

Subsidized Housing and Urban Development: Evidence from South Africa*

Benjamin Bradlow[†]
Stefano Polloni[‡]
William Violette[§]

This Version: October 31, 2018

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, housing density, and census data. To identify project impacts, we compare constructed and unconstructed projects at fine geographic levels using a difference-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 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 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

In developing countries, 30% of urban populations live in crowded slums where households often suffer from unsanitary conditions and high crime rates [United Nations, 2015]. Combined with low infrastructure access and insecure property rights, slum 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. These programs aim to provide not only direct health and economic benefits to recipients, but also greater incentives for neighbors to invest in their homes and communities. Local investments may reduce negative externalities and steering residents away from poverty traps. At the same time, subsidized housing may attract slums by improving access to bulk services like water and sanitation as well as providing under-utilized developable land within project areas. These indirect housing opportunities are especially salient in South Africa where *backyarding* is a common and widely documented form of land use.¹ In this way, public housing programs may ultimately exacerbate the same negative externalities they were designed to remediate. While much of the existing literature has focused on estimating direct recipient impacts [Cattaneo et al., 2009, Franklin, 2016, Galiani et al., 2017], the broader consequences of subsidized housing for urban development remain poorly understood.

We estimate economic spillovers from a large-scale housing program in South Africa, focusing on the quality, density, and price of nearby housing. Providing over 3 million houses, this program is one of the largest in the developing world and continues to respond to large backlogs in demand [Department of Human Settlements, 2012, 2015]. According to the Department of Human Settlements, this program is designed not only to serve as “a key strategy for poverty alleviation” for direct beneficiaries, but also to generate community-wide benefits, “leveraging growth in the economy, [...] combatting 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].

We combine administrative records for 68 completed housing projects in the province

¹See recent work by Brueckner et al. [2018].

of Gauteng, with data on property transactions, formal and informal housing densities, and dwelling characteristics to measure the local impact 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.

Our empirical estimates rely on a difference-in-differences design leveraging both the timing of housing project construction as well as the precise geography of treated 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 the projects [Diamond and McQuade, 2016, Tissington, 2011]. To address the potential endogenous placement of housing projects, we analyze 65 planned but unconstructed projects as counterfactuals and detect no impacts of these uncompleted projects on local housing markets.

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 interpret the results. Section 4 describes the data used to measure outcomes and details our approach to identifying housing projects while Section 6 provides descriptive evidence. We present spillover results for residential home prices in Section 9 and demographic outcomes in Section 7.1. Section 10 includes a discussion of our findings before providing some concluding thoughts.

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,² the program has consistently sought to redress the economic and geographic

²Public 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.

legacy of apartheid by providing formal housing to low-income households. Subsidized housing efforts at the national level have focused on constructing and allocating 40m², single-story, two-room dwellings in groups of 50 to 500 per project. According to government figures, upwards of 3 million housing units have been delivered between 1994 and 2015.

2.1. Planning and Delivery

Housing projects are primarily located on undeveloped state-owned land although in some cases, municipalities work with private developers to purchase inexpensive, vacant private land for these projects. Finding suitable land plots often requires policymakers to locate these 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 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 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 citizenship, no previous property ownership, being married or having financial dependents, and having a monthly household income below R3,500 [Durojaye et al., 2013].³ 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. 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 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, which often serve alternative land uses. The construction of informal backyard shacks is a common occurrence within project boundaries. Recipients typically enter agreements with tenants (at times relatives) to lease the shacks for cash or in-kind

³The Gauteng Province has implemented their own waiting list since 2008 in order to exert greater control over the allocation process.

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

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 *backyard-ing* 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. Conceptual Framework

Stay tuned.

4. 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 area 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 Gauteng's housing policies, deeds data on housing transactions, household-level census data, and building-based land use information. Each is described in turn below.

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

4.1. 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 Johannesburg 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.

