

# Subsidized Housing & Urban Development: Evidence from South Africa

Ben Bradlow  
Stefano Polloni  
Will Violette

October 2018

# Introduction

- ▶ Housing is a pressing issue in the developing world due to rapid urbanization.
  - ▶ 30% of urban pop living in informal housing (UN, 2015)

# Introduction

- ▶ Housing is a pressing issue in the developing world due to rapid urbanization.
  - ▶ 30% of urban pop living in informal housing (UN, 2015)
- ▶ Slums thought to be lasting poverty traps (Marx, 2013):
  - ▶ Poor infrastructure, High crime, Health externalities
  - ▶ Weak incentives to invest in housing/public goods

# Introduction

- ▶ Housing is a pressing issue in the developing world due to rapid urbanization.
  - ▶ 30% of urban pop living in informal housing (UN, 2015)
- ▶ Slums thought to be lasting poverty traps (Marx, 2013):
  - ▶ Poor infrastructure, High crime, Health externalities
  - ▶ Weak incentives to invest in housing/public goods
- ▶ Government response → **Public Housing Provision**

# What do we know?

## ① Direct Recipient Impacts

- ▶ Health, Well-Being:  
Cattaneo et al. [2009], Galiani et al. [2017]
- ▶ Employment, Income:  
Barnhardt et al. [2015], Picarelli [2017], Franklin [2018]

# What do we know?

## ① Direct Recipient Impacts

- ▶ Health, Well-Being:  
Cattaneo et al. [2009], Galiani et al. [2017]
- ▶ Employment, Income:  
Barnhardt et al. [2015], Picarelli [2017], Franklin [2018]

## ② Indirect Recipient Impacts

- ▶ Informal housing possibilities within project areas.
- ▶ Amenity value to neighboring areas:  
Diamond & McQuade [2016], Baum-Snow & Marion [2008]

# Indirect Recipients

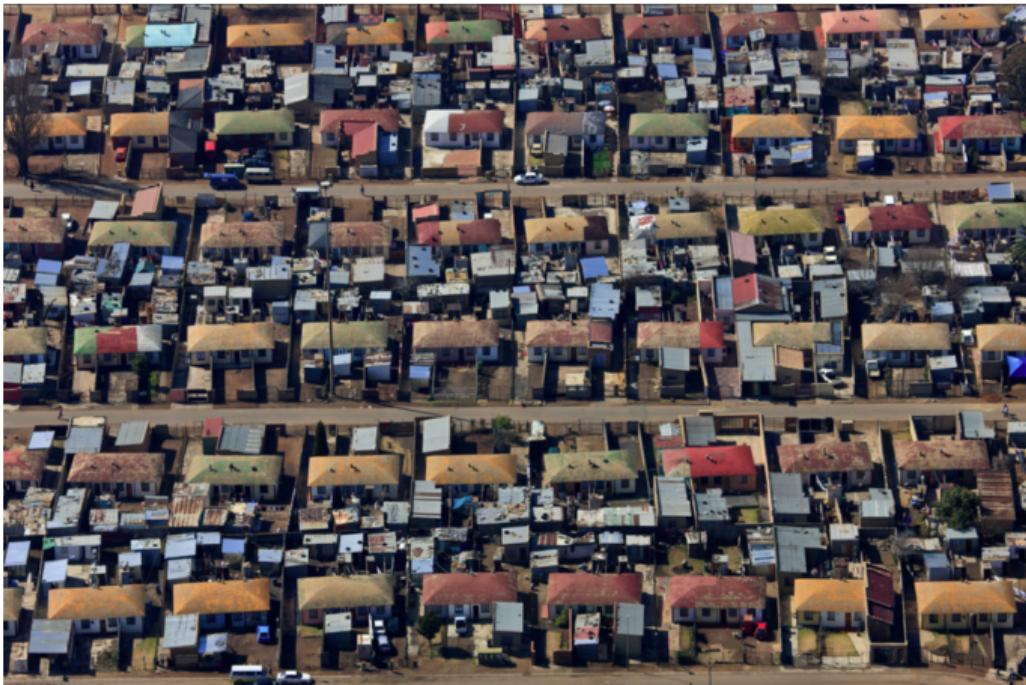


Figure: Government Housing with Backyard Shacks in Soweto

# This Paper

► **Question:**

What are the urban development impacts from subsidized housing in developing contexts?

- ▶ Evaluate household's access to services within projects and nearby.
- ▶ Examine composition of building growth (formal and informal structures) in subsidized vs non-subsidized areas.
- ▶ Assess spillover effects on nearby home values in the private housing market.

# This Paper

- ▶ **Approach:**

- ① Leverage granular spatial data with precise geography of housing projects.
- ② Use planned but unconstructed projects as additional counter-factual.

- ▶ **Data and Setting:**

~ 60 public housing projects in Gauteng province combined with GPS property transactions, building-based land information, and census data.

## Preview of Findings

- ① Significant improvement in households' access to water, sanitation, and electricity within project areas.

