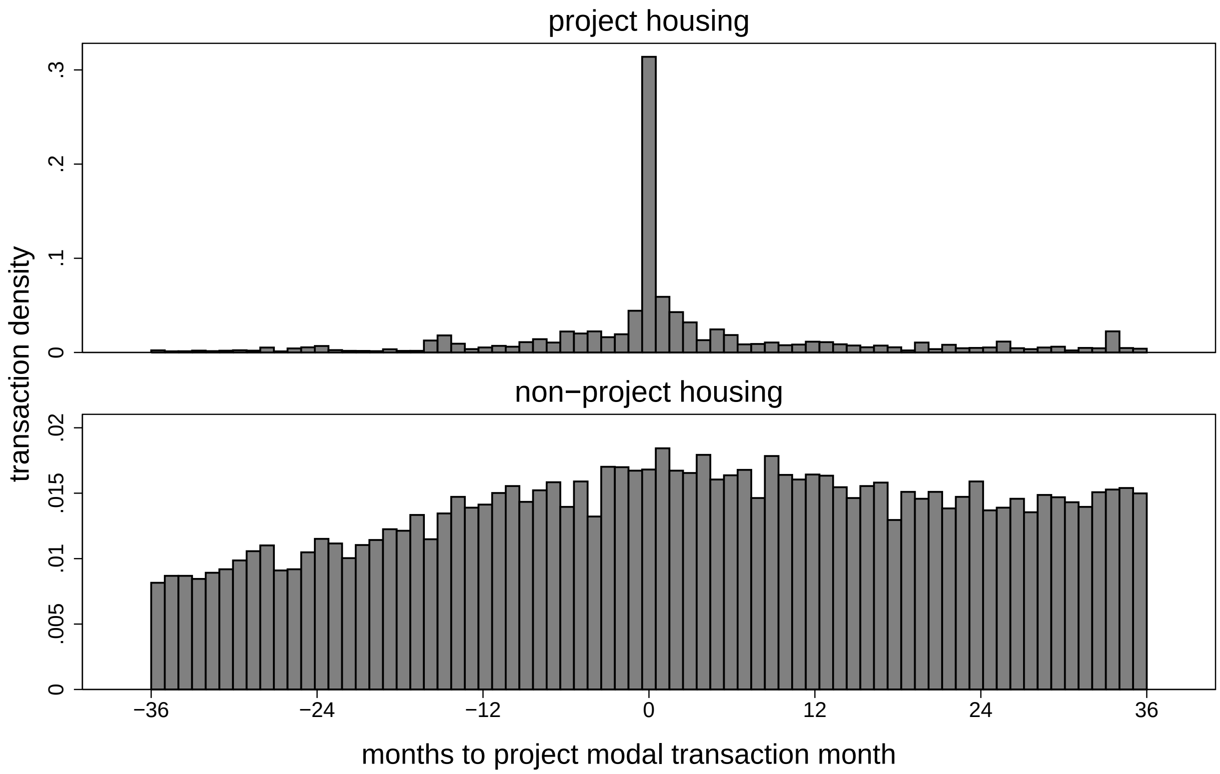
Figure 2: Transaction Price Histogram




ure 2 provides a histogram of deed prices under R100,000 for transactions that meet our project definition as well as those that do not. Government agencies often record either zero price or the value of the subsidy in the deeds, producing substantial bunching at these values compared to non-project transactions. At the same time, there exists a small density of projects smoothly distributed across the lower range of prices for project housing, which may be evidence of some misclassification of our project housing measure. We assign a completion date for each project according to the modal year of these deeds transactions. Figure 3 indicates the distribution of transaction dates for properties within a 4 km buffer around housing projects (above panel) and within the selected project areas (below panel). Project areas exhibit substantial bunching during a single month when projects were completed. There are also more transactions after the modal year than before the modal year, consistent with either a gradual roll-out for some project areas or immediate resale of projects houses after construction, which would be counter to housing regulations.

Evidence of similar bunching around the modal year for transactions coded as non-project transactions (above panel) would suggest that we may be miscoding project transactions as non-project transactions; instead, we find a smooth pattern relative to the modal year for these non-

Figure 3: Transaction Densities Relative to Modal Project Month




The upper panel includes all transactions within the final sample of clusters. The lower panel includes transactions within a 4 kilometer buffer around a project cluster relative to the modal year of that cluster.

project transactions. The slight increase in density around and just following the modal year may also be consistent with housing projects having an immediate impact on local housing markets, which we will explore further below.

### 4.2. Planned but Unconstructed Housing Projects

Identifying 68 constructed housing projects leaves 99 planned but possibly unconstructed housing projects in the administrative map. We propose using these projects as counterfactuals, capturing the level of the development that would have occurred in the absence of full construction.

In order to determine which of the unconstructed projects were planned to be completed between 2001 and 2011, we construct an estimated completion date using National Treasury budget reports. We were able to digitize data for 132 projects from budget reports spanning 2004 to 2009, which detail the name, start date, expected completion date, and cost of each housing project. We then use a fuzzy-string matching algorithm with bigrams to link project names from the budget reports to the administrative maps. We keep all matches with over 60% similarity with 44 projects matching exactly. Appendix 6 compares unmatched and matched projects finding that matched projects have much higher densities of informal housing and project housing density, but similar densities of formal housing and project dates. One reason may be that the budget reports only include larger, more expensive projects.

We find that the project start date (from the budget reports) is three years before the project completion date (from the deeds data) on average. In other words, beneficiaries receive title to their new houses about three years after the housing program is announced in the budget. Using this lag, we assign an expected completion date for unconstructed projects that is three years after the announced start-date in the budget reports. Since the budget data does not include month of planning, we assign a random month for unconstructed projects, which ensures that this date is uncorrelated with seasonality but also introduces measurement error in the timing of unconstructed projects. This error is only likely to affect the housing transaction estimates, which are the only estimates performed at the monthly level. This procedure results in a final sample of 65 unconstructed projects.

Table 1 tabulates project descriptions from the administrative data according to our definitions of completed and uncompleted. We find that constructed projects are much more likely to be classified as "completed," "planning," or in "implementation" while unconstructed projects are more likely to fall into "Uncertain" or "Investigating" categories. These correlations provide an external validation of our classification of constructed and unconstructed projects with deeds records.

## 5. Descriptive Evidence

To contextualize our analysis, we provide descriptive evidence of the location, size, and surrounding areas of housing projects. Figure 4 maps our sample of unconstructed and constructed housing projects. We find that projects fall exclusively within affordable housing areas (defined as enumer-

Table 1: Project Descriptions

| | Unconstructed | Constructed |
|---|---|---|
| Implementation, Completed | 5 | 18 |
| Planning | 12 | 8 |
| Future, Investigating, Proposed | 18 | 5 |
| Uncertain | 2 | 0 |
| No Description | 28 | 37 |

Descriptions from administrative project map.

