Subsidized Housing and Urban Development: Evidence from South Africa*

Benjamin Bradlow†
Stefano Polloni‡
William Violette§

This Version: January 23, 2019

Abstract

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

*We are grateful to Andrew Foster, Daniel Bjorkegren, Matthew Turner, Jesse Shapiro, Brian Knight, and John Friedman for their feedback and advice. We also thank Adelaide Steedley and the Centre for Affordable Housing and Finance in Africa as well as GeoTerraImage for generous research guidance and for providing data without which this project would not have been possible. 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.

†Dept. of Sociology, Brown University, Providence, RI E-mail: benjamin_bradlow@brown.edu
‡Dept. of Economics, Brown University, Providence, RI E-mail: stefano_polloni@brown.edu
§Federal Trade Commission, Washington, DC. E-mail: william.j.violette@gmail.com

1. Introduction

30% of urban populations in the developing world live in crowded slums characterized by low infrastructure access, unsanitary conditions, and high crime rates [United Nations, 2015]. These living conditions are thought to pose long-lasting obstacles to upward mobility and local economic development [Marx et al., 2013]. We study a common policy response where governments replace slums with serviced formal structures and move slum dwellers into public housing projects.

The promoted goals of these programs are twofold. First, policymakers stress how better housing improves household wellbeing along a variety of dimensions.[1] This is well supported by existing studies on direct recipient impacts, which link housing programs to improvements in child health, women's employment, and quality of life [Cattaneo et al., 2009, Franklin, 2016, Galiani et al., 2017]. A second set of policy goals pertains to broader urban development impacts, driven by spillovers to neighboring areas. For instance, subsidized housing in South Africa is proposed for "leveraging growth in the economy, [...] combating crime, promoting social cohesion, [...] and utilizing housing as an instrument for the development of sustainable human settlements, in support of spatial restructuring" [Department of Human Settlements, 2004].

Spillover effects may account for a significant share of the total welfare impacts of housing projects. In the US, Diamond and McQuade [2016] calculate that affordable housing in low-income areas generates substantial societal benefits on surrounding areas, often exceeding the total costs of the programs. Welfare effects may be even greater in developing settings, where preexisting slums impose negative externalities on neighboring areas. At the same time, the informal housing market may respond to the improved infrastructure, thereby mitigating positive amenity spillovers.[2]

Both the policy importance and theoretical ambiguity of these mechanisms motivate empirical analysis in a developing context. Until recently, however, data limitations have largely precluded researchers from constructing granular measurements

[1] The United Nations Habitat Program articulates how improved housing can provide "social, cultural and economic benefits in the form of improved quality of life, poverty alleviation, environmental protection and improved health and safety" [United Nations Habitat Program, 2018].

[2] For example, Brueckner et al. [2018] document the common practice of "backyarding", where households rent informal dwellings erected on the same land plots as government-sponsored houses.

on the quantity, quality, and price of housing, both formal and informal. Using novel data sources, this paper contributes to filling this gap by identifying impacts on local neighborhoods at high spatial resolutions.

We analyze the direct and spillover effects of a large public housing program in South Africa. We combine administrative records for 57 constructed and 60 planned but unconstructed projects with spatial data on property transactions, formal and informal housing densities, and dwelling characteristics. To measure spillovers, we implement a difference-in-differences design at varying distances from project boundaries. At every distance, we compare changes in outcomes for constructed projects to changes for planned but unconstructed projects. When feasible, our estimates include a third layer of differencing, comparing effects near versus far from project boundaries. In this way, we are able to net-out systematic time-varying differences between treated and control areas, absent project construction. Our econometric approach allows us to flexibly capture how spillover effects change across space. Planned but unconstructed project areas provide useful counterfactuals since governments may select unobservably worse land plots to build subsidized housing projects. These control areas resemble treated areas at baseline across a range of measures, including preexisting housing density.

Within project boundaries, we find large improvements in access to services and home quality, but little net effects on total housing development. Compared to unconstructed areas, households living within projects (either in subsidized formal or informal housing) are more than 10 percentage points more likely to access flush toilets, indoor water taps, and electricity. The small net effects on total housing density mask a compositional change from informal to formal structures. We find greater densities of informal *backyard* dwellings, but not enough to offset steep declines in other types of informal housing. On the whole, these results suggest that subsidized housing projects provide positive local amenities through better house quality and improved infrastructure.

Outside project boundaries, and in spite of our previous findings, we fail to detect spillover effects across our full range of outcomes. Housing projects do not appear to stimulate additional construction of formal or informal dwellings at all distances outside project boundaries, up to 1200 meters away. Similarly, dwelling characteristics in neighboring areas remain largely the same after project construction. Leveraging the exact timing of construction, we do not detect discrete changes in formal housing prices just after construction, though the data admit a wide range of treatment effects. Finally, we provide some evidence that housing markets anticipate projects, and find no indication of residents sorting into or out of project areas, as measured by changes in the composition of dwellers.

Our findings do not attribute an important role to neighborhood externalities in housing markets in developing countries. Another interpretation is that informal backyard houses negate the positive benefits from government-sponsored homes. From a policy perspective, these results suggest that spillover effects merit less emphasis than direct recipients effects in designing housing policy.

Our analysis contributes to a recent literature focusing on the broader urban development impacts of housing policy in developing countries. Michaels et al. [2017] study sites-and-services schemes implemented on the urban fringe of Tanzanian cities. Using comparable greenfields as counterfactuals, they find that initial infrastructure investments have long-term positive impacts on nearby housing quality. Harari and Wong [2018] examine the impacts of a slum upgrading program in Jakarta, Indonesia. The authors also use untreated comparable areas as counterfactuals, and find negative effects on land values and levels of formalization. Importantly, in a boundary discontinuity design, the authors note no evidence of spillovers from the program to neighboring areas. In ongoing work, Gechter and Tsivanidis [2018] study spillovers from formal housing development using sharp temporal changes in Mumbai's zoning laws. Through the lens of a quantitative general equilibrium model, the authors find high higher property prices, more formal construction, and lower population density in the vicinity of targeted areas. Finally, outside of the developing world, our research draws upon an established literature on housing externalities [Rossi-Hansberg et al., 2010, Hornbeck and Keniston, 2017, Diamond and McQuade, 2016].

We proceed by first providing background on the South African housing program in Section 2. Section 3 describes the data used to measure outcomes and details our approach to identifying housing projects. Section 4 lays out the estimation strategy and identifying assumptions, which we validate with descriptive evidence in section 5. Section 6 presents and discusses our main results, and Section 7 concludes.

2. Subsidized Housing in South Africa

The housing projects studied in this paper were implemented as part of a large national housing subsidy scheme enacted in 1994. Though periodically revised and renamed,[3] the program has consistently sought to redress the economic and geographic legacy of apartheid by providing formal housing to low-income households. Subsidized housing efforts at the national level have focused on constructing and allocating $40m^2$, single-story, two-room dwellings. According to government figures, upwards of 3 million housing units have been delivered nationally between 1994 and 2015.

2.1. Planning and Delivery

Housing projects are often located on undeveloped state-owned land although in some cases, municipalities work with private developers to purchase inexpensive, vacant private land. Finding suitable land plots often requires policymakers to locate the projects far from city centers and economic opportunities [Department of Human Settlements, 2012, 2015]. Undeveloped land plots may contain 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 maintain yearly housing allocations [Department of Human Settlements, 2012, 2015]. While the location and size 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

[3]Public housing in South Africa has been delivered under the *Reconstruction and Development Program* (RDP) starting in 1994 and subsequently by its successor, *Breaking New Grounds*, as of 2004.

projects have been delayed for upwards of 10 years [DA-GPL, 2017].

2.2. Recipients

The National Department of Human Settlements issues guidelines for eligibility and maintains an official waiting list for eligible households. Eligibility requires South African citizenship, no previous property ownership, being married or having financial dependents, and having a monthly household income below R3,500 [Durojaye et al., 2013]. The share of households reporting at least one member on the waiting list has remained stable at over 13% from 2009 to 2013.[4] Before construction, each project is assigned beneficiaries in a first-come, first-served basis according to the waiting list in their province or municipality. However, when projects replace existing informal settlements, previous inhabitants receive renovated houses while remaining units are allocated according to the housing waiting list.

In practice, these guidelines are 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 as being still occupied by their original beneficiaries within five years of construction.[4] Anecdotal evidence suggests that project managers are aware of active secondary markets but have difficulty policing these transactions [Matsena, 2018].

2.3. Backyard Shacks

By design, subsidized housing in South Africa offers recipients generous amounts of yard space, often diverted to alternative land uses. The construction of informal backyard shacks is a common occurrence within project boundaries. Recipients typically

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

enter agreements with tenants (at times relatives) to lease the shacks for cash or in-kind payments. According to the 2011 census, backyard tenants represent 7.5% of South African households.[5] Though the structures are not as durable as their government-provisioned counterparts, living conditions are typically superior to conventional informal settlements. Tenants gain indirect access to their landlord's services such as water taps, electricity connections, and toilets, and benefit from reduced eviction threats and greater personal safety [Beall et al., 2003]. Brueckner et al. [2018] note that *backyarding* is suggestive of inefficiencies in the program delivery, and allows to correct potential misallocations of resources. Importantly for our analysis, backyard shacks reveal that more households may benefit from subsidized housing than the listed number of direct recipients. According to the General Household Surveys (2009-2013), more than 33% of government-provisioned homes have backyard shacks within two years of the delivery date.

3. Data Sources

Our analysis focuses on the South African province of Gauteng, both the smallest geographically and most populated province in the country. Gauteng's boundaries roughly correspond to the greater metropolitan areas of Johannesburg and Pretoria. Understanding the local development impacts of public housing requires (1) outcomes measured at high spatial resolutions, and (2) a precise measure of the location, timing, and size of housing projects. To this end, we rely on four main data sources: administrative maps describing housing policies in Gauteng, deeds data on housing transactions, household-level census data, and building-based land use information. Each is described in turn below.

3.1. Deeds Data and Location of Housing Projects

We locate housing projects using a combination of administrative policy maps and deeds data. We obtain the former from the Gauteng City Regional Observatory, a research unit composed of the Gauteng Provincial Government and two Johanesburg

[5] This figure jumps to 12.5% for the province of Gauteng, where our study is conducted.

universities.[6] The maps describe 192 planned housing projects as of 2008, including notes on their completion status for some projects.[7] The latter is sourced from the South African National Deeds Office, covering the 2001 to 2011 universe of transactions in *affordable areas*, defined as census enumeration areas with 2010 mean house prices below R500,000. These data were kindly provided by the Affordable Land and Housing Data Centre, which tracks affordable housing markets. Transaction information includes sale price, GPS location, plot size, buyer name, and seller name.