Table 1. Top-Five Sellers in Housing Transactions Sample

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

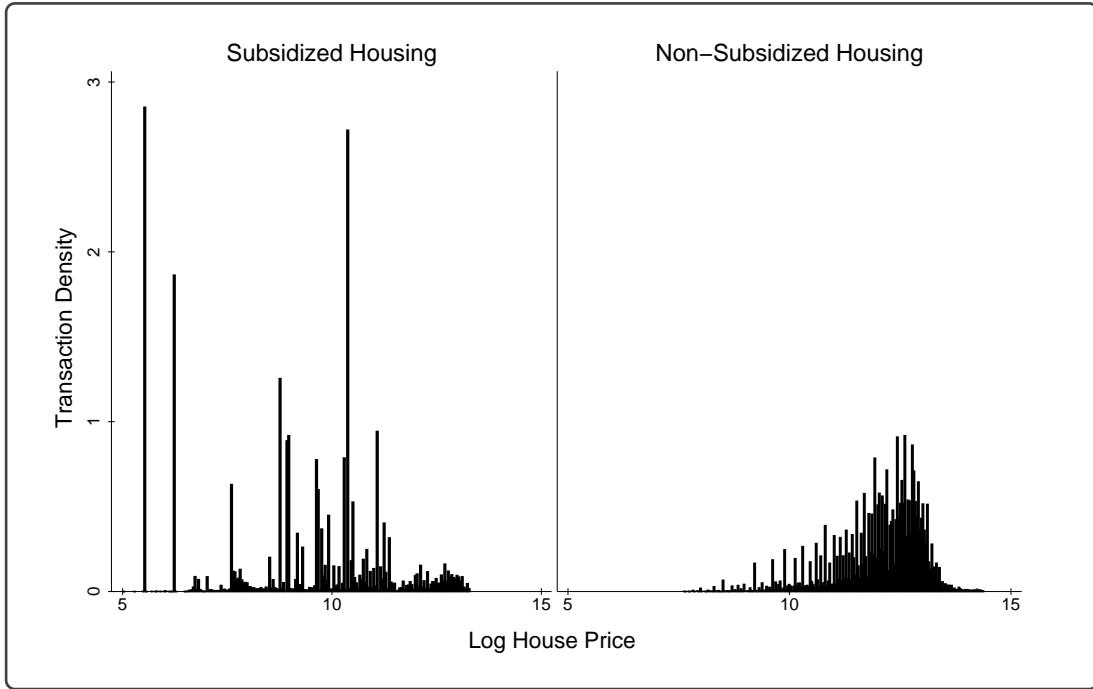
Constructed Housing Projects

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 our 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. In Table 1, we show that these institutional sellers represent an important share of our sample. We then exclude deeds flagged as large buildings used for commercial purposes (less than 2% of transactions),

⁶www.gcro.ac.za/

⁷Since this data comes in the form of overlapping shapefiles, we use the union of intersecting shapes, excluding small shapes below 0.5 km².

