Public Housing Spillovers in a Developing Country*

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

This Version: September 30, 2019

Abstract

We estimate the spillover effects of large public housing projects in South Africa using aerial measures of housing density, property transaction records, and census data. We compare (1) constructed projects versus planned but unconstructed projects, (2) before versus after scheduled construction, and (3) areas nearby versus far from project footprints. Projects increase the quantity and quality of formal housing within their footprints. Between 0 and 500 meters outside of their footprints, projects generate declines in informal housing quantity and formal housing prices especially in higher income neighborhoods. Estimates from a structural model of housing markets suggest that projects are negative amenities generating net welfare losses on the order of 6.6 million USD per project.

Keywords: housing policy; place-based policy; urban development.
JEL Codes: O18; O22; H4; R3.

---

*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
[‡]Mapdwell, Cambridge, MA. E-mail: stefano_polloni@alumni.brown.edu
[§]Federal Trade Commission, Washington, DC. E-mail: william.j.violette@gmail.com

I. Introduction

30% of urban populations in developing countries live in slums often characterized by substandard dwellings, overcrowding, unsanitary conditions, and high crime rates [United Nations, 2015]. While many cities lack the resources to target slum growth directly, some cities have responded by replacing slums with large public housing projects. These projects aim not only to improve local housing conditions, but also to spur investment in surrounding neighborhoods.[1] By providing new formal houses and improved infrastructure, housing projects may generate positive neighborhood amenities. Projects may also have unintended consequences by attracting growth of low-quality, informal dwellings in the backyards of project houses [Brueckner et al., 2018]. Dense backyard dwellings may recreate slum conditions and depress neighborhood property investment. The net spillover effects of housing projects remain an empirical question that has important implications for designing place-based policies to address slum growth.

This paper analyzes the spillover effects of large-scale public housing projects both within their footprints and within their surrounding neighborhoods. We focus on South Africa's National Housing Program, which is unique among developing countries for its massive scale having allocated over 3 million houses across hundreds of projects since 1994. This large volume of projects allows us to identify spatial impacts, building on existing research which is largely limited to considering outcomes for direct project recipients.[2]

To measure these impacts, we compile detailed spatial data for the Johannesburg metro-area including aerial measures of formal and informal houses, census data, as well as deeds records of formal housing transactions. We find that housing projects more than double the density of formal housing within their footprints and improve access to basic services like piped water. While housing projects eradicate preexisting

[1]For example, the South African Department of Housing promotes "accelerating the delivery of housing as a key strategy for poverty alleviation," while also "utilizing housing as an instrument for the development of sustainable human settlements, in support of spatial restructuring" [Department of Human Settlements, 2004].

[2]Cattaneo et al. [2009] estimate child health improvements from housing upgrades in Mexico, Franklin [2016] finds greater women's employment for public housing recipients in South Africa, and Galiani et al. [2017] combine housing projects from throughout Latin America, finding large improvements in reported well-being of households.

informal housing, they also crowd-in enough backyard houses to keep the total density of informal housing unchanged. Our estimates show that for neighboring areas between 0 and 500 meters from project boundaries, informal housing density declines and formal housing prices drop with the strongest reductions concentrated in relatively higher income neighborhoods.

In order to identify these effects, we leverage three comparisons: First, we compare constructed projects versus planned but unconstructed projects to control for how governments strategically locate projects on undeveloped land with preexisting slums. Second, we compare projects before versus after scheduled construction to control for how neighborhoods with constructed projects often have more preexisting housing than neighborhoods with unconstructed projects. Third, we compare areas nearby versus far from a project footprint within the same neighborhood to control for housing market trends specific to each neighborhood. Our main estimates combine these three comparisons into a triple-differences estimator, which extends standard difference-in-differences approaches by allowing for the endogenous placement of projects within neighborhoods.[3]

To assess welfare impacts, we build an equilibrium model of housing markets that maps housing density effects into household preferences for living within and nearby housing projects. In the model, developers choose to build houses on land plots when the value to households from renting and enjoying local amenities exceeds the construction costs. Housing projects may affect both local construction costs as well as the quality of local amenities. By mapping the number of houses to welfare, this approach may have useful applications in developing country settings because price data are often scarce while housing quantity measures are increasingly available through remote sensing.[4]

Structural estimates suggest that housing projects are likely to reduce aggregate welfare. By subsidizing home construction, projects generate welfare gains for formal housing markets within their footprints; however, these gains are overwhelmed by the

[3]See Rossi-Hansberg et al. [2010], Hornbeck and Keniston [2017], and Diamond and McQuade [2016] for difference-in-differences estimators of place-based policies.

[4]Marx et al. use satellite images to measure density and improvements in the quality of roofs in a Kenyan slum. Brueckner et al. [2018] use the same satellite data as this paper to study backyard houses in South Africa. Donaldson and Storeygard [2016] review several research projects using building-level remote sensing data to map urban growth especially in developing countries.

costs of project implementation as well as negative spillover effects on nearby formal and informal housing markets. In total, simulation exercises estimate a net welfare loss of 6.6 million USD per project, which is economically large given implementation costs of 3.2 million USD per project.[5] Since this framework assumes that land owners are able to capture all rental surplus, the distributional consequences of these policies depend on which households are also land owners. Without data on ownership over time especially for informal housing, we instead explore distributional consequences by examining effects separately by neighborhood income quartile.

We proceed by first providing background on public housing and the South African housing program in Section II. In Section III, we build an equilibrium model of housing markets. Section IV describes the data used to measure outcomes and details our approach to identifying housing projects. Motivated with descriptive evidence in section V, Section VI lays out the estimation strategy and identifying assumptions. Section VII presents our main results, and Section VIII concludes.

II. Public Housing Projects

European cities began investing in social housing programs in the early 20th century, and national governments in the West continued to grow these investments in the post-WWII era. After the wave of de-colonization in the 1950s, many developing nations also began public investments in housing, which were progressively phased out in order to secure international financing for national debt crises in the 1970s and 1980s. Beginning in the 1990s large middle-income countries took on subsidized housing programs that focused on channeling public funds to private developers to build full top-structure dwellings. Chile, China, South Africa, and Brazil are examples of this more recent approach.

[5]Implementation costs are drawn from the National Treasury's Annual Budget Reports.

II.A. Public Housing in South Africa

The housing projects studied in this paper were implemented as part of a large national housing scheme enacted in 1994. Though periodically revised and renamed,[6] the program has consistently sought to redress the economic and geographic legacy of apartheid by providing formal housing to low-income households.

This program allocates yearly funds to local housing authorities to construct and allocate 40m^2, single-story, two-room houses on individual plots with bulk services.[7] Local authorities identify many large parcels of land that are able to each accommodate several hundred houses. To meet budgetary targets, authorities often identify inexpensive land located on the urban periphery [Bradlow, 2018, Department of Human Settlements, 2012, 2015]. Since project construction requires passing an environmental impact assessment, receiving zoning approval from local municipalities, and providing bulk service, authorities implement the most feasible subset of candidate projects each year.[8] In many cases, authorities combine housing projects with slum eradication goals by replacing preexisting informal settlements with new housing projects (Hofmeyr [2008]). Housing authorities subcontract with local housing developers for housing construction. According to government figures, upwards of 3 million housing units have been delivered nationally between 1994 and 2015.

Eligible households are required to sign up for an official waiting list maintained by the National Department of Housing.[9] 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]. Each project house is assigned to a beneficiary in a first-come, first-served basis according to the the be beneficiary's place in the waiting list for the province or municipality. When projects

[6] 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.

[7] Local housing authorities include provincial housing departments and municipal housing departments for urban areas like Johannesburg [Department of Human Settlements, 2012, 2015].

[8] Existing policy research characterizes South African housing as an opaque, top-down system where government authorities locate and allocate houses with minimal participation from local communities [Durojaye et al., 2013]. Tissington [2012] documents how local community calls for housing improvements in Guateng were left unanswered for over a decade by local authorities. Durojaye et al. [2013] report how the Gauteng government refused to provide information about plans for housing subsidies and allocation to the Soweto Form, a local community based organization (p. 67).

[9] According to the General Household Surveys, 2009 to 2013, the share of households reporting at least one member on the waiting list has remained stable at over 13% from 2009 to 2013.

replace existing informal settlements, previous inhabitants have priority for receiving project houses. 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.

In practice, these guidelines are loosely followed. Recent reports point to cases of corruption in the allocation of houses [Durojaye et al., 2013].[10] Also, only 82% of project houses are reported as being still occupied by their original beneficiaries within five years of construction.[11]

III. Modeling the Housing Market

We develop a simple model of residential housing markets to capture the welfare effects of local housing projects within their neighborhoods. This model infers welfare changes not only through changes in rents, but also through changes in the quantities of housing supplied.

Each time period, developers choose how many formal and informal houses to construct and maintain on each plot of land as well as the level of rent for each house. Each house can serve up to one household, and houses have identical quality. Then households choose whether to rent a house receiving utility given by

$$U = \delta_{hlt} - \lambda_h(k_{hjt}) - \theta R_{hjt} + \epsilon_{hjt}$$

where h indexes housing sector (either formal or informal), l indexes location, j indexes the specific land plot within location l, and t indexes the time period. δ_{hlt} indicates the amenity value of location, l to housing sector, h in time period, t. k_{hjt} measures the number of houses of a sector on the land plot. $\lambda_h(k_{hjt})$ captures the congestion disutility from having multiple houses of the same sector on the same plot and is assumed to be increasing and convex in the number of houses per plot. $\lambda_h(k_{hjt})$ is also assumed to be independent of the number of houses from another sector on the same plot. R_{hjt} is

[10] Research suggests that beneficiaries are often selected over the course of project construction and sometimes even after construction has finished [Durojaye et al., 2013].

[11] This figure is calculated from the General Household Surveys from 2009 to 2013. Anecdotal evidence suggests that project managers are aware of active secondary markets but have difficulty policing these transactions [Matsena, 2018].

the rent charged for each house on plot j and θ measures the marginal disutility of rent. ϵ_{hjt} is the plot and time-specific amenity shock. The distribution of ϵ_{hjt} may be correlated across housing sectors to capture whether land plots that are suitable for formal housing may also be suitable for informal housing. Households can also choose to live outside the city and receive reservation value, $\overline{U}$.

This approach assumes that households have identical preferences across houses and locations, which implies that housing demand is perfectly elastic. This assumption is consistent with households having low moving costs within the city as well as no idiosyncratic preferences for particular neighborhoods. These assumptions may be reasonable in a context like South Africa where high levels of unemployment mean that different households face similar strong incentives to locate near employment centers.

Given household preferences, developers maximize profits for each housing sector by setting rents just high enough to ensure that households are indifferent between renting each house on plot j and their reservation utility $\overline{U}$. Profit-maximizing rents as a function of the number of houses on a plot is given by

$$R^*_{hjt}(k_{hjt}) = \frac{\delta_{hlt} - \lambda_h(k_{hjt}) + \epsilon_{hjt} - \overline{U}}{\theta} \tag{1}$$

From a distributional perspective, this approach assumes that developers are able to extract all social surplus by setting rents relative to household reservation utility. Although some households may also be developers, welfare for non-developer households remains unaffected by shifts in local amenities because any shifts are immediately capitalized into rents. Without detailed measures of home ownership, this approach is unable to disentangle the distributional consequences of public housing policies across the population of households.

Given profit-maximizing rents, developers then choose how many houses to build on each plot by maximizing total profit per plot given by

$$max_{\,k_{hjt}} \quad k_{hjt}\left[\, R^*_{hjt}(k_{hjt}) - C_{hjt} \,\right]$$

where the C_{hjt} measures the cost of constructing and maintaining each house, which may vary based on location.

When $\lambda(k_{hjt})$ is convex, the profit-maximizing number of houses in each sector, k^*_{hjt} can be expressed as a function of the random amenity shock ϵ_{hjt} given by

$$
\begin{aligned}
k^*_{hjt} &= \begin{cases} 0 & \text{if } \epsilon_{hjt} + V_{hjt} \leq \Lambda_h(1) \\ k & \text{if } \Lambda_h(k) < \epsilon_{hjt} + V_{hjt} \leq \Lambda_h(k+1) \end{cases} \\
&\text{where} \\
V_{hjt} &= \delta_{hlt} - \theta C_{hjt} \\
\Lambda_h(k) &= k\lambda_h(k) - (k-1)\lambda_h(k-1) + \overline{U} \\
\lambda(0) &= 0 \text{ and } \lambda(.) \text{ is increasing and convex}
\end{aligned}
\tag{2}
$$

The profit-maximizing number of plots with houses and the number of houses per plot by sector are increasing in amenity values, δ_{hlt}, decreasing in congestion disutility, $\Lambda_h(k)$, and decreasing in construction costs, C_{hjt}. According to this framework, observed increases in the number of plots with houses and/or the number of houses per plot indicate an improvement in welfare holding constant reservation utility. The variance of amenity shocks, ϵ_{hjt} may be interpreted as a measure of the sector-specific housing market elasticity.

Given optimal rents in equation (1) and optimal houses in equation (2), let the price of each plot equal profit for each plot as follows

$$
P^*_{hjt} = k^*_{hjt}\Big[\frac{\delta_{hlt} - \lambda_h(k^*_{hjt}) + \epsilon_{hjt} - \overline{U}}{\theta} - C_{hjt}\Big] \tag{3}
$$

Like the profit-maximizing number of houses per plot, plot prices are increasing with local amenities, decreasing in congestion disutility, and decreasing in construction costs. Changes in plot prices fully capture changes in social welfare because equilibrium rents ensure that household utility is always equal to reservation utility, leaving developers to enjoy all additional surplus.

III.A. Modeling Public Housing Projects

Public housing projects enter the model by shifting housing costs and amenity values both within as well as nearby project footprints. Let location index l indicate whether plots fall within project footprints, w or nearby spillover areas, s. Similarly, let housing sector index h indicate whether amenities/costs apply to formal, f or informal, i housing sectors.

Housing projects allocate subsidies to local developers to construct houses within project footprints. These subsidies directly lower formal housing construction costs in project footprints, C_{fwt}. These houses are then allocated to beneficiary households who can choose to live in these houses or rent them at efficient prices. Housing projects also remove preexisting informal housing and rezone land for formal housing development, leaving limited land for new informal developments. These practices may increase costs for informal housing, C_{iwt} in project footprints.

Housing projects may also affect local amenity values for formal housing, δ_{fwt} and informal housing, δ_{iwt} within project footprints. Housing projects may boost local amenities by increasing access to piped water, electricity, and sewerage services while providing better roads and community centers.

Housing projects may also have effects on amenities for neighboring areas just outside of project footprints. On one hand, improvements in infrastructure in project footprints as well as greater densities of formal housing may provide positive amenities for surrounding neighborhoods, δ_{hst}. On the other hand, greater densities of backyard housing within projects as well as shifts toward lower income demographics in project areas may be perceived as negative amenities by surrounding housing markets, δ_{hst}.

Collapsing these effects into amenity terms, δ_{hlt} provides a reduced-form method of capturing the spillover effects of housing projects; however, this approach implies a series of restrictions on behavior. First, the amenity term, δ_{hlt} enters household utility linearly and independent of the number of houses per plot, which nests the assumption that the all households enjoy local amenities equally and independently of the density of neighboring housing. This assumption does not allow for congestion costs from overusing public services.

Second, this approach imposes a limited structure for amenity externalities to extend across space: amenities in project areas are allowed to affect amenities for spillover areas while amenities in all other areas are assumed to be spatially independent of each other. This assumption is consistent with housing projects generating large, first-order changes that precipitate across neighborhoods while each individual's development decisions produce negligible externalities across neighborhoods. Alternative approaches that explicitly model housing externalities across space would likely suffer from multiple equilibria as well as challenges for identification in this setting.[12]

IV. Data

Understanding the local development impacts of public housing requires (1) a precise measure of the location, timing, and size of housing projects, and (2) outcomes measured at high spatial resolutions.

We locate projects using a map of housing projects obtained from the Gauteng City Regional Observatory, a research unit composed of the Gauteng Provincial Government and two Johanesburg universities.[13] The map describes 642 housing projects as of 2008, including notes on their completion status. We classify 172 projects with descriptions including "current," "under implementation," or "complete" as constructed projects. We then classify 145 projects with descriptions including "proposed," "under planning," "future," "investigating," and "uncertain" as planned but unconstructed projects. The remaining 325 projects either do not include descriptions or their descriptions are difficult to classify as either constructed or unconstructed. Appendix Table VII includes tabulations of the full list of descriptions.

For constructed projects, we assign a completion date to each project according to the date when recipients received deeds to their project houses. To recover this date, we overlay project locations onto deeds data from the South African National Deeds Office. The deeds cover the universe of housing transactions from 2001 to 2011 in *affordable areas*, which are defined as census enumeration areas with 2010 mean house prices below R500,000.[14] Deeds include price, GPS location, plot size, buyer name, and seller name. Although the data do not identify whether deeds belong to housing

[12]Rossi-Hansberg et al. [2010] develop a model of spatial housing externalities over a single dimension, "housing services" where identification draws from the functional form of the housing externalities. Our setting would require households to consider externalities across housing quantity as well as sectors.

[13]www.gcro.ac.za/

[14]These data were provided by the Affordable Land and Housing Data Centre, which tracks affordable housing markets.