Constructed Housing Projects

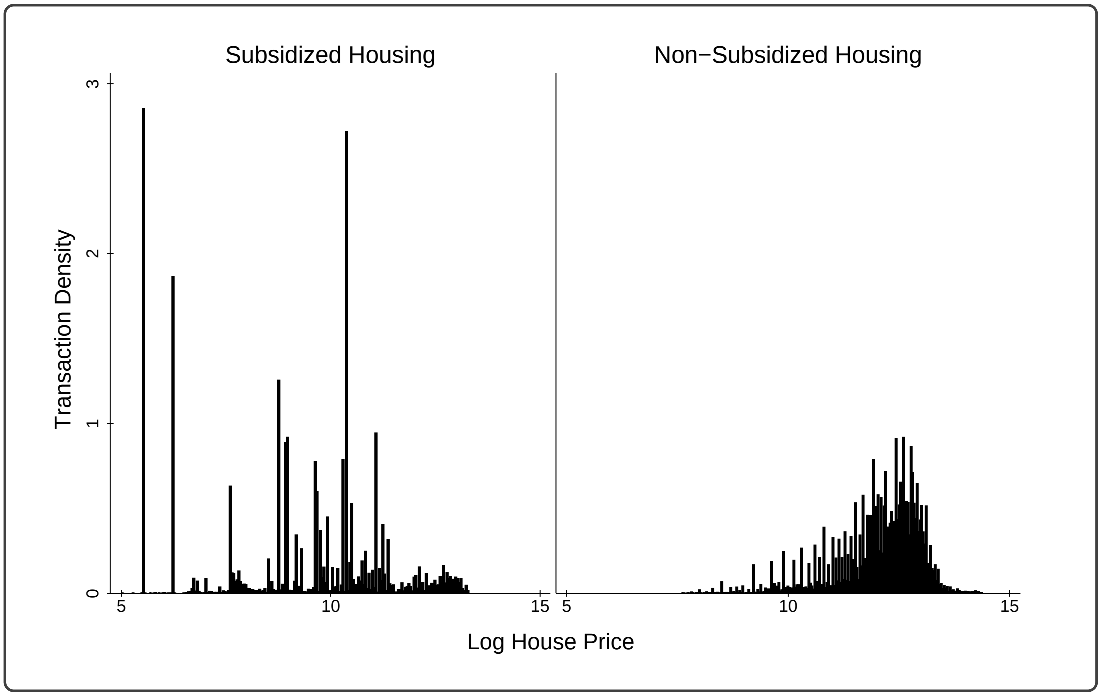
Using the administrative policy maps, we define projects as successfully constructed when we observe recipients receiving deeds to their government-sponsored house within the projects' boundaries. To identify state-sponsored housing transactions, we follow a filtering procedure established by the data provider. Properties are first assumed to belong to a housing project if the seller name includes a government, municipality, or large developer when first transacted. We then exclude deeds flagged as large buildings used for commercial purposes (less than 2% of transactions), as well as purchase prices more than R50,000 above the yearly nominal subsidy values (less than 4% of remaining transactions). As shown in the right panel of Figure 1, transaction prices for properties deemed to belong to public housing programs exhibit significant mass-points in their distribution, consistent with many properties being recorded with these subsidy values. Finally, we exclude transactions occurring in Gauteng's historic townships, because the bulk of residential construction within these areas precedes the start of South Africa's housing program.[8] Altogether, this filtering procedure identifies over 127,000 properties as government-sponsored housing. Overlaying theses properties' GPS coordinates with the housing policy maps, we are able to identify 57 constructed housing projects. Delivery dates for the projects are then inferred from the distribution of transactions over time. Specifically, considering transactions within each project separately, we set the modal transaction month as the delivery month for

[6] www.gcro.ac.za/

[7] Since this data comes in the form of overlapping shapefiles, we use the union of intersecting shapes, excluding small shapes below 0.5 km^2.

[8] Urban townships are apartheid-era residential areas created for black migrant workers. Township boundaries were also provided by the Gauteng City-Region Observatory.

Figure 1. Transaction Price Histogram



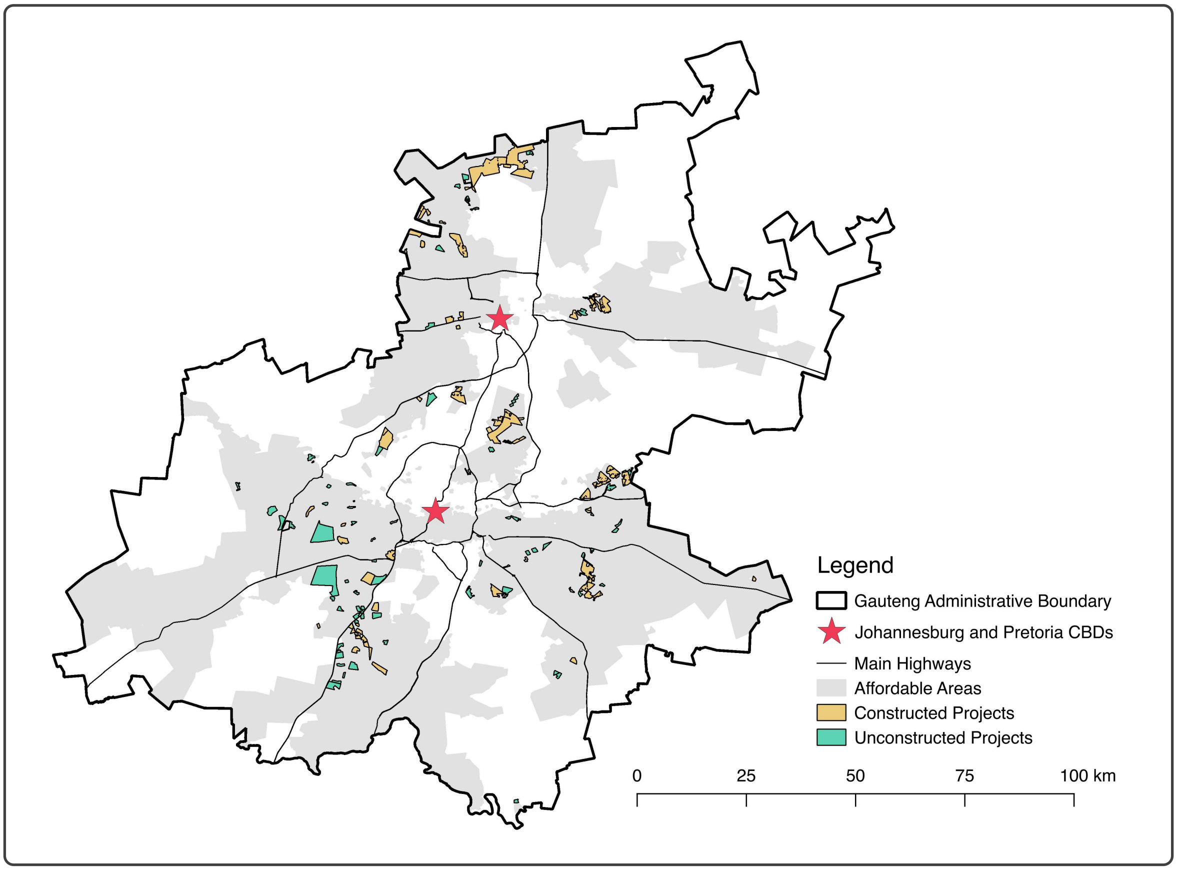
each project. Within projects, most government-sponsored properties are transacted in the same month.[9]

Planned but Unconstructed Housing Projects

Identifying 57 constructed housing projects leaves 135 planned but possibly unconstructed housing projects in the administrative map. We propose using these projects as counterfactuals, capturing the level of urban development that would have occurred in the absence of construction. Because our deeds data cover the 2001-2011 period and the policy maps pertain to 1994-2008, a concern is that some of the remaining 135 projects were planned and delivered prior to 2001. In order to determine which of the unconstructed projects were planned between 2001 and 2011, we make use of National Treasury budget reports. We are 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 string-matching algorithm to link project names from the budget reports to the administrative maps. This proce-

[9] Appendix A.3 plots the distribution of subsidized and non-subsidized transactions around the modal transaction month.

Figure 2. Housing Project Map



dure results in a final sample of 60 unconstructed projects. We provide further details about the digitization and string-matching algorithm in Appendix A.1. For attributing construction dates to these counterfactual projects, we use expected completion dates from the budget reports.[10]

Figure 2 shows the geographic setting of our analysis, displaying our final sample of constructed and unconstructed projects as polygons. Importantly, Figure 2 also shows that the affordable areas – the coverage area for our deeds data – contain every project boundary. 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 isolated projects of both types. Projects are generally located at relatively great distances from central business districts (CBDs), suggesting that vacant or inexpensive land plots are especially tar-

[10]This will matter only for our analysis on housing prices, since we observe two time periods, pre and post, for all other data sources.

geted by housing authorities. Despite their distance from the CBDs, housing projects are often next to arterial highways, easing commuting costs for recipients.

Table 1. Project Descriptions

	Constructed	Unconstructed
Proposed	5	20
Planning	8	12
Under Implementation	12	4
Complete	6	1
No Description	26	23
Total	57	60

We note that our approach is not without limitations, and may introduce measurement error insofar as we are misattributing deeds to housing projects (false-positives), or wrongly assuming a project is unconstructed (false-negatives). To provide some validation for our classification, we tabulate in Table 1 project descriptions from the 2008 administrative policy maps according to whether projects are classified as constructed or unconstructed by 2011. As of 2008, we find that constructed projects are more likely to be classified as "completed" or "under implementation", while unconstructed projects are more likely to fall into "proposed" or "planning" categories. Among constructed projects, we find 5 "proposed" and 8 "planning" indicators, which could imply either (1) these projects are false-positives or (2) while planned at some point prior to 2008, these projects were eventually completed by 2011.[11] We also identify a single "complete" project as unconstructed, likely because its houses do not appear in our deeds records. Figure 5a in section 5 further validates our sample, showing how formal housing increases disproportionately in projects classified as constructed.

Nearby housing transactions

To examine spillovers nearby housing projects, we include the remainder of properties that are located outside of the project areas and are not sold by government housing agencies or large developers. We identify these developers as sellers appearing more than 30 times in the transaction sample. Excluding these sellers limits the external validity of our results only to small developers and individual homeowners. We focus

[11] We are unable to assess the frequency at which these descriptions are updated in the data.

on properties located within 4 kilometers of constructed and unconstructed housing projects, forming a sample of over 140,000 transactions. We exclude the top 1% of prices as well as prices below 2,500 Rand, which are likely to be composed of mismeasured prices, or titles exchanged between family members. The price distribution of this sample is displayed in the right panel of Figure 1, which stands in contrast to the distribution of subsidized transactions in the left panel.

3.2. Census Data

To measure impacts on dwelling characteristics, we use the 2001 and 2011 National Censuses of Population and Housing. Specifically, we analyze household-level responses describing the quality of their living quarters. Our outcomes are mainly binary indicators and pertain to the household's access to services (flush toilets, water tap, electricity access), housing durability, and tenure arrangements. We identify households at the *small-area* level, the smallest available census geography. The province of Gauteng is divided between approximately 11,000 small-areas in 2001, and 17,000 small-areas in 2011.[12] On average, each census area contains 170 households. Though we observe responses from every surveyed household in both census waves, the data does not allow to link households across time periods.