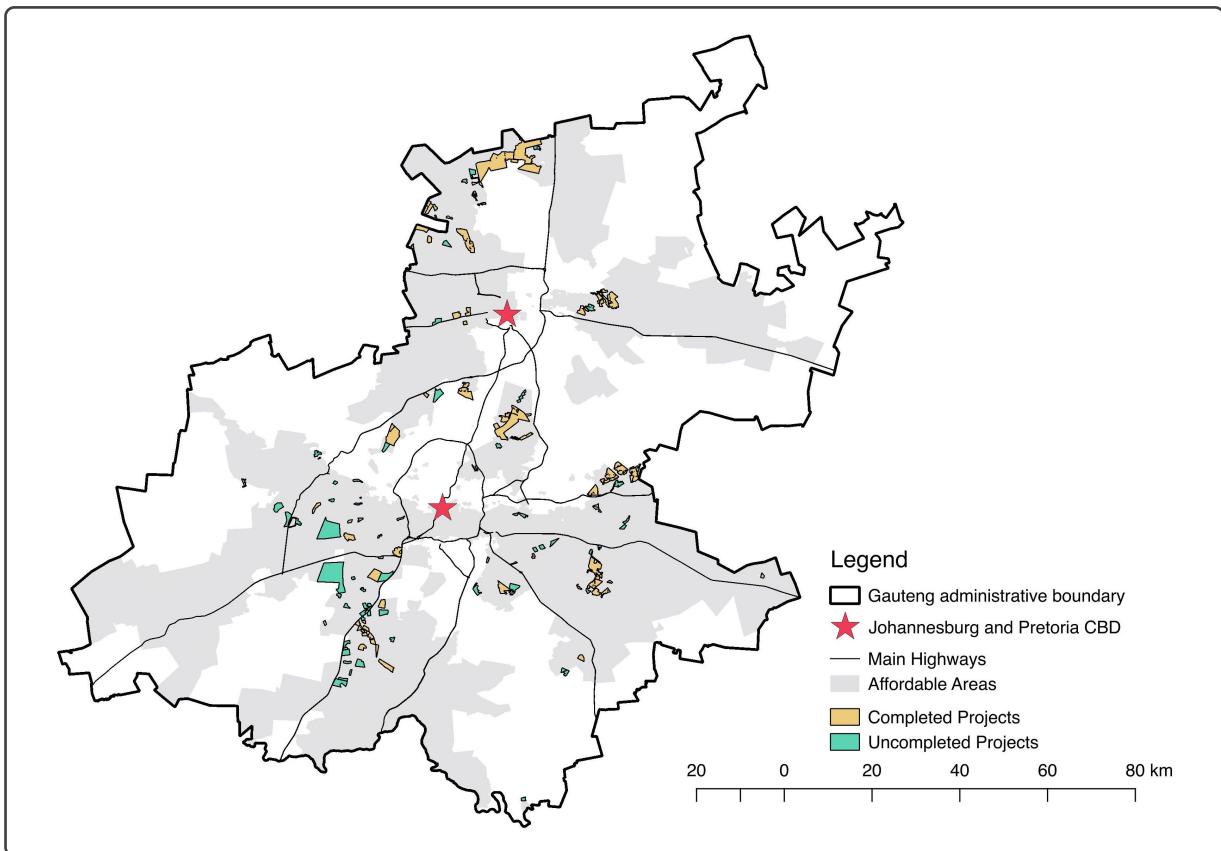
Figure 1. Transaction Price Histogram



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 these properties’ GPS coordinates with the housing policy maps, we are able to identify 68 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 each project. Within projects, most government-sponsored properties are transacted in the same month.

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

Figure 2. Housing Project Map



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 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 99 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 procedure results in a final sample of 65 unconstructed projects. We provide further details

about the digitization and string-matching algorithm in appendix Appendix 7. Figure 2 shows the geographic setting of our analysis where our final sample of constructed and unconstructed project boundaries appear as polygons. Importantly, Figure 2 also shows that the affordable areas – the coverage areas 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), validating 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. Project Descriptions

	Constructed	Unconstructed
Proposed	5	20
Planning	8	12
Under Implementation	12	4
Complete	6	1
No Description	37	28
Total	68	65

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 2 project descriptions from the administrative policy maps according to whether projects are classified as constructed or unconstructed. 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.⁹

⁹We cannot verify that every project description, when available, is indicative of the most recent project status.

Nearby housing transactions

To examine spillovers on nearby housing projects, we include properties that are located outside of the project areas and are not sold by government housing agencies or large developers. We focus 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 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.