## Preview of Findings

- ① Significant improvement in households' access to water, sanitation, and electricity within project areas.
- ② Public housing is direct substitute for informal housing:  
Total structure growth no different than control areas, but significant composition effects.

## Preview of Findings

- ① Significant improvement in households' access to water, sanitation, and electricity within project areas.
- ② Public housing is direct substitute for informal housing:  
Total structure growth no different than control areas, but significant composition effects.
- ③ No evidence of amenity spillovers to neighboring areas.

## Background

## Public Housing in South Africa

- ▶ Large national subsidy scheme providing housing opportunities to eligible households.
- ▶ 3 million houses delivered since program inception in 1994.
  - ▶ free-standing, single-story, two-bedrooms, 30-40m<sup>2</sup> dwellings
- ▶ Annual expenditure of 6bn Rands (US\$500M).
- ▶ Supply planned by national → provincial → municipal housing agencies, project construction outsourced to private developers.
- ▶ Constraints on costs per unit, services access, and rooms/lot sizes.

# Where are the houses built?

- ▶ Vacant private or government-owned land-plots with varying levels of (mostly informal) residential structures.



- ▶ Projects are (should be) fully serviced: roads, water, sanitation, electricity.

# Who gets a house?

- ▶ **Official Policy:**

- ▶ Must be eligible: South African Citizen, Married or with dependents, Monthly income < R3,500.
- ▶ National/provincial waiting lists.
- ▶ No resale within 7 years.

# Who gets a house?

- ▶ **Official Policy:**

- ▶ Must be eligible: South African Citizen, Married or with dependents, Monthly income < R3,500.
- ▶ National/provincial waiting lists.
- ▶ No resale within 7 years.

- ▶ **In Practice:**

- ▶ Wait-lists and eligibility weakly enforced, with many noted cases of corruption.
- ▶ Developers often fail to meet quality requirements.
- ▶ 20% of houses occupied not occupied by initial owners after 5 years.
- ▶ More than a third have backyard shacks after 2 or less years.

# Canceled/Delayed Projects

*"R95m down the tubes as housing project picked apart brick by brick"*  
– Timeslive, 2017

*"MEC Mashatile delays Munsieville Ext 5 multimillion housing project"*  
– DA-GPL, 2017

*"Objections put R242m housing project on hold"*  
– IOL News, 2016

## Why?

- ▶ Disputes over beneficiaries; disagreement with security contractor.
- ▶ Lack of approval/coordination from all agencies.
- ▶ Failed environmental impact assessment.

## Data and Set-up

## Data Sources

- ▶ Focus on Gauteng Province (includes Johannesburg and Pretoria)

# Data Sources

- ▶ Focus on Gauteng Province (includes Johannesburg and Pretoria)
- ▶ Four main data sources:
  - ① **Administrative Data on project location and planning**
  - ② **Deeds-Based Housing Transactions**
  - ③ **Building-Based Land Use**
  - ④ **Household Census**

## Administrative Data on Subsidized Housing.

- ▶ Data on location of government housing initiatives as of 2008.
  - ▶ Includes planned but unconstructed projects.
- ▶ Annual Budget Reports from National Treasury.