3.3. Building Based Land Use

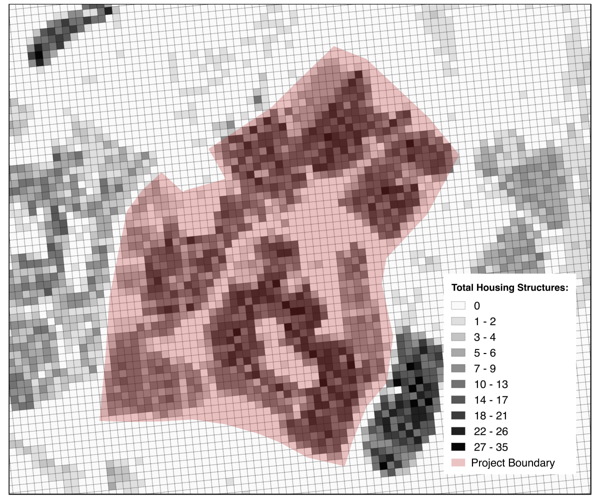
Our final data source consists of hand-coded building surveys derived from high-resolution aerial and satellite imagery. We obtain these data from GeoTerraImage (Pty) ltd., a local remote-sensing specialist.[13] The data differentiates structures across over 30 categories, including formal and informal residential dwellings. Informal housing structures are easily identified from their temporary nature, often made of materials such as recycled wood and corrugated metal. In contrast, formal housing structures are permanent, generally made out of brick, and may have a pitched or a flat roof with tiles, zinc panels, or other materials. Importantly for our analysis, the data encodes

[12] Despite many boundaries being very similar, census geographies are not constant in both time periods. Most of the 2001-2011 growth in census small-areas is due to geographies being split into more part in the 2011 census.

[13] http://www.geoterraimage.com/

Figure 3. Building-Based Land Use Data

(a) raw data identifying (in)formal structures



(b) 50×50m grid cells data aggregation

backyard shacks as informal structures, but distinguishes between backyard and other types of informal buildings. We track changes in residential development by using two available survey waves in Gauteng: 2001 and 2012. As a validation exercise, we compute aggregated building counts at the census small-area level, and check the resulting figures against the number of households reporting to live in formal/informal housing in the census. The two data sources paint a consistent picture of housing in Gauteng, with high correlations (≥ 0.85) between formal and informal building quantities in both available comparison years.[14] For estimation purposes, we transform this data into 50×50m grid cell aggregates, thereby creating five measures of housing density: (1) total residential structures split between (2) formal residential structures and (3) informal residential structures, which can be further decomposed into (4) backyard informal and (5) non-backyard informal structures. These grid cells will be the main units of observations for our housing density regressions, described in section 6.2. In Figure 3, we provide an example of the raw data and a depiction of our griding procedure, using the 2012 data wave.

[14]We correlate the 2001 and 2012 building surveys with the 2001 and 2011 censuses, respectively.

4. Empirical Methodology

To estimate the urban development impacts of subsidized housing both within project boundaries as well as nearby, we implement a difference-in-differences strategy assessing changes in outcomes for constructed projects, using unconstructed projects as counterfactuals. Given the varying temporal and spatial resolution of our measured outcomes, we estimate different forms of the following specification:

$$y_{ipt} \ = \ \lambda_p + \sum_{d \in D} I_{ip}^d \Big(\alpha^d \textsc{Post}_{pt} \times \textsc{Const}_p + \beta^d \textsc{Post}_{pt} + \gamma^d \textsc{Const}_p + \theta^d \Big) + \delta X_{ipt} + \varepsilon_{ipt} \tag{1}$$

where y_{ipt} is the outcome for unit i in vicinity of project p observed at time t, with idiosyncratic error ε_{ipt}. Depending on the outcome, unit i is either a census small-area, 50×50m grid-cell, or formal housing property. $\textsc{Post}_{pt}$ equals one if period t is after project p's assigned construction date, and $\textsc{Const}_p$ equals one if project p is constructed (treated). We interact the terms $\textsc{Post}_{pt}$, $\textsc{Const}_p$, and $\textsc{Post}_{pt} \times \textsc{Const}_p$ with a set of distance dummies I_{ipt}^d, where I_{ipt}^d equals one if unit i is located at distance d from the boundary of project p. Our estimating equation may also include project fixed-effects λ_p and a vector X_{ipt} of additional controls. The set of distances D partitions the land within and surrounding housing projects and varies across data sources. This approach therefore allows for difference-in-differences estimates to vary flexibly according to the geographic exposure to housing projects.

We interpret the coefficients of interest α^d as the causal impacts of subsidized housing on outcome y at distance d from a housing project. Specifically, α^d measures the differential change in outcomes for constructed projects, relative to unconstructed projects. A causal interpretation requires the assumption that absent the completion of a housing project, constructed and unconstructed areas would experience the same changes in outcomes. We discuss caveats below.

By comparing unconstructed versus constructed project areas, this method builds on previous studies of housing spillovers, which instead focus on comparing areas nearby to areas just further away from projects (Diamond and McQuade [2016]). The advantage of our method is that it does not require assuming how spillover effects dissipate with distance ex-ante; instead, we directly estimate these spillovers. It is particularly helpful to relax this assumption for evaluating South African housing programs since (1) the large sizes of projects suggest that spillover effects may extend for long distances and (2) program recipients may be drawn from wide areas around housing projects.

A central concern for our approach is that project construction may be driven by changes in local political or economic conditions that may also affect local housing development. For example, growing neighborhoods may generate greater political resistance to housing projects, delaying project construction despite improvements in local housing. Alternatively, worsening infrastructure quality may both increase the costs of finishing projects while also worsening local housing conditions. To the extent that these trends are reflected in average levels of local characteristics, we can evaluate potential sources of endogeneity by comparing baseline characteristics across constructed and unconstructed areas as in Section 5. However, we are unable to test for these factors if changes in local economic and political conditions happen at the same time as project construction.

Another potential concern is that forward-looking households anticipate investments in these areas and alter their housing investment decisions accordingly. For example, squatter households may be more likely to construct temporary, informal structures or locate elsewhere rather than invest in long-term dwellings in these areas. In this case, comparisons between constructed and uncontructed areas would overestimate the role of project construction in boosting formal housing development. Similarly, nearby housing markets may anticipate projects by either appreciating or depreciating in advance (depending on the anticipated amenity value of the projects). We are able to empirically examine these effects to some extent by closely observing dynamics in housing prices around the construction date. For the majority of our empirical approaches, we compare changes in outcomes between 2001 and 2011 so that our baseline time-period is likely occur well before households are aware of future project construction. This long-term comparison helps to ensure that anticipatory effects play a minimal role in biasing results. Qualitative evidence further suggests that these effects may be limited by the substantial uncertainty in project timing due to difficulties coordinating stakeholders and sources of funding [Tissington, 2011].[15]

5. Descriptive Statistics

To contextualize our main analysis, we first provide descriptive evidence on the location, size, and surrounding housing stock of our 117 sampled housing projects. Table 2 reveals that constructed projects provide generous quantities of new housing, delivering on average more than 890 units per project. Constructed projects are also substantially larger in area than unconstructed projects. As mentioned in section 3, these discrepancies may be due to many unconstructed projects representing smaller planned extensions to previously implemented projects. It may also reflect that larger land plots are first targeted by housing authorities to maximize delivery. Though unconstructed projects are located in slightly more remote locations, both types of projects impose long commutes on residents, as measured by the average distance to central business districts. These distances are consistent with housing authorities targeting inexpensive, vacant land for these projects. Finally, prices for non-subsidized housing in the vicinity of project boundaries are slightly lower for unconstructed projects, perhaps driven in part by their more distant locations.

Table 2. Housing Projects Areas Description

	Constructed	Unconstructed
Number of Projects	57	60
Area (km2)	3.37	2.45
Median Construction Year	2005	2007*
Delivered Houses	894.87	.
House Price within 1km (Rands†)	215,172.77	204,490.48
Distance to CBD‡ (km)	31.08	33.43

*Calculated from *expected* completion dates using Gauteng National Treasury budget reports.
† The USD averaged to about 7.70 Rands during the 2001-2011 period.
‡Measured as the average minimum distance with respect to Johannesburg and Pretoria CBDs.

Next, we turn to household census data to assess how both types of projects differ in terms of dwelling characteristics prior to project construction. We identify census small-areas as pertaining to a housing project when over 30% of their land areas over-

[15] Diamond and McQuade [2016] also leverage uncertainty in project timing for affordable housing in the US.

laps with a project boundary.[16] This procedure results in 1,210 small-areas containing a total of 257,997 households. Table 3 compares means for household indicators aggregated at the small-area level based on whether the census boundaries overlap with constructed (first column) or planned but unconstructed projects (second column). As a reference, the third column provides averages for each variable using all census small-areas in Gauteng in 2001.

Table 3. Dwelling Characteristics at Baseline from 2001 Census

	Constructed	Unconstructed	All Households
Flush Toilet	0.55	0.49	0.81
Piped Water in Home	0.12	0.21	0.47
Electricity for Cooking	0.32	0.39	0.73
Electricity for Heating	0.29	0.37	0.70
Electricity for Lighting	0.58	0.49	0.82
Number of Rooms	2.85	2.59	3.72
Household Size	3.53	3.25	3.32
N	924	286	10,711

Table 4. Dwelling Characteristics at Baseline from 2001 Census Het

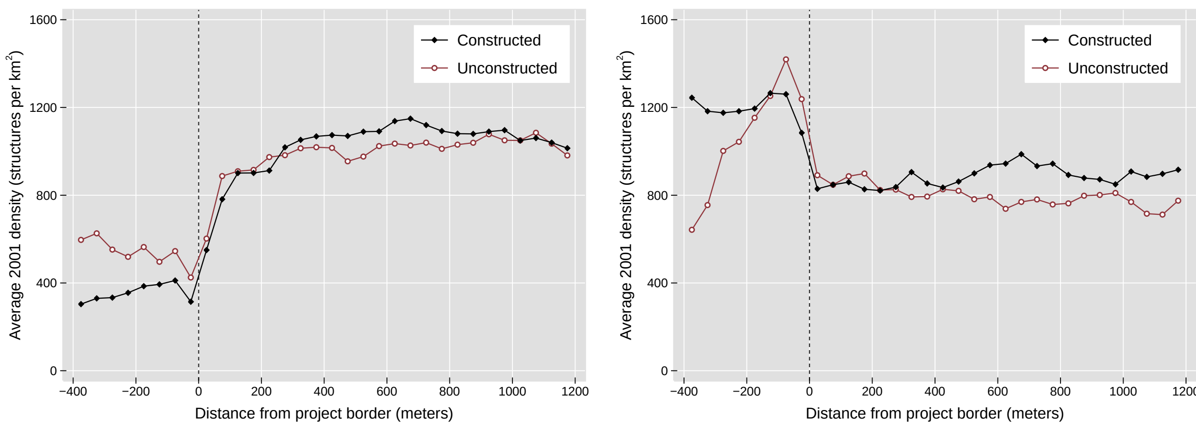
	Const Near	Unconst Near	Const Far	Unconst Far	All
Flush Toilet	0.66	0.41	0.45	0.56	0.81
Piped Water in Home	0.16	0.12	0.07	0.29	0.47
Electricity for Cooking	0.34	0.28	0.30	0.49	0.73
Electricity for Heating	0.31	0.28	0.27	0.46	0.70
Electricity for Lighting	0.54	0.31	0.63	0.64	0.82
Number of Rooms	2.42	2.23	3.25	2.90	3.72
Household Size	3.30	3.06	3.74	3.42	3.32
N	445	135	479	151	10,711