4.2. Census Data

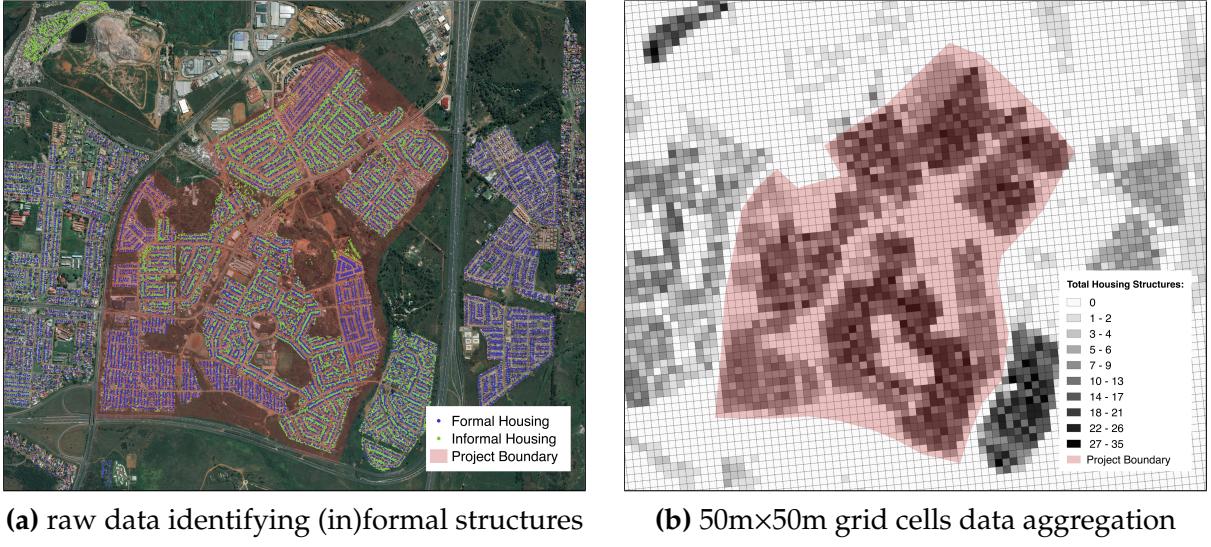
To measure impacts on dwelling characteristics, we use the 2001 and 2011 National Censuses of Population and Housing. Specifically, we analyze household-level responses describing the quality of their living quarters. Our outcomes are mainly binary indicators, and pertain to the household's access to 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 2011.¹⁰ On average, each census area contains 170 household. Though we observe responses from every surveyed household in both census waves, the data does not allow to link households across time periods.

4.3. 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. 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

¹⁰Despite 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.

Figure 3. Building-Based Land Use Data



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 backyard shacks as informal structures, but distinguishes between backyard and other types of informal buildings. We track changes in residential development by using all 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 highly correlations (≥ 0.85) between formal and informal building quantities in both available comparison years.¹¹ For estimation purposes, we transform this data into 50m×50m grid cells aggregates, thereby creating five measures of housing density and the built environment: (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 ???. In Figure 3, we provide an example of the raw data and a depiction of our gridding procedure, using the 2012 data wave.

¹¹We correlate the 2001 and 2012 building surveys with the 2001 and 2011 censuses, respectively.

5. 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 granularities of our measured outcomes, we estimate different forms of the following specification:

$$y_{ipt} = \lambda_p + \sum_d I_{ipt}^d (\alpha^d D_{pt} C_p + \beta^d D_{pt} + \gamma^d C_p + \theta^{pt}) + \delta X_{ipt} + \varepsilon_{ipt} \quad (1)$$

where y_{ipt} is the outcome for unit i in vicinity of project p observed at time t . D_{pt} equals one if period t is after project p was constructed, and C_p equals one if project p is constructed. We interact the post D_{pt} , constructed C_p , and interaction $D_{pt}C_p$ terms 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 . This specification also includes project fixed-effects λ_p , a vector X_{ipt} of additional control variables, and an idiosyncratic error term ε_{ipt} . Distances d account for impacts both inside and outside of project boundaries. This approach allows for difference-in-differences estimates to vary flexibly according to geographic exposure to housing projects.

We interpret the coefficients of interest α^d as the causal impact 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. Interpreting this coefficient as the causal effect of project construction requires the assumption that absent the completion of a housing project, constructed and unconstructed areas would experience the same changes in outcomes.