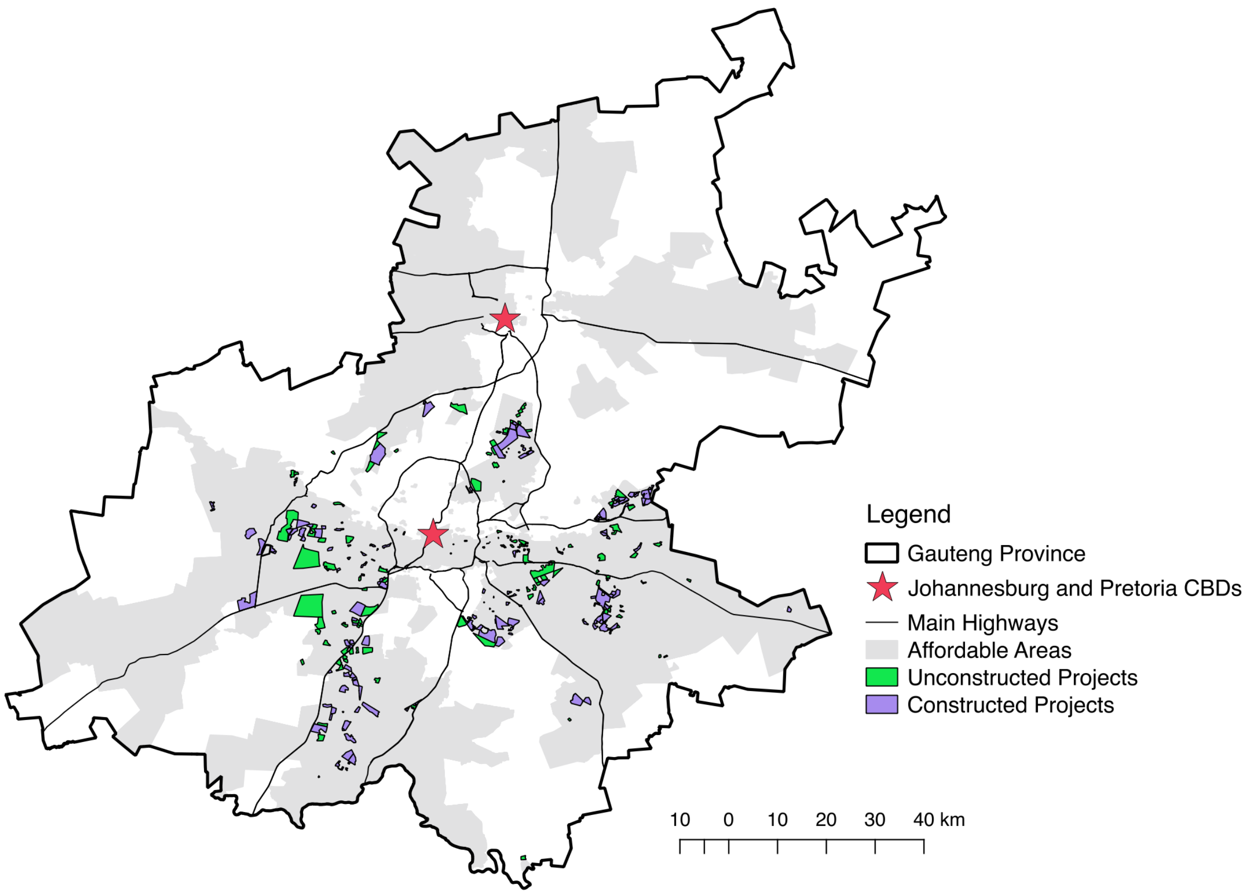
projects, we are able to infer project status according to whether the seller name includes a government or municipality when first transacted. This project definition also excludes deeds flagged as large buildings used for commercial purposes (less than 2% of transactions) as well as purchase prices more than R50,000 above the yearly nominal subsidy values (less than 4% of remaining transactions), resulting in a sample of over 48,000 deeds. Appendix Figure VII plots the histogram of sale prices according to this project definition, finding substantial bunching around subsidy values for project deeds and a smooth distribution of prices for non-project deeds. We assign a completion date according to the modal year and month for deeds within each project. Within projects, most government-sponsored properties are transacted in the same month.[15]

For planned but unconstructed projects, we approximate expected completion dates by matching project names to National Treasury budget reports, which list expected project completion dates. We digitized data for 422 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. Out of the 642 total project names from the administrative map, we are able to match 322 projects, including 126 constructed projects and 73 planned but unconstructed projects. We provide further details about the digitization and string-matching algorithm in Appendix A.B. Because of the low rate of matches, we use expected completion dates only for a subsection of the empirical analysis that focuses on estimating the immediate price impacts of housing projects. The majority of the analysis compares outcomes in 2001 and 2011 for constructed and unconstructed projects under the assumption that these projects either were constructed or would have been constructed between 2001 and 2011.

Figure I maps the sample of constructed and unconstructed projects as polygons within the Gauteng province. Importantly, Figure I also shows that the affordable areas – the coverage area for our deeds data – contain nearly every project boundary. While constructed and unconstructed projects are often adjacent to each other, possibly indicating cases where authorities were unable to complete different phases of planned projects, there are also many examples of isolated projects of both types. Projects are

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

Figure I. Housing Project Map



generally located at relatively great distances from central business districts (CBDs), suggesting that vacant or inexpensive land plots are targeted by housing authorities. Despite their distance from the CBDs, housing projects are often next to arterial highways, potentially easing commuting costs for recipients.

To examine housing price spillovers nearby projects, we include properties that are located outside of the project areas and are not sold by government housing agencies or large developers identified as sellers with more than 30 transactions. Excluding these sellers limits the external validity of our results only to small developers and individual homeowners. We focus on properties located within 1.5 kilometers of constructed and unconstructed housing projects, forming a sample of over 67,000 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.

Figure II. Building Data: Example Housing Project

We measure impacts on dwelling and household characteristics with the full 1996, 2001, and 2011 Censuses of Population and Housing. We analyze average outcomes at the smallest available census geography, which is referred to as *small-areas* by the South African census and which we refer to as census blocks for the purposes of this research. The province of Gauteng is divided between approximately 17,000 blocks in 1996, 11,000 blocks in 2001, and 17,000 blocks in 2011, which are not constant over time periods. On average, each block contains 170 households.

To measure housing growth, we analyze hand-coded building data derived from high-resolution aerial and satellite imagery. We obtain these data from GeoTerraImage (Pty) ltd., a local remote-sensing specialist.[16] The data differentiate structures across 30 categories, including formal and informal residential dwellings. Informal housing structures are easily identified from their temporary nature, often made of materials

[16] `http://www.geoterraimage.com/`

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 two available survey waves in Gauteng: 2001 and 2012. As a validation exercise, we compute aggregated building counts at the census block 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.[17] To measure building density, we transform this data into 25×25m grid cell aggregates, 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. We exclude grid cells where development is infeasible because these cells intersect with rivers, lakes, recreational locations (tennis courts, pools, basketball courts, etc.), and mining excavation areas (10% of total grid cells). In Figure II, we provide an example of the raw data and a depiction of our griding procedure, using the 2012 data wave.

We also construct a measure of construction costs as a function of land slope. First, recent research by the Center for Affordable Housing and Finance (CAHF) suggests that construction costs account for the majority of the value of an average property, which is consistent with a competitive market in property development (Gardner and Pienaar [2019]). The average property value in our data is R231,000. To benchmark this value, a 2016 study by AECOM engineering estimates construction costs to be around R3,500 per square meter of floor space for the lower end of the housing market (AECOM [2016]). Dividing construction costs by costs per square meter implies that the floor space of the average house in the low end of the housing market is around 66 m2, which is comparable to the 40 m2 of floor space available for each project house. The CAHF also estimates that the average construction cost for a fully serviced, 46 m2 house (with a 9 m2 balcony on a 120 m2 plot) in Pretoria is R336,000, which is

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

also consistent with our estimate given that our data includes the bottom 20% of the housing market.

Second, South Africa's National construction guidelines indicate that constructing homes on slopes with 6% to 12% gradients increases infrastructure costs by 25% and building costs by 5%, and constructing on slopes with greater than 12% gradients increases infrastructure costs by 50% and building costs by 15% (CSIR Building and Construction Technology [2005]). Recent research from the CAHF in Pretoria estimates that out of total construction costs, around 62% is attributable to building costs while around 12% is attribute to infrastructure costs (Gardner and Pienaar [2019]). To measure land slope, we use 200m elevation contour lines provided by the National Geo-Spatial Information component of the Department of Rural Development and Land Reform of South Africa. We calculate a proxy for slope by finding the highest and lowest points in a $500m^2$ square and dividing their elevation difference by the euclidean distance between them. The dataset includes 9,770 squares of which 93% have slopes of less than 6%, 5% have slopes between 6% and 12%, and the remaining 2% have slopes greater than 12%.

V. Descriptive Statistics

We provide descriptive evidence to validate our measures of constructed and unconstructed projects. We then examine differences between neighborhoods with constructed projects and neighborhoods with unconstructed projects. We also examine where projects are located within neighborhoods. Finally, we document how project areas and nearby areas evolve differentially over time.

Table I compares average attributes for constructed and unconstructed projects. The first row counts the number of project houses recorded in the deeds data within constructed and unconstructed projects, providing an independent measure of which projects were constructed. Projects categorized as constructed receive a large influx of 374 deeds project houses between 2001 and 2011, indicating that many of these projects were successfully completed in this interval. By contrast, projects categorized as unconstructed receive just 11 deeds project houses on average, which suggests that completed projects are unlikely to be misclassified as unconstructed projects.

Housing prices within 1 km of projects are lower than the average price of 252,000 Rand for areas over 1 km from projects, indicating that projects are located in neighborhoods with less expensive land. With average land areas exceeding one km^2, housing projects represent a significant change in local neighborhoods. Size, timing, location, and nearby house prices are broadly similar between unconstructed and constructed projects.

Table I. Housing Project Areas Description

	Constructed	Unconstructed
Deeds Project Houses	374	11
Median Construction Year	2006	2006
Area (km2)	1.17	1.16
House Price within 1km (Rands†)	188,441	218,635
Distance to CBD‡ (km)	32.5	27.7
Number of Projects	172	145

† The USD averaged around 7.70 Rands during the 2001-2011 period.
‡Measured as the average minimum distance with respect to Johannesburg and Pretoria CBDs.

To further assess where projects are located, Table II uses pooled 1996 and 2001 census data to examine preexisting housing quality in project footprints before construction. Since census blocks are often larger than housing projects, Table II considers census blocks as pertaining to a housing project when over 50% of the land area of the block overlaps with a project footprint. Table II compares means for household indicators aggregated at the block 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 all census blocks in the Gauteng province.

Table II finds that before project construction, preexisting houses in project footprints are smaller and less formal with worse access to services compared to other houses in the province. This finding is consistent with efforts by the Gauteng government to locate housing projects in inexpensive, poorer neighborhoods to both upgrade these neighborhoods as well as save costs. Comparing preexisting houses in constructed and unconstructed projects, we do not observe systematic differences in housing attributes. While constructed project footprints are more likely to have flush toilets, electricity, and large household sizes, households in unconstructed project foot-

Table II. Mean Housing Characteristics from the 1996 and 2001 Censuses

	Constructed	Unconstructed	All blocks
Formal House	0.28	0.47	0.62
Flush Toilet	0.55	0.41	0.65
Piped Water in Home	0.19	0.30	0.45
Electricity	0.49	0.43	0.66
Number of Rooms	2.44	2.84	3.60
Household Size	3.24	3.07	3.13
% Area Overlap with Projects	0.82	0.75	0.06
N	2,163	412	18,018

"Constructed" and "Unconstructed" include census blocks with over 50% area overlap with constructed and unconstructed projects respectively. "All" includes all blocks. Means are weighted by land area.

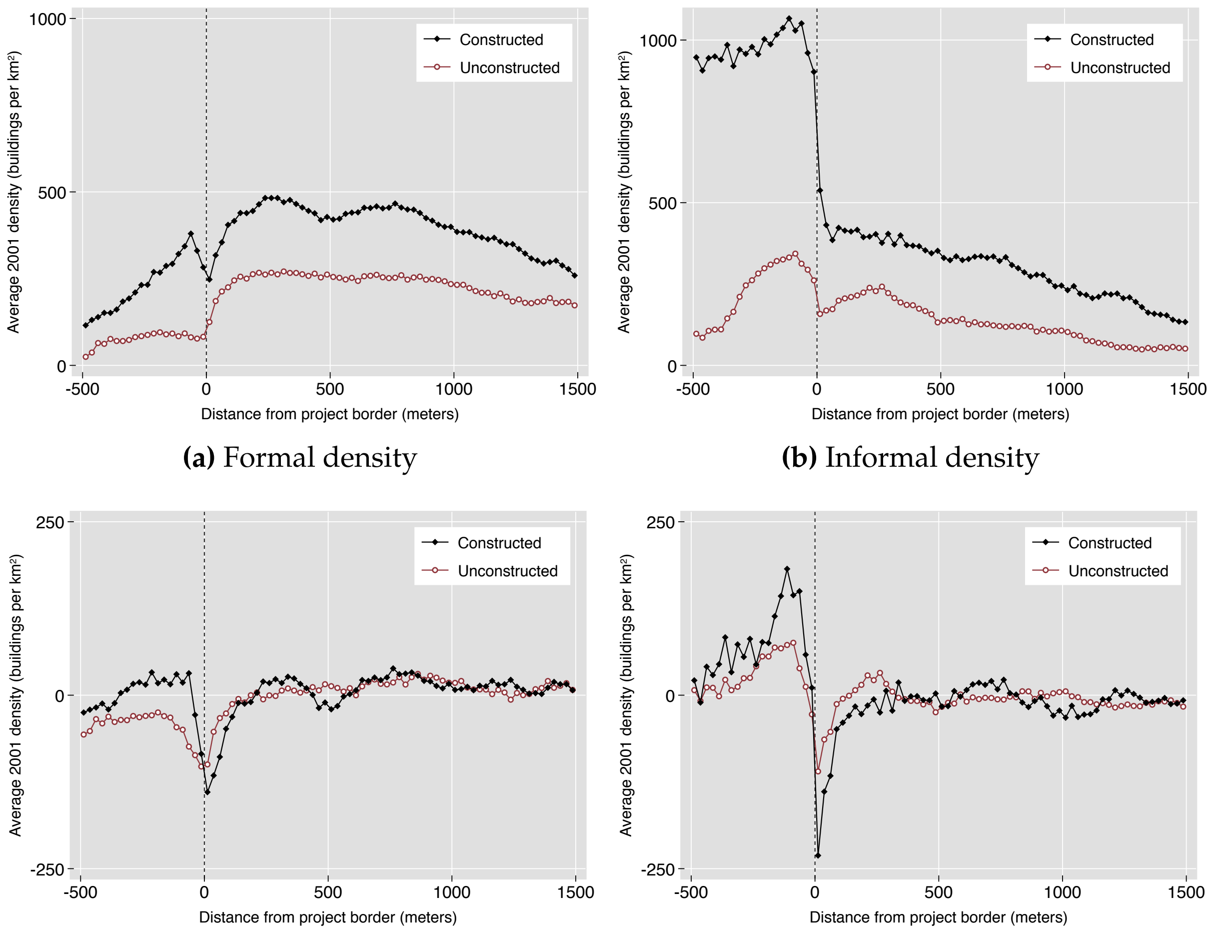
prints report having better access to formal houses and piped water in their homes. With a small number of blocks, average characteristics are not statistically different between constructed and unconstructed project footprints except for the formal housing measure.[18]

The building-based land-use data complement the census data by providing a spatially precise measure of preexisting development in areas prior to project construction. We calculate euclidean distances from the centroid of 25×25m grid-cells to the nearest project boundary, assigning negative distances to grid-cells inside project footprints. We then compute the average density at 25m intervals, reported in houses per square kilometer. Greater negative distances exclude smaller projects in computing average densities, which may account for some of the variation in building density within projects, particularly for informal housing. Appendix A.D illustrates this compositional change in more detail.

Figures IIIa and IIIb respectively plot average preexisting densities of formal and informal houses in 2001. We find low densities of formal housing and very high densities of informal housing within both unconstructed and constructed project footprints (at negative distances). Crossing just outside of project footprints (to positive distances), formal housing densities quickly increase while informal housing densities sharply jump downwards. These trends are consistent with housing projects being precisely targeted within poorer neighborhoods with few existing formal dwellings. Formal and

[18]Statistical significance is calculated at 5% confidence level weighted by land area.

Figure III. 2001 Housing Densities in Constructed and Unconstructed Projects



(a) Formal density

(b) Informal density

(c) Formal density within neighborhood

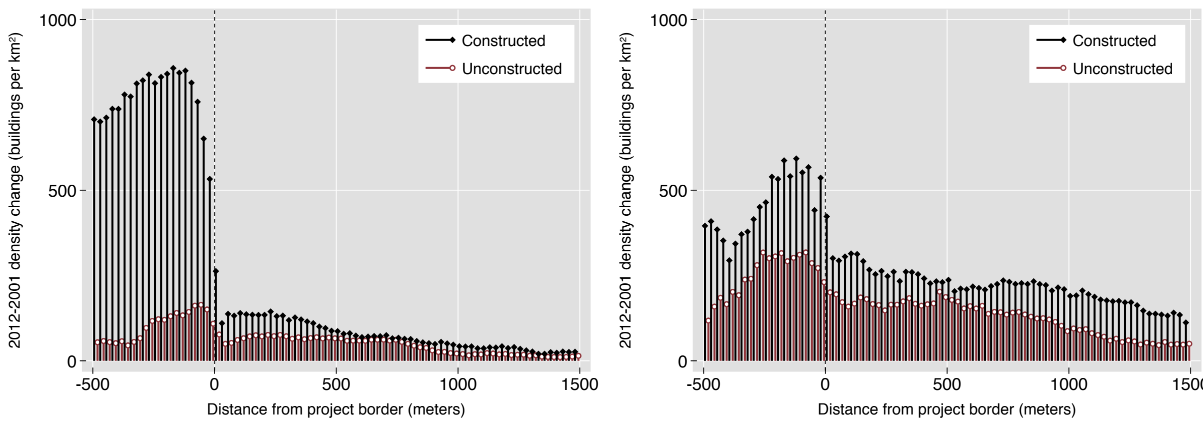
(d) Informal density within neighborhood

Negative distances indicate areas within project boundaries while positive distances measure areas away increasingly outside of project boundaries. Within neighborhood densities are calculated by subtracting mean densities at the 2001 census subplace level from each 25×25m grid cell.

informal housing densities gradually decay moving further from project footprints. On average, constructed projects are located in neighborhoods with greater housing densities than unconstructed projects although housing densities following similar trends with respect to distance from project boundaries. Informal densities within project footprints are significantly higher for constructed projects compared to unconstructed projects, which is consistent with housing authorities prioritizing project completion in neighborhoods with large slums (Hofmeyr [2008]).