Table 2: Housing Projects at Baseline

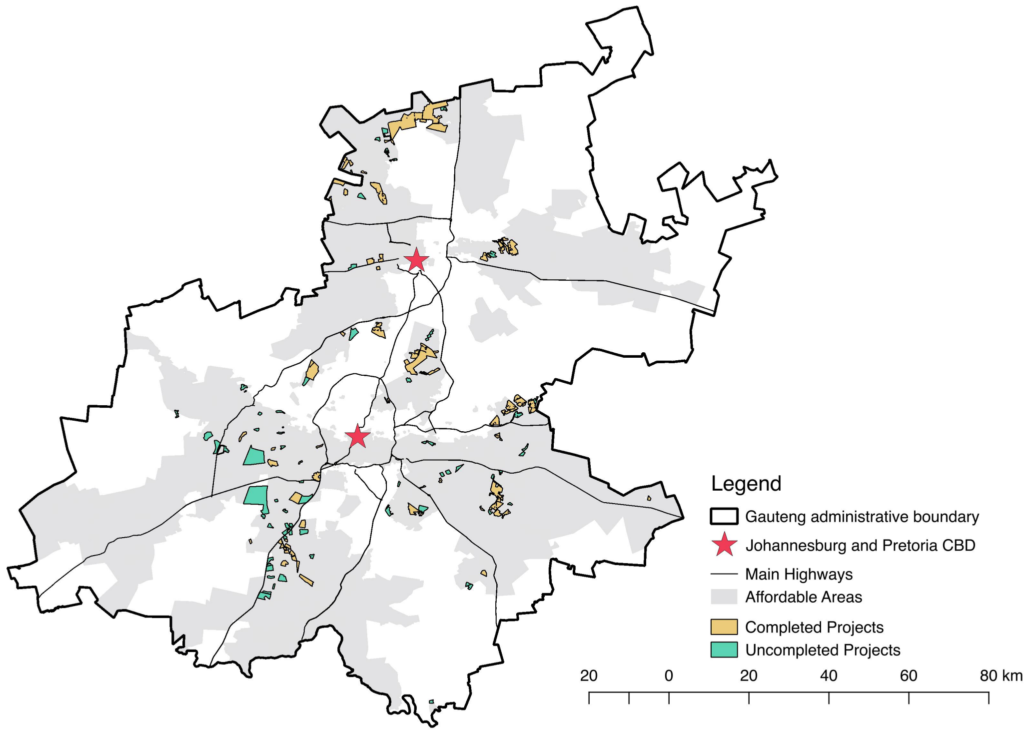
| | Unconstructed | Constructed | Within 4 km. | T-Stat: Cons. = Uncons. |
|---|---|---|---|---|
| Number of Projects | 65 | 68 | | |
| Area (km2) | 2.21 | 5.15 | | -1.61 |
| Project Houses (per km2) | | 508 | | |
| Informal Buildings (per km2) | 1,194 | 997 | 696 | 0.90 |
| Formal Buildings (per km2) | 355 | 428 | 836 | -0.63 |
| House Price (Rand) | 113,956 | 88,781 | 153,972 | 1.13 |
| Distance to CBD (km) | 33 | 30 | | 1.27 |

The third column includes areas outside but within 4 km of the project areas.
Central Business Districts (CBD) are measured with respect to Johannesburg and Tshwane.
T-stats are computed from project averages.

ator areas in the bottom quintile of house prices in 2003). While constructed and unconstructed projects are often adjacent to each other possibly indicating cases where authorities were unable to complete final phases of planned projects, there are also many examples of stand-alone projects of both types. Projects are generally located at relatively great distances from central business districts (CBDs), which suggests that vacant or inexpensive land plots are especially targeted by housing authorities. Despite their distance from the CBDs, housing projects are often next to arterial highways easing commuting costs for recipients.

Table 2 provides descriptive statistics at baseline for our sample of 68 constructed and 65 unconstructed projects. Constructed projects are larger in area than unconstructed projects likely because some unconstructed project represent smaller extensions to previously implemented projects. These differences are also exacerbated by a few very large constructed projects. Compared to unconstructed projects at baseline, constructed projects have smaller densities of informal housing

Figure 4: Housing Project Map




and slightly larger densities of formal housing, indicating better housing stock on average. At the same time, deeds records indicate somewhat higher house prices for homes in unconstructed areas. Both types of projects share similarly long distances to Central Business Districts consistent with authorities targeting inexpensive, vacant land for these projects. With an average of around 500 project houses per square kilometer, these housing programs have the potential to strongly impact the local housing stock for constructed project areas, in many cases doubling the existing stock of formal housing.

Areas within 4 km of housing projects have greater total building density and higher ratios of formal housing to informal housing compared to both the unconstructed and constructed projects. Similarly, housing prices are much higher for properties outside of project areas. These findings suggest that at similar distances away from central business districts, governments are still selecting less developed and less valuable land for housing projects. At least in terms of building composition, density, and value, unconstructed projects seem to provide a better comparison group

Table 3: Infrastructure and Demographics at Baseline

| | Unconstructed | Constructed | Within 4 km. | T-Stat: Cons. = Uncons. |
|---|---|---|---|---|
| Flush Toilet | 0.525 | 0.762 | 0.762 | -3.60 |
| Piped Water in Home | 0.312 | 0.477 | 0.486 | -2.75 |
| Household Size | 3.11 | 3.27 | 3.30 | -1.91 |
| Owns House | 0.378 | 0.495 | 0.514 | -2.62 |
| Single House | 0.424 | 0.631 | 0.618 | -3.57 |
| Number of Rooms | 3.10 | 3.27 | 3.67 | -0.98 |
| Pop. Density (per km2) | 6,882 | 8,510 | 6,772 | -1.56 |

Averages include the 15,454 census blocks that share over 30% area of overlap with projects areas. T-stats are computed from project averages.

for constructed projects than simply examining nearby areas.

Table 3 compares census measures at baseline across constructed and unconstructed projects as well as areas within 4 km of housing projects. We identify census block as pertaining to a housing project if over 30% of their land areas overlap, which results in 15,454 intersecting census blocks. We find that unconstructed project areas lag behind constructed project areas across all demographic measures. At the same time, constructed project areas are almost indistinguishable from nearby, non-project areas despite having fewer structures and greater population densities.

Taken together, these findings are consistent with governments (1) selecting low-cost land to implement projects and (2) facing high costs of finishing projects in dense slums. These types of selection are important to consider when interpreting our results and to help motivate the differences-in-differences empirical strategy detailed below.

## 6. Empirical Strategy and Results

To estimate the impacts of housing projects both within their project areas as well as just nearby, we use a differences-in-differences strategy comparing changes in outcomes for constructed project areas to unconstructed project areas. Given the varying spatial and temporal dimensions of the outcome measures, we use slightly different specifications of the differences-in-differences approach for each dataset.

### 6.1. Infrastructure and Demographic Effects

To estimate impacts for census measures, we implement the following estimating equation, which estimates a standard differences-in-differences model both within the project area and just around the project area:

$$y_{hpt} = \lambda_p + \sum_e I^e_{hpt}\left(\alpha^e D_t C_p + \beta^e D_t + \gamma^e C_p + \theta^e\right) + \varepsilon_{hpt}$$

$y_{hpt}$ measures the outcome for household $h$ living in the vicinity of project $p$, observed in census year $t$. $I^e_{hpt}$=1 if household $h$ is in the exposure area $e$ of project $p$. Exposure area $e$ includes (1) project areas defined as census blocks that overlap over 30% of their area with a housing project and (2) spillover areas defined as census blocks with fewer than 30% area of overlap but whose centroids still fall within 1.5 km of the project boundary. Appendix Figure 12 maps an example of the project and spillover definitions. The summation includes a standard differences-in-differences specification for each exposure area where $D_t$ =1 if year $t$ is census year 2011 (post period) and $C_p$ =1 if project $p$ has been constructed. $\lambda_p$ includes a project fixed-effect to control for fixed differences between project areas that may be correlated with local demographic changes.