One concern would be that the project construction may be driven by changes in local political or economic conditions that may also drive changes in 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 in declining neighborhoods may both increase the costs of finishing projects while decreasing local housing conditions. To the extent that these trends are reflected in average levels of local de-

velopment, we can evaluate potential sources of endogeneity by comparing baseline characteristics across project areas as in Section 6. 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 in unconstructed areas may anticipate the eventual construction (or conversion) of these areas and therefore alter their housing investment decisions. Knowing that these areas may be soon redeveloped, households may choose to construct temporary, informal structures instead of investing in long-term, formal houses.

Given these potential limitations to comparing unconstructed and constructed areas, we also consider a complementary approach that compares nearby areas to areas just further away from project boundaries. Under the assumption that areas just nearby projects would have evolved in the same way as areas just further away in the absence of the housing project.

Anticipation (unconstructed knowing something will pop-up)

Sorting

6. Descriptive Statistics

To contextualize our main analysis, we first provide descriptive evidence on the location, size, and housing stock of areas surrounding our sampled projects. Table 3 reveals that the policy provides substantial quantities of housing, delivering on average upwards of 850 new units per project. Constructed projects are larger in area than unconstructed projects, likely due to some unconstructed project representing smaller extensions to previously implemented projects.

givesprovides descriptive statistics at baseline for our sample of 68 constructed and 65 unconstructed 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 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

Table 3. Housing Projects at Baseline

	Constructed	Unconstructed
Number of Projects	68	65
Area (km2)	3.61	2.21
Delivered Houses	858.75	.
House Price within 1km (Rands)	215,172.77	204,490.48
Distance to CBD* (km)	30.38	32.92

*measured with respect to Johannesburg and Pretoria.

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.

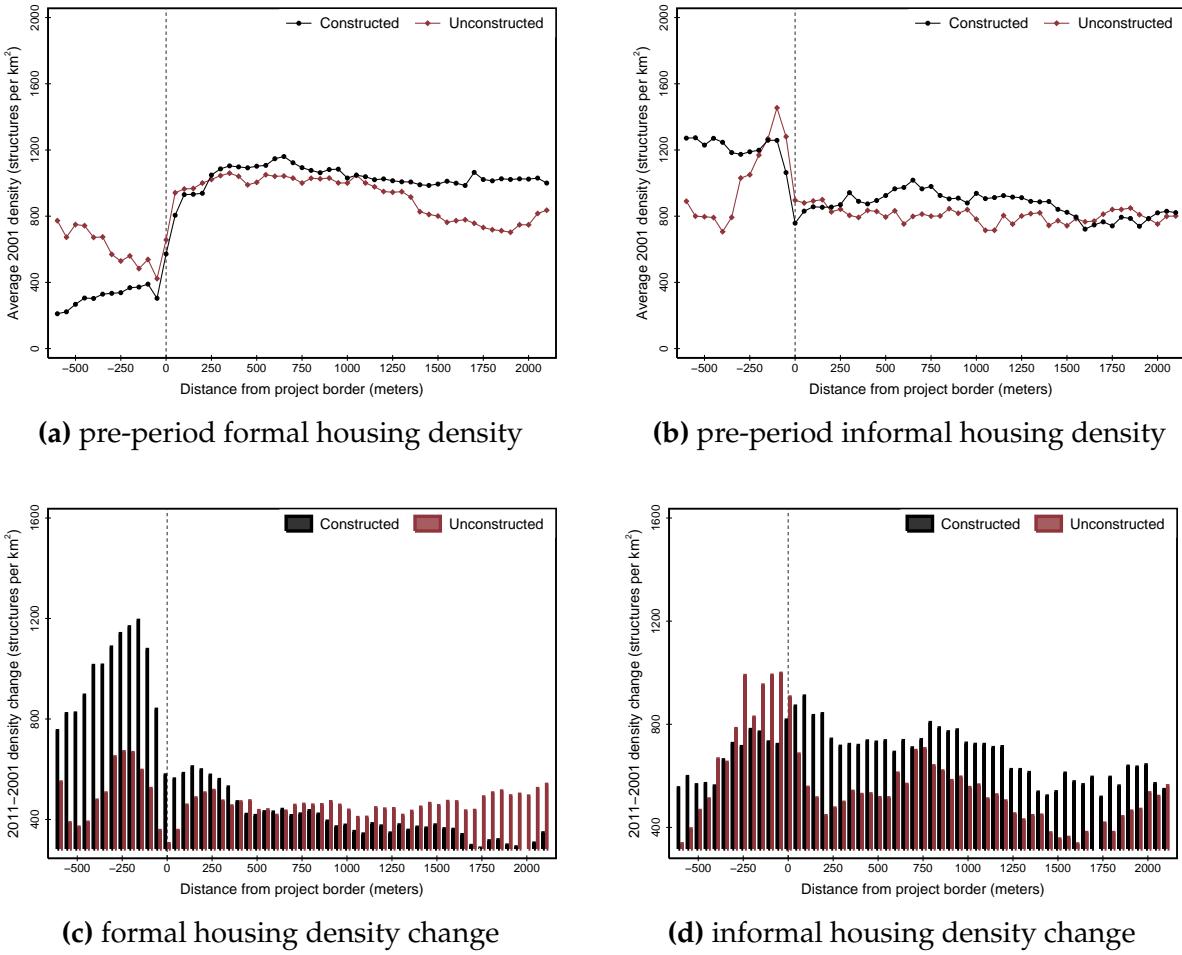
Areas within 4 km of housing projects have greater total building density and higher ratios of formal housing to informal housing compared to both the unconstructed and constructed projects. Similarly, housing prices are much higher for properties outside of project areas. These findings suggest that 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 for constructed projects than simply examining nearby areas.