A clear takeaway from table 3 is that dwelling quality at baseline for both constructed and unconstructed project areas is far worse than the average Gauteng dwelling. Given the quality of land plots that were targeted by this policy, we expect project areas to lag behind other areas, and reflect more precarious housing conditions. We find that access to sanitation, water and electricity are between 30 and 70% lower in both types of project areas relative to the province as a whole. Comparing constructed to unconstructed projects, we find no systematic differences in dwelling quality. While

[16]Our census geography selection method is depicted in more detail in appendix A.2.

constructed areas are more likely to have flush toilets, electricity for lighting, and a larger number of rooms, households in unconstructed areas report having better access to piped water and electricity for heating and cooking. Despite some similarities between averages in constructed and unconstructed areas, a simple t-test rejects the null hypothesis that means are equal for every tabulated variable.

Figure 4. Housing Densities in Constructed and Unconstructed projects

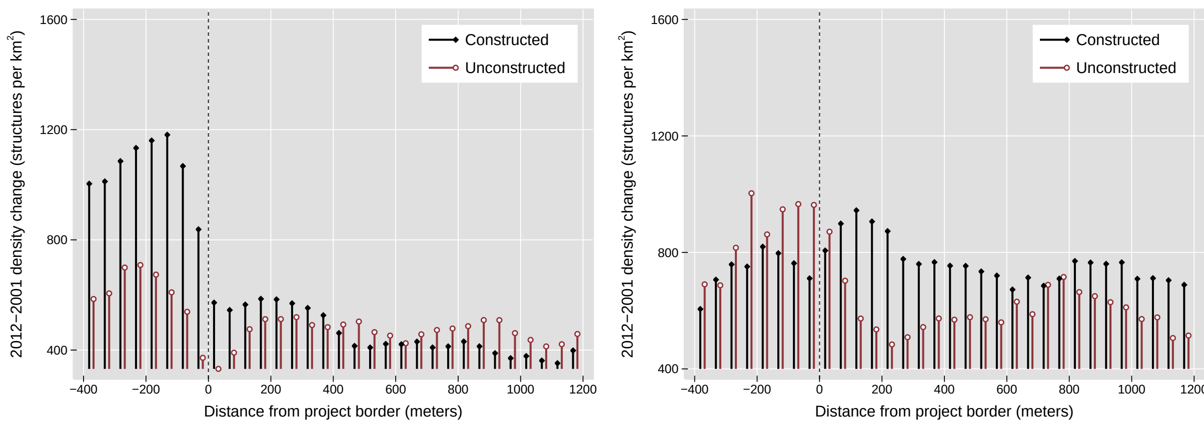


(a) pre-period formal housing density **(b)** pre-period informal housing density

Using the building-based land-use measures described in section 3.3, we then examine housing densities across constructed and unconstructed project areas. To inspect how densities change with proximity to housing projects, we calculate euclidean distances from the centroid of each 50×50m grid-cell to the nearest project boundary. We assign negative distances to grid-cells lying inside these boundaries. We then compute the average density at 50m intervals, reported in structures per square kilometer.

Figures 4a and 4b respectively plot average densities of formal and informal houses in the pre-period. In Figure 4a, we find low densities of formal housing within both unconstructed and constructed project areas. This finding is consistent with governments locating projects on underdeveloped land. Within project boundaries, unconstructed projects exhibit around 20 to 50% higher densities of formal housing, which suggests possibly higher costs of purchasing and converting this land into projects. Moving outwards, formal housing densities jump sharply at project boundaries and stabilize at higher levels across positive distances. Figure 4b reports an inverse pattern for informal housing, with high levels of informal housing sharply declining just outside of

Figure 5. Housing Densities in Constructed and Unconstructed projects



(a) formal housing density change

(b) informal housing density change

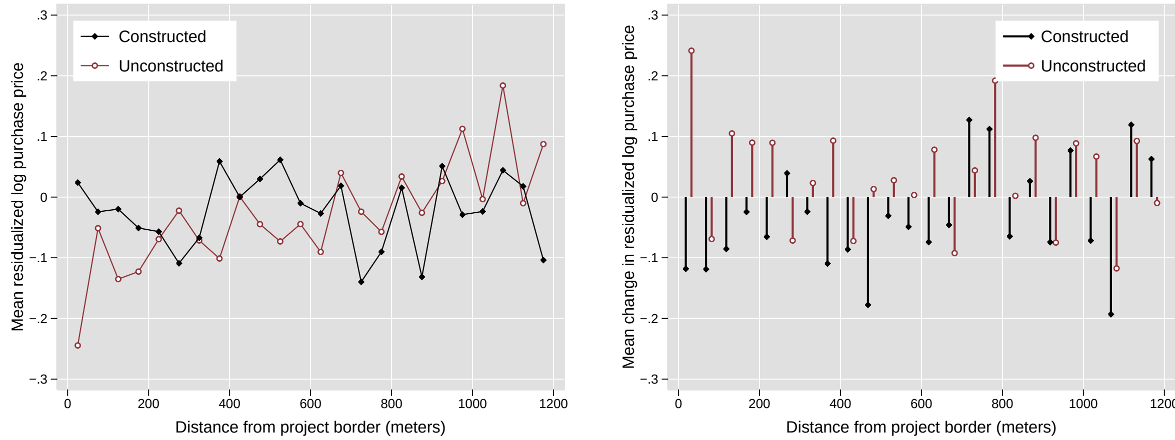
project boundaries.[17] Relative to the steep changes occurring at around distance zero, we find reasonably flat trends in housing density outside of project boundaries for all four data series. Similarities at baseline across treated and control projects at positive distances provide suggestive evidence that housing markets in both areas were following similar trajectories before the policy designation. This finding supports our identifying assumptions that later deviations in housing markets between constructed and unconstructed areas are solely due to project completion.

Figures 5a and 5b extend this exercise by plotting average changes in housing densities between 2001 and 2012. The effect of the policy is clearly evidenced in figure 5a, where tall black bars at negative distances show increases in the density of formal housing of around 1,000 structures per km^2 within constructed project areas. Compared to a baseline mean density of about 400 structures per km^2 in Figure 4a, this shift represents a nearly 250% increase. In contrast, unconstructed areas (red bars) experience increases in formal housing density at much smaller rates. At positive distances, both series are almost overlapping. This result previews our empirical analysis by showing little evidence of spillover effects on neighboring formal housing markets.

In figure 5b, we observe large increases in informal density within project boundaries for both constructed and unconstructed project areas. At some distances, un-

[17] Greater negative distances exclude smaller projects in computing average density, which may account for some of the variation in building density within projects, particularly for informal housing. Appendix A.4 illustrates this compositional change in more detail.

Figure 6. House Prices outside Constructed and Unconstructed projects



(a) pre-period log-purchase prices **(b)** change in log-purchase prices

constructed projects appear to overtake constructed areas in informal housing growth. Outside of project boundaries, constructed areas experience larger increases in informal density than unconstructed areas. With the exception of 0-200m, these increases do not appear to be systematically different between the treated and control ares. Relative to baseline levels in Figure 4b, informal housing increases are large, increasing by at least 50% across all distances and nearly doubling in some cases.

Figure 6 performs an analogous exercise with property transaction data from the deeds records. These records are likely to be composed of formal houses because informal housing transactions are rarely recorded. Since there are very few non-subsidized housing transactions within project footprints, we focus only on transactions at positive distances. Using the date of each transaction, we group prices before and after the expected completion dates for both constructed and unconstructed projects. To make comparisons within a normalized timeline, we first residualize log purchase prices from year×month and project fixed effects.

Figure 6a plots average residualized log purchase prices before the expected completion date for both constructed and unconstructed areas. We document an increase in prices of about 0.2 log-points or roughly 20% for unconstructed project areas. The picture is less clear for constructed projects, which show an increasing price gradient within the first 500m, but exhibit noisy patterns further away. Overall, both series show large variations in house prices, which make it difficult to identify distinct patterns. The steady increase observed for unconstructed project areas is consistent with findings in table 3 that project areas (both constructed and unconstructed) fare worse in terms of indicators of dwelling quality.

Finally, figure 6b graphs average changes in log-purchase prices before and after the expected completion of projects. Both series are once more quite noisy. Under 400 meters, we observe some evidence that unconstructed areas experience a boom in nearby housing prices that was not mirrored by constructed areas. Yet, this gap quickly closes, further suggesting that housing projects play a minimal role in driving nearby housing markets.

6. Estimation Results

To draw statistical inference on the patterns depicted in section 5, we proceed to estimate variants of equation (1) for each outcome of interest.