Given that housing projects appear to be located in inexpensive neighborhoods with high preexisting densities of informal housing, we next examine whether hous-

Figure IV. Changes in Housing Densities in Constructed and Unconstructed projects



(a) formal housing density change

(b) informal housing density change

ing projects are further targeted to particular locations within neighborhoods. Figures IIIc and IIId are calculated by first subtracting the average housing density in each neighborhood from the density in each grid cell, then calculating averages over distances to project footprints. Neighborhoods are defined according to 1,001 census 2001 subplaces, which identify suburbs, wards, villages, or informal settlements (Statistics South Africa [2001]). Figures IIIc and IIId can be interpreted as the difference in housing density from the neighborhood average associated with each distance from a project boundary.

Accounting for neighborhood differences removes much of the variation in pre-existing housing densities across distances to housing projects, especially outside of project footprints. Yet, low levels of formal housing and high levels of informal housing persist just within project footprints, consistent with housing authorities targeting projects toward inexpensive, poorer plots of land even within neighborhoods. Similar patterns within neighborhoods for constructed and unconstructed areas suggest that both project types are located using similar criteria. Accounting for neighborhood densities collapses average differences in densities between constructed and unconstructed projects although constructed projects have higher building densities within project footprints.

To examine how project areas evolve over time, Figures IVa and IVb are constructed by computing the change in building density for each grid cell from 2001 to 2012 and then averaging these changes according to distances to project boundaries. Both formal and informal housing densities increase over this time period, reflecting rapid urbanization throughout the Guateng province.

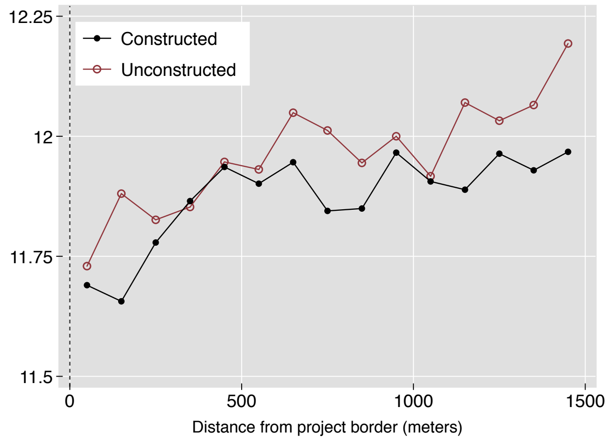
Growth in formal housing is concentrated within constructed project footprints, dropping abruptly at project boundaries (at zero distance). This trend is consistent with housing projects producing large amounts of new, formal housing (Department of Human Settlements [2012, 2015]). Sharp changes in formal housing growth at project boundaries are also consistent with having a precise measure of housing project boundaries.[19] Compared to constructed projects, unconstructed projects experience much less growth in formal housing within their footprints. This finding confirms their status as planned, but unconstructed. Moving from within footprints to outside of footprints, unconstructed projects observe declining formal housing growth. Since projects are often located on undeveloped land, private housing development may crowd into these areas in the absence of housing project construction. Moving further from project boundaries, formal housing growth decays similarly between constructed and unconstructed areas.

Growth in informal housing density is higher within than outside of project footprints as shown by Figure IVb. This result is unsurprising for unconstructed projects: informal settlements grow in underdeveloped, inexpensive land in the absence of project construction. Yet, for constructed projects, this result suggests that even after authorities have cleared the existing informal housing stock to make way for housing projects, total informal housing growth still exceeds growth in unconstructed project areas. Moving further from footprints, growth in informal housing decays with distance at a faster rate for unconstructed projects than for constructed projects.

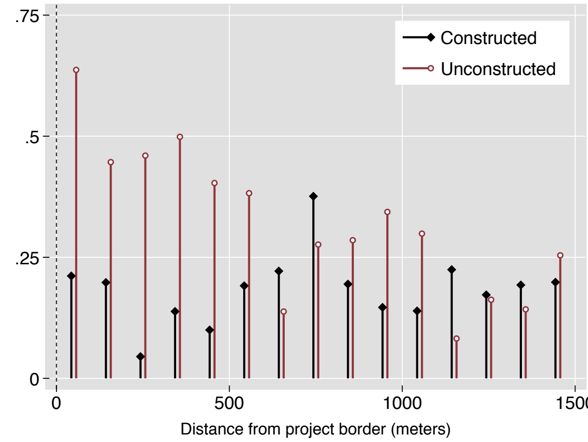
Figure V performs an analogous exercise with property transaction data from the deeds records. These records are likely to be composed of formal houses because informal housing transactions are rarely recorded. Since there are very few non-subsidized housing transactions within project footprints, we focus only on transactions at positive distances. Using the date of each transaction, we group prices before and after the expected construction dates for both constructed and unconstructed projects.

[19] In contrast, an imprecise measure that misclassifies some non-project areas as project areas (and vice versa) would predict a smooth average trend in formal housing changes across footprint boundary.

Figure V. House prices outside constructed and unconstructed projects



(a) pre-period log-prices

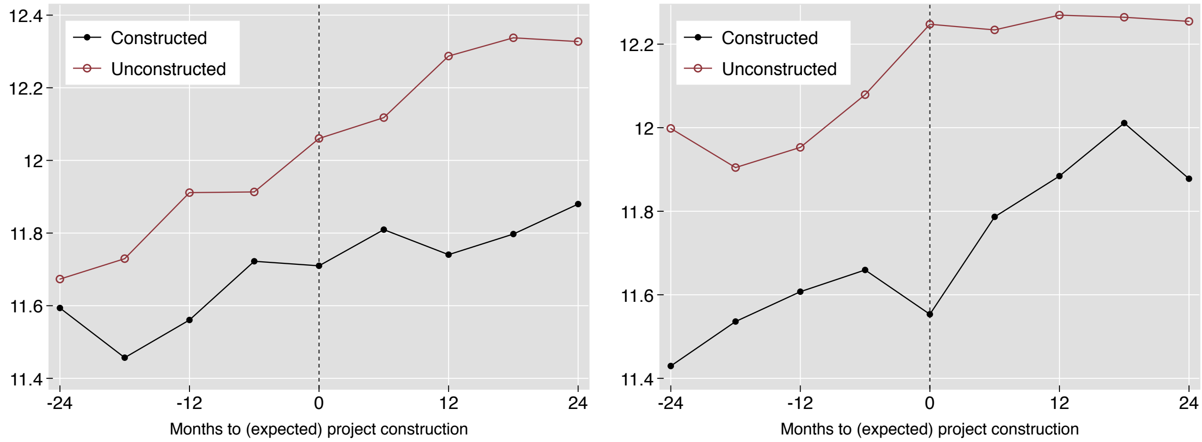


(b) change in log-prices

Figure Va plots average log purchase prices before the expected construction date for both constructed and unconstructed areas with respect to distance to project footprints. Moving from 0 to 1,500 meters from project footprints, prices increase by about 0.3 log-points or roughly 30% for both constructed and unconstructed projects. Low prices near project boundaries are consistent with reports of public housing officials targeting low-cost land for housing development. Figure Vb graphs average changes in log-purchase prices before and after the expected construction of projects with respect to distance to project footprints. While prices increase overall, prices near unconstructed footprints increase more than prices for constructed footprints, especially within 500 meters from footprints.

Deeds data contain the exact date of each housing transaction, which allows us to examine how local housing prices vary with respect to the completion date for constructed projects and the expected completion date for unconstructed projects. Figure VIa plots average log-prices for housing transactions between 0 and 500 meters outside project footprints from two years before and after expected construction. Within 500 meters, prices for constructed and unconstructed projects follow roughly parallel trends prior to expected construction. Prices diverge after expected construction with prices near unconstructed projects rising more quickly than prices near constructed projects. Figure VIb includes the analogous plot for housing transactions located 500 to 1,500 meters from project footprints. Again, both constructed and unconstructed

Figure VI. Average house prices before and after expected construction



(a) log-prices 0 to 500m from projects **(b)** log-prices 500 to 1,500m from projects

prices follow parallel trends prior to expected construction, following similar slopes to their corresponding plots in Figure VIa. After expected construction, price growth slows for areas near unconstructed projects, shrinking the gap between prices near constructed and unconstructed projects. Taken together, Figure VIa and Figure VIb do not provide strong evidence for pretrends in prices in anticipation of project construction.

VI. Empirical Methodology

To estimate the local impacts of housing projects, we implement a difference-in-differences-in-differences (or triple-differences) strategy comparing changes in outcomes nearby projects to areas further away from projects across constructed project areas and planned but unconstructed project areas. This triple differences strategy overcomes limitations of two standard differences-in-differences approaches: (1) comparing changes in outcomes nearby housing projects to areas further way from projects and (2) comparing changes in outcomes for constructed projects to planned but unconstructed projects.

First, previous research identifies housing project impacts by comparing changes in outcomes nearby projects to changes in outcomes further away from projects under two identifying assumptions: (1) project impacts dissipate with distance, and (2) areas nearby projects would have evolved in the same way as areas further away from projects in the absence of project construction (Diamond and McQuade [2016], Harari and Wong [2018]). In our context, projects are specifically located in areas with inexpensive land (Table I) and dense, preexisting slums (Table II and Figure IIIb). Therefore, economic shifts such as changes in labor market demand for informal workers or demand for informal housing would likely produce differential impacts for these nearby areas, driving differential trends in development that are unrelated to housing project construction. Moreover, local governments may select areas for housing projects according to local trends in development, using projects to either reinvigorate deteriorating areas or build on local growth. These factors suggest that a difference-in-differences estimator that relies on distances to projects may generate biased estimates in this context.

Second, an alternative approach would be to compare changes in areas with constructed projects to areas with planned but unconstructed projects under the assumption that constructed project areas would have evolved in the same way as unconstructed project areas in the absence of construction. Constructed and unconstructed project areas appear to be located according to similar criteria: both have sharply higher levels of informal housing and lower levels of formal housing compared to their immediate neighborhoods (Figures IIIc and IIId). Controlling for changes in outcomes for unconstructed projects may help to capture any location-specific trends that constructed projects would have experienced in the absence of construction. However, differences in neighborhood characteristics between constructed and unconstructed projects suggest that these projects may be exposed to differential trends at the neighborhood-level. For example, constructed project areas have higher preexisting densities of both informal and formal housing on average (Figures IIIa and IIIb), which may make these areas more sensitive to broad housing market fluctuations than unconstructed project areas. In fact, constructed projects experience greater growth in informal housing density than unconstructed projects, even at distances over 1 km from project boundaries where any spillover effects of project construction are likely to have dissipated (Figures IVa and IVb). Therefore, comparing changes in outcomes between constructed and constructed project areas may conflate differential neighborhood trends with the true effect of housing construction.

We propose combining near and far comparisons with constructed and unconstructed comparisons in the following triple-differences specification:

$$\begin{aligned} y_{ipnt} = & \textsc{Proj}_{ip} \left(\alpha_1 \textsc{Post}_t \times \textsc{Const}_p + \alpha_2 \textsc{Post}_t + \alpha_3 \textsc{Const}_p + \alpha_4\right) + \\ & \textsc{Spill}_{ip} \left(\beta_1 \textsc{Post}_t \times \textsc{Const}_p + \beta_2 \textsc{Post}_t + \beta_3 \textsc{Const}_p + \beta_4\right) + \\ & \gamma_1 \textsc{Post}_t \times \textsc{Const}_p + \gamma_2 \textsc{Post}_t + \gamma_3 \textsc{Proj}_p + \theta X_{ipnt} + \lambda_n + \varepsilon_{ipnt} \end{aligned} \tag{4}$$

where y_{ipnt} is the outcome for unit i in vicinity of project p in neighborhood n observed at time t with idiosyncratic error ε_{ipnt}. We consider three units of observation: a 25×25m grid-cell, a census block, and a formal housing property. Units are weighted according to their spatial area to ensure that project effects are representative across space. Each unit is assigned to its nearest project according to the euclidean distances to project boundaries. $\textsc{Post}_t$ equals one in the period after construction or expected construction and zero otherwise. $\textsc{Const}_p$ equals one if project p is constructed and zero if project p remains unconstructed. λ_n includes neighborhood fixed effects and X_{ipnt} includes a vector of additional controls. The terms $\textsc{Post}_t$, $\textsc{Const}_p$, and $\textsc{Post}_t \times \textsc{Const}_p$ represent a standard differences-in-differences framework, controlling for differential changes for constructed and unconstructed projects before and after construction.

We augment this standard framework by also estimating differential changes nearby project footprints (0 to 500m) relative to further away from footprints (500m to 1,500m). $\textsc{Proj}_{ip}$ measures the percentage of unit i's area that overlaps with the footprint of project p. When unit i is measured in terms of GPS points such as the deeds data, $\textsc{Proj}_{ip}$ is either one (if unit i is within footprint p) or zero (if unit i is outside footprint p). Other outcomes are measured at geographical areas such as census blocks, which are often larger than project footprints or only partially overlap with project footprints. In these cases, $\textsc{Proj}_{ip}$ takes a value between zero and one depending on the percentage of unit i's area that overlaps with the project footprint. Likewise, $\textsc{Spill}_{ip}$ measures the percentage of unit i's area that overlaps with a 0 to 500m buffer around the footprint of project p.

While our choice of buffer area is arbitrary, we aim to select a buffer area that is sufficiently wide to capture any potential spillover effects of housing projects. With average areas of 3 km^2, buffers from 0 to 500m are 2.5 times larger than average project footprints. To the extent that spillover effects dissipate across space, these buffers are likely to include spillover effects.[20] Similarly, we limit the control group to areas between 500m to 1,500m under the assumption that spillover effects from housing projects have fully dissipated by 500m from project footprints. This control group provides a useful way to control for any economic or housing market trends that may affect these neighborhoods. In this framework, since each project includes two treatment groups (project and spillover) and one control group, we cluster all standard errors at the project by group level.

Interpreting coefficients on $\textsc{Proj}_{ip}$ and $\textsc{Spill}_{ip}$ terms as capturing the effect of moving from 0% to 100% spatial exposure requires two additional assumptions: (1) the spatial effects are proportional to spatial area, and (2) the distribution of project and spillover areas within outcome units is uncorrelated with the distribution of outcomes within these units.

Our coefficients of interest, α_1 and β_1, capture differential changes in outcomes for constructed projects relative to unconstructed projects in footprints and buffer areas (<500m) respectively relative to areas further away from footprints (500 to 1,500m). A causal interpretation requires the assumption that absent the completion of a housing project, changes within and nearby constructed and unconstructed areas would have evolved similarly relative to areas far away from projects. This triple-difference framework allows for constructed and unconstructed projects to experience different trends in economic development over time as long as these trends follow similar trajectories near and far from projects. This framework also permits differential economic trends in areas near and far from projects given that these differential trends are parallel between constructed and unconstructed projects.

Despite these advantages, this approach also has several limitations. A central concern is that forward-looking households anticipate project construction 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 areas that are likely to receive housing projects. In this case, comparisons between constructed and uncontructed areas would overestimate

[20] As shown in Figures IIIc and IIId, pre-period differences in building density within neighborhoods are uncorrelated with distance to project footprint at 500m and greater. This finding suggests that any preexisting spillover effects from living near project areas disappear by 500m.

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). Qualitative evidence suggests that these effects may be limited by the substantial uncertainty in project location and timing due to difficulties coordinating stakeholders and sources of funding as well as a lack of transparency among housing authorities [Tissington, 2011].[21]

We are able to empirically examine anticipatory effects to some extent by observing dynamics in housing prices around the the time of construction. Descriptive Figures VIa and VIb find little evidence of differential trends in the years leading up to project construction. Moreover, the majority of our empirical approaches 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 any possible anticipatory effects play a minimal role in biasing results.

Other concerns include local economic and housing market trends correlated both with project location within neighborhoods and with whether projects are constructed. For example, if housing authorities target constructed projects to both declining or improving neighborhoods and declining or improving areas within these neighborhoods, then the triple-differences estimator would be unable to disentangle these trends from the effects of housing projects. We are unable to empirically test for the presence of these trends.

[21] Diamond and McQuade [2016] also leverage uncertainty in project timing to study affordable housing in the US.