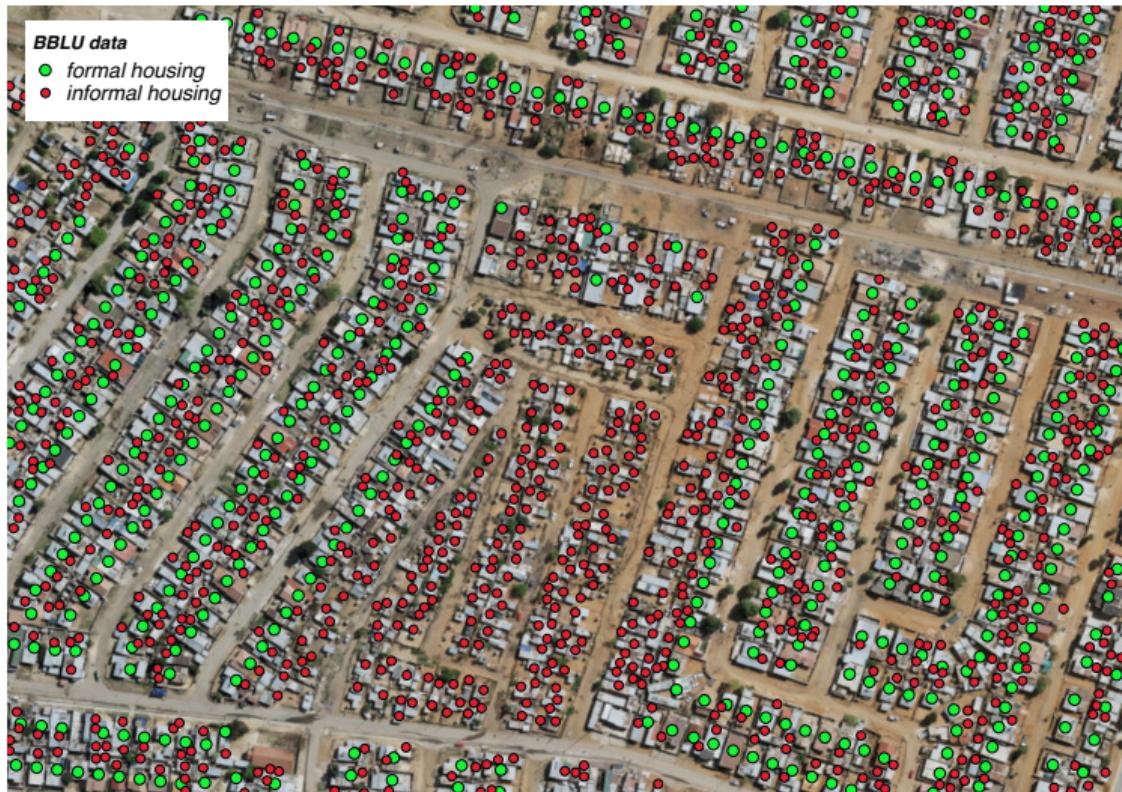
## Housing Transactions

- ▶ Sourced from South African deeds registry.
- ▶ "Universe" of formal housing transactions recorded during 2001-2011 in **affordable areas**. (550K transactions)
- ▶ Exact geographic location of traded property, but limited information on characteristics other than price and lot size.
- ▶ Includes buyer and seller name.

## Building Based Land-Use

- ▶ GeoTerralmage<sup>©</sup>: Exhaustive hand-coded building identification from aerial imagery.
- ▶ 2 cross-sections: 2001 and 2011.
- ▶ Building type differentiated by category: residential, commercial, industrial, etc.
- ▶ Within residential, ability to differentiate formal from informal housing, including backyard shacks.
- ▶ High Correlation ( $>80\%$ ) with reported dwelling type from census data at coarser spatial resolution.

# Building Based Land-Use



## Census Data

- ▶ Full coverage from 2001 and 2011 censuses at the household level.
- ▶ Smallest identifiable geography is **small area**.
  - ▶ Gauteng: 11,000 small areas in 2001, 17,000 in 2011.
  - ▶ 170 household per small area, on average.
- ▶ Summary information on dwelling characteristics and access to services.

# Identifying Housing Projects

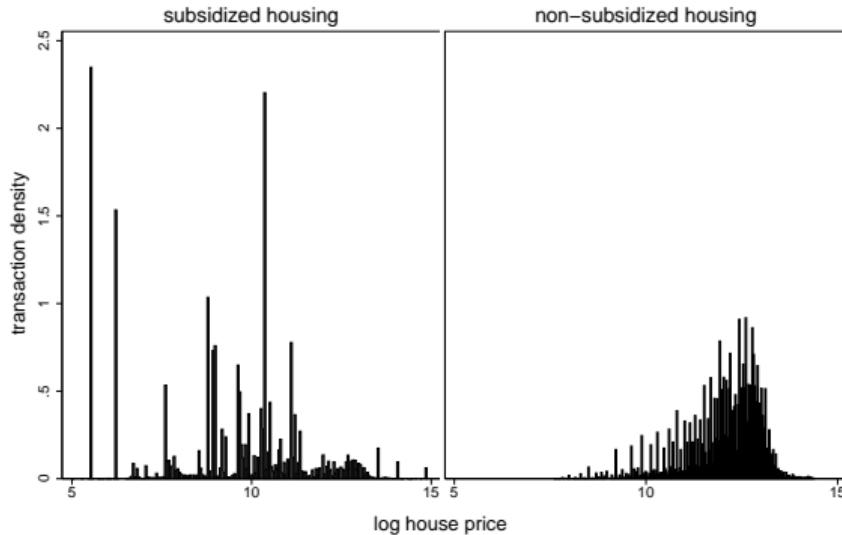
- ① **Filter on Seller Identity:** Find Governments, Housing authorities and Large Sellers from seller-names in transactions.

Figure: Top 5 Seller Names

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

# Identifying Housing Projects

- ① **Filter on Seller Identity:** Find Governments, Housing authorities and Large Sellers from seller-names in transactions.
- ② **Filter on Price:** Exclude purchase prices R50,000 above yearly subsidy value for government houses.