The coefficients of interest are $\alpha^e$, which measure the differential change in outcomes for constructed relative to unconstructed projects (either in the project or spillover areas). Interpreting this ceofficient as the causal effect of housing projects requires assuming that the areas within and nearby constructed projects would have evolved in the same way as corresponding areas for unconstructed projects in the absence of the program. Descriptive evidence from Section 5 indicates differences between constructed and unconstructed projects at baseline, particularly in terms of census measures. These differences would likely bias any cross-sectional comparisons between projects.

One potential concern is that whether a project is ultimately constructed depends on changes in the housing market or local development, which may separately drive changes in census outcomes. For example, residents of quickly growing neighborhoods may be more likely to petition local governments to stall or stop housing projects. Similarly, it may be easier to finish housing projects in areas where informal housing grows slowly. Both of these factors may lead this approach to overestimate the impacts of public housing. At the same time, the extent to which constructed projects are misclassified as unconstructed would lead this approach to underestimate the true effects of this program.

Table 4 provides difference-in-differences estimates for a range of household-level outcomes. The first four rows contain estimates within the project areas. The "Project" coefficient in the fourth row compares unconstructed project areas to spillover areas. Matching earlier descriptive evidence, these areas tend to have significantly worse outcomes than average. Adding the coefficient in the third row produces average outcomes for constructed project areas, partially attenuating this gap with spillover areas, although these terms are statistically insignificant. The second row finds some improvement in outcomes for constructed projects in the post period such as electricity measures and whether there is water inside the house; however, the differences-in-differences coefficient of interest in the first row documents larger and more statistically significant improvements in outcomes for flush toilets, water inside the house, living in a single house, as well as electricity access. The economic magnitudes of these findings are large, consistent with a substantial improvement in housing conditions following these projects. While population and household density increases substantially in unconstructed project areas, it noisily decreases in constructed project areas possibly because project houses can fit fewer people in the same building footprint.

The bottom three rows test for spillover effects in areas within 1.5 kilometers from projects. Row six finds a broad improvements across all measures in spillover areas, but row five finds no economically significant differences for constructed projects compared to unconstructed projects. These results provide little evidence that finishing a housing project differentially benefits housing quality in nearby areas.

Table 4: Census Household-level Estimates

| | (1) Flush Toilet | (2) Water Inside | (3) Water Utility | (4) Own House | (5) Single House | (6) Elec. Cooking | (7) Elec. Heating | (8) Elec. Lighting | (9) HH Density | (10) Pop. Density |
|---|---|---|---|---|---|---|---|---|---|---|
| Project X Post X Const. | 0.121* | 0.136*** | 0.035 | -0.085 | 0.142*** | 0.290*** | 0.202*** | 0.109 | 32.094 | 46.908 |
| | (0.063) | (0.050) | (0.039) | (0.074) | (0.053) | (0.069) | (0.066) | (0.080) | (782.776) | (1,464.384) |
| Project X Post | 0.080* | 0.099** | -0.030 | 0.041 | 0.075** | 0.179*** | 0.154*** | 0.155** | 1,084.115* | 2,217.685** |
| | (0.047) | (0.038) | (0.037) | (0.038) | (0.036) | (0.059) | (0.056) | (0.066) | (588.958) | (1,112.445) |
| Project X Const. | 0.088 | -0.015 | 0.040* | 0.120 | 0.044 | -0.057 | -0.050 | 0.132 | -648.420 | -531.525 |
| | (0.111) | (0.084) | (0.022) | (0.087) | (0.100) | (0.107) | (0.092) | (0.127) | (774.234) | (1,684.820) |
| Project | -0.316*** | -0.212*** | -0.021 | -0.182*** | -0.245*** | -0.344*** | -0.308*** | -0.354*** | 293.853 | 434.061 |
| | (0.074) | (0.050) | (0.022) | (0.054) | (0.063) | (0.066) | (0.060) | (0.079) | (291.835) | (667.762) |
| Spillover X Post X Const. | 0.037 | 0.031 | -0.012 | -0.008 | 0.015 | 0.042 | -0.031 | -0.018 | 306.171 | 483.836 |
| | (0.035) | (0.033) | (0.012) | (0.028) | (0.029) | (0.033) | (0.045) | (0.030) | (207.380) | (470.049) |
| Spillover X Post | 0.049* | 0.136*** | 0.012 | -0.039** | 0.048** | 0.126*** | 0.100*** | 0.085*** | 278.146* | 668.571** |
| | (0.025) | (0.025) | (0.009) | (0.015) | (0.022) | (0.027) | (0.024) | (0.027) | (144.106) | (330.608) |
| Spillover X Const. | -0.026 | -0.040 | 0.030* | 0.019 | -0.000 | -0.057 | -0.032 | 0.025 | -142.147 | 91.949 |
| | (0.057) | (0.053) | (0.016) | (0.043) | (0.045) | (0.048) | (0.047) | (0.062) | (617.352) | (1,308.905) |
| Constant | 0.798*** | 0.399*** | 0.925*** | 0.517*** | 0.599*** | 0.727*** | 0.669*** | 0.777*** | 2,491.825*** | 7,460.057*** |
| | (0.045) | (0.042) | (0.013) | (0.032) | (0.035) | (0.040) | (0.034) | (0.050) | (459.510) | (974.713) |
| N | 2020549.000 | 2020549.000 | 2020549.000 | 1958335.000 | 1936911.000 | 2020549.000 | 2020549.000 | 2020549.000 | 9,669.000 | 9,669.000 |
| Mean2001 | 0.699 | 0.300 | 0.947 | 0.507 | 0.535 | 0.570 | 0.541 | 0.709 | 2,229.575 | 7,364.545 |
| Mean2011 | 0.807 | 0.481 | 0.941 | 0.442 | 0.622 | 0.811 | 0.690 | 0.829 | 3,188.246 | 9,162.186 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Project census blocks have >30% project overlap. Spillover census blocks have $\leq$30% overlap but lie within 1.5 km of a project.
Total census blocks: 9,668