6.1. Infrastructure and Demographic Effects

To assess impacts on dwelling quality, we pool the 2001 and 2011 census cross-sections and test for differential 10-year changes between constructed and unconstructed project areas. Because these data have coarser spatial resolutions, we perform this comparison at two levels of exposure to the policy: $D = \{project, spillover\}$. As in table 3, *project* small-areas overlap with more than 30% of their area with project boundaries. The remaining *spillover* small-areas have less than 30% area overlap, but their centroids fall within 1.5 km of a project boundary. Appendix A.2 illustrates an example of these two definitions for a given project. Our estimating equation is:

$$y_{ipt} \; = \; \lambda_p + \sum_{\substack{d \in \{project, \\ spillover\}}} I_{ip}^d \Big(\alpha^d \text{POST}_{pt} \times \text{CONST}_p + \beta^d \text{POST}_{pt} + \gamma^d \text{CONST}_p \Big) + \theta I_{ip}^{project} + \varepsilon_{ipt} \tag{2}$$

where y_{ipt} denotes the census outcome for small-area i in vicinity of (either constructed or planned but unconstructed) project p, observed in census $t \in \{2001, 2011\}$. The parameter $\alpha^{project}$ ($\alpha^{spillover}$) expresses the difference-in-differences effect of the public

Table 5. Census Household-level Estimates

	(1) Flush Toilet	(2) Water Indoors	(3) Electricity Cooking	(4) Electricity Heating	(5) Electricity Lighting	(6) Number of Rooms	(7) Household Size	(8) Population Density
project × post × constr	0.115^{c}	0.164^{a}	0.259^{a}	0.181^{a}	0.066	-0.074	-0.266^{b}	99.563
	(0.069)	(0.050)	(0.072)	(0.066)	(0.084)	(0.173)	(0.107)	(1448.879)
project × post	0.064	0.090^{b}	0.176^{a}	0.151^{a}	0.144^{b}	0.408^{a}	-0.039	2320.431^{b}
	(0.046)	(0.039)	(0.060)	(0.055)	(0.066)	(0.129)	(0.056)	(1111.722)
project × constr	0.070	-0.024	-0.066	-0.061	0.129	0.195	0.460^{a}	-611.013
	(0.111)	(0.090)	(0.105)	(0.093)	(0.125)	(0.284)	(0.128)	(1733.494)
project	-0.321^{a}	-0.239^{a}	-0.354^{a}	-0.320^{a}	-0.362^{a}	-1.065^{a}	-0.403^{a}	416.265
	(0.077)	(0.053)	(0.069)	(0.064)	(0.080)	(0.175)	(0.082)	(691.463)
spillover × post × constr	0.013	0.045	0.028	-0.037	-0.038	0.053	-0.153^{a}	444.064
	(0.034)	(0.033)	(0.034)	(0.043)	(0.027)	(0.078)	(0.049)	(470.258)
spillover × post	0.039	0.133^{a}	0.101^{a}	0.082^{a}	0.058^{b}	0.249^{a}	-0.156^{a}	652.913^{b}
	(0.024)	(0.028)	(0.024)	(0.023)	(0.023)	(0.055)	(0.036)	(325.594)
spillover × constr	-0.029	-0.056	-0.065	-0.041	0.015	-0.252^{c}	0.143^{b}	-23.442
	(0.049)	(0.055)	(0.044)	(0.042)	(0.050)	(0.145)	(0.059)	(1353.816)
p-val, h_0: project=spill.	0.119	0.029	0.001	0.001	0.208	0.439	0.250	0.811
Mean Outcome 2001	0.72	0.31	0.60	0.57	0.75	3.31	3.59	7,313.26
Mean Outcome 2011	0.79	0.51	0.80	0.69	0.82	3.57	3.22	9,118.36
R^2	0.336	0.318	0.363	0.315	0.294	0.368	0.443	0.398
# projects	117	117	117	117	117	117	117	117
N project areas	3,631	3,631	3,631	3,631	3,631	3,630	3,632	3,632
N spillover areas	5,978	5,978	5,978	5,978	5,978	5,964	5,977	5,980
N	9,609	9,609	9,609	9,609	9,609	9,594	9,609	9,612

All regressions include project Fixed-Effects. Standard errors clustered at the project level in parenthesis. ^{c} p<0.10, ^{b} p<0.05, ^{a} p<0.01

Table 6. Census Household-level Estimates Het

	(1) Flush Toilet	(2) Water Indoors	(3) Electricity Cooking	(4) Electricity Heating	(5) Electricity Lighting	(6) Number of Rooms	(7) Household Size	(8) Population Density
near × proj	0.118	0.094	0.215^{c}	0.119	0.029	-0.198	-0.374^{b}	623.207
	(0.092)	(0.081)	(0.110)	(0.098)	(0.115)	(0.264)	(0.152)	(2214.128)
near × spill	0.021	-0.006	-0.046	-0.124^{b}	-0.090^{b}	0.018	-0.177^{a}	873.022
	(0.043)	(0.042)	(0.038)	(0.050)	(0.036)	(0.115)	(0.060)	(654.256)
far × proj	0.088	0.220^{a}	0.280^{a}	0.218^{a}	0.084	-0.002	-0.171^{c}	-275.692
	(0.081)	(0.044)	(0.082)	(0.079)	(0.114)	(0.222)	(0.102)	(295.380)
far × spill	0.016	0.122^{b}	0.141^{a}	0.098^{b}	0.033	0.097	-0.118^{c}	-221.692
	(0.057)	(0.054)	(0.037)	(0.045)	(0.036)	(0.092)	(0.070)	(390.402)
p-val, h_0: near proj = near spill	0.256	0.238	0.020	0.016	0.297	0.381	0.146	0.916
p-val, h_0: far proj = far spill	0.354	0.126	0.088	0.137	0.640	0.641	0.637	0.909
R^2	0.344	0.324	0.374	0.329	0.300	0.370	0.445	0.407
N near proj areas	1,909	1,909	1,909	1,909	1,909	1,908	1,910	1,910
N near spill areas	3,672	3,672	3,672	3,672	3,672	3,659	3,670	3,673
N far proj areas	1,722	1,722	1,722	1,722	1,722	1,722	1,722	1,722
N far spill areas	2,306	2,306	2,306	2,306	2,306	2,305	2,307	2,307

All regressions include project Fixed-Effects. Standard errors clustered at the project level in parenthesis. ^{c} $p<0.10$, ^{b} $p<0.05$, ^{a} $p<0.01$

housing policy in *project* (*spillover*) small-areas, using unconstructed projects as the control group. We estimate equation (2) using the same binary outcomes tabulated in table 3, with the addition of a population density variable. We construct this additional outcome by computing the sum of household sizes and dividing it by the square-kilometer area of the census small-area. To allow for arbitrary correlations between observations within the same project area, we cluster standard errors at the project level.

Table 5 provides estimates for each outcome. Rows 1 through 4 describe effects for households in our *project* exposure definition, while rows 5 through 7 describe effects for our *spillover* definition. Rows 3 and 7 respectively show that for both *project* and *spillover* exposures, the model cannot reject the null that treated and control areas are equal at baseline (in 2001). The difference-in-differences coefficients in the first row (project $\times$ post $\times$ const) document statistically significant improvements in access to services such as flush toilets, indoor water taps, and electricity.[18] With access rates increasing by over 10 percentage points, the economic magnitudes of these estimates are large, consistent with the policy successfully providing better housing. We note that households living within *project* exposure areas need not be direct recipients of government houses, since our definition is inclusive of indirect recipients such as backyard shack dwellers as well.

We do not find corresponding effects for households in the vicinity of constructed projects. With the exception of number of rooms in column (6), row 5 of table 5 detects no statistically significant effects for outcomes in *spillover* areas. Though 95% confidence intervals admit sizeable positive effects, we can rule in most cases that this effect equals its *project* counterpart.[19] These results suggest that households adjacent to constructed projects are not incentivized to invest in higher housing quality.

6.2. Housing Density Effects

To assess impacts on housing densities, we pool both cross-sections of our gridded density measures, and test for differential 11-year changes between constructed and unconstructed project areas. Drawing from figure 4, we define the set of distances D as

[18] With the exception of electricity for lighting, which falls short of the 10% significance level.

[19] Row 8 of table 5 tests the hypothesis that the *project* and *spillover* DID effects are equal, and reports p-values from the associated F-test.

containing every 50m distance intervals from 400m inside project boundaries to 1200m outside project boundaries. We estimate:

$$\begin{aligned} y_{ipt} = \lambda_0 &+ \sum_{d \in D \setminus \{1200m\}} I_{ip}^d \Big(\alpha^d \textsc{Post}_{pt} \times \textsc{Const}_p + \beta^d \textsc{Post}_{pt} + \gamma^d \textsc{Const}_p + \theta^d \Big) \\ &+ \alpha^o \textsc{Post}_{pt} \times \textsc{Const}_p + \beta^o \textsc{Post}_{pt} + \gamma^o \textsc{Const}_p + \varepsilon_{ipt} \end{aligned} \tag{3}$$

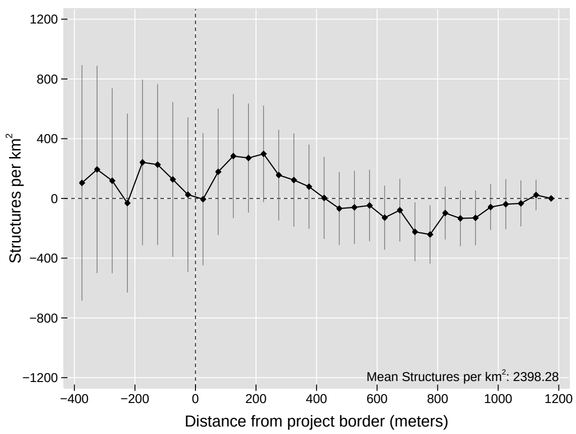
where units i now denote 50×50m grid cells and $t \in \{2001, 2012\}$. Equation (3) departs from (1) in that coefficients α^d are expressed relative to the DID effect in the furthest distance bin, α^{1200m}, which is normalized to zero. This rearrangement introduces a third layer of differencing, comparing double-differences near versus far from project boundaries. A triple-differences strategy allows for differential trends between constructed and unconstructed projects, so long as these trends are identical $\forall d \in D$ in the absence of project construction. It further requires both that (1) any spillover effects fully dissipate within 1,200 meters from project boundaries and (2) parallel trends hold between 1,200 meters and all other closer distance bins.

We estimate impacts on five housing density measures: 1) total houses, 2) formal houses, 3) informal houses, which are divided into 4) backyard shacks and 5) informal non-backyard houses. Figure 7 plots the α^d coefficients across all outcomes for distance bins within projects (negative distances) and outside projects (positive distances). We accompany figure 7 with table 7, which provides point-estimates and standard errors for wider 400m distance bins relative to the omitted category [400m,1200m]. As in our previous estimation, we cluster standard at the project level for all regressions.

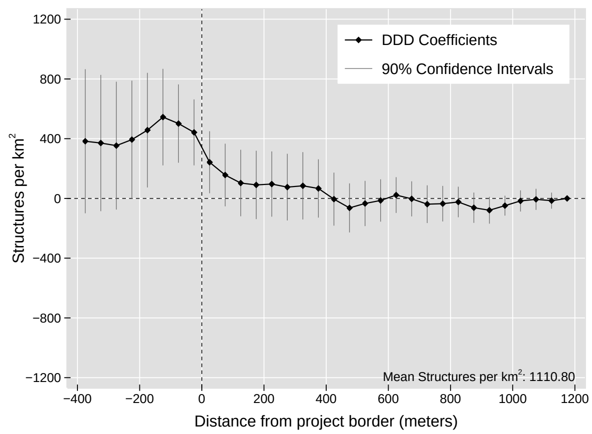
Figure 7a finds weak increases in total housing within and just outside ($\leq$ 200 meters) of projects before settling to zero at all greater distances. Predicting increases of around 200 houses per kilometer near the project boundary, these coefficients are both statistically and economically insignificant given an average density of nearly 2,400 structures per kilometer.