VI.A. Estimating Heterogeneity by Income

We also examine how effects may vary by neighborhood income quartile according to the following specification:

$$\begin{aligned} y_{ipnt} = \sum_{q=1}^{4} Q_n^q \Big[& \textsc{Proj}_{ipq} \left(\alpha_1^q \textsc{Post}_t \times \textsc{Const}_{pq} + \alpha_2^q \textsc{Post}_t + \alpha_3^q \textsc{Const}_{pq} + \alpha_4^q\right) + \\ & \textsc{Spill}_{ipq} \left(\beta_1^q \textsc{Post}_t \times \textsc{Const}_{pq} + \beta_2^q \textsc{Post}_t + \beta_3^q \textsc{Const}_{pq} + \beta_4^q\right) + \\ & \gamma_1^q \textsc{Post}_t \times \textsc{Const}_{pq} + \gamma_2^q \textsc{Post}_t + \gamma_3^q \textsc{Proj}_{pq} \Big] + \theta X_{ipnt} + \lambda_n + \varepsilon_{ipnt} \end{aligned} \tag{5}$$

where q indexes neighborhood income quartiles, and Q_s^q takes a value of one if average household income for neighborhood n falls within quartile q. Average income is measured by taking average household incomes across census subplaces as reported in the 2001 Census of Population and Housing.

VI.B. Estimating Housing Demand

The triple-difference approach also provides a useful framework for estimating residential housing demand given by equation (2). Our empirical setting allows us to identify the causal effect of housing projects on local amenities net of unobserved housing costs. Variation in housing costs also identify the marginal utility of income, which helps translate amenity effects of housing projects into monetary welfare estimates.

To leverage housing cost variation, we divide housing costs into an unobserved portion, c_{hlt}^u that varies at the location level and an observed portion, c_j^o so that total costs are $C_{hjt} = c_{hlt}^u + c_j^o$. To measure the welfare effects of local housing projects, let $\overline{V}_{hlt}$ equal the amenity value net of unobserved housing costs, $(\delta_{hlt} - \theta c_{hlt}^u)$, which may vary over time, project construction status, and location in the following triple-differences framework

$$\begin{aligned} \overline{V}_{hlt} = & \textsc{Proj}_{lp} \left(\alpha_1^{Vh} \textsc{Post}_t \times \textsc{Const}_p + \alpha_2^{Vh} \textsc{Post}_t + \alpha_3^{Vh} \textsc{Const}_p + \alpha_4^{Vh}\right) + \\ & \textsc{Spill}_{lp} \left(\beta_1^{Vh} \textsc{Post}_t \times \textsc{Const}_p + \beta_2^{Vh} \textsc{Post}_t + \beta_3^{Vh} \textsc{Const}_p + \beta_4^{Vh}\right) + \\ & \gamma_1^{Vh} \textsc{Post}_t \times \textsc{Const}_p + \gamma_2^{Vh} \textsc{Post}_t + \gamma_3^{Vh} \textsc{Proj}_p \end{aligned}$$

α_1^{Vh} and β_1^{Vh} measure the effect of housing projects on local amenities net of unobserved housing costs by housing sector in project and spillover areas respectively. Without direct measures of housing costs and amenities, we are unable to disentangle the relative contributions of cost and amenity changes; however, changes in amenities net of housing costs summarize the total local welfare effects of housing projects.

We assume that idiosyncratic amenity shocks, ϵ_{hjt} for formal and informal housing sectors follow a bivariate normal distribution with a two-by-two variance matrix Σ with diagonal terms capturing the variance of formal housing, σ^2_{for} and informal housing, σ^2_{inf} and off-diagonal terms both capturing the correlation, ρ between formal and informal amenity shocks in the same plot. We also assume that amenity shocks are uncorrelated with all other determinants of household utility: (1) local fixed amenities net of housing costs, $\overline{V}_{hlt}$, (2) observed housing costs, c^o_j, (3) disamenities from crowding, $\Lambda(k)$, as well as (4) reservation utility, $\overline{U}$. Under these assumptions, α_1^{Vh} and β_1^{Vh} measure the causal effects of housing projects in an analogous way to equation (4). θ measures the marginal utility of income from observed variation in housing costs. As described in Section IV, we proxy for observed housing costs using average land slope because steep land slopes increase the costs of home construction. Identification of θ requires assuming that land slope only affects the number of buildings through its effect on construction and maintenance costs. This assumption may be violated particularly in cases where households have particular preferences for living on sloped land. Disamenities from crowding, $\Lambda(k)$ are identified from the distribution of the number of houses per plot given the bivariate normal distribution of amenity shocks. The frequency of plots with both formal and informal housing identify the correlation term, ρ. All estimates are normalized with respect to σ^2_{for} and σ^2_{inf}, which are not separately identified in this framework.

Under these assumptions, the probability of observing a given number of formal and informal houses per plot according to equation (2) can be written in terms of a seemingly unrelated bivariate ordered probit model as follows [Sajaia, 2008]

$$\begin{aligned}
Pr[k^*_{for,jt} = k_{for}, k^*_{inf,jt} = k_{inf}] \; &= \Phi_2\left(\frac{\Lambda_{for}(k_{for}+1) - \overline{V}_{for,jt} - \theta c^o_j}{\sigma_{for}}, \frac{\Lambda_{inf}(k_{inf}+1) - \overline{V}_{inf,jt} - \theta c^o_j}{\sigma_{inf}}, \rho\right) \\
&- \Phi_2\left(\frac{\Lambda_{for}(k_{for}) - \overline{V}_{for,jt} - \theta c^o_j}{\sigma_{for}}, \frac{\Lambda_{inf}(k_{inf}+1) - \overline{V}_{inf,jt} - \theta c^o_j}{\sigma_{inf}}, \rho\right) \\
&- \Phi_2\left(\frac{\Lambda_{for}(k_{for}+1) - \overline{V}_{for,jt} - \theta c^o_j}{\sigma_{for}}, \frac{\Lambda_{inf}(k_{inf}) - \overline{V}_{inf,jt} - \theta c^o_j}{\sigma_{inf}}, \rho\right) \\
&+ \Phi_2\left(\frac{\Lambda_{for}(k_{for}) - \overline{V}_{for,jt} - \theta c^o_j}{\sigma_{for}}, \frac{\Lambda_{inf}(k_{inf}) - \overline{V}_{inf,jt} - \theta c^o_j}{\sigma_{inf}}, \rho\right)
\end{aligned}$$

where $\Phi_2(.)$ indicates a standard normal bivariate distribution with correlation terms, ρ. This expression can be estimated by maximizing the following log-likelihood function

$$LL = \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{m_{for}=1}^{K_{for}} \sum_{m_{inf}=1}^{K_{inf}} \mathbb{1}\{k_{for,jt} = m_{for}, k_{inf,jt} = m_{inf}\} \;\; log\left(Pr[k^*_{for,jt} = k_{for}, k^*_{inf,jt} = k_{inf}]\right) \tag{6}$$

where J is the total number of plots observed, T is the total number of time periods, and K_{for} and K_{inf} are the maximum observed number of formal and informal houses observed per plot.

VII. Estimation Results

VII.A. Housing Density

We first investigate how housing projects impact housing density within and around their footprints. We estimate the reduced form equation (4) where the outcomes are building density measures for 25×25m grid cells in 2001 and 2012 and Post_t equals one for observations in 2012. Table III reports the triple-difference coefficients of interest for (1) within project footprints and (2) 0-500m outside of project footprints.[22] Re-

[22]Appendix Table IX includes the full set of coefficients.

sults are reported in terms of houses/km^2. Appendix Table X reports triple-difference coefficients separately by neighborhood income quartile following the specification in equation (5).

Column (1) of Table III finds that total housing density within project footprints statistically significantly increases by 618 houses/km^2, which is over twice the baseline average of 579 houses/km^2. This finding suggests that government housing projects are able to substantially out-pace any private housing development in similar but unconstructed project footprints. At the same time, total housing density 0-500m outside of projects decreases by around 217 houses/km^2 with statistical significance at the 10% level. Given an average project area of 1.2km^2 and an average buffer area of 3km^2, the magnitudes of these estimates suggest that each project generates around 742 new houses inside project footprints while crowding out around 651 houses in nearby areas.[23] Appendix Table X includes corresponding estimates by neighborhood income quartile. Column (1) of Table X finds that quartiles with larger increases in total housing (ie. Q2 and Q3) are also accompanied by greater decreases in housing density 0-500m outside of housing projects. This finding provides further suggestive evidence that housing projects crowd-out nearby housing growth.

To examine the composition of housing growth, Columns (2) through (5) of Table III disaggregate total housing into formal housing, informal non-backyard housing, and informal backyard housing. Column (2) finds that the increase in total housing within projects is almost entirely driven by formal housing with a statistically significant, positive coefficient of 620 houses/km^2. This finding matches earlier descriptive evidence from Figure IV. Large reductions in informal non-backyard housing (Column 4) within projects are consistent with qualitative evidence that housing authorities cleared preexisting slums before constructing new projects [Hofmeyr, 2008]. Column (4) documents a statistically significant decrease of 573 informal non-backyard houses/km^2 within projects, which suggests that these slum clearance efforts were particularly aggressive given low preexisting densities of around 184 houses/km^2. Column (5) indicates that project areas also experience substantial growth in backyard informal housing on the

[23] 742 new houses per footprint is much larger the 374 houses identified as project houses within footprints using the deeds transaction data in Table I. This discrepancy is likely due to under-reporting of housing projects in deeds records [Durojaye et al., 2013].

order of 542 houses/km^2, which almost exactly offsets reductions in informal non-backyard houses.

Appendix Table X finds similarly zero net changes in informal housing within projects across all neighborhood income quartiles. Yet, compared to richer quartiles (Q3 and Q4), poorer quartiles (Q1 and Q2) experience larger decreases in non-backyard informal housing that are offset by larger increases in backyard informal housing.

According to Columns (2) through (5) of Table III, total reductions in housing density 0-500m outside of project footprints are driven by decreases in informal non-backyard housing and informal backyard housing although these coefficients are statistically insignificantly different from zero. Column (2) shows a statistically insignificant decrease in formal housing nearby housing project boundaries, which appears to contrast with descriptive evidence in Figure IV. This discrepancy is likely due to differences in weighting observations: while Figure IV computes average changes in formal housing according to distance bins, Table III reports differential changes that are spatially representative, which results in applying relatively greater weights to areas at further distances from project boundaries since further distances cover greater land areas. These findings suggest that housing projects primarily crowd-out nearby informal housing supply with little impact on nearby formal housing supply.

Table III. Housing Density Estimates

	(1) Total Housing	(2) Total Formal Housing	(3) Total Informal Housing	(4) Informal Non-Bkyd. Housing	(5) Informal Backyard Housing
inside project	618.4^a	649.7^a	-31.3	-572.9^a	541.6^a
	(209.7)	(108.4)	(153.2)	(144.1)	(136.8)
0-500m outside	-216.7^b	-34.3	-182.4^b	-73.1	-109.4
	(110.4)	(61.9)	(92.7)	(80.7)	(80.7)
Mean Pre	579.36	288.49	290.87	184.19	106.68
Mean Post	922.93	424.41	498.52	182.46	316.06
R^2	0.322	0.292	0.276	0.272	0.200
N	6,143,756	6,143,756	6,143,756	6,143,756	6,143,756

Standard errors clustered at the project level in parenthesis. ^{c} p<0.10,^{b} p<0.05,^{a} p<0.01
Outcomes are measure in terms of houses per km^2.
Estimates control for 2001 census sub-place fixed effects.
Coefficients are interactions with constructed, $\textsc{Const}_p$ and post, $\textsc{Post}_t$ indicators.
Appendix Table IX includes all corresponding triple-difference coefficients.

VII.B. Housing Demand

Structural estimates of housing demand provide a mapping from reduced-form changes in housing density into estimates of the local welfare effects of housing projects. We estimate equation (6) where the unit of analysis is a 25m×25m plot in 2001 and 2012 and the outcome variables are the counts of formal and formal houses on each plot. Outcome variables count the number of buildings from 0 to 9 with a last category indicating 10 or more buildings. 86% of plot observations have zero formal buildings and 91% have zero informal buildings. Less than 0.05% of plots have more than 10 total buildings. Appendix Table XIV provides a histogram of building counts for plots with at least zero formal or informal houses.

Table IV includes estimates for the parameters of interest while Appendix Table XIII includes the full set of estimates. According to Table IV, estimates of the marginal utility of income, θ, are similar across formal and informal sectors suggesting that higher costs of housing from steeply sloped land have a similar impact on housing construction across sectors. These estimates allow us to convert net amenity effects from units of utility into units of income by dividing net amenity coefficients by θ. For this exercise, we refer to the estimate of θ for formal housing since the variation in housing costs underlying this estimate are relevant to formal housing construction (CSIR Building and Construction Technology [2005]). This exercise shows that for the formal housing sector, average net amenities increase by R33,831 per plot within project areas and decrease by R2,119 within spillover areas. For the informal housing sector, average net amenities decrease by R20,351 per plot within project areas and decrease by R12,657 within spillover areas. Given average housing prices of R231,000, these results translate into price changes of between -9% for informal housing in project areas to +15% for formal housing in project areas.

Positive estimates of ρ in Table IV indicate that unobserved plot-specific amenities for formal and informal housing are positively correlated, which is likely driven by prevalent backyard housing. Appendix Table XIII includes estimates for cut-points that determine breaks in the distribution of unobserved amenities that correspond to each equilibrium number of houses. The lowest cut point for formal housing is slightly more negative than the lowest cut point for informal housing while the highest cut point for formal housing is substantially more positive than the highest cut point for informal housing. These cut point estimates parallel the histogram of building counts in Appendix A.C: plots are more likely to have any formal housing than any informal housing; however, conditional on having any housing, plots are more likely to have many informal houses than many formal houses. These findings provide suggestive evidence that the supply of informal housing is more elastic than the supply of formal housing in this context.

Results for formal housing in Table IV mirror reduced-form results from Table III. Project areas promote statistically significant growth of formal housing within their footprints with little change in spillover areas. Structural estimates for informal housing contrast with reduced form results: both specifications suggest declines in informal housing nearby project footprints, but structural estimates further predict significant declines in informal housing within project areas while reduced form estimates find little change in nearby informal housing. One explanation for this discrepancy is that structural estimates may prioritize extensive margins of whether plots have buildings while reduced-form estimates simply measure the average change in building density. Therefore, the structural approach may interpret the compositional shift away from non-backyard toward backyard informal housing in project areas as a decline in the local net amenities enjoyed by the informal housing sector.

We also allow for heterogeneity in housing demand by neighborhood income quartile with results presented in Appendix Table XIV. While effects for formal housing do not appear correlated with neighborhood income quartile, informal housing markets exhibit a clear negative gradient with respect to neighborhood income: lower income neighborhoods experience improvements in local amenities while high income neighborhoods experience steep declines in local amenities in both project and spillover areas.

Changes in local net amenities are likely to overstate the full welfare effects of housing projects as renters substitute between project areas, spillover areas, and outside areas. To compute the full welfare effects associated with changes in local amenities, we use estimates from Table IV to simulate the effect of housing projects on local building growth. Given simulated choices, we calculate the change in net rents extracted by landlords, which fully summarize total welfare effects. Appendix A.J includes details on the simulation exercise and Table XV includes predicted net rental changes by housing sector, project and spillover areas, and neighborhood income quartile.

Within housing projects, this exercise predicts that net rents from formal housing increase by R17,772 on average per plot; at the same time, net rents from informal housing decrease by R16,743 on average per plot. Together, these findings suggest that average rents per plot increase by a total of R1,030 within project footprints. High average densities of informal housing magnify the effect of declines in local informal housing amenities on welfare, almost offsetting welfare gains from formal housing. Patterns by neighborhood income quartile reflect housing demand coefficient estimates in Appendix Table XIV.

In spillover areas, housing projects drive a slight decrease in nearby net rents from formal housing of R869 per plot and a more substantial decrease in nearby net rents from informal housing of R5,021 per plot, which together lead to a total decline of R5,890. Effects for spillover areas are smaller per plot than corresponding effects for project areas likely because project areas have higher overall densities of housing.

To compute the overall average change in net rents per plot as a result of housing projects, we weight effects for project areas and spillover areas by their corresponding land areas (1.2km^2 and 3km^2 respectively). We compute an average decline of R3,913 per plot. This change represents around a 1.7% decline in average housing prices.

To provide a rough estimate of the aggregate welfare effects of housing projects, we compare changes in net rents at the project-level to project implementation costs. Given around 6,720 plots surrounding project areas, the project-level decline in rents totals 26 million Rand (3.4 million USD). National budget reports suggest an average cost per project of 25 million Rand (3.2 million USD). In total, these results suggest that housing projects generate an aggregate decline in welfare of 51 million Rand (6.6 million USD).

VII.C. Infrastructure and Demographics

Given large shifts in quantities and types of housing within and around projects, we next use census data to understand whether housing projects also affect infrastructure access, housing quality, and residents' demographics. We reestimate equation (4) using

Table IV. Housing Demand Estimates

	Formal Housing	Informal Housing
Net amenities: $(\delta_{hl} - \theta c^u_{hlt})$		
inside project	0.319[c]	-0.192[c]
	(0.193)	(0.103)
0-500m outside	-0.020	-0.119
	(0.069)	(0.094)
Marginal utility of income: $-\theta_h$	-0.0000094[a]	-0.0000103[a]
	(0.0000035)	(0.0000038)
Correlation of amenity shocks: $\rho_{for,inf}$	0.620[a]	
	(0.018)	
N	6,143,756	

Standard errors clustered at the project level in parenthesis. [c] $p<0.10$,[b] $p<0.05$,[a] $p<0.01$ Outcome categories by plot and year include categories for formal and informal building counts from 0 to 9 and a category at least 10. Table XIII includes the full set of coefficients. Net amenities equal local amenities minus local housing costs.

census data from 1996, 2001, and 2011. Outcomes include housing and demographic characteristics averaged across households in each census block and year. POST_t equals one for observations from the 2011 census. PROJ_i and SPILL_i measure the fraction of spatial overlap of each census block with project areas and spillover areas (0-500m) respectively.[24] Specifications control for the spatial area of census blocks up to a cubic term and interacted with year in order to account for the possibility that household sampling may vary according to census block size. All estimates are weighted by the spatial area of census blocks so that the estimates are representative of the spatial impacts of the projects. Weighting according to spatial areas further ensures that the census results are directly comparable to the building density results. Table V reports the triple-difference coefficients of interest for (1) within project footprints and (2) 0-500m outside of project footprints.[25]

Columns (1) to (6) of Table V capture changes in housing characteristics within and around housing project footprints. Inside projects, the share of households living in

[24] 3,118 blocks partially overlap with project areas while 1,138 blocks fully overlap with project areas. Likewise, 7,456 blocks partially overlap with spillover areas while 1,402 blocks fully overlap with spillover areas.