## 7. Impacts on Housing Density

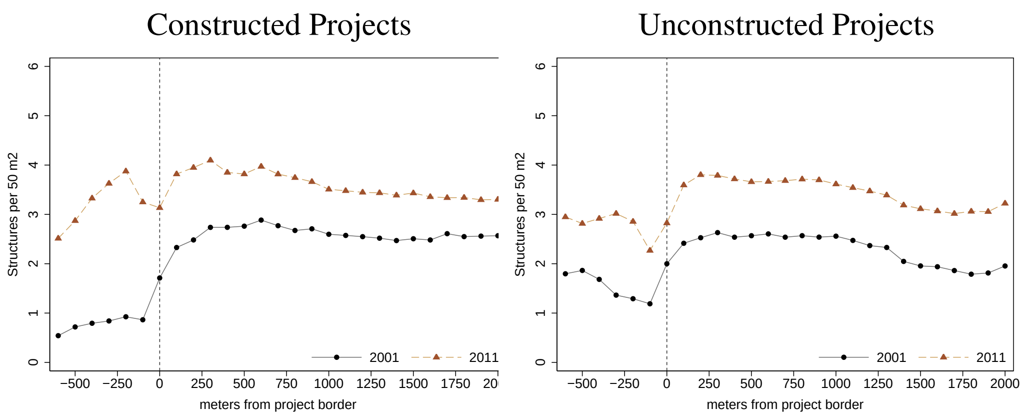
The high geographic resolution of our housing density measure allows us to take a more non-parametric approach to estimating the impacts of housing projects on local building density. This geographic resolution is instrumental in detecting possible spillover effects just outside of housing project boundaries. Figure 14 in Appendix A.4 demonstrates this resolution by depicting all structures within and around an example housing project. To construct a measure of housing density, we take the average number of houses within a 50 meter by 50 meter grid cell. Figure 15 in Appendix A.4 provides an example of this density measure over a housing project. Using these measures, we estimate the following equation:

$$y_{ipt} = \lambda_i + \sum_d I^d_{ipt} \Big( \alpha^d D_t C_p + \beta^d D_t \Big) + \varepsilon_{ipt} \tag{1}$$

$y_{itdp}$ measures the average number of houses for cell $i$ in vicinity of project $p$ observed in year $t$. We consider four housing density measures: 1) total buildings, 2) formal buildings, 3) informal buildings, and 4) informal backyard shacks. In this specification, we compute the differential change in density at a variety of distances from project boundaries. $I^d_{ip}$ =1 if cell $i$ is at distance $d$ of project $p$. $D_t$ =1 if year $t$ is 2011 again referencing the post period. $C_p$ =1 if project $p$ has been constructed. In this specification, $\lambda_i$ is a grid-cell level fixed effect. Interpreting the coefficients of interest, $\alpha^d$, as causal effects of project construction requires a parallel trends assumption at each distance from the project boundary.

In this context, we are able to perform a more non-parametric comparison of outcomes at baseline. Figure 5 plots simple means of formal housing density with respect to the housing project boundaries (where negative distances track housing density within the boundaries). Comparing averages for 2001 (black circles), between constructed (left panel) and unconstructed projects (right panel), we find slightly greater formal housing within unconstructed project borders, but nearly identical levels at all distances growing further from boundaries. Moreover, there does not appear to be a strong gradient in formal housing density moving away from project boundaries. This flat

Figure 5: Average Formal Housing Density



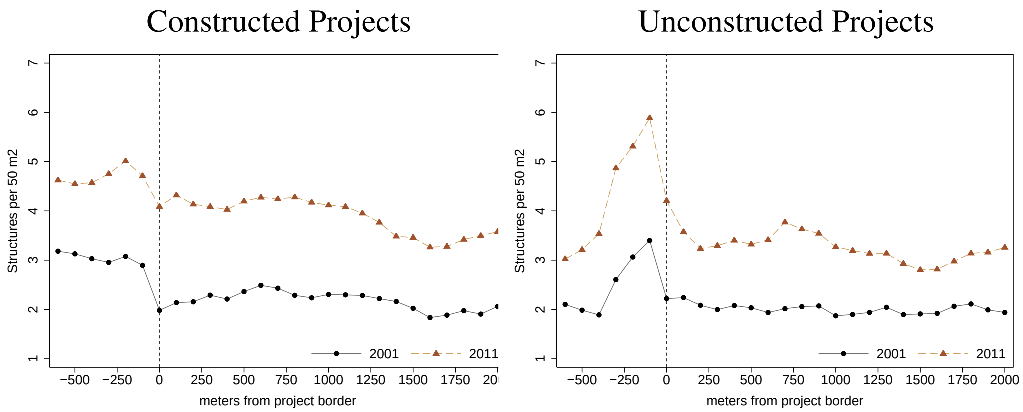

gradient suggests that while project areas themselves have less development, proximity to these projects is relatively uncorrelated with development. Therefore, areas further away from projects may additionally serve as plausible comparison groups for areas just outside of housing projects, which may be more likely to experience spillover effects from these projects.

Figure 5 also previews the empirical analysis by plotting average formal density for 2011 (orange triangles). For all positive distances in both panels, we observe a parallel upward shift in formal building density, which does not appear to be correlated with proximity to the project border. Within the project area (at negative distances), formal housing density increases more substantially for constructed projects than unconstructed projects, consistent with successful project construction.

Figure 6 exhibits similar patterns for informal housing density. Within project areas at baseline (2001), we observe similar average levels of informal housing densities although they are more unevenly distributed for unconstructed projects. Again, there appears to be little gradient in housing density moving further away from both types of projects.

After construction (in 2011), informal housing density increases across the board although increases are smaller within project areas for constructed projects. By contrast, informal housing spikes substantially just within unconstructed projects. Outside of project areas, both types of projects experience parallel shifts in informal housing density that appear uncorrelated with distance from projects.

Figure 6: Average Informal Housing Density




Taken together, these descriptive results broadly exhibit balance on observable housing densities at baseline. To the extent that baseline levels in housing density may be correlated with future changes in density, these results provide some additional support for our parallel trends assumptions.

We then estimate equation (1) to examine the differential changes in housing density between unconstructed and constructed projects. We present the non-parametric results for formal houses in Figure 7, informal houses in Figure 8, and aggregating to total houses in Figure 9. These figures plot differences-in-differences coefficients for each distance threshold relative to housing project boundaries. Standard errors are clustered at the project-level for the 68 constructed and 65 unconstructed areas. Estimates are especially noisy within project areas because the varying sizes of project footprints mean that observations per distance threshold decline quickly moving further within the footprints.

Figure 7 finds a substantial increase in formal houses within the project boundary which abruptly shifts to zero moving outside of these projects. Estimates suggest an housing projects generate increase of 1 formal house per 2500 m2 which is equivalent to around 400 houses per km2, roughly in line with average project density in Table 2. These results indicate zero spillovers in terms of formal developments close to these projects, suggesting that these housing projects have minimal local amenity effects. Figure 8 provides the opposite pattern for informal housing with a decrease in density within constructed relative to unconstructed projects. Positive distances