After measuring substantial numbers of houses per project in the descriptive evidence (Table 2), it is surprising to find little evidence of housing spikes in these areas; however, the small effects on aggregate housing density mask larger shifts in the composition of housing. Figure 7b estimates over 30% increases in formal housing density

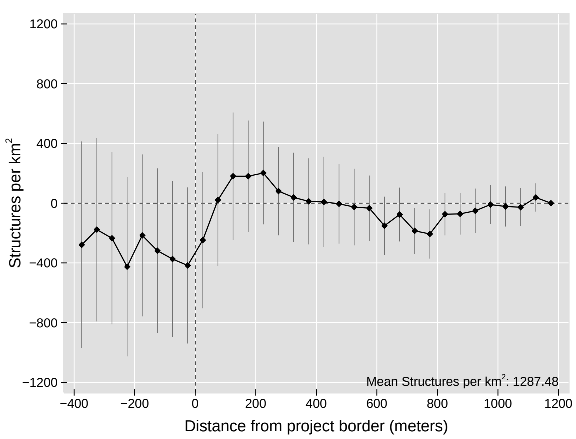
(a) Total Houses



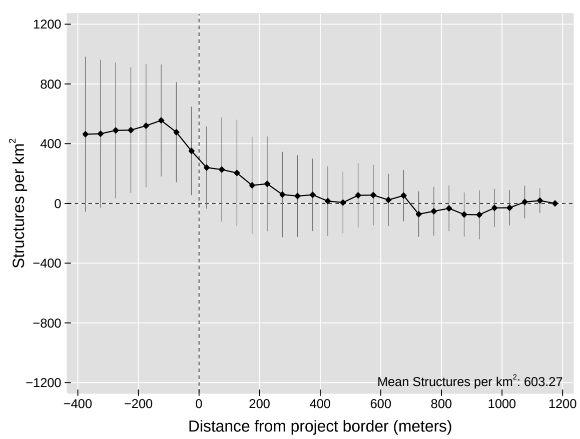
(b) Formal Houses



(c) Informal Houses



(d) Backyard Informal Houses



(e) Non-Backyard Informal Houses

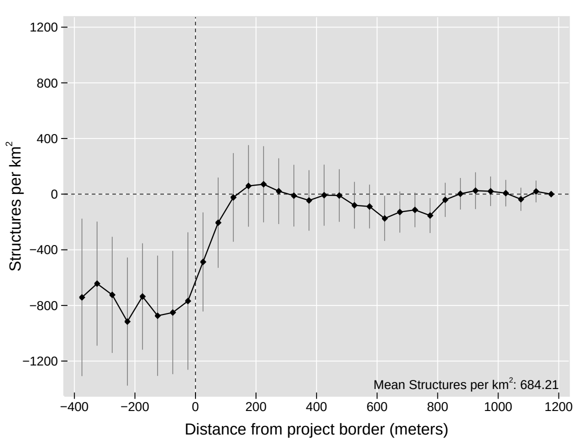
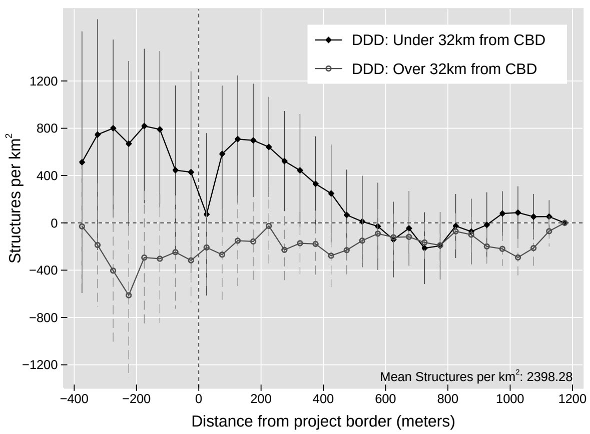
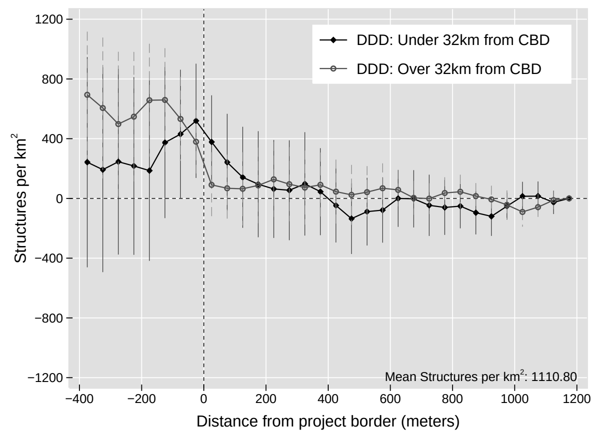


Figure 7. DDD coefficients (equation 3) for fives types of housing densities.

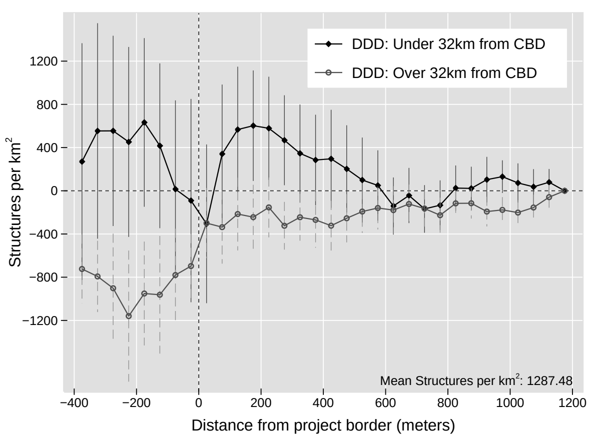
(a) Total Houses



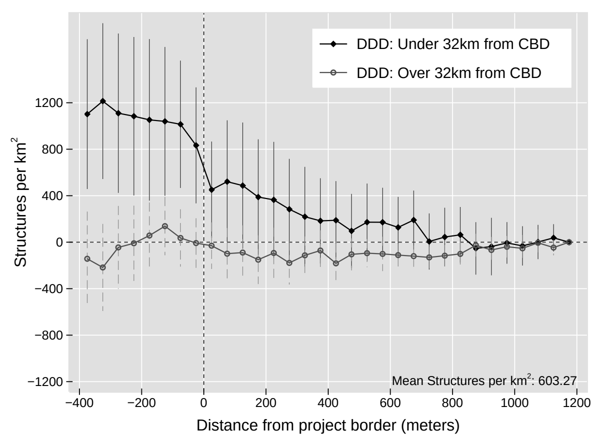
(b) Formal Houses



(c) Informal Houses



(d) Backyard Informal Houses



(e) Non-Backyard Informal Houses

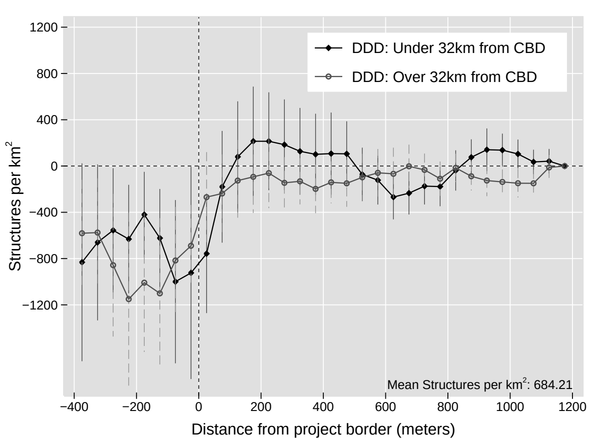


Figure 8. DDD coefficients (equation 3) for fives types of housing densities.

within project areas, which quickly dissipates just beyond their boundaries. Coefficients within 200 meters inside the boundary are statistically different from those furthest away at the 10% level. Gains in formal housing density are almost entirely offset by losses in informal housing density as demonstrated by Figure 7c. Informal housing density also seems to increase slightly just outside of project areas before stabilizing at zero growth further away. Disaggregating informal housing further between backyard houses (Figure 7d) and non-backyard houses (Figure 7e) indicates that massive declines in non-backyard housing (of around 100% of the average) are only partially offset by an almost doubling of the densities of backyard housing within projects. Backyard housing growth extends slightly outside of the project boundary, accounting for the small increases in total housing observed in these regions (Figure 7a).

To assess the magnitude and statistical significance of these effects, Table 7 provides estimates for a simple specification that recovers triple-difference coefficients for areas up to 400 meters within and outside of project boundaries. This specification pools all areas over 400 meters outside project boundaries in the control group. We find small, statistically insignificant changes in total housing and informal housing but economically large effects that are statistically significant at the 1% level for the three remaining outcomes: formal housing increases by around 45%, backyard housing increases by nearly 80%, and non-backyard informal housing decreases by over 100% of its overall mean.

Table 7. Triple Difference Estimates

	(1) Total Housing	(2) Formal Housing	(3) Informal Housing	(4) Backyard Housing	(5) Non-Bkyrd Housing
-400m to 0m	204.94	497.78^a	-292.85	477.22^a	-770.07^a
	(275.63)	(185.68)	(272.37)	(178.33)	(231.27)
0m to 400m	229.11	128.62	100.49	107.76	-7.27
	(139.03)	(95.00)	(140.45)	(117.71)	(120.67)
Mean dep. var.	2,398.28	1,110.80	1,287.48	603.27	684.21
# Projects	117	117	117	117	117
R^2	0.098	0.113	0.057	0.100	0.048
N	307,196	307,196	307,196	307,196	307,196

Standard errors clustered at the project level in parenthesis. ^{c} $p<0.10$, ^{b} $p<0.05$, ^{a} $p<0.01$

Table 8. Triple Difference Estimates

	(1) Total Housing	(2) Formal Housing	(3) Informal Housing	(4) Backyard Housing	(5) Non-Bkyrd Housing
-400m to 0m near CBD	637.71[c]	432.37	205.34	973.45[a]	-768.11[a]
	(344.29)	(290.02)	(417.34)	(288.32)	(286.63)
0m to 400m near CBD	516.36[a]	200.50	315.86[c]	283.26	32.60
	(164.51)	(147.61)	(187.20)	(178.12)	(179.56)
-400m to 0m far CBD	-170.91	514.55[b]	-685.46[b]	88.11	-773.57[b]
	(285.42)	(212.19)	(268.32)	(145.24)	(344.01)
0m to 400m far CBD	-64.38	41.62	-106.01	-66.09	-39.92
	(155.49)	(103.82)	(126.08)	(107.36)	(139.06)
Mean dep. var.	2,398.28	1,110.80	1,287.48	603.27	684.21
# Projects Near	68	68	68	68	68
# Projects Far	59	59	59	59	59
R^2	0.126	0.123	0.082	0.124	0.056
N	307,196	307,196	307,196	307,196	307,196

"Near" is within 32 km from the CBD and "Far" is greater than 32km from the CBD.
Standard errors clustered at the project level in parenthesis. [c] $p<0.10$,[b] $p<0.05$,[a] $p<0.01$