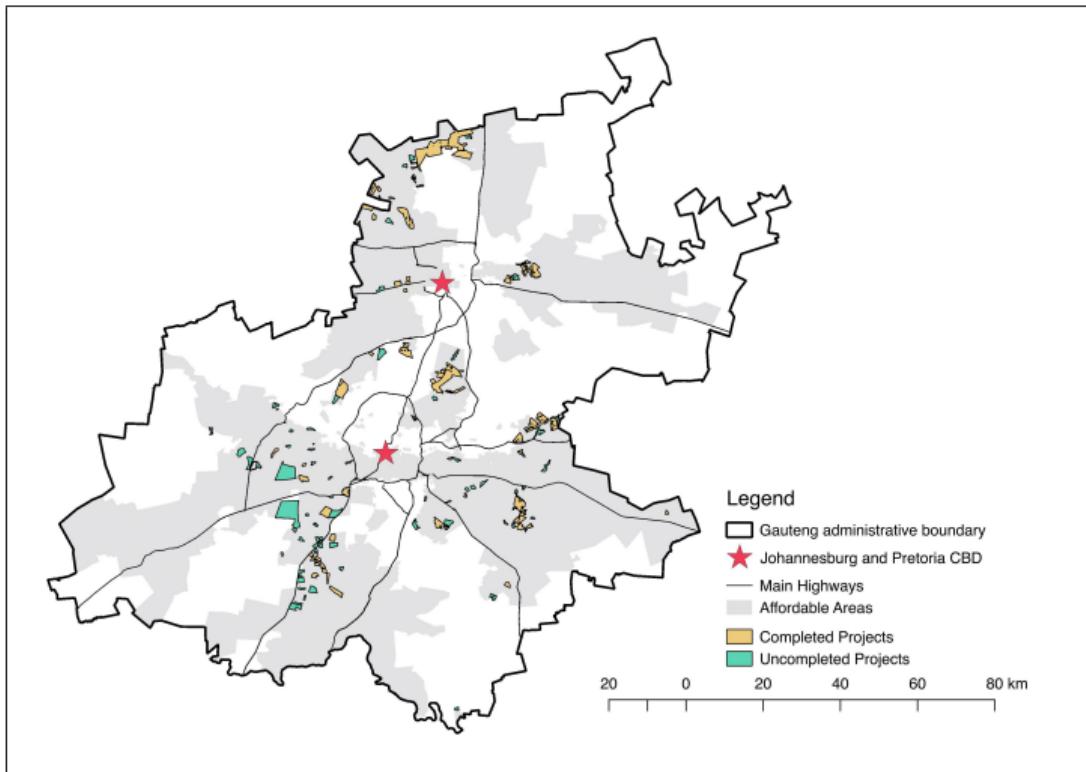
# Identifying Housing Projects

- ① **Filter on Seller Identity:** Find Governments, Housing authorities and Large Sellers from seller-names in transactions.
- ② **Filter on Price:** Exclude purchase prices R50,000 above yearly subsidy value for government houses.
- ③ **Pre-Existing Formal Dwellings:** Exclude land plots containing formal structures from 2001 building-based land-use data.

# Identifying Housing Projects

- ① **Filter on Seller Identity:** Find Governments, Housing authorities and Large Sellers from seller-names in transactions.
- ② **Filter on Price:** Exclude purchase prices R50,000 above yearly subsidy value for government houses.
- ③ **Pre-Existing Formal Dwellings:** Exclude land plots containing formal structures from 2001 building-based land-use data.
- ④ **Spatial Clustering:** collect nearby houses into projects with density-based clustering algorithm. (DBSCAN)

# Title



# Identifying Planned but Unconstructed Projects

- ① Admin. data have “planned,” “proposed,” “implementing” projects
    - ▶ Exclude projects with identified project transactions
  - ② Assign projects an expected completion date
    - ▶ Fuzzy-string match budget data (with start-dates) on project names
    - ▶ Add avg. diff. between transaction-date and start-date for completed projects
- ▶ Why are projects canceled/delayed?
- ▶ Legal disputes, service delivery backlogs, funding complications
  - ▶ Delays often exceed 12 years