Table 4 compares census measures at baseline across constructed and unconstructed projects as well as areas within 4 km of housing projects. We identify census blocks as

Table 4. Infrastructure and Demographics at Baseline

	Constructed	Unconstructed	All Households
Flush Toilet	0.55	0.45	0.75
Piped Water in Home	0.13	0.19	0.34
Electricity for Cooking	0.32	0.36	0.64
Electricity for Heating	0.29	0.35	0.62
Electricity for Lighting	0.56	0.45	0.75
Number of Rooms	2.74	2.49	3.32
Household Size	3.43	3.19	3.46
N	196,723	61,779	1,278,901

Figure 4. Housing Densities in Constructed and Unconstructed projects



pertaining to a housing project if over 30% of their land areas overlap, which results in 15,454 intersecting census blocks. While constructed areas are more likely to have flush toilets, electricity for lighting, and households, unconstructed areas have better access to piped water and electricity for eating and cooking. On average, constructed and unconstructed areas have roughly similar levels of dwelling quality while lagging substantially behind households outside of the project areas across all measures. Taken together, these findings suggest both that (1) governments select low-cost land to implement projects and (2) the decision to finish project construction does not appear to be strongly correlated with pre-existing levels of development in the project areas.

7. Estimation Results

To estimate the impacts of housing projects both within their project areas as well as just nearby, we use a difference-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 difference-in-differences approach for each dataset.

7.1. Infrastructure and Demographic Effects

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

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

y_{hpt} measures the outcome for household h living in the vicinity of project p , observed in census year t . $I_{hpt}^e = 1$ 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 8 maps an example of the project and spillover definitions. The summation includes a standard difference-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. λ_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 α^e , which measure the differential change in outcomes for constructed relative to unconstructed projects (either in the project or spillover areas). Interpreting this coefficient 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 6 indicates differences between constructed and unconstructed projects at baseline, particularly in terms of census mea-

sures. 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 5 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 difference-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.