6.3. Price Effects

To identify the spillover effects of public housing on local housing markets, we adopt the specification from Section 6.2, instead using transaction price as an outcome:

$$\begin{aligned} y_{ipt} &= \lambda_0 + \sum_{d \in D \setminus \{1200m\}} I^d_{ip}\Big(\alpha^d \textsc{Post}_{pt} \times \textsc{Const}_p + \beta^d \textsc{Post}_{pt} + \gamma^d \textsc{Const}_p + \theta^d\Big) \\ &\quad + \alpha^o \textsc{Post}_{pt} \times \textsc{Const}_p + \beta^o \textsc{Post}_{pt} + \lambda_p + \eta_t + \delta X_i + \varepsilon_{ipt} \end{aligned} \tag{4}$$

The outcome, y_{itp}, is measured in terms of the log-purchase price of property i sold at time t in the vicinity of project p. Since we observe the exact transaction date, we set $\textsc{Post}_{pt}$ =1 if date t is after the month of project implementation and zero otherwise. To account for smaller sample sizes than in the density estimation, we consider wider 200 meter distance bins from 0 to 1200 meters. Additionally, λ_p includes a project fixed effect, controlling for any fixed, unobserved drivers of house prices that vary between projects.[20] η_t controls for year, year-month, and calendar month fixed effects

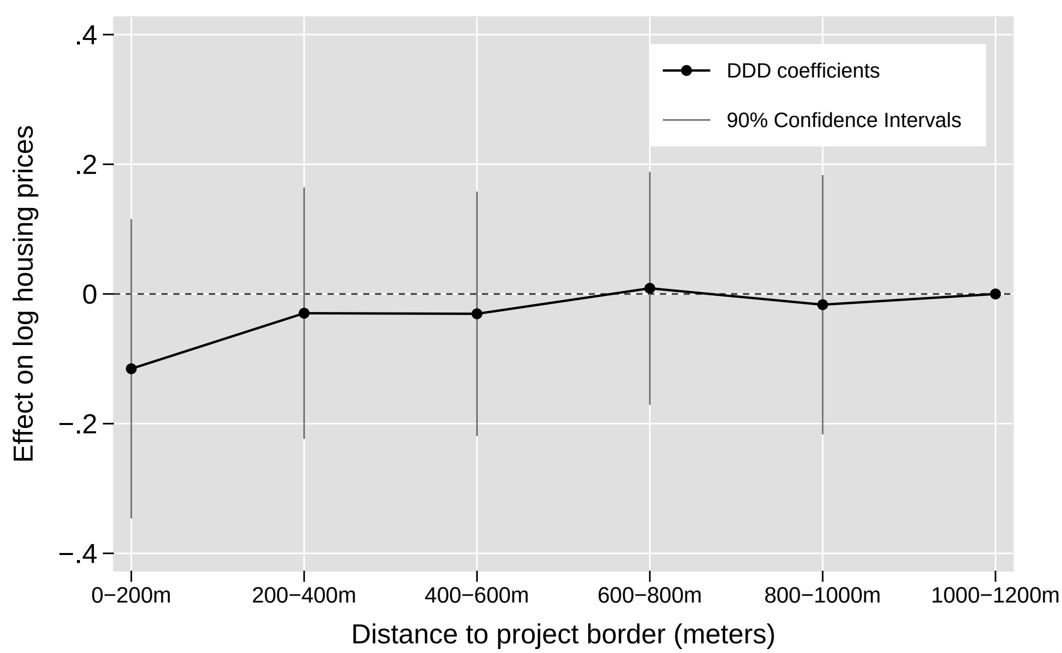
[20]The project fixed effect subsumes the $\textsc{Const}_p$ effect since construction status does not vary within project over time.

depending on the specification to account for any factors such as shifts in aggregate housing demand that may be correlated with prices and the timing of housing projects. To control for property characteristics, X_i includes up to a cubic polynomial in lot size and a control for latitude-longitude cells in some specifications. Similarly to section 6.2, coefficients α^d express double-difference effects relative to effects within 1,000 to 1,200 meters of projects, again allowing for a triple-differences interpretation.

Figure 9 plots coefficients α^d and associated standard errors for every $d \in D$. This specification includes project as well as calendar month fixed effects and includes only transactions that occur within three years of project construction. Examining the point estimates, we find evidence of a drop in prices of around 0.1 log points or around 10% within 200 meters of the project boundary. Beyond 200 meters, this negative price effect disappears with all point estimates closely tracking zero. Clustering standard errors at the project-level leaves wide confidence intervals at all distances, meaning that we are unable to reject equivalence between any coefficient and the price effect for the reference category at the 10% confidence level.

Table 9 provides point estimates for distance bins up to 600 meters across a variety of controls and using a broader reference category including all distances from 600

Figure 9. Price Estimates over Distance from Project



meters to 1,200 meters. The first row confirms a price drop of around 10% that is relatively stable as we include additional controls for project, time, and location. We find much smaller but similarly negative price effects between 200 and 600 meters. Wide standard errors ensure that no coefficients are statistically different from zero. Due to the noisiness of the data, we are unable to rule out large effects at all distances.

Table 9. Triple Difference Estimates on Log-Prices

	(1)	(2)	(3)	(4)
0 to 200m	-0.093	-0.116	-0.067	-0.142
	(0.131)	(0.120)	(0.134)	(0.149)
200m to 400m	-0.026	-0.031	-0.002	-0.065
	(0.116)	(0.097)	(0.106)	(0.113)
400m to 600m	-0.029	-0.031	-0.020	-0.042
	(0.125)	(0.105)	(0.114)	(0.089)
Cubic in lot size	✓	✓	✓	✓
Project FE	.	✓	.	.
Year × Project FE	.	.	✓	.
Year × Lat-Lon cell FE	.	.	.	✓
Year-Month FE	✓	✓	.	.
Month FE	.	.	✓	✓
R^2	0.177	0.331	0.354	0.279
N	87,253	87,253	87,253	87,253

Standard errors clustered at the project level in parenthesis. ^{c} $p<0.10$, ^{b} $p<0.05$, ^{a} $p<0.01$

We look into the dynamics of these effects by substituting the $\textsc{Post}_{pt}$ term with year dummies relative to the date of construction. We estimate these trends separately for four different distance bins relative to the project boundary. Figure 10 plots these coefficients, excluding the confidence intervals. Further splitting the sample for this exercise widens the confidence intervals so that all 90% intervals contain zero. We set the reference category to be two-years prior to project construction for the furthest distance bin. Figure 10 shows some evidence of anticipation, with prices level declining one year before construction at all distances. After project construction, prices appear to gradually recover.

Table 10. Triple Difference Estimates on Log-Prices Het

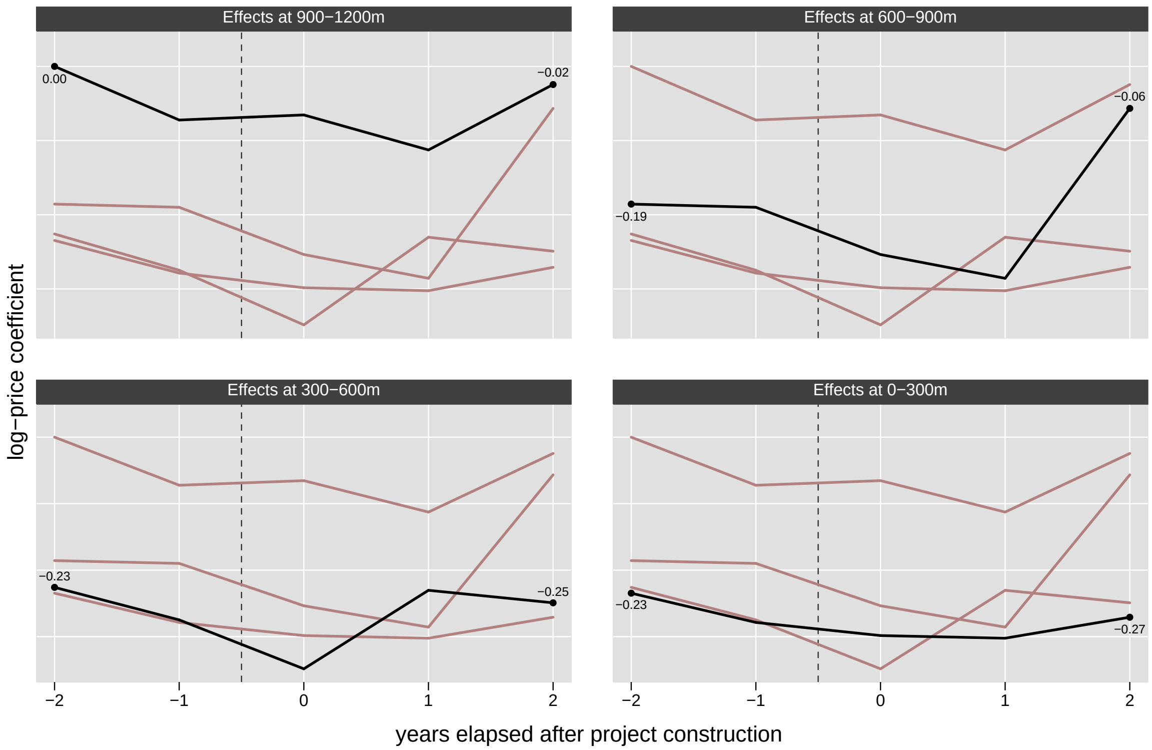
	(1)	(2)	(3)	(4)
Near CBD 0 to 200m	-0.067	-0.073	-0.059	-0.178
	(0.198)	(0.197)	(0.218)	(0.201)
Near CBD 200m to 400m	-0.093	-0.105	-0.100	-0.206
	(0.162)	(0.119)	(0.122)	(0.132)
Near CBD 400m to 600m	-0.220	-0.228	-0.242	-0.291
	(0.172)	(0.125)	(0.131)	(0.125)
Far from CBD 0 to 200m	-0.538	-0.183	0.020	-0.413
	(0.252)	(0.196)	(0.223)	(0.292)
Far from CBD 200m to 400m	-0.179	0.188	0.306	0.019
	(0.251)	(0.191)	(0.195)	(0.130)
Far from CBD 400m to 600m	-0.162	0.248	0.358	0.043
	(0.264)	(0.272)	(0.250)	(0.297)
Cubic in lot size	✓	✓	✓	✓
Project FE	.	✓	.	.
Year × Project FE	.	.	✓	.
Year × Lat-Lon cell FE	.	.	.	✓
Year-Month FE	✓	✓	.	.
Month FE	.	.	✓	✓
R^2	0.202	0.334	0.356	0.289
N	87,253	87,253	87,253	87,253

Standard errors clustered at the project level in parenthesis. ^{c} p<0.10,^{b} p<0.05,^{a} p<0.01

6.4. Population Composition

To inspect whether housing projects induce population sorting within and across project areas, we test how average sociodemographics change as a result of their construction. Re-estimating equation (2) using person-level responses from the 2001 and 2011 censuses, table 11 provides the differences-in-differences results for years of age, an indicator for place of birth outside of Gauteng province, an indicator for unemployment, years of education, and monthly income in Rands. We detect no statistically significant effects on average age and being born outside of Gauteng for both *project* and *spillover* areas. However, we find that unemployment rates improve by 3.5 and 5.4 percentage points for project and spillover areas respectively. Similarly, average years of education and monthly income both increase at comparable magnitudes in both exposure definitions. The unemployment and education coefficients are statistically significant for locations within and adjacent to project boundaries. These findings suggest that

Figure 10. Time-to-Event Price Estimates



more educated, higher income households are moving to constructed areas relative to unconstructed project areas. To check indirectly whether specific segments of the population move into new projects from locations nearby, row 8 of table 11 tests the equality of the *project* and *spillover* DID effects. If these "lateral relocations" are happening at high rates, one would expect to find coefficients of opposite signs. The model cannot reject the equality of both coefficients for every outcome.