Figure 7: Formal Building Density

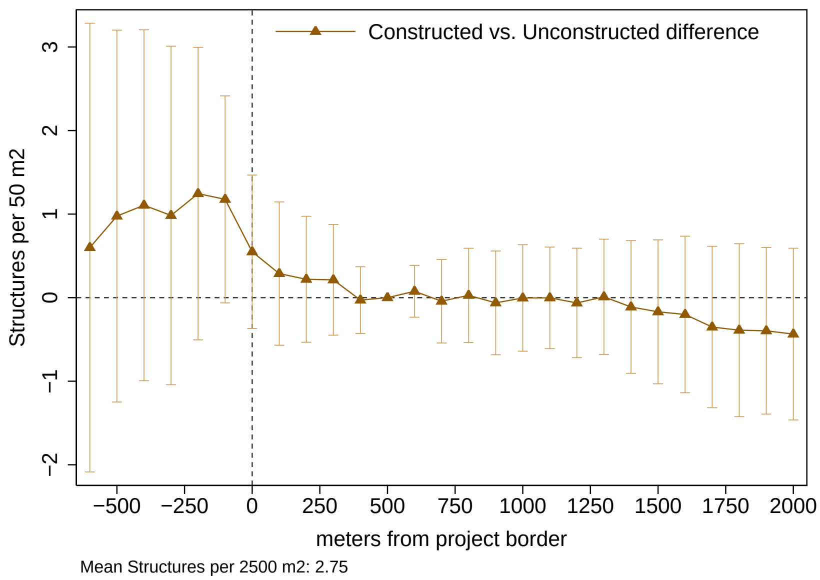



Figure 8: Informal Building Density

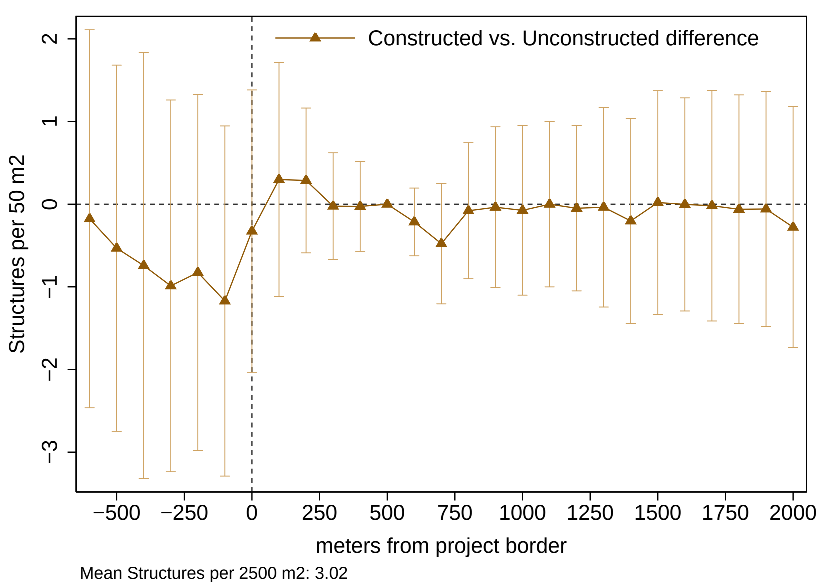
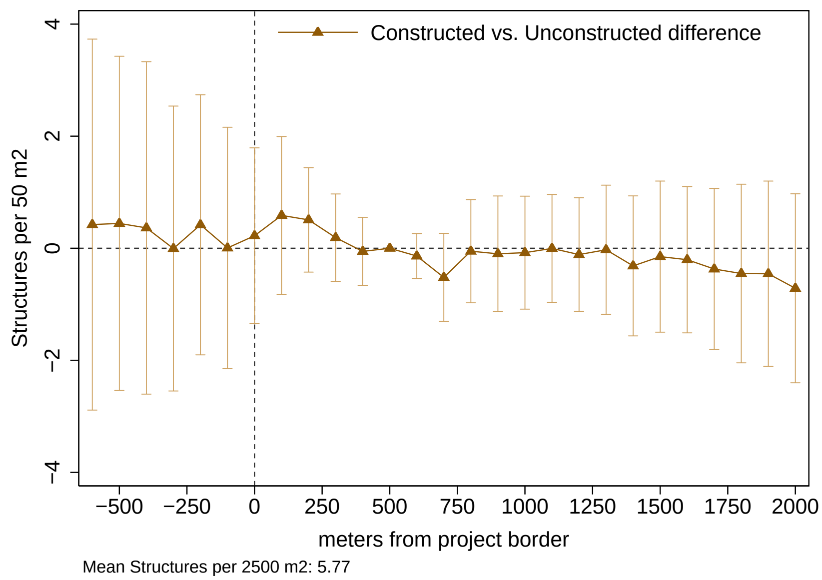


Figure 9: Total Housing Density




from project areas exhibit zero change in informal housing development indicating a similar lack of spillover effects to neighboring housing stock.

Taken together, these results suggest that formal project houses crowd-out growth in informal housing leading to zero total change in houses within project areas as demonstrated by Figure 9. Repeating this exercise for a subset of informal houses coded as backyard shacks, Figure 13 in Appendix A.3 provides additional evidence that stand-alone informal houses are replaced by backyard shacks in constructed project areas. These results match both the improvements in housing quality as well as the little change in population density detected in the previous census results in Table 4.

## 8. Price Effects

To identify the spillover effects of public housing on residential home prices in the formal market, we use a similar difference-in-differences approach comparing prices for areas close and far from project areas before and after project implementation. Our main empirical strategy takes the following specification:

$$P_{itp} = \alpha D_{tp} T_{ip} + \theta_1 D_{tp} + \theta_2 T_{ip} + +\lambda_p + \eta_t + \varepsilon_{itp}$$

The outcome, $P_{itp}$, is measured in terms of the log-purchase price of property $i$ sold at time $t$ in the vicinity of project $p$. To capture changes in prices over time within project areas, $D_{tp}$ is equal to one if date $t$ is after the month of project implementation and zero otherwise. $T_{ip}$ takes a value of one if property $i$ is within 700m of the project boundary (zero otherwise). The coefficient of interest, $\alpha$ captures the differential change in prices between near and far properties before and after project construction. Additionally, $\lambda_p$ includes a project fixed affect controlling for any fixed, unobserved drivers of house prices that vary between projects. Likewise, $\eta_t$ controls for calendar month (year $\times$ month) fixed-effects to account for any factors such as shifts in aggregate housing demand that may be correlated with prices and the timing of housing projects.

Interpreting the coefficient, $\alpha$, as the causal effect of housing projects on nearby home prices requires the following differences-in-differences assumption:

$$E[\varepsilon_{itp} | X_i, T_{ip}, D_{tp}, \lambda_p, \eta_t] = 0$$

This assumption implies that there are no other factors occurring in the same time and place as the housing projects which may otherwise impact home prices. One possibility is that housing markets anticipate the construction of these projects so that transactions in the pre-period may be partially treated by the advent of a housing project. Anecdotal evidence suggests that completion dates for these projects are very uncertain due to the large coordination of stakeholders needed for each project, making it difficult to accurately anticipate implementation. Another concern would be that housing projects are accompanied by other social programs that would stimulate investments in neighborhoods near project areas. In order to isolate market anticipation or accompanying social programs from the actual impacts of housing projects, we estimate an identical model for planned but unconstructed projects to test the robustness of the results. Similarly, in targeting housing projects to particular areas, governments may be responding to local trends in housing markets or economic conditions.