## Impact on Dwelling Characteristics

## Census Areas Exposure Measures

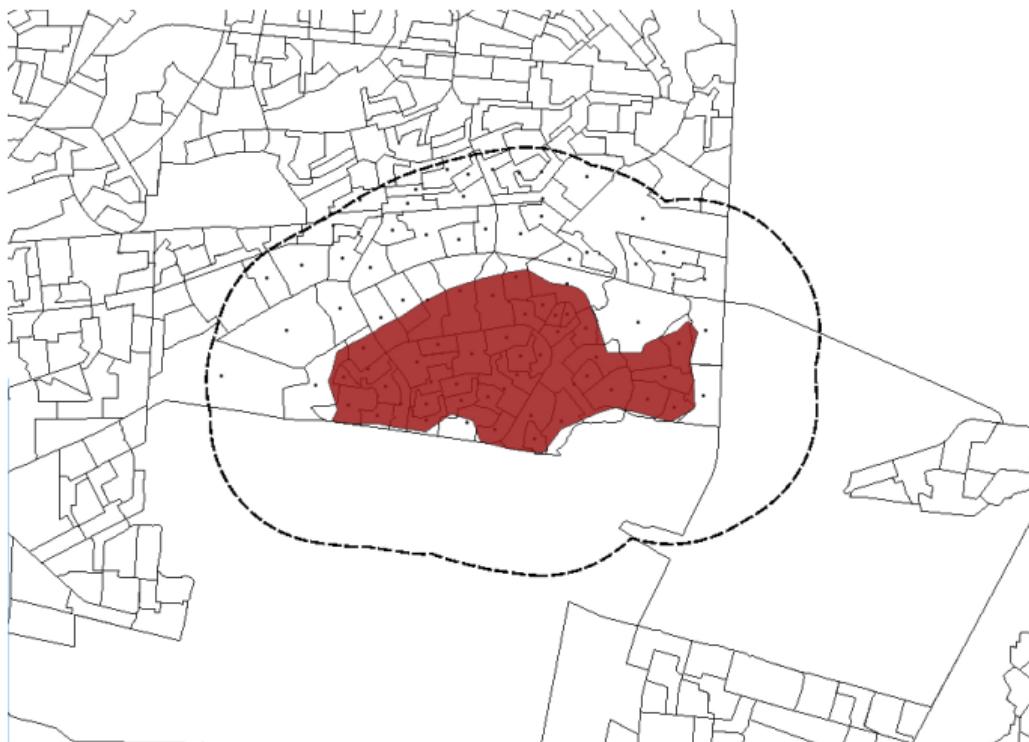


Figure: Census "Small Areas" with centroid within 1500m of project border

## Census Areas Exposure Measures



Figure: "project" exposure: small areas with  $> 30\%$  area overlap

# Census Areas Exposure Measures



Figure: "spillover" exposure: small areas with < 30% area overlap

# Empirical Specification

$$y_{hpt} = \lambda_p + \alpha D_t C_p + \beta D_t + \gamma C_p + \varepsilon_{hpt}$$

with:

- ▶  $y_{hpt}$ : Outcome for household  $h$  living in vicinity of project  $p$ , observed in census year  $t$ .
- ▶  $D_t = 1$  if year  $t$  is census year 2011 (post period).
- ▶  $C_p = 1$  if project  $p$  has been constructed.
- ▶  $\lambda_p$ : project fixed-effect.
- ▶  $\varepsilon_{hpt}$ : error term.

# Empirical Specification

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

with:

- ▶  $y_{hpt}$ : Outcome for household  $h$  living in vicinity of project  $p$ , observed in census year  $t$ .
- ▶  $I_{hpt}^e = 1$  if household  $h$  is in exposure area  $e$  of project  $p$ .
- ▶  $D_t = 1$  if year  $t$  is census year 2011 (post period).
- ▶  $C_p = 1$  if project  $p$  has been constructed.
- ▶  $\lambda_p$ : project fixed-effect.
- ▶  $\varepsilon_{hpt}$ : error term.

# Empirical Specification

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

- ▶  $\alpha^e$  is the DD effect of housing projects on outcome  $y$  at exposure  $e$ , comparing constructed vs. non-constructed projects, before and after construction.
- ▶ Assumes no differential trends between treatment and control areas, absent of project construction.
- ▶ concerns:
  - ▶ correlation between what stopped project.
  - ▶ endogenous census geography.
  - ▶ "dosage".

# Empirical Specification

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

## Outcomes:

- ▶ water and sewers: whether households have a flush toilets, access to an indoors water tap, water from utility company.
- ▶ main energy sources for cooking, heating and lighting.
- ▶ whether household live in a formal house and whether they own it.
- ▶ population and household density.

# Effect On Dwelling Characteristics I. (Nearby Projects)

	flush toilet	water tap	utility water	owner	house
spillover_DD	0.017 (0.037)	-0.011 (0.033)	-0.005 (0.013)	-0.019 (0.029)	-0.012 (0.030)
spillover_post	0.069** (0.027)	0.177*** (0.024)	0.006 (0.010)	-0.030* (0.017)	0.077*** (0.023)
spillover_complete	-0.016 (0.052)	0.004 (0.049)	0.013 (0.017)	0.029 (0.032)	0.067 (0.043)
spillover	-0.023 (0.039)	-0.041 (0.038)	-0.003 (0.014)	-0.008 (0.023)	-0.029 (0.033)
			...		
$\bar{y}$ 2001	0.747	0.345	0.948	0.511	0.550
$\bar{y}$ 2011	0.822	0.518	0.936	0.457	0.623
N	3,214,472	3,214,472	3,214,472	3,112,425	3,067,560
R <sup>2</sup>	0.231	0.147	0.062	0.064	0.110

All specifications include project Fixed Effects. Standard errors clustered at the project level.

# Effect On Dwelling Characteristics I. (Inside Projects)

	flush toilet	water tap	utility water	owner	house
...					
project_DD	0.137** (0.065)	0.143*** (0.052)	0.036 (0.038)	-0.085 (0.073)	0.158*** (0.055)
project_post	0.068 (0.048)	0.087** (0.040)	-0.028 (0.036)	0.036 (0.039)	0.065* (0.036)
project_completed	0.100 (0.120)	0.030 (0.082)	0.036 (0.033)	0.124 (0.081)	0.105 (0.092)
project	-0.343*** (0.089)	-0.251*** (0.059)	-0.038 (0.028)	-0.187*** (0.056)	-0.263*** (0.076)
$\bar{y}$ 2001	0.747	0.345	0.948	0.511	0.550
$\bar{y}$ 2011	0.822	0.518	0.936	0.457	0.623
N	3,214,472	3,214,472	3,214,472	3,112,425	3,067,560
R <sup>2</sup>	0.231	0.147	0.062	0.064	0.110

All specifications include project Fixed Effects. Standard errors clustered at the project level.