[25] Appendix Tables XI and XII include the full set of coefficients. Census blocks lack the spatial precision needed to estimate effects separately by quartile of neighborhood income.

Table V. Census Household-level Estimates

	(1) Is a Formal House	(2) Has a Flush Toilet	(3) Has Piped Water Indoors	(4) Has Electricity	(5) Total Rooms	(6) Owns House
inside project	0.614[a]	0.257	0.381[a]	0.103	1.887[a]	0.084
	(0.137)	(0.158)	(0.128)	(0.158)	(0.680)	(0.120)
0-500m outside	0.120	-0.053	0.062	-0.022	0.146	0.063
	(0.080)	(0.071)	(0.089)	(0.077)	(0.356)	(0.073)
Mean Pre	0.62	0.71	0.47	0.70	3.48	0.41
Mean Post	0.65	0.77	0.54	0.77	3.79	0.36
R^2	0.571	0.639	0.607	0.607	0.642	0.610
N	20,641	20,639	20,639	20,639	20,500	20,631

	(7) HH Size	(8) People per km^2	(9) Age HoH	(10) HoH is African	(11) HoH is Employed	(12) HH Log Income
inside project	0.586[b]	1610.3	2.785	0.007	-0.060	0.123
	(0.275)	(1288.8)	(2.095)	(0.094)	(0.092)	(0.312)
0-500m outside	0.222	378.3	0.010	-0.035	-0.042	0.213
	(0.185)	(708.0)	(1.265)	(0.061)	(0.037)	(0.176)
Mean Pre	3.25	1,793.5	41.26	0.73	0.82	7.30
Mean Post	2.89	2,690.4	42.34	0.81	0.83	8.25
R^2	0.734	0.583	0.606	0.737	0.590	0.756
N	20,509	21,197	20,505	20,641	20,501	20,495

Standard errors clustered at the project level in parenthesis.[c] $p<0.10$,[b] $p<0.05$,[a] $p<0.01$
Outcomes in Columns (1) to (4), (6), and (9) to (11) are indicator variables.
Specifications include census 2001 subplace fixed effects.
Results are weighted by unit area and control for up to a cubic in unit area by time.
"HH" stands for household and "HoH" stands for head of household.

formal housing jumps by 61 percentage points consistent with the large growth in formal housing density documented in Table III. The shares of households in project areas with flush toilets, piped water indoors, and/or electricity also increase by economically meaningful magnitudes although only the coefficient for access to piped water indoors is statistically significant. The average number of rooms per house increases by 1.9 likely because new project houses are often larger than informal backyard and non-backyard houses. The coefficient for the share of households in project areas that report owning their home is positive but statistically insignificant.

In spillover areas (0-500m outside of projects), the share of households that report living in formal houses (Column (1)) increases by a statistically insignificant 12 percentage points. Any increase in the ratio of formal to informal housing nearby projects may reflect crowding-out of nearby informal housing as described by housing density estimates from Table III. Economically small and statistically insignificant coefficients for infrastructure access, number of rooms, and homeownership in spillover areas in Columns (2) to (5) suggest little scope for infrastructure and housing stock investments in project areas to spillover to surrounding neighborhoods.

Complementing changes in housing characteristics, Columns (7) to (12) of Table V estimate changes in demographics within and around project footprints. Within projects, household ("HH") sizes increase by 0.59 people on average given a baseline average of 3.25 people per household, which is consistent with recipients of project houses inviting additional family members to live in their new houses. Population density increases by 1,610 people per km^2, which is statistically insignificant, but economically meaningful — this effect almost doubles the baseline average population density of 1,793. Coefficients for average age, race, and employment status of the head of households ("HoH") as well as household log-income are economically small and statistically insignificant, suggesting little demographic change as a result of these housing projects. The point estimate for household income (Column (12)) suggests a 12% increase in average household incomes in project areas although this result is statistically insignificant. Taken together, these results are consistent with housing authorities ensuring that preexisting slum residents are able to benefit from new project houses. The results are less consistent with the hypothesis that political corruption steers project houses toward wealthier and more politically connected households [Durojaye et al., 2013].

Spillover areas report positive, statistically insignificant coefficients for both household size (Column (7)) and population growth (Column (8)) that are much smaller in magnitude and significance than corresponding estimates inside projects. Spillover estimates for household demographics in Columns (8) to (11) parallel estimates within projects by having low statistical significance and small magnitudes. Estimates predict a large, statistically insignificant increase of 21% in average household income.

VII.D. Prices

We examine how housing projects may affect the prices of nearby formal housing using deeds data on housing transactions. Because few private housing transactions occur within project footprints and because it is possible that project house allocations may be miscoded as private housing transactions, we restrict our analysis to properties located outside of project footprints. We estimate the following modification of equation (4):

$$\begin{aligned} y_{ipnt} \;=\; & \textsc{Spill}_{ip}\left(\beta_1\,\textsc{Post}_t \times \textsc{Const}_p \;+\; \beta_2\,\textsc{Post}_t \;+\; \beta_3\,\textsc{Const}_p \;+\; \beta_4\right) + \\ & \gamma_1\,\textsc{Post}_t \times \textsc{Const}_p \;+\; \gamma_2\,\textsc{Post}_t \;+\; \gamma_3\,\textsc{Proj}_p \;+\; \theta\, X_{ipnt} \;+\; \lambda_n \;+\; \varepsilon_{ipnt} \end{aligned} \tag{7}$$

where the unit of analysis, i is a property transaction, time, t is monthly, and the outcome, y_{ipnt} is the log-price of each property transaction. To utilize monthly transaction data, we limit the sample to include constructed housing projects where the month of completion is available and unconstructed housing projects where we are able to project an expected month of construction. We refer to the completion month for constructed projects and the expected completion month for unconstructed projects jointly as the scheduled construction month. $\textsc{Post}_t$ takes a value of one for transactions occurring after the scheduled construction month. $\textsc{Spill}_{ip}$ takes a value of one for properties between 0 and 500 meters outside of their nearest project boundary. Some specifications include additional controls, X_{ipnt} for the calendar month of each transaction and up to a cubic polynomial in the plot size of each property. All specifications control for neighborhood fixed effects, λ_n.

To examine how price effects may vary by neighborhood income quartile, we also estimate the following modification of equation (5):

$$
\begin{aligned}
y_{ipnt} = \sum_{q=1}^{4} Q_n^q \Bigg[& \textsc{Spill}_{ipq} \left(\beta_1^q \, \textsc{Post}_t \times \textsc{Const}_{pq} + \beta_2^q \, \textsc{Post}_t + \beta_3^q \, \textsc{Const}_{pq} + \beta_4^q \right) + \\
& \gamma_1^q \, \textsc{Post}_t \times \textsc{Const}_{pq} + \gamma_2^q \, \textsc{Post}_t + \gamma_3^q \, \textsc{Proj}_{pq} \Bigg] + \theta X_{ipnt} + \lambda_n + \varepsilon_{ipnt}
\end{aligned}
\tag{8}
$$

where Q_n^q indexes neighborhood income quartile.

Table VI reports estimation results for all neighborhoods following equation (7) and separately by neighborhood quartile following equation (8). Table VI includes triple difference coefficients, β_1 which capture the differential price changes for properties between 0 and 500 meters from project boundaries before and after scheduled construction for constructed versus unconstructed projects. With log-prices as an outcome, coefficients can be interpreted as in terms of percent of transaction prices.

Column (1) of Table VI documents a statistically insignificant 7% decrease in housing prices, which shrinks to 6% in column (2) after accounting for year-month fixed effects as well as up to a cubic polynomial in plot size. Examining price effects by quartile of neighborhood income, we find that price effects become increasingly negative in higher quartile neighborhoods. The poorest quartile (Q1) exhibits a small, statistically insignificant 3% change in prices (column (2)) while the richest quartile (Q4) experiences a 17% decrease in prices, which is statistically significant at the 5% level. Comparing columns (1) and (2), we find that additional controls have little impact on the magnitudes or statistical significance of the estimates.

Appendix A.F estimates differential changes in prices before and after scheduled construction for constructed versus unconstructed projects non-parametrically according to distance to the nearest project boundary. The strongest decreases in prices appear concentrated closest to project boundaries — especially for the richest income quartile (Q4) — although not all price gradients are monotonically increasing with distance to project boundaries.

To examine the dynamics of these price effects, Appendix A.G provides corresponding estimates of differential changes in prices for properties near versus far from project boundaries for constructed versus unconstructed projects non-parametrically

Table VI. Triple-Difference Estimates of Log-Prices 0 to 500 Meters Outside of Housing Projects

	(1)	(2)
All Neighborhoods	-0.072	-0.061
	(0.051)	(0.051)
R2	0.48	0.49
By Neighborhood Income Quartile		
Q1	0.026	0.032
	(0.127)	(0.129)
Q2	0.063	0.077
	(0.097)	(0.096)
Q3	-0.076	-0.074
	(0.085)	(0.090)
Q4	-0.198[b]	-0.166[b]
	(0.085)	(0.084)
Year-Month FE		✓
Plot Size (up to cubic polynomial)		✓
R2	0.49	0.49
N	62,349	62,349

Estimates capture differential price effects for properties 0 to 500 meters outside of project boundaries relative to properties 500 to 1,500 meters from boundaries before and after scheduled construction for constructed and unconstructed projects. Income quartiles are from average household 2001 incomes in census 2001 subplaces. All estimates include neighborhood fixed effects. Standard errors clustered at the project level in parenthesis.
[c] $p<0.10$,[b] $p<0.05$,[a] $p<0.01$

by months to scheduled construction. Estimates indicate some evidence of pretrends with prices declining in anticipation of project completion particularly for the bottom income quartile. Pretrends are close to zero for the top two income quartiles. Declines in prices occur sharply after scheduled construction and are most pronounced in the 48 months following scheduled construction for the top two income quartiles.

One concern with this approach is that our method of imputing the expected month of project completion for planned, but unconstructed may introduce noise into the estimates. As a robustness exercise, Appendix A.H provides differences-in-differences estimates of price effects using only constructed projects and comparing properties near and far from project boundaries before and after construction. Results from this exercise (Table XVI) closely match triple-difference estimates (Table VI).

VIII. Discussion and Conclusion

Our results provide a framework for evaluating whether housing projects succeed in attaining three broad goals of South African housing policy as articulated in the National Housing Act of 1997.

As a first goal, "every municipality must... ensure that the inhabitants of its area of jurisdiction have access to adequate housing on a progressive basis."[26] Our results provide evidence that housing projects boost the quantity and quality of dwellings available to households within project footprints. Building density estimates find increases in formal housing with no changes in total informal housing, reinforcing census estimates that document improvements in infrastructure access and housing quality.

As a second goal, "government must use public money available for housing development in a manner which stimulates private investment in, and the contributions of individuals to, housing development."[27] We find little evidence that housing projects are able to attract local investment in housing markets within and outside of their footprints. Instead, formal housing growth nearby remains unchanged while prices for formal houses drop especially in relatively higher income areas.

Finally as a third goal, "government must promote the establishment, development

[26]See Government of South Africa [2001].
[27]Ibid.

and maintenance of socially and economically viable communities and of safe and healthy living conditions to ensure the elimination and prevention of slums and slum conditions."[28] We find mixed support for this goal. On one hand, growth in nearby informal housing slows as a result of housing projects, which supports this goal to the extent that informal houses reflect government criteria for slum conditions. On the other hand, housing projects are unable to slow growth in informal houses within project footprints. Instead, backyard shacks replace preexisting informal houses. Our model of the housing market suggests that net reductions in informal housing translate into substantial decreases in social welfare. Although our approach does not explicitly model possible negative externalities of slums such as poor sanitation, crime, and congestion, our results suggest that these externalities would have to be very large in order to offset net welfare losses on the order of 6.6 million USD per project. Taken together, our results suggests that housing projects are unlikely to provide a cost-effective policy for slum reduction.

Our results also provide broader lessons for considering informal housing markets in the design of place-based policies in developing countries. We find that the supply of informal housing may be particularly responsive to place-based policies. While projects generate price decreases in formal housing markets for high-income areas just as in developed country settings (Diamond and McQuade [2016]), informal markets experience reductions in the quantity of dwellings. In future research, housing quantity data may complement formal housing price data for measuring the total welfare impacts of these policies.

Active informal housing markets also allow for the spillover effects of housing projects to extend not only just outside of project footprints, but also within project footprints. We document substantial construction of informal houses in the backyards of project houses, which leave the total number of informal houses unchanged after project construction. To measure the spillover effects of place-based policies, researchers and policymakers have largely focused on changes outside of policy footprints, treating any changes within these footprints as exogenously driven by the policies (Diamond and McQuade [2016], Rossi-Hansberg et al. [2010], Hornbeck and Keniston [2017]). Our analysis suggests that future research of place-based policies in devel-

[28] Ibid.

oping contexts may benefit broadening the area of analysis to include outcomes in the project footprints themselves.

This analysis abstracts away several features that may provide a more nuanced understanding of the welfare effects of housing projects. First, we do not directly model housing externalities, which may be especially salient in slum areas. Poor sanitation, crime, and congested infrastructure may each depend differently on the quantity and quality of nearby houses, which may be affected by local housing projects (Marx et al.). Precise data on these outcomes would allow for disentangling the importance of each of these channels for the total welfare effects of these policies. Second, we abstract away from any potential interactions between housing projects and local labor markets. Housing projects may generate local agglomeration economies: housing projects increase the supply of local labor, which may induce firms to locate nearby these projects. Finally, without panel data on households over time, we are unable to provide a precise accounting of policy effects on individual households as well as the overall distributional consequences of housing policies in this context. Instead, our estimates are constrained to considering aggregate social welfare.

References

AECOM. Africa property and construction cost guide. Technical report, 2016. URL https://www.aecom.com/za/wp-content/uploads/2018/07/africa_property-construction_cost-guide_2018.pdf.

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.

B. Bradlow. Interview with manny sotomi. Technical report, 2018.

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.

City of Johannesburg Metro Municipality. Spatial development framework 2040, 2016. URL https://www.joburg.org.za/.

CSIR Building and Construction Technology. Guidelines for human settlement planning and design. Technical report, Department of Human Settlements, 2005.

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/.

D. Donaldson and A. Storeygard. The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4):171–98, 2016.

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.

D. Gardner and J. Pienaar. Benchmarking housing construction costs across africa. Technical report, Center for Affordable Housing Finance, May 2019. URL http://housingfinanceafrica.org/dashboards/benchmarking-housing-construction-costs-africa/.

Gauteng Department of Human Settlements. Mixed housing programme. Technical report, 2019. URL http://www.gauteng.gov.za/services/housing/Pages/Projects.aspx.

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

Government of South Africa. Housing act 107 of 1997, 2001. URL http://thehda.co.za/pdf/uploads/multimedia/Housing_Act_107_of_1997.pdf.

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.

J. Hofmeyr. *Risk and Opportunity*. Transformation audit. Institute for Justice and Reconciliation, 2008. ISBN 9781920299101. URL https://books.google.com/books?id=uOfEZf__85oC.

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. M. Stoker, and T. Suri. There is no free house: Ethnic patronage in a kenyan slum. *American Economic Journal: Applied Economics*.

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.

S. Rosen. Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, 82(1):34–55, 1974.

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

Z. Sajaia. Maximum likelihood estimation of a bivariate ordered probit model: implementation and monte carlo simulations. *The Stata Journal*, 4(2):1–18, 2008.

Statistics South Africa. Census 2001 metadata. Technical report, 2001. URL http://www.statssa.gov.za/census/census_2001/metadata/Geography.pdf.

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.

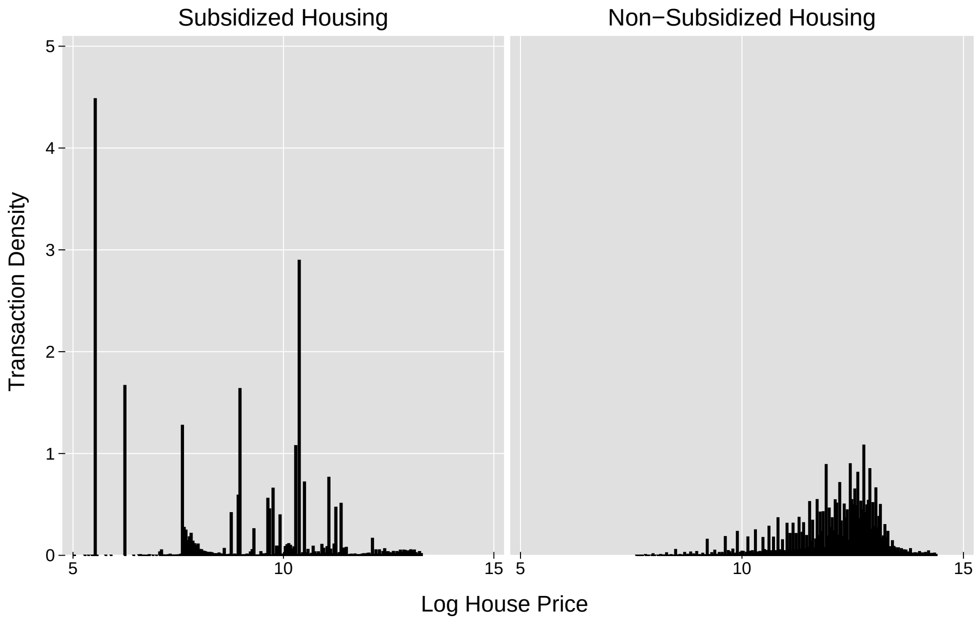
K. Tissington. Towards greater community participation in informal settlement upgrading: A case study from slovo park, johannesburg.