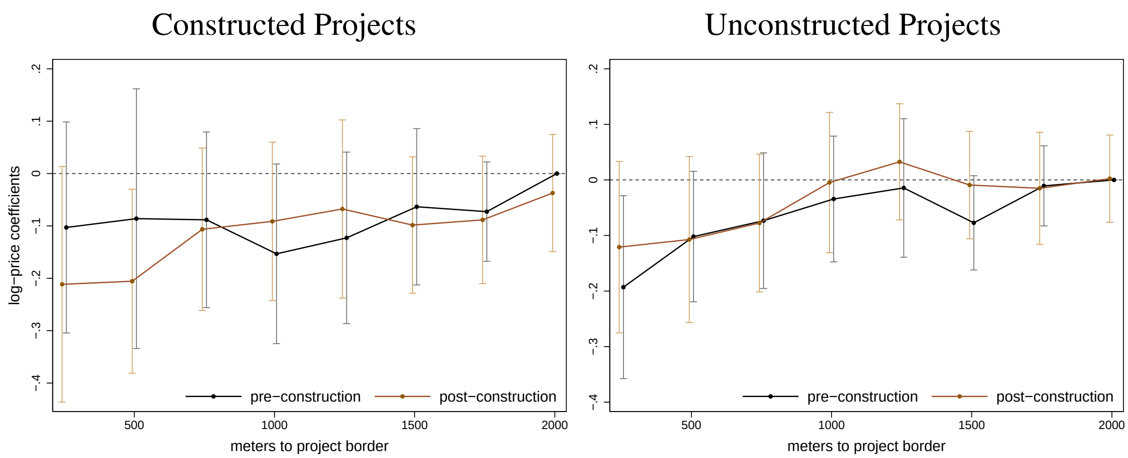
To separate project impacts from secular market trends, we leverage the sudden roll-out of housing projects under the assumption that market trends are relatively smooth over space and time relative to the construction of housing projects. In order to non-parametrically assess identification in this way, we also estimate a more flexible model both in terms of distance to project, $D_{tp}$ and time relative to project construction, $T_{ip}$. Specifically, we estimate separate treatment effects for each 250 meters of distance, $\sum_{d=1}^{8} \alpha_d D_{tpd} T_{ip}$ up to 2000 meters away from the project boundary. We do not estimate price effects within project boundaries both because (1) there are very few formal transfers of deeds within project boundaries and (2) there is a greater risk of miscoding housing project transactions as outcome transaction within project areas (although we address this mismeasurement below). We also allow effects to vary according to 12 month intervals relative to construction from three years before and three years after the expected construction date, $\sum_{l=-3}^{3} \alpha_l D_{tp} T_{ipl}$. All regressions cluster standard errors at the project level in order to account for potential correlation in prices due to unobserved factors within very localized housing markets.

One concern is that our ability to distinguish between deeds pertaining to project houses and deeds associated with normal house transactions may be imperfect. Without addressing this concern, we would mistakenly include project house transfers in our outcome measure of normal house transactions, which would substantially bias price effects. To account for this possibility, we limit our sample to include only seller-names that appear less than 30 times in the data, excluding 28% of transactions.[5] Alongside government housing projects, this approach also mainly excludes large housing developers. Therefore, the results take a more narrow interpretation as the effect of housing projects on prices for small-scale property owners. Removing these observations leaves a total sample of 87,253 deeds transactions within 2000 meters of a project boundary.

Figure 10 separately plots coefficients for average home prices before construction, $\alpha_{d,pre}$ and after construction, $\alpha_{d,post}$ for each 250m distance ring from housing project boundaries. The reference category is unconstructed projects 2000m from the project boundary. The left panel presents

[5]Results are robust to different thresholds.

Figure 10: Price Estimates over Distance from Project




price gradients for constructed projects. For constructed projects, the pre-project gradient slopes slightly upwards consistent with housing projects being located in undeveloped land or preexisting informal settlements, which may have negative amenity effects. After implementation, the price gradient slopes downwards slightly within 700 meters of the project area suggesting that either housing projects provide a negative amenity value or serve as a local housing supply shock. The large standard errors on all estimates imply that almost all estimates are no statistically different from zero or each other.

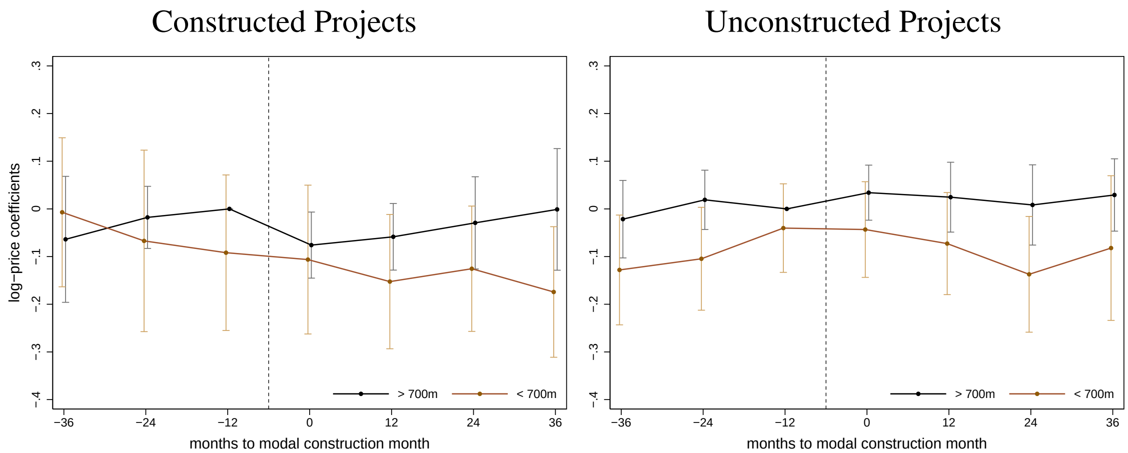
Unconstructed projects in the right panel of Figure 10 exhibit a similarly increasing price gradient before project implementation, again consistent with project land having a negative amenity effect on nearby prices. This gradient remains unchanged in periods after the expected construction date. These results provide suggestive evidence that the exact location of housing project selection does not appear to be strongly correlated with local housing market trends. The change in price gradients following project construction (left panel of Figure 10) is observed to be greatest at distances less than 700 meters from the project boundary. Given this result, we then use this threshold to define distance rings of “near” and “far” to examine how prices might change over time relative to date of project construction.

Figure 11 plots coefficients for average price changes over time relative to the modal construction month for properties within and beyond 700m of a project boundary. The reference group for this regression is properties further than 700m from a project border in the year before the expected construction date.

Focusing on constructed projects in the left panel, areas far from project boundaries follow a stable zero trend with respect to the completion of projects. Meanwhile, areas close to the project boundaries experience a slight but stable decline in prices over the period. According to the point estimates, the decline is economically significant reaching 15% below average prices (given that the outcome is log-prices). This decline may provide evidence of anticipatory behavior: local homeowners hear news about a new housing project and decide to sell their houses, anticipating a price drop due to the housing project.