## Effect On Dwelling Characteristics II. (Nearby Projects)

	elec cooking	elec heat	elec light	hh density	pop density
spillover_DD	-0.004 (0.039)	-0.065 (0.047)	-0.053 (0.037)	345.6* (205.3)	579.8 (445.8)
spillover_post	0.170*** (0.031)	0.131*** (0.027)	0.118*** (0.033)	270.2** (118.8)	625.1** (278.1)
spillover_complete	-0.017 (0.042)	0.005 (0.044)	0.019 (0.043)	-118.7 (302.5)	-69.4 (753.3)
spillover	-0.044 (0.035)	-0.036 (0.033)	-0.033 (0.036)	284.4 (204.1)	654.5 (505.2)
			...		
$\bar{y}$ 2001	0.644	0.617	0.751	2,167.9	7,192.7
$\bar{y}$ 2011	0.821	0.716	0.842	2,869.2	8,354.4
N	3,214,472	3,214,472	3,214,472	15,454	15,454
R <sup>2</sup>	0.193	0.152	0.168	0.301	0.303

All specifications include project Fixed Effects. Standard errors clustered at the project level.

## Effect On Dwelling Characteristics II. (Inside Projects)

	elec cooking	elec heat	elec light	hh density	pop density
...					
project_DD	0.294*** (0.070)	0.209*** (0.067)	0.116 (0.081)	-127.3 (864.7)	-328.5 (1,631.7)
project_post	0.168*** (0.060)	0.142** (0.056)	0.143** (0.066)	1,265.5* (714.9)	2,635.3* (1,376.0)
project_completed	-0.038 (0.099)	-0.022 (0.088)	0.111 (0.110)	-928.7* (495.1)	-1,384.9 (1,254.1)
project	-0.359*** (0.071)	-0.329*** (0.068)	-0.369*** (0.081)	835.1** (383.7)	1,649.6 (1,005.9)
$\bar{y}$ 2001	0.644	0.617	0.751	2,167.9	7,192.7
$\bar{y}$ 2011	0.821	0.716	0.842	2,869.2	8,354.4
N	3,214,472	3,214,472	3,214,472	15,454	15,454
R <sup>2</sup>	0.193	0.152	0.168	0.301	0.303

All specifications include project Fixed Effects. Standard errors clustered at the project level.

## Impact on Building Density

# Gridded Building Density Measure

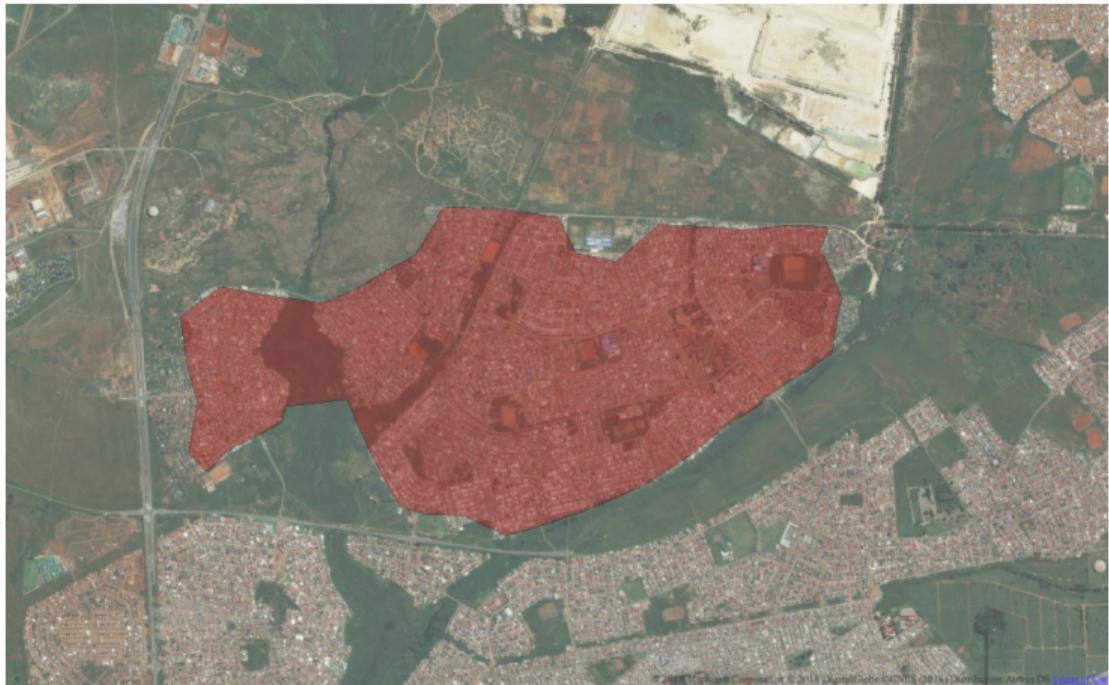


Figure: Empirical design for building density regressions

# Gridded Building Density Measure



Figure: BBLU data showing formal (black) and informal (green) residential structures

# Gridded Building Density Measure



Figure: 50m  $\times$  50m grid overlay.