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/.

Figure VII. Transaction Price Histogram



Note: Transactions are censored at R100,000.

A. Appendix

A.A. Project Descriptions

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

We then use a fuzzy-string matching algorithm with bigrams to link project names from the budget reports to the administrative maps. We keep all matches with over 60% similarity, with 37 projects matching exactly. Appendix VIII compares unmatched and matched projects finding that matched projects have much higher numbers of project houses, lower nearby housing prices, and are further from the central business district compared to unmatched projects. One reason may be that the budget reports only include larger, more expansive projects, which are likely to take place further from city centers.

Table VII. Project Descriptions

	Counts
Constructed	
current	89
current/overflow	24
under implementation	45
complete	14
Total constructed	172
Unconstructed	
proposed	71
under planning	40
future	16
investigating	10
uncertain	8
Total unconstructed	145
Other	
informal	87
hostel	23
new	27
upg	6
p	1
no description	181
Total other	325
Total	642

A.C. Transaction Frequency

A.D. Project counts

A.E. Full Housing Demand Estimates

A.F. Price Estimates by Distance to Project Boundary

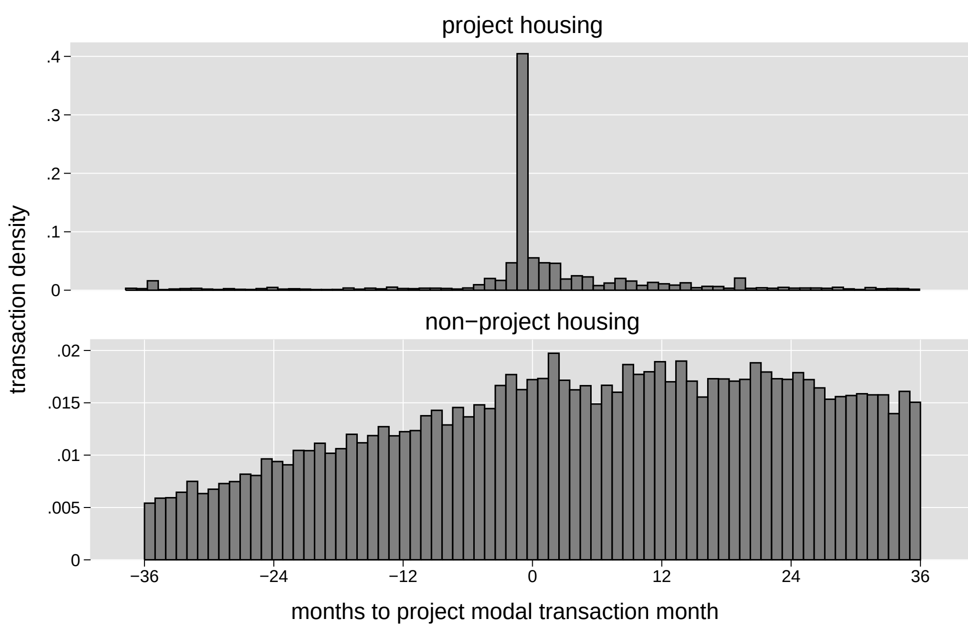
We examine how prices change differentially over time for constructed relative to unconstructed projects at different distances from project boundaries using the following

Table VIII. Assessing Name Matching between Budget Reports and Administrative Project Records

	Matched		**Unmatched**	
	Const.	Unconst.	Const.	Unconst.
Number of Projects	126	73	46	72
Area (km2)	1.22	1.02	1.05	1.30
Delivered Houses	416	15	259	7
House Price in 1 km (R†)	183,930	206,496	199,923	232,166
Distance to CBD‡ (km)	34.3	31.3	27.5	24.0

Const. refers to constructed projects and unconst. refers to unconstructed projects.
*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.

Figure VIII. Transaction Frequency Relative to Project Construction Month



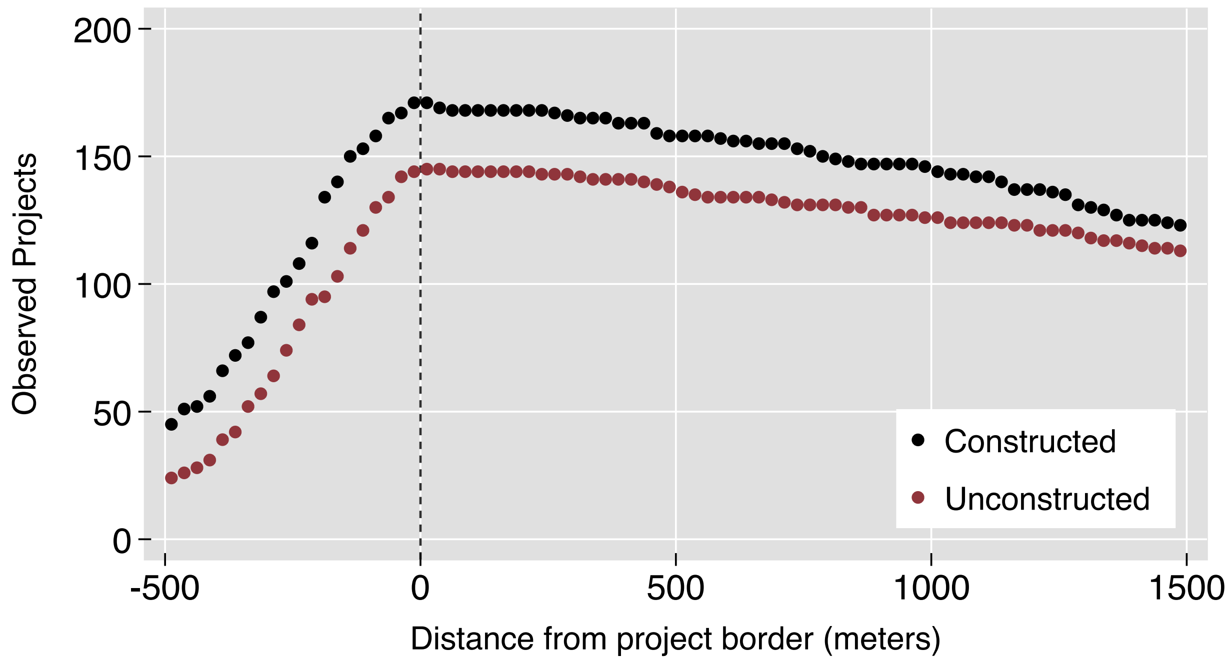
specification:

$$
\begin{aligned}
\text{Log-Price}_{ipnt} &= \beta_1\,\textsc{Post}_t \times \textsc{Const}_p \;+\; \beta_2\,\textsc{Post}_t \;+\; \beta_3\,\textsc{Const}_p \\
&\quad + \sum_{d=1}^{5} D_i^d\left(\alpha_1^d\,\textsc{Post}_t \times \textsc{Const}_p \;+\; \alpha_2^d\,\textsc{Post}_t \;+\; \alpha_3^d\,\textsc{Const}_p \;+\; \alpha_4^d\right) \\
&\quad + \theta X_{ipnt} \;+\; \lambda_n \;+\; \varepsilon_{ipnt}
\end{aligned}
$$

where

$$D_i^d = \mathbb{1}\Big\{-250 \le dist_i - (250 * d) \le 250\Big\}$$

Figure IX. Project Counts by Distance to Project Boundary



where distance to project boundary, d, ranges from 0 to 1,200 meters in intervals of 250 meters. $\textsc{Const}_p$ is an indicator for being near a constructed project, and $\textsc{Post}_t$ is an indicator that takes a value of one for properties sold after scheduled construction date for each project. Coefficient, β_1 captures average differential changes in log-prices for constructed and unconstructed projects pre and post scheduled construction located in neighborhood income quartile q. D_i^d takes a value of one for properties, i, in 250 meter intervals, d, from 0 to 1,250 meters. The coefficients of interest, $\alpha_1^{q,d}$ capture the differential changes in log-prices between constructed and unconstructed projects pre and post scheduled construction for each distance bin d relative to similar changes for properties located in the omitted distance category, which is the furthest distance bin (1,250 to 1,500 meters). α_1^d coefficients track how log-price changes evolve over distance to project boundaries. Additional controls X_{ipnt} include up to a cubic polynomial in the lot size of properties as well as a fixed effects for the year-month of each transaction. The specification also includes neighborhood fixed effects λ_n. Figure X plots α_1^d

Table IX. Triple Difference Estimates of Housing Density

	(1) Total Housing	(2) Formal Housing	(3) Informal Housing	(4) Informal Backyard Housing	(5) Informal Non-Bkyd. Housing
inside × constr × post	618.4[a]	649.7[a]	-31.3	-572.9[a]	541.6[a]
	(209.7)	(108.4)	(153.2)	(144.1)	(136.8)
0-500m away × constr × post	-216.7[b]	-34.3	-182.4[b]	-73.1	-109.4
	(110.4)	(61.9)	(92.7)	(80.7)	(80.7)
inside × post	373.7[b]	80.8	292.9[b]	277.1[b]	15.8
	(146.0)	(62.0)	(122.0)	(116.1)	(65.6)
0-500m away × post	473.2[a]	175.5[a]	297.7[a]	38.5	259.1[a]
	(92.9)	(57.1)	(77.5)	(75.6)	(65.3)
constr × post	77.4[c]	6.2	71.2[b]	-17.2	88.4[a]
	(41.6)	(13.4)	(36.0)	(12.6)	(33.4)
inside × constr	45.6	5.7	39.9	170.6[b]	-130.7[b]
	(103.9)	(48.9)	(87.5)	(78.1)	(57.3)
0-500m away × constr	-17.7	-29.7	11.9	-19.3	31.3
	(61.4)	(43.3)	(57.3)	(55.0)	(37.7)
post	171.2[a]	46.2[a]	125.0[a]	33.8[a]	91.2[a]
	(25.8)	(9.6)	(20.9)	(8.6)	(18.7)
inside	-167.8[a]	-163.8[a]	-3.9	66.1	-70.1[b]
	(61.4)	(34.4)	(47.0)	(42.5)	(29.0)
0-500m away	-116.2[b]	-57.4	-58.8	68.1	-126.9[a]
	(52.0)	(40.5)	(50.5)	(51.1)	(30.4)
constr	3.7	25.8	-22.0	-2.1	-19.9
	(33.7)	(17.3)	(24.0)	(14.2)	(17.7)
Mean Pre	579.36	288.49	290.87	184.19	106.68
Mean Post	922.93	424.41	498.52	182.46	316.06
R^2	0.322	0.292	0.276	0.272	0.200
N	6,143,756	6,143,756	6,143,756	6,143,756	6,143,756

Standard errors clustered at the project level in parenthesis. [c] p<0.10,[b] p<0.05,[a] p<0.01
Project areas are 25×25m grid cells within 1,500m of project footprints.
Housing densities are measure in terms of houses per km^2.
"inside" means within project footprint. "constr" means constructed.

coefficients with respect to distance to the nearest project boundary.

In Figure X, we observe low price changes nearby project boundaries, which rise moving to 750 meters from project boundaries before falling back to zero price change at further distances. Overall, this pattern does not provide evidence of a clear, monotonic gradient in housing price changes with respect to distance to housing projects.

We also explore how these effects vary according to neighborhood income quartile

Table X. Building Density by Neighborhood Income Quartile

	(1) Total Housing	(2) Formal Housing	(3) Informal Housing	(4) Informal Non-Bkyd. Housing	(5) Informal Bkyd. Housing
Q1: inside project	158.786	645.656[a]	-486.870	-1076.663[a]	589.793[a]
	(332.489)	(148.783)	(295.986)	(318.545)	(205.212)
Q1: 0-500m outside project	-127.239	-53.468	-73.771	20.876	-94.646
	(200.414)	(159.558)	(187.859)	(202.674)	(186.594)
Q2: inside project	1072.658[b]	641.216[a]	431.442	-181.306	612.748[b]
	(456.610)	(189.734)	(297.703)	(129.887)	(287.505)
Q2: 0-500m outside project	-348.734	51.902	-400.636	-9.340	-391.296[c]
	(256.546)	(41.531)	(246.177)	(60.775)	(228.942)
Q3: inside project	546.381[b]	652.869[a]	-106.488	-380.784[b]	274.297[a]
	(272.069)	(197.426)	(199.100)	(171.097)	(97.758)
Q3: 0-500m outside project	-391.067[b]	-176.028[c]	-215.039[c]	-59.031	-156.008
	(195.362)	(100.938)	(126.066)	(80.438)	(103.811)
Q4: inside project	524.434	497.350[c]	27.084	-88.771	115.855
	(518.930)	(253.939)	(302.993)	(64.540)	(267.761)
Q4: 0-500m outside project	-245.420	-21.493	-223.927	-169.971	-53.956
	(266.217)	(83.955)	(190.562)	(163.233)	(81.636)
Mean Pre	579.36	288.49	290.87	184.19	106.68
Mean Post	922.93	424.41	498.52	182.46	316.06
R^2	0.323	0.294	0.278	0.274	0.205
N	6,143,756	6,143,756	6,143,756	6,143,756	6,143,756

in the following specification:

$$\begin{aligned}\text{Log-Price}_{ipnt} = &\sum_{q=1}^{4} Q_n^q \Big[\beta_1^q \,\textsc{Post}_t \times \textsc{Const}_{pq} + \beta_2^q \,\textsc{Post}_t + \beta_3^q \,\textsc{Const}_{pq} \\ &+ \sum_{d=1}^{5} D_i^d \left(\alpha_1^{q,d} \,\textsc{Post}_t \times \textsc{Const}_{pq} + \alpha_2^{q,d} \,\textsc{Post}_t + \alpha_3^{q,d} \,\textsc{Const}_{pq} + \alpha_4^{q,d} \right) \Big] \\ &+ \theta X_{ipnt} + \lambda_n + \varepsilon_{ipnt}\end{aligned}$$

where

$$D_i^d = \mathbb{1}\Big\{ -250 \le dist_i - (250 * d) \le 250 \Big\}$$

where q indexes quartiles from 1 to 4 and Q_n^q takes a value of one if average income in

Table XI. Full Census Household-level Estimates

	(1) Is a Formal House	(2) Has a Flush Toilet	(3) Has Piped Water Indoors	(4) Has Electricity	(5) Total Rooms	(6) Owns House
inside × constr × post	0.614[a]	0.257	0.381[a]	0.103	1.887[a]	0.084
	(0.137)	(0.158)	(0.128)	(0.158)	(0.680)	(0.120)
0-500m away × constr × post	0.120	-0.053	0.062	-0.022	0.146	0.063
	(0.080)	(0.071)	(0.089)	(0.077)	(0.356)	(0.073)
inside × post	-0.276[b]	0.002	-0.219[c]	0.062	-1.609[b]	-0.147
	(0.125)	(0.148)	(0.119)	(0.147)	(0.652)	(0.113)
0-500m away × post	-0.057	0.059	-0.039	0.063	-0.219	-0.056
	(0.071)	(0.064)	(0.079)	(0.072)	(0.329)	(0.069)
constr × post	0.003	0.079[b]	0.047	0.035	0.048	-0.065[b]
	(0.033)	(0.031)	(0.038)	(0.036)	(0.157)	(0.027)
inside × constr	-0.276[b]	-0.108	-0.041	-0.140	-0.716	0.056
	(0.114)	(0.125)	(0.097)	(0.116)	(0.605)	(0.101)
0-500m away × constr	-0.054	0.006	-0.063	-0.017	-0.157	-0.045
	(0.065)	(0.064)	(0.071)	(0.067)	(0.304)	(0.064)
post	0.030	-0.030	0.020	0.036	0.328[a]	-0.083[a]
	(0.023)	(0.022)	(0.023)	(0.025)	(0.111)	(0.018)
inside	0.035	-0.046	-0.145	-0.070	0.081	-0.057
	(0.112)	(0.121)	(0.095)	(0.108)	(0.611)	(0.095)
0-500m away	0.011	-0.019	-0.007	-0.040	0.040	0.024
	(0.059)	(0.058)	(0.066)	(0.064)	(0.276)	(0.062)
constr	-0.045	-0.066[c]	-0.025	-0.031	-0.263	-0.031
	(0.030)	(0.034)	(0.037)	(0.038)	(0.187)	(0.025)
Mean Pre	0.62	0.71	0.47	0.70	3.48	0.41
Mean Post	0.65	0.77	0.54	0.77	3.79	0.36
R^2	0.571	0.639	0.607	0.607	0.642	0.610
N	20,641	20,639	20,639	20,639	20,500	20,631

Standard errors clustered at the project level in parenthesis.[c] $p<0.10$,[b] $p<0.05$,[a] $p<0.01$
Outcomes in Columns (1) to (4), (6), and (9) to (11) are indicator variables.
Specifications include census 2001 subplace fixed effects.
Results are weighted by unit area and control for up to a cubic polynomial in unit area by time.
"HH" stands for household and "HoH" stands for head of household.

neighborhood n falls within quartile q. Figure XI plots $\alpha_1^{d,q}$ coefficients with respect to distance to the nearest project boundary across income quartiles.