For unconstructed projects in the right panel, areas far and near from project boundaries follow similar parallel trends to constructed projects leading up to the expected completion date and remain indistinguishable for the entire duration after this date. This null result helps to exclude competing explanations that rely on the timing of housing projects such as targeting housing programs according to hyper-local housing market trends or pairing housing projects with other place-based policies. One caveat for these results is that measurement error both in matching project names to budget statements and in assigning the proper date of completion for each project may bias results toward zero. It is unclear the extent to which anticipation effects might have a role in this case; if nearby homeowners have little information on when a project will be constructed, then we should expect to see the same anticipation effects as for constructed projects. However, if local homeowners are well informed about the local dynamics determining project construction, then anticipation effects would likely be more muted in the case of unconstructed projects. To test the significance of these effects in a regression framework, we first run two simple differences-in-differences regressions separately for constructed and unconstructed projects where "Post" is defined as anytime after expected project construction and "Near" includes properties within 700m of project boundaries. We also conduct a triple-differences analysis which adds the indicator $C_{it}$ for whether a project is constructed as well as the full set of interaction terms with this indicator in the following

Figure 11: Price Estimates over Time




equation:

$$P_{itp} = \alpha D_{tp}T_{ip}C_{it} + \theta_1 D_{tp} + \theta_2 T_{ip} + \theta_3 C_{it} + \theta_4 D_{tp}T_{ip} + \theta_5 C_{it}T_{ip} + \theta_6 C_{it}T_{ip} + \lambda_p + \eta_t + \varepsilon_{itp}$$

This triple-differences specification allows for a slightly less restrictive form of parallel trends: instead of requiring price changes to be parallel close and far from a project, this approach simply requires that differential price changes close and far from a project for constructed areas would be the same in unconstructed project areas in the absence of construction.

Table 5 provides the estimation results with results for constructed projects in the first column, unconstructed projects in the second column, and finally triple-difference estimates in the third column. The coefficient of interest in the first column, *PostXNear*, translates the declining price gradient for nearby properties in constructed project areas in Figure 11 to a 7.5% decrease in house prices which is significant at the 10% level. The corresponding difference-in-differences coefficient for planned but unconstructed projects in the second column predicts a much smaller and statistically insignificant decline in prices. Pooling samples in the third column, the triple-difference coefficient of interest captures roughly the midpoint between the constructed and unconstructed difference-in-differences coefficients, resulting in a noisy 6% decline in house prices. Meanwhile to put this effect into perspective, being within 700m or in a constructed area both

Table 5: Price Estimates for Completed Projects

| VARIABLES | (1) Const | (2) Unconst | (3) All |
|---|---|---|---|
| Post X Near X Const. | | | -0.0597 |
| | | | (0.0665) |
| Post X Const. | | | 0.00354 |
| | | | (0.0469) |
| Near X Const. | | | 0.0662 |
| | | | (0.0676) |
| Post X Near | -0.0747* | -0.0325 | -0.0253 |
| | (0.0414) | (0.0413) | (0.0495) |
| Post | 0.0332 | 0.0407* | 0.0277 |
| | (0.0219) | (0.0239) | (0.0306) |
| Near | -0.0280 | -0.151*** | -0.126** |
| | (0.0421) | (0.0453) | (0.0546) |
| Const. | | | -0.129** |
| | | | (0.0543) |
| Constant | 10.36*** | 7.868*** | 10.59*** |
| | (0.265) | (0.310) | (0.237) |
| Observations | 41,881 | 34,146 | 87,253 |
| R-squared | 0.348 | 0.348 | 0.330 |
| Year-Month FE | Yes | Yes | Yes |
| Project FE | Yes | Yes | Yes |

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Clustered at the project level.
Outcomes are in log-purchase prices. "Const" refers to constructed while "Unconst" refers to unconstructed housing projects.

shift mean prices by 12%, which are statistically significant at the 5% level. Taken together, these results provide weak evidence of a slight decline in home prices, which seem to be more driven by anticipation of project completion than by the direct impact of nearby formal housing.

## 9. Discussion and Conclusion

We find that South African housing projects greatly improve the housing conditions within their footprints. Housing projects successfully deliver large numbers of new formal structures (Figure 7) and much better access to bulk services (Table 4). At the same time, formal project housing appears to crowd-out growth in informal settlements so that the total density of housing remains the same as in unconstructed project areas (Figure 9). Moreover, results in Table 4 indicate that the average quality of new project houses are at least as good as neighboring houses at baseline. Therefore consistent with the model of amenity effects in Diamond and McQuade [2016], we would expect to find that these programs increase nearby home prices; however, results in Table 5 report if anything the opposite trend. Similarly, we do not observe large new developments in either informal or formal housing just nearby these projects.

The lack of significant spillover effects in price as well as in quantity is consistent with countervailing amenity and housing supply effects. While an influx of new houses threatens to decrease prices and housing supply nearby, positive amenities provided by housing projects may also attract greater housing investment, increasing prices. While both amenity and supply effects predict increases in local population, there may be spatial or geographic constraints on population growth given that these settlements in South Africa are almost exclusively single-story and informal settlements often fill any remaining surface area.

Our interpretation is encouraging from a policy perspective: housing projects can produce enough positive amenities to mediate any housing supply effects. Additionally, these projects effectively replace informal settlements with high quality formal houses, achieving a signature goal of South Africa's policy. At the same time, there is little evidence supporting housing policy as an important engine for local economic growth and development. These results may at least mediate any concerns policymakers may have for adversely affecting communities with housing projects, despite sometimes loud complaints from nearby community members.

It is important to emphasize that our analysis does not currently speak to the overall welfare implications of these policies. For example, public housing may lead to an overall net reduction in slums in the city, which may generate total welfare gains through fewer congestion externalities or other mechanisms. In light of South Africa's emphasis on using housing policy to stimulate neighborhood development, we hope that this paper will be useful for policymakers in considering the range of informal housing market responses as they design new housing policies.

# References

M. Busso, J. Gregory, and P. Kline. Assessing the incidence and efficiency of a prominent place based policy. *American Economic Review*, 103(2):897–947, 2013.

M. D. Cattaneo, S. Galiani, P. J. Gertler, S. Martinez, and R. Titiunik. Housing, health, and happiness. *American Economic Journal: Economic Policy*, 1(1):75–105, 2009.

DA-GPL. Mec mashatile delays munsieville ext 5 multimillion housing project. Jul 2017. URL http://da-gpl.co.za/mec-mashatile-delays-munsieville-ext-5-multimillion-housing-project/.

Department of Human Settlements. Breaking new ground: A comprehensive plan for the development of sustainable human settlements. Technical report, 2004. URL http://abahlali.org/files/Breaking%20new%20ground%20New_Housing_Plan_Cabinet_approved_version.pdf.

Department of Human Settlements. Annual reports. Technical report, 2012,2015. URL http://www.dhs.gov.za/content/annual-reports.

F. Devoto, E. Duflo, P. Dupas, W. Parienté, and V. Pons. Happiness on tap: piped water adoption in urban morocco. *American Economic Journal: Economic Policy*, 4(4):68–99, 2012.

R. Diamond and T. McQuade. Who wants affordable housing in their backyard? an equilibrium analysis of low income property development. Technical report, National Bureau of Economic Research, 2016.

P. Dlmini. R95m down the tubes as housing project picked apart brick by brick. 2017. URL https://www.timeslive.co.za/news/south-africa/2017-09-12-r95m-down-the-tubes-as-housing-project-picked-apart-brick-by-brick/.

E. Durojaye, G. Mirugi-Mukundi, N. Munshi, and K. Tissington. Jumping the queue, waiting lists and other myths: Perceptions and practice around housing demand and allocation in south africa. Technical report, Socio-Economic Rights Institute of South Africa, Community Law Centre University of the Western Cape, 2013. URL http://www.seri-sa.org/images/Jumping_the_Queue_MainReport_Jul13.pdf.