8. Impacts on Housing Density

Table 5. Census Household-level Estimates

	(1) Flush Toilet	(2) Water Inside	(3) Water Utility	(4) Own House	(5) Single House	(6) Elec. Cooking
Project X Post X Const.	0.121* (0.063)	0.136*** (0.050)	0.035 (0.039)	-0.085 (0.074)	0.142*** (0.053)	0.290*** (0.069)
Project X Post	0.080* (0.047)	0.099** (0.038)	-0.030 (0.037)	0.041 (0.038)	0.075** (0.036)	0.179*** (0.059)
Project X Const.	0.088 (0.111)	-0.015 (0.084)	0.040* (0.022)	0.120 (0.087)	0.044 (0.100)	-0.057 (0.107)
Project	-0.316*** (0.074)	-0.212*** (0.050)	-0.021 (0.022)	-0.182*** (0.054)	-0.245*** (0.063)	-0.344*** (0.066)
Spillover X Post X Const.	0.037 (0.035)	0.031 (0.033)	-0.012 (0.012)	-0.008 (0.028)	0.015 (0.029)	0.042 (0.033)
Spillover X Post	0.049* (0.025)	0.136*** (0.025)	0.012 (0.009)	-0.039** (0.015)	0.048** (0.022)	0.126*** (0.027)
Spillover X Const.	-0.026 (0.057)	-0.040 (0.053)	0.030* (0.016)	0.019 (0.043)	-0.000 (0.045)	-0.057 (0.048)
Constant	0.798*** (0.045)	0.399*** (0.042)	0.925*** (0.013)	0.517*** (0.032)	0.599*** (0.035)	0.727*** (0.040)
N	2020549	2020549	2020549	1958335	1936911.000	2020549
Mean2001	0.699	0.300	0.947	0.507	0.535	0.570
Mean2011	0.807	0.481	0.941	0.442	0.622	0.811

Standard errors in parentheses

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

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 ?? in Appendix ?? 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 ?? in Appendix ?? provides an example of this density measure over a housing project. Using these measures, we estimate the following equation:

$$y_{ipt} = \lambda_i + \sum_d I_{ipt}^d (\alpha^d D_t C_p + \beta^d D_t) + \varepsilon_{ipt} \quad (2)$$

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. $I_{ip}^d = 1$ 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, λ_i is a grid-cell level fixed effect. Interpreting the coefficients of interest, α^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 ?? 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 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 ?? 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 ?? 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.

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 (2) to examine the differential changes in housing density between unconstructed and constructed projects. We present the non-parametric results for formal houses in Figure ??, informal houses in Figure ??, and aggregating to total houses in Figure ?? . These figures plot difference-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 ?? 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 3. 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 ?? provides the opposite pattern for informal housing with a decrease in density within constructed relative to unconstructed projects. Positive distances 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 ???. Repeating this exercise for a subset of informal houses coded as backyard shacks, Figure ?? in Appendix ?? 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 5.

9. 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, α captures the differential change in

prices between near and far properties before and after project construction. Additionally, λ_p includes a project fixed affect controlling for any fixed, unobserved drivers of house prices that vary between projects. Likewise, η_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, α , as the causal effect of housing projects on nearby home prices requires the following difference-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.

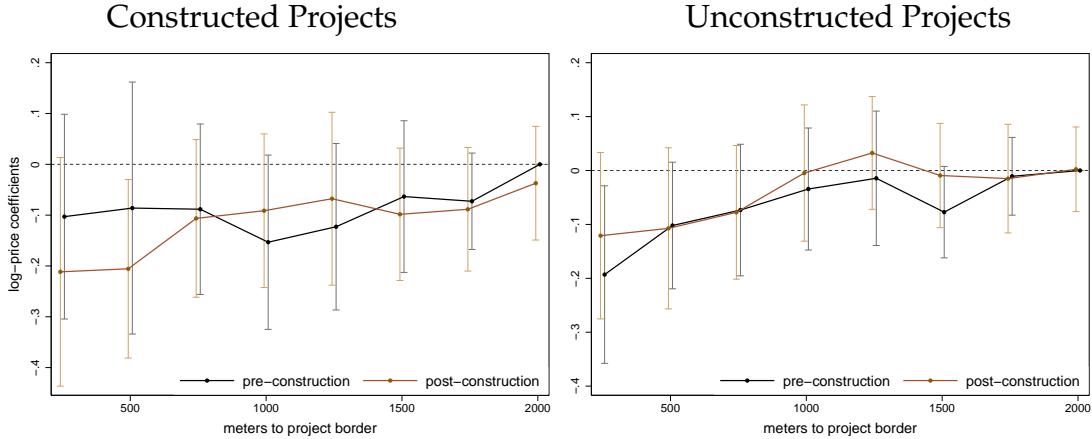
One concern is that our ability to distinguish between deeds pertaining to project houses and deeds associated with normal house transactions may be imperfect. Without addressing this concern, we would mistakenly include project house transfers in our outcome measure of normal house transactions, which 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.¹² 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 6 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 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 6 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