In Figure XI, the poorest quartile (Q1) exhibits a decreasing gradient consistent with an appreciation in prices close to project boundaries. The second and third quartiles (Q2 and Q3) provide little evidence of a clear pattern betwene prices and distance. The richest quartile (Q4) finds an increasing gradient consistent with large drops in prices nearby constructed projects.

Table XII. Full Census Household-level Estimates Cont.

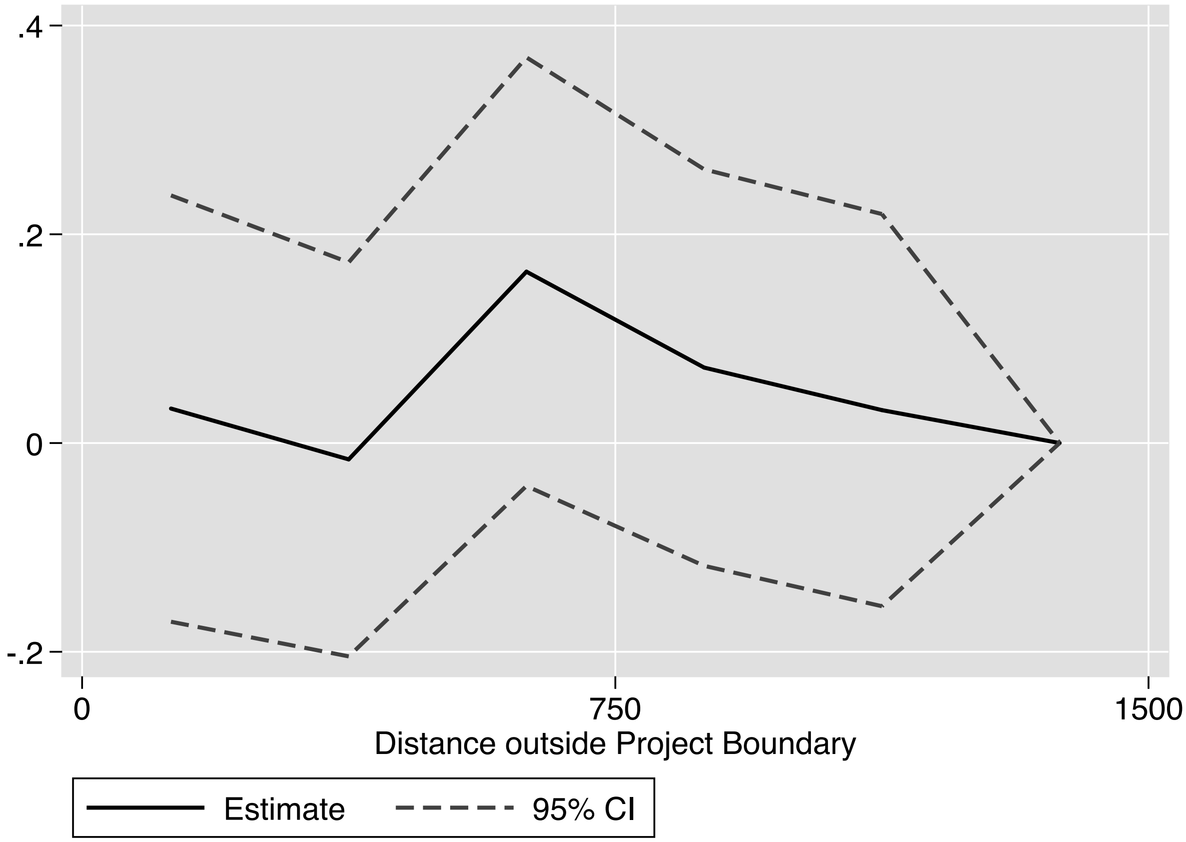
	(7) HH Size	(8) People per km^2	(9) Age HoH	(10) HoH is African	(11) HoH is Employed	(12) HH Log Income
inside × constr × post	0.586[b]	1610.323	2.785	0.007	-0.060	0.123
	(0.275)	(1288.772)	(2.095)	(0.094)	(0.092)	(0.312)
0-500m away × constr × post	0.222	378.259	0.010	-0.035	-0.042	0.213
	(0.185)	(708.021)	(1.265)	(0.061)	(0.037)	(0.176)
inside × post	-0.331	707.134	-3.285[c]	0.012	0.069	-0.373
	(0.243)	(1171.449)	(1.956)	(0.088)	(0.089)	(0.295)
0-500m away × post	-0.073	845.731	0.015	0.027	0.057	-0.394[b]
	(0.168)	(633.891)	(1.196)	(0.054)	(0.035)	(0.156)
constr × post	-0.151[b]	-303.780	0.138	-0.032	0.041[b]	-0.087
	(0.075)	(196.213)	(0.504)	(0.029)	(0.018)	(0.085)
inside × constr	-0.114	2052.603[a]	-0.691	-0.055	0.060	0.225
	(0.240)	(753.955)	(1.776)	(0.074)	(0.079)	(0.251)
0-500m away × constr	-0.180	588.430	0.682	0.026	0.037	-0.125
	(0.160)	(537.429)	(0.854)	(0.052)	(0.030)	(0.129)
post	-0.268[a]	849.690[a]	1.629[a]	0.096[a]	0.019	1.009[a]
	(0.049)	(152.432)	(0.361)	(0.018)	(0.015)	(0.055)
inside	-0.077	-620.895	-0.385	0.171[b]	-0.117	-0.487[c]
	(0.218)	(654.993)	(1.744)	(0.070)	(0.078)	(0.251)
0-500m away	0.094	466.651	-0.883	0.018	-0.057[b]	0.077
	(0.151)	(497.352)	(0.840)	(0.048)	(0.029)	(0.122)
constr	0.021	-260.744	-0.321	0.045	-0.027[c]	-0.134
	(0.081)	(170.227)	(0.475)	(0.035)	(0.015)	(0.083)
Mean Pre	3.25	1,793.45	41.26	0.73	0.82	7.30
Mean Post	2.89	2,690.42	42.34	0.81	0.83	8.25
R^2	0.734	0.583	0.606	0.737	0.590	0.756
N	20,509	21,197	20,505	20,641	20,501	20,495

Standard errors clustered at the project level in parenthesis.[c] $p<0.10$,[b] $p<0.05$,[a] $p<0.01$
Outcomes in Columns (1) to (4), (6), and (9) to (11) are indicator variables.
Specifications include census 2001 subplace fixed effects.
Results are weighted by unit area and control for up to a cubic polynomial in unit area by time.
"HH" stands for household and "HoH" stands for head of household.

A.G. Price Estimates by Months to Project Construction

We examine how prices change differentially for properties less than 500 meters outside of boundaries compared to projects between 500 and 1,500 meters outside of boundaries and between constructed and unconstructed project areas. We further plot these differences with respect to months to the (expected) construction month of each

Figure X. Log-Price Estimates by Distance to Nearest Project Boundary



Estimates include differential changes in prices before and after (construction) for constructed relative to unconstructed project areas. Estimates are normalized with respect to price changes in the furthest distance group — 1,250 to 1,500 meters from project boundaries. The y-axis is in terms of changes in log-prices.

project across neighborhood income quartiles using the following specification:

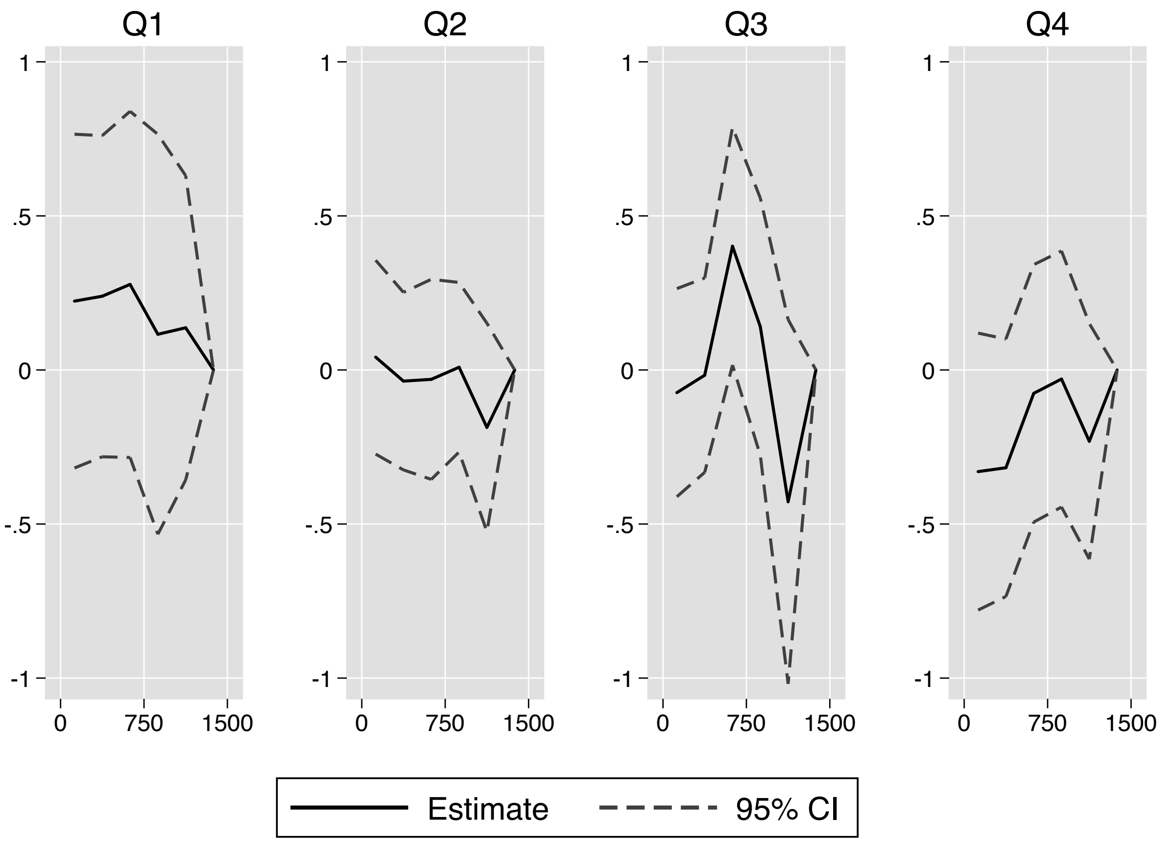
$$\begin{aligned}\text{Log-Prices}_{ipnt} &= \beta_1\,\textsc{Spill}_i \times \textsc{Const}_p \;+\; \beta_2\,\textsc{Spill}_i \;+\; \beta_3\,\textsc{Const}_p \\ &\quad + \sum_{l=1}^{5} T_l \left(\alpha_1^l\,\textsc{Spill}_i \times \textsc{Const}_p \;+\; \alpha_2^l\,\textsc{Spill}_i \;+\; \alpha_3^l\,\textsc{Const}_p \;+\; \alpha_4^l \right) \\ &\quad + \theta X_{ipnt} \;+\; \lambda_n \;+\; \varepsilon_{ipnt}\end{aligned}$$

Where

$$T_1 = \mathbb{1}\Big\{ t \leq -48 \Big\},\;\; T_2 = \mathbb{1}\Big\{ -24 \leq t < -48 \Big\},$$

$$T_3 = \mathbb{1}\Big\{ 0 < t <= 24 \Big\},\;\; T_4 = \mathbb{1}\Big\{ 24 < t <= 48 \Big\},\;\; T_5 = \mathbb{1}\Big\{ 48 \leq t \Big\}$$

Figure XI. Log-Price Estimates by Distance to Nearest Project Boundary by Neighborhood Income Quartile



Estimates include differential changes in prices before and after scheduled construction relative to unconstructed project areas. Estimates are normalized with respect to price changes in the furthest distance group — 1,250 to 1,500 meters from project boundaries. Q1 to Q4 index neighborhood income quartiles measured using the 2001 census of population and housing. The y-axis is in terms of changes in log-prices and the x-axis measures meters outside of the nearest housing project boundary.

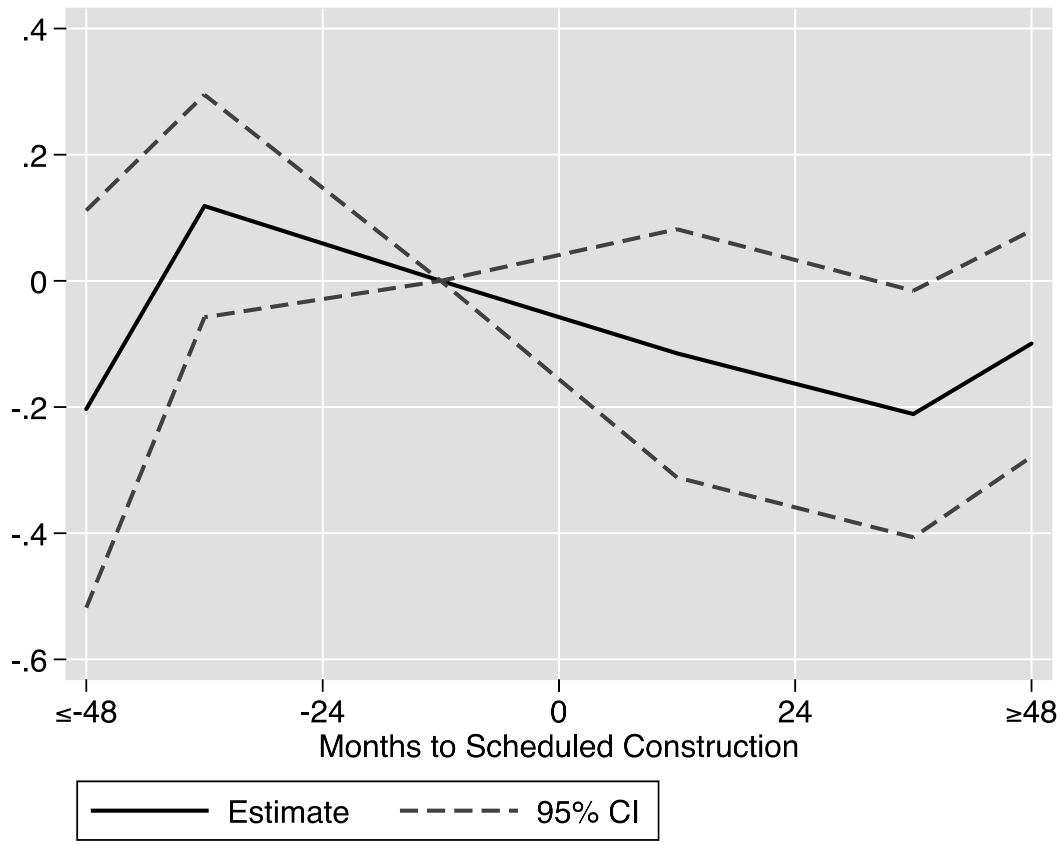
where SPILL_i is an indicator for properties within 500 meters outside of a project boundary and CONST_p is an indicator for being nearby a project that is successfully constructed. t indexes months to scheduled construction with the scheduled construction month normalized to zero. Differential prices near and far from constructed projects compared to unconstructed projects are estimated separately according to (1) all transactions more than 48 months before scheduled construction, (2) transactions between 48 and 24 months prior to scheduled construction, (3) transactions within 24 months after scheduled construction, (4) transactions within 24 months to 48 months after scheduled construction, and (5) all transactions over 48 months after scheduled construction. This specification leaves transactions occurring in the 24 months immediately before scheduled construction as the omitted category. Coefficient, β_1 captures differential changes in log-prices near and far from constructed and unconstructed projects across all time periods. The coefficients of interest, α_1^l track how prices change differentially in the 5 time windows relative to the omitted category — 24 months before scheduled construction. Additional controls X_{ipnt} include up to a cubic polynomial in the lot size of properties as well as a fixed effects for the year-month of each transaction. The specification also includes neighborhood fixed effects λ_n. Figure XII plots α_1^l coefficients with respect to months to scheduled construction.

Figure XII finds a negative price differential 48 or more months before construction, which jumps above zero 48 to 24 months before scheduled construction. The results support a declining pretrend with prices dropping by 10% in the 48 months before scheduled construction. This trend continues with a similar slope in the 48 months after scheduled construction, reaching a 20% decline in prices between 24 and 48 months after construction. Over 48 months after construction the price difference remains negative.

Figure XII finds a negative price differential 48 or more months before construction, which recovers to above zero in 48 to 24 months before scheduled construction. In the 48 months after scheduled construction, the price difference dips increasingly below zero. Over 48 months after construction the price difference remains negative.