7. Conclusion

Table 11. Effect of Housing Projects on Socio-demographics

	(1) Age	(2) P.O.B. not Gauteng	(3) Unemployed	(4) Years of Education	(5) Monthly Income
project × post × constr	0.064	-0.053	-0.035[c]	0.547[a]	597.627
	(0.299)	(0.059)	(0.021)	(0.149)	(406.408)
project × post	0.889[a]	0.003	-0.110[a]	1.105[a]	21.065
	(0.259)	(0.017)	(0.014)	(0.097)	(307.375)
project × constr	-0.777	0.013	0.034	-0.652[b]	-388.495
	(0.601)	(0.055)	(0.027)	(0.257)	(770.437)
project	-1.629[a]	0.126[a]	0.093[a]	-0.966[a]	-1206.783[a]
	(0.297)	(0.033)	(0.019)	(0.159)	(387.258)
spillover × post × constr	0.150	-0.017	-0.054[a]	0.411[a]	388.582
	(0.220)	(0.029)	(0.017)	(0.120)	(426.674)
spillover × post	1.173[a]	0.017	-0.080[a]	0.891[a]	1697.582[a]
	(0.170)	(0.011)	(0.013)	(0.107)	(351.209)
spillover × constr	-0.746[b]	0.014	0.079[a]	-0.663[a]	-1212.462[b]
	(0.361)	(0.027)	(0.016)	(0.127)	(558.348)
p-val, h_0: project=spill.	0.801	0.335	0.366	0.393	0.694
Mean Outcome 2001	26.57	0.42	0.48	8.08	2,225.65
Mean Outcome 2011	27.72	0.44	0.33	9.57	4,044.49
R^2	0.395	0.529	0.291	0.511	0.315
# projects	117	117	117	117	117
N project areas	3,632	3,632	3,631	3,631	3,631
N spillover areas	5,980	5,976	5,975	5,976	5,975
N	9,612	9,608	9,606	9,607	9,606

Standard errors clustered at the project level in parenthesis. [c] $p<0.10$, [b] $p<0.05$, [a] $p<0.01$
P.O.B. means "place of birth." Monthly income is in Rands.

References

J. Beall, O. Crankshaw, and S. Parnell. Social differentiation and urban governance in greater soweto. *Emerging Johannesburg: Perspectives on the postapartheid city*, page 197e214, 2003.

J. K. Brueckner, C. Rabe, and H. Selod. Backyarding in south africa. *Working Paper*, 2018.

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.

M. Gechter and N. Tsivanidis. The welfare consequences of formalizing developing country cities: Evidence from the mumbai mills redevelopment. 2018.

M. Harari and M. Wong. Slum upgrading and long-run urban development: Evidence from indonesia. In *2018 Meeting Papers*, number 367. Society for Economic Dynamics, 2018.

R. Hornbeck and D. Keniston. Creative destruction: Barriers to urban growth and the great boston fire of 1872. *American Economic Review*, 107(6):1365–98, 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/.

G. Michaels, D. Nigmatulina, F. Rauch, T. Regan, N. Baruah, and A. Dahlstrand-Rudin. Planning ahead for better neighborhoods: Long run evidence from tanzania. Technical report, CEPR Discussion Papers, 2017.

E. Rossi-Hansberg, P.-D. Sarte, and R. Owens III. Housing externalities. *Journal of political Economy*, 118(3):485–535, 2010.

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 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.

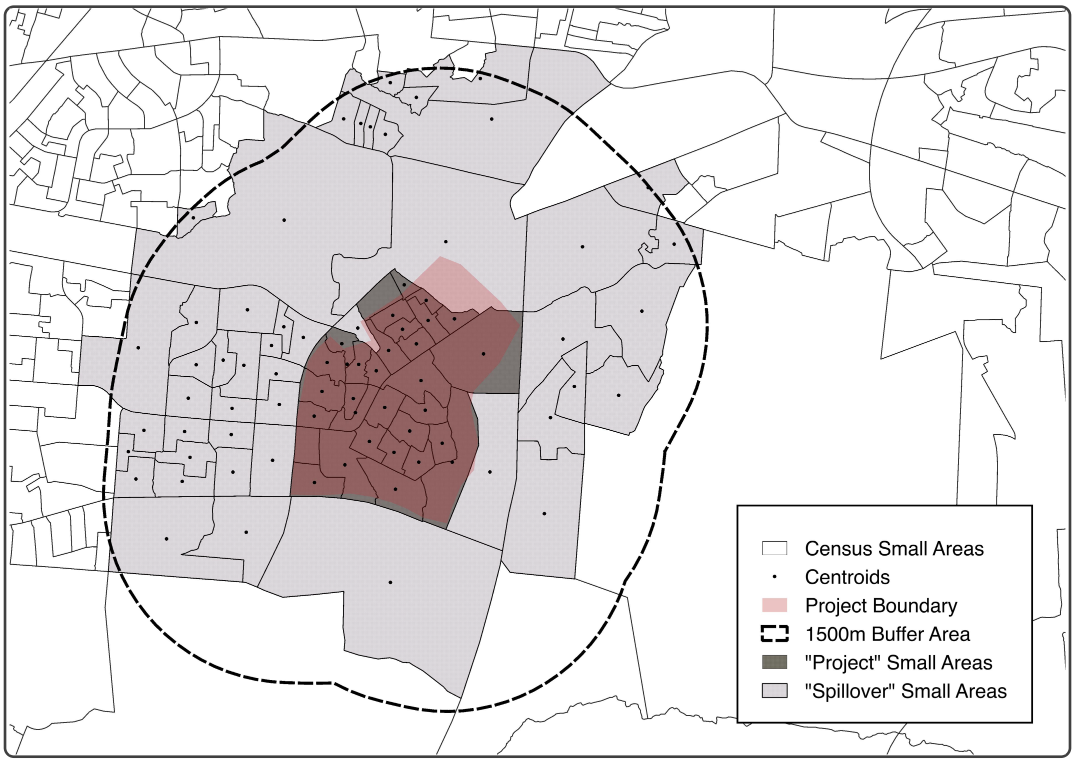
United Nations Habitat Program. Housing and slum upgrading. Technical report, 2018. URL https://unhabitat.org/urban-themes/housing-slum-upgrading/.

A. Appendix

A.1. String Matching for Project Names and Delivery Dates

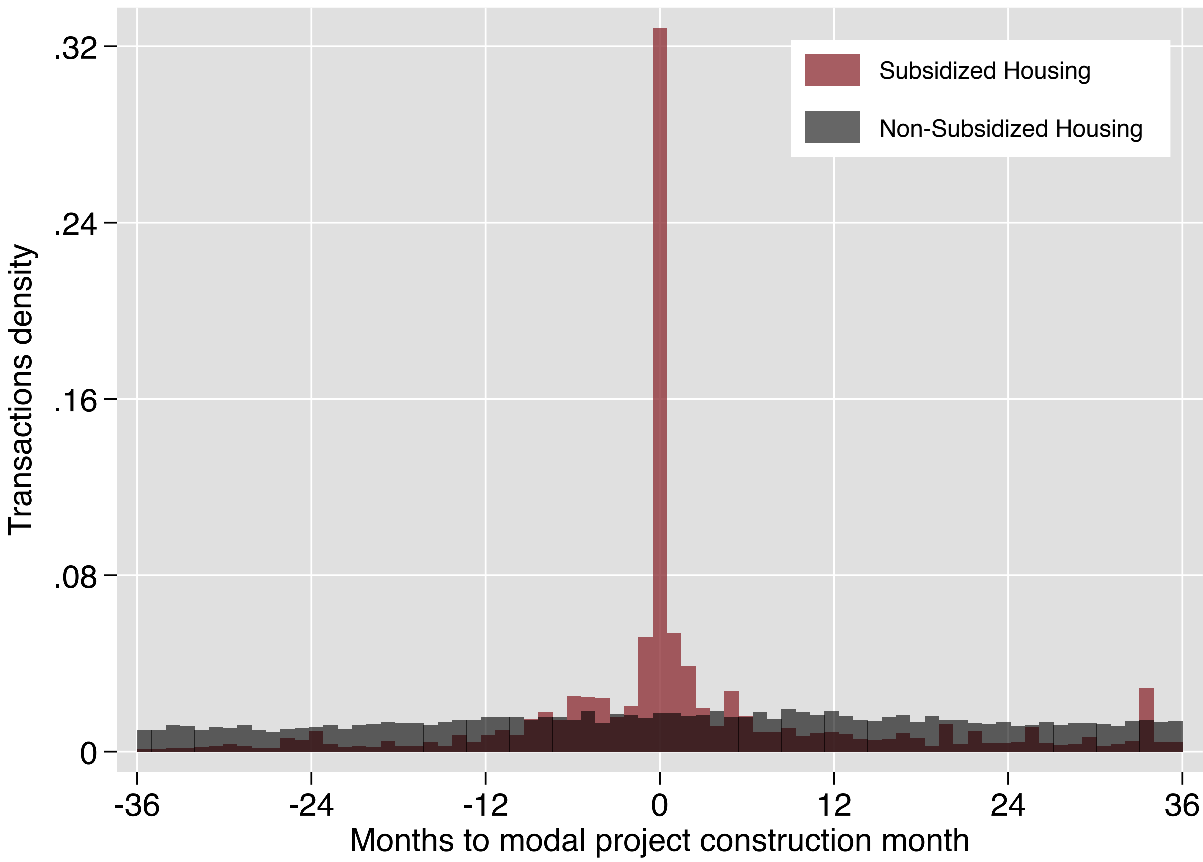
[...]

A.2. Buffer Design for Census Small Areas



Note: We consider every census small-areas whose centroid fall within 1500m of a project boundary. We denote small-areas with over 30% area overlap with project boundaries as ″project″ small-areas, and use household responses within these geographies to capture dwelling characteristics inside project boundaries. All remaining small-areas are denoted as ″spillover″ geographies, and are used to capture dwelling characteristics nearby.

A.3. Transaction Frequency



A.4. Project Counts