¹²Results are robust to different thresholds.

Figure 6. Price Estimates over Distance from Project



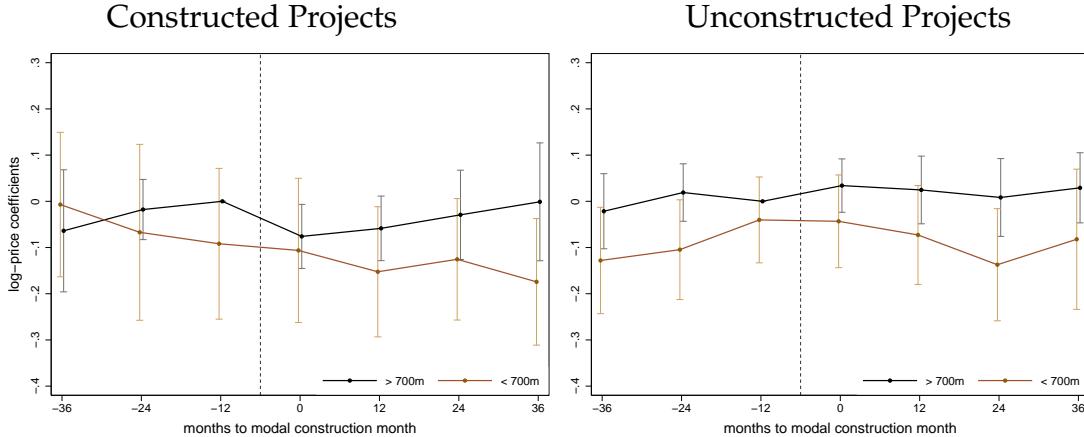
in periods after the expected construction date. These results provide suggestive evidence that the exact location of housing project selection does not appear to be strongly correlated with local housing market trends. The change in price gradients following project construction (left panel of Figure 6) 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 7 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

Figure 7. Price Estimates over Time



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 difference-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 equation:

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

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 6 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 7 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 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.

10. 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 ??) and much better access to bulk services (Table 5). 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 ??). Moreover, results in Table 5 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 6 report if anything the opposite trend. Similarly, we do not observe large new

Table 6. 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* (0.0414)	-0.0325 (0.0413)	-0.0253 (0.0495)
Post	0.0332 (0.0219)	0.0407* (0.0239)	0.0277 (0.0306)
Near	-0.0280 (0.0421)	-0.151*** (0.0453)	-0.126** (0.0546)
Const.			-0.129** (0.0543)
Constant	10.36*** (0.265)	7.868*** (0.310)	10.59*** (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.

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

- J. Beall, O. Crankshaw, and S. Parnell. Social differentiation and urban governance in greater soweto. *Emerging Johannesburg e 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>
- 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 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>.

Table 7. Assessing Name Matching between Budget and Spatial Administrative Data

	Unconstructed Matched	Unconstructed Unmatched	Constructed Matched	Constructed 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 km².

A. Appendix

A.1. String Matching for Project Names

A.2. Census Block Analysis

Figure 8. Buffer Design Census Block Example