# Gridded Building Density Measure



Figure: Aggregated data at the cell level.

# Raw Density Means

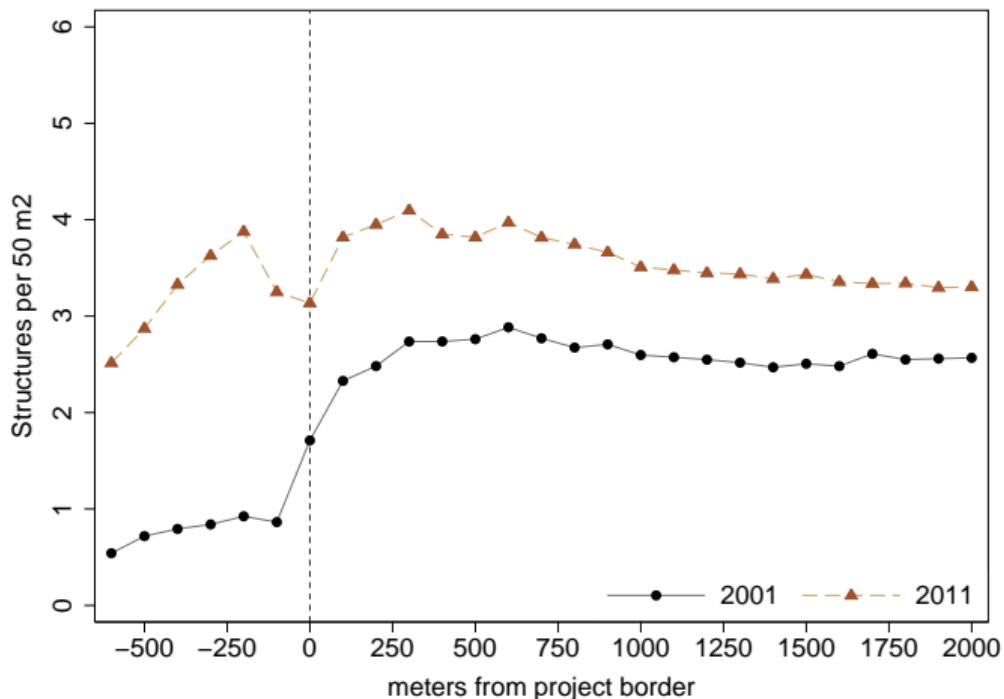


Figure: formal structures, constructed projects.

# Raw Density Means

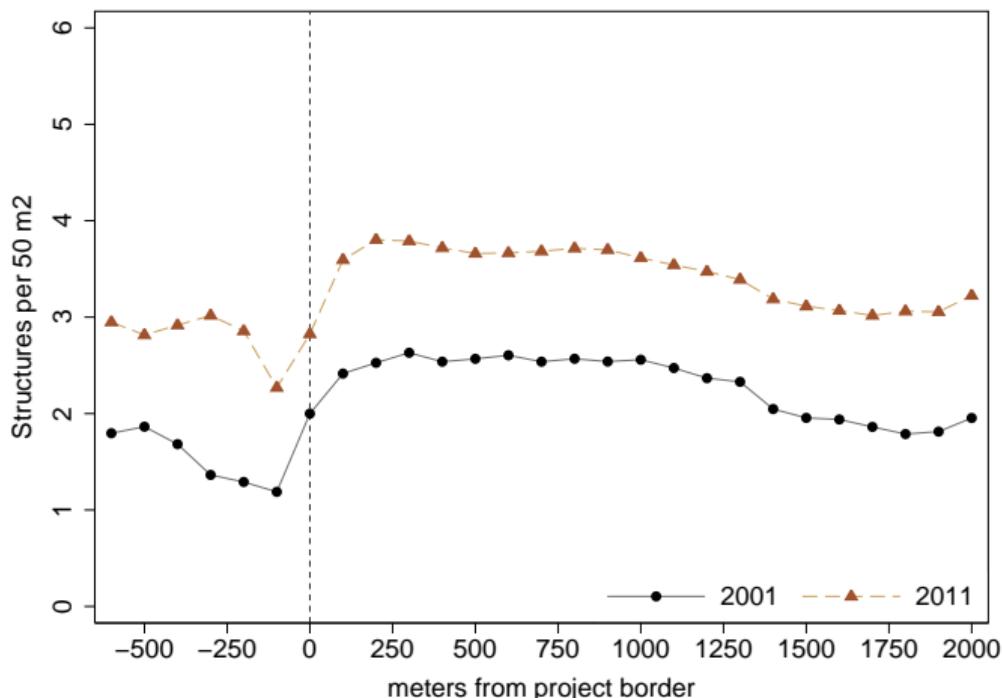


Figure: formal structures, non-constructed projects.

# Raw Density Means

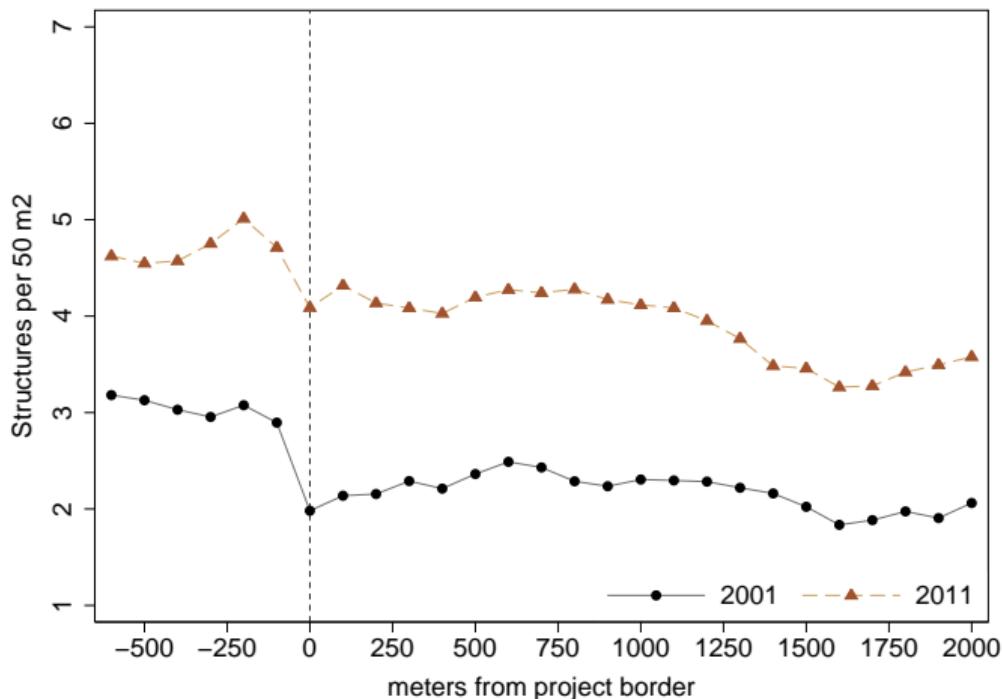


Figure: informal structures, constructed projects.

# Raw Density Means

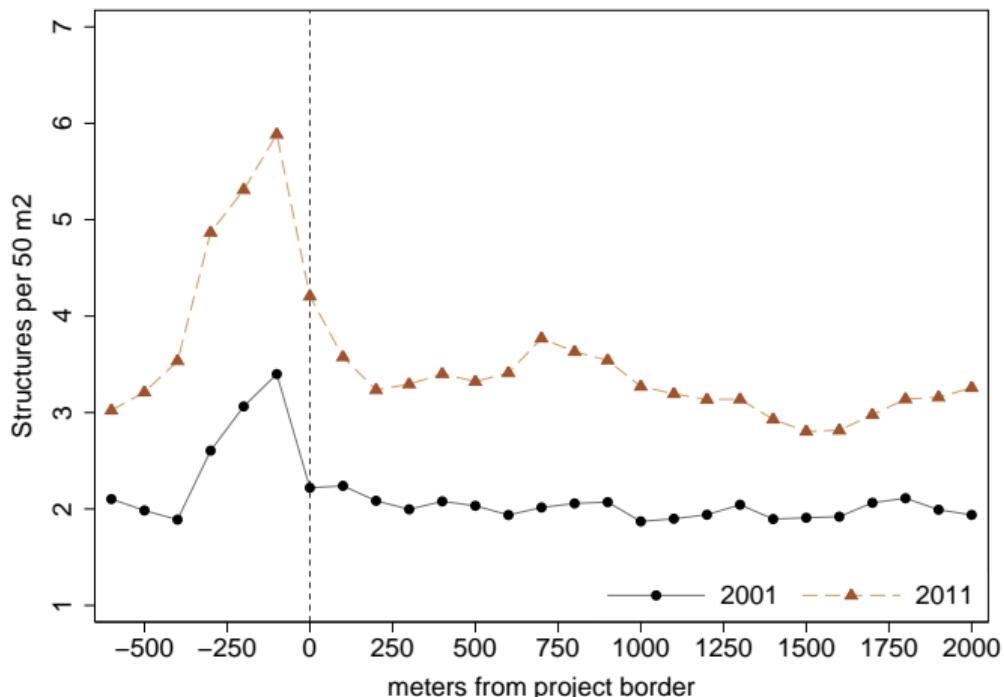


Figure: informal structures, non-constructed projects.

# Empirical Specification

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

with:

- ▶  $y_{itdp}$ : building density for cell  $i$  in vicinity of project  $p$  observed in year  $t$ .
- ▶  $I_{ip}^d = 1$  if cell  $i$  is at distance  $d$  of project  $p$ .
- ▶  $D_t = 1$  if year  $t$  is 2011 (post period).
- ▶  $C_p = 1$  if project  $p$  has been constructed.
- ▶  $\lambda_i$ : cell fixed-effect.
- ▶  $\varepsilon_{ipt}$ : error term