S. Franklin. Enabled to work: the impact of government housing on slum dwellers in south africa. 2016.

S. Galiani, P. J. Gertler, R. Undurraga, R. Cooper, S. Martínez, and A. Ross. Shelter from the storm: Upgrading housing infrastructure in latin american slums. *Journal of Urban Economics*, 98:187–213, 2017.

B. Marx, T. Stoker, and T. Suri. The economics of slums in the developing world. *Journal of Economic Perspectives*, 27(4):187–210, November 2013. doi: 10.1257/jep.27.4.187. URL http://www.aeaweb.org/articles?id=10.1257/jep.27.4.187.

M. Mathoho. A case study of participation in the thinasonke ext 4 fedup pledge project. 2010. URL http://www.planact.org.za/participation-and-development-from-the-perspective-of-the-poor-2008-2009/.

D. Matsena. Premier launches rdp housing project in north. 2018. URL https://rekordnorth.co.za/121824/premier-launches-rdp-housing-project-in-north/.

A. Takeuchi, M. Cropper, and A. Bento. Measuring the welfare effects of slum improvement programs: The case of mumbai. *Journal of Urban Economics*, 64(1):65–84, 2008.

K. Tissington. A resource guide to housing in south africa 1994-2010: Legislation, policy, programmes and practice. Technical report, Socio-Economic Rights Institute of

Table 6: Assessing Name Matching between
Budget and Spatial Administrative Data

| | Unconstructed | | Constructed | |
|---|---|---|---|---|
| | Matched | Unmatched | Matched | Unmatched |
| Formal Density: 2001 | 188.4 | 142.5 | 481.7 | 546.1 |
| Formal Density: 2011 | 505.0 | 284.0 | 1,527.2 | 1,575.4 |
| Informal Density: 2001 | 998.6 | 457.2 | 1,611.1 | 985.2 |
| Informal Density: 2011 | 1,619.1 | 701.3 | 2,379.3 | 1,064.2 |
| Project House Density | 0.0 | 0.0 | 558.6 | 348.4 |
| Project Mode Year | . | . | 2004 | 2005 |
| Area (km2) | 2.2 | 2.9 | 5.4 | 1.4 |
| Observations | 67 | 32 | 75 | 18 |

Density is measured in structures per km$^2$.

South Africa, 2011. URL https://www.seri-sa.org/images/stories/SERI_Housing_Resource_Guide_Feb11.pdf.

United Nations. Annex millennium development goals, targets and indicators, 2015: statistical tables. 2015. URL http://mdgs.un.org/unsd/mdg/Resources/Static/Products/Progress2015/StatAnnex.pdf.

# A. Appendix

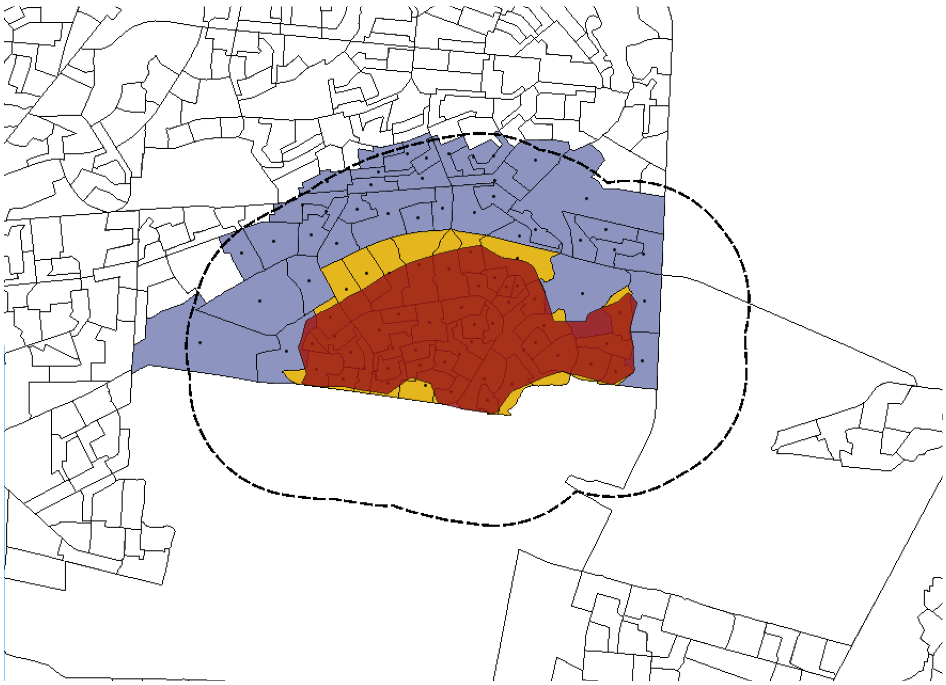
## A.1. String Matching for Project Names

## A.2. Census Block Analysis

## A.3. Backyard Shack Density

## A.4. Housing Density Design

Figure 12: Buffer Design Census Block Example



- Census Blocks with ≥30% project area overlap : yellow polygons
- Census Blocks with <30% project area overlap : blue polygons
- Buffer Area (1.5 km from clusters) : dotted line
- Project Area : red polygon

Figure 13: Total Housing Density

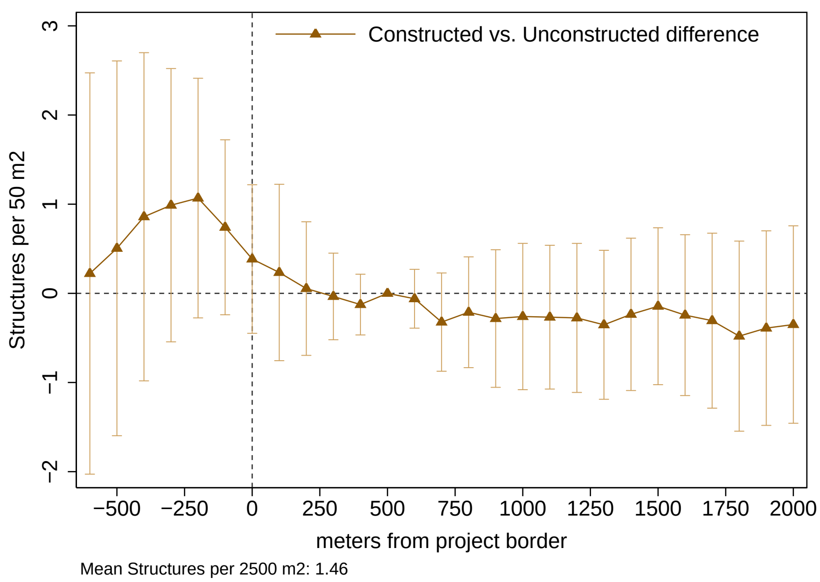


Figure 14: House Census Example

- Red Polygon : Completed Housing Project
- Green Points : Informal Housing
- Black Points : Formal Housing
- Blue Grid : 50m by 50m Grid Cells

Figure 15: House Density Example

- Building density per 50m by 50m cell
- Darker = greater density