We also explore how these effects vary according to neighborhood income quartile

Figure XII. Log-Prices By Months to Scheduled Construction



Estimates capture differential changes in log-prices for properties within 500m of project boundaries compared to properties located between 500m to 1,500m from project boundaries for constructed relative to unconstructed projects. These estimates are computed at different time periods relative to the scheduled construction month of each project with price changes in the 24 months before scheduled construction normalized to zero. The y-axis is in terms of changes in log-prices.

in the following specification:

$$\begin{aligned}\text{Log-Prices}_{ipnt} = &\sum_{q=1}^{4} Q_n^q \Big[\ \beta_1^q \,\textsc{Spill}_{iq} \times \textsc{Const}_{pq} + \beta_2 \,\textsc{Spill}_{iq} + \beta_3 \,\textsc{Const}_{pq} \\ &+ \sum_{l=1}^{5} T_l \left(\alpha_1^{q,l} \,\textsc{Spill}_{iq} \times \textsc{Const}_{pq} + \alpha_2^{q,l} \,\textsc{Spill}_{iq} + \alpha_3^{q,l} \,\textsc{Const}_{pq} + \alpha_4^{q,l} \right) \Big] \\ &+ \theta X_{ipnt} + \lambda_n + \varepsilon_{ipnt}\end{aligned}$$

Where

$$T_1 = \mathbb{1}\Big\{ t \leq -48 \Big\},\ \ T_2 = \mathbb{1}\Big\{ -24 \leq t < -48 \Big\}\ ,$$

$$T_3 = \mathbb{1}\Big\{0 < t <= 24\Big\}\ ,\ \ T_4 5\not\equiv\ \mathbb{1}\Big\{24 < t <= 48\Big\},\ \ T_5 = \mathbb{1}\Big\{48 \leq t\Big\}$$

where q indexes quartiles from 1 to 4 and Q_n^q takes a value of one if average income in neighborhood n falls within quartile q. Figure XIII plots $\alpha_1^{d,q}$ coefficients with respect to time to scheduled construction across income quartiles.

Figure XIII provides evidence of a strong decreasing pretrend in prices for the first quartile, dropping from price increases of around 100% over 48 months before scheduled construction. By contrast, the fourth quartile exhibits little pretrend in anticipation of scheduled construction. Following scheduled construction, prices following a slight upward trend in the first quartile. Moving from the second to the fourth quartiles, trends in prices after scheduled construction become more and more negative. Prices for the third and fourth quartiles drop substantially in the 48 months reaching just after scheduled construction before partially rebounding more than 48 months after scheduled construction.

A.H. Difference-in-Differences Estimates on Log-Prices

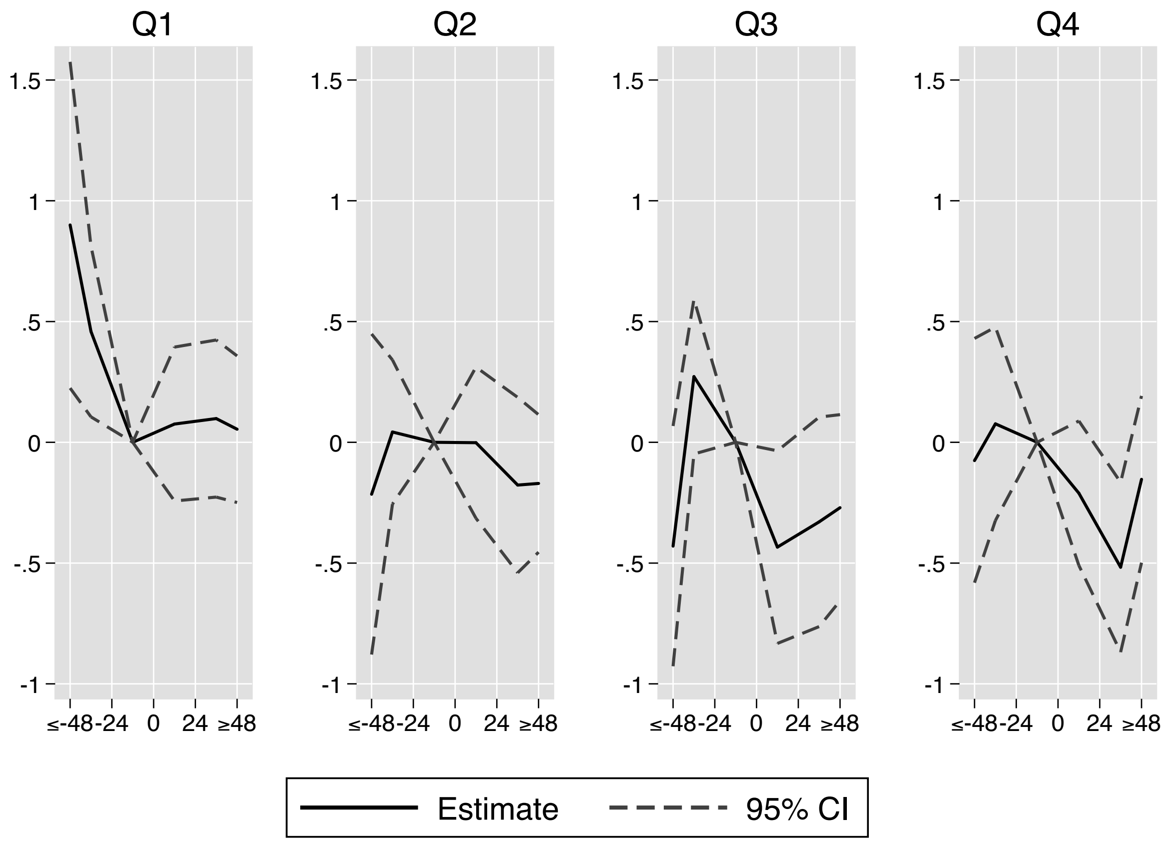
As a robustness exercise, we estimate a differences-in-differences specification comparing properties close and far from project boundaries before and after scheduled construction only for projects that were successfully constructed. We implement the following specification:

$$y_{ipnt} = \beta_1 \textsc{Spill}_{ip} \times \textsc{Post}_t + \beta_2 \textsc{Post}_t + \beta_3 \textsc{Spill}_{ip} + \theta X_{ipnt} + \lambda_n + \varepsilon_{ipnt} \tag{9}$$

We also estimate this differences-in-differences specification separately according to neighborhood income quartile with the following specification:

$$y_{ipnt} = \sum_{q=1}^{4} Q_n^q \left[\beta_1^q \textsc{Spill}_{ipq} \times \textsc{Post}_t + \beta_2^q \textsc{Post}_t + \beta_3^q \textsc{Spill}_{ipq} \right] + \theta X_{ipnt} + \lambda_n + \varepsilon_{ipnt} \tag{10}$$

Figure XIII. Log-Prices By Months to Scheduled Construction by Neighborhood Income Quartile



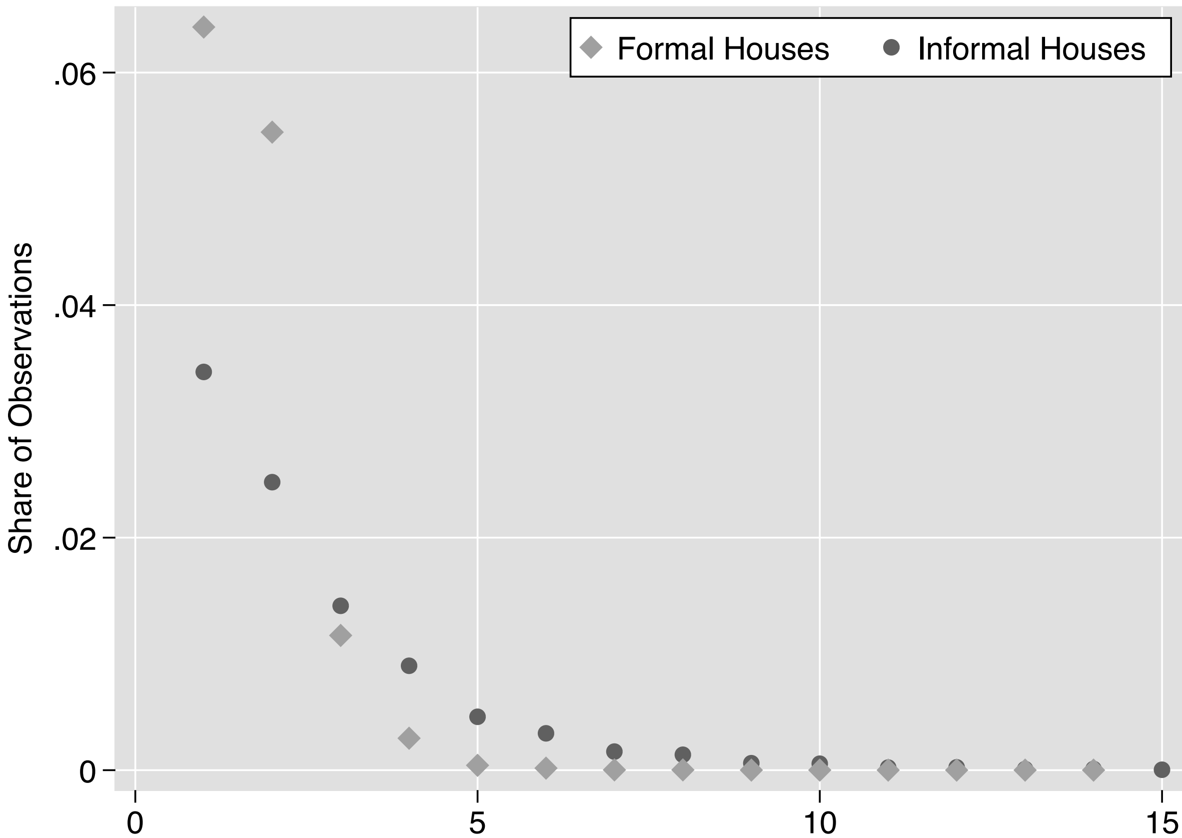
Estimates capture differential changes in log-prices for properties within 500m of project boundaries compared to properties located between 500m to 1,500m from project boundaries for constructed relative to unconstructed projects. These price effects are computed at different time periods relative to the scheduled construction month of each project with price effects in the 24 months before scheduled construction normalized to zero. Q1 to Q4 index neighborhood income quartiles measured using the 2001 census of population and housing. The y-axis is in terms of changes in log-prices and the x-axis measures months to scheduled construction.

where q indexes quartiles from 1 to 4 and Q_n^q takes a value of one if average income in neighborhood n falls within quartile q.

Table XVI provides estimates across all neighborhoods and by income quartile separately, finding similar results to triple-difference results in Table VI.

A.I. Building Histogram

Figure XIV. Histogram of Buildings per Plot with at least 1 Building



Plots with zero buildings are not shown. 86% of plots have zero formal houses, and 91% of plots have zero informal houses. This figure also excludes plots with over 15 buildings, which compose less than 0.001% of the sample.

A.J. Welfare Simulation

To simulate counterfactual housing choices, we begin by taking 10,000 random draws from a standard normal distribution, which represent plot-specific random amenity shocks. The simulation sets housing costs, C equal to the average housing price, R231,000, and sets θ equal to 0.0000094, which is the marginal utility of income estimate for formal housing. Estimated cutpoints, $\Lambda_h(k)$ allow us to recover the implied congestion cost $\lambda_h(k)$, assuming a reservation utility, $\overline{U}$ equal to zero. Net amenities are are given by

$$
\begin{aligned}
\overline{V}_{hlt} = \textsc{Proj}_{lp} \; &\left(\alpha_1^{Vh}\,\textsc{Post}_t \times \textsc{Const}_p \;+\; \alpha_2^{Vh}\,\textsc{Post}_t \;+\; \alpha_3^{Vh}\,\textsc{Const}_p \;+\; \alpha_4^{Vh}\right) + \\
&\textsc{Spill}_{lp} \left(\beta_1^{Vh}\,\textsc{Post}_t \times \textsc{Const}_p \;+\; \beta_2^{Vh}\,\textsc{Post}_t \;+\; \beta_3^{Vh}\,\textsc{Const}_p \;+\; \beta_4^{Vh}\right) + \\
&\gamma_1^{Vh}\,\textsc{Post}_t \times \textsc{Const}_p \;+\; \gamma_2^{Vh}\,\textsc{Post}_t \;+\; \gamma_3^{Vh}\,\textsc{Proj}_p
\end{aligned}
$$

where the change in welfare is computed by shifting α_1^{Vh} and β_1^{Vh} from zero to their empirical estimates for project and spillover areas respectively. Given the full set of parameters, we are able to simulate optimal building choices, k_h^* as well as prices, P_h^* before and after project construction.

Table XIII. Full Housing Demand Estimates

	Formal	Informal
net amenities: $(\delta_{hl} - \theta c^{u}_{hlt})$		
inside × constr × post	0.32[c]	-0.19[c]
	(0.19)	(0.10)
0-500m away × constr × post	-0.02	-0.12
	(0.07)	(0.09)
inside × post	0.50[a]	0.15
	(0.16)	(0.09)
0-500m away × post	0.16[b]	0.10
	(0.06)	(0.09)
constr × post	-0.02	-0.02
	(0.02)	(0.04)
inside × constr	0.48[b]	0.35
	(0.20)	(0.21)
0-500m away × constr	-0.26[c]	-0.19
	(0.15)	(0.17)
post	0.08[a]	0.22[a]
	(0.02)	(0.03)
inside	-0.73[a]	0.43[b]
	(0.16)	(0.17)
0-500m away	0.35[a]	0.57[a]
	(0.12)	(0.14)
constr	0.31[a]	0.30[a]
	(0.10)	(0.10)
marginal utility of income: $-\theta$	-0.0000094[a]	-0.0000103[a]
	(0.0000035)	(0.0000038)
cut point: $\Lambda_h(1)$	-0.815	-0.580
	(0.834)	(0.902)
cut point: $\Lambda_h(2)$	-0.428	-0.321
	(0.838)	(0.901)
cut point: $\Lambda_h(3)$	0.284	-0.060
	(0.837)	(0.900)
cut point: $\Lambda_h(4)$	0.797	0.158
	(0.837)	(0.900)
cut point: $\Lambda_h(5)$	1.270	0.365
	(0.836)	(0.899)
cut point: $\Lambda_h(6)$	1.534[c]	0.523
	(0.833)	(0.898)
cut point: $\Lambda_h(7)$	1.887[b]	0.683
	(0.830)	(0.897)
cut point: $\Lambda_h(8)$	1.998[b]	0.802
	(0.828)	(0.897)
cut point: $\Lambda_h(9)$	2.245[a]	0.947
	(0.823)	(0.897)
cut point: $\Lambda_h(10)$	2.320[a]	1.045
	(0.822)	(0.898)
correlation of amenity shocks: $\rho_{for,inf}$	0.620[a]	
	(0.018)	
N	6,143,756	

Standard errors clustered at the project level in parenthesis. [c] p<0.10,[b] p<0.05,[a] p<0.01
Outcome categories by plot and year include categories for formal and informal building counts from 0 to 9 and a category at least 10. Table XIII includes the full set of coefficients. Net amenities equal local amenities minus local housing costs.

Table XIV. Housing Demand Estimates by Income Quartile

	Formal Housing	Informal Housing
Net amenities: $(\delta_{hl} - \theta c^{u}_{hlt})$		
Q1: inside project	0.204	0.322[c]
	(0.207)	(0.166)
Q1: 0-500m outside	-0.094	0.510[a]
	(0.210)	(0.143)
Q2: inside project	0.921[a]	0.318[c]
	(0.223)	(0.193)
Q2: 0-500m outside	0.186	0.159
	(0.146)	(0.161)
Q3: inside project	0.538	-0.567[c]
	(0.465)	(0.294)
Q3: 0-500m outside	-0.102	-0.546[a]
	(0.155)	(0.191)
Q4: inside project	0.346	-0.782
	(0.476)	(0.606)
Q4: 0-500m outside	0.138	-1.243[a]
	(0.149)	(0.376)
Marginal utility of income: $-\theta_h$	-0.0000089[b]	-0.0000047
	(0.0000036)	(0.0000037)
Correlation of amenity shocks: $\rho_{for,inf}$	0.620[a]	
	(0.018)	
N	6,143,756	

Standard errors clustered at the project level in parenthesis. [c] $p<0.10$,[b] $p<0.05$,[a] $p<0.01$
Outcome categories include formal and informal building counts from 0 to 9
and at least 10 by plot and year. Table XIII includes the full set of coefficients.
Net amenities equal local amenities minus local housing costs.

Table XV. Simulated Effects Housing Projects on Net Prices by Neighborhood Income Quartile

	Area	Formal Housing	Informal Housing
All	inside project	17,772	-16,743
	0-500m outside	-869	-5,021
Quartile			
Q1	inside project	12,145	28,977
	0-500m outside	-4,360	27,978
Q2	inside project	20,619	21,918
	0-500m outside	7,920	9,522
Q3	inside project	21,299	-34,861
	0-500m outside	-3,045	-24,882
Q4	inside project	9,833	-36,683
	0-500m outside	2,442	-45,657

All results are in Rands per plot. Average housing price is 231,000 Rands. Quartiles are measured with respect to average neighborhood household income in the 2001 census. Appendix A.J describes the simulation in more detail.

Table XVI. Differences-in-Differences Estimates of Log-Prices 0 to 500 Meters Outside of Housing Projects

	(1)	(2)
All Neighborhoods	-0.068[c]	-0.062[c]
	(0.037)	(0.037)
R2	0.49	0.50
By Neighborhood Income Quartile		
Q1	-0.068	-0.061
	(0.066)	(0.067)
Q2	0.045	0.049
	(0.093)	(0.090)
Q3	-0.075	-0.074
	(0.058)	(0.060)
Q4	-0.209[a]	-0.195[a]
	(0.052)	(0.052)
Year-Month FE		✓
Plot Size (up to cubic)		✓
r2	0.50	0.50
N	34,839	34,839

Estimates capture differential price effects for properties 0 to 500 meters outside of project boundaries relative to properties 500 to 1,500 meters from boundaries before and after scheduled construction.
Income quartiles are from average household 2001 incomes in census 2001 subplaces. All estimates include neighborhood fixed effects.
Standard errors clustered at the project level in parenthesis.
[c] $p<0.10$, [b] $p<0.05$, [a] $p<0.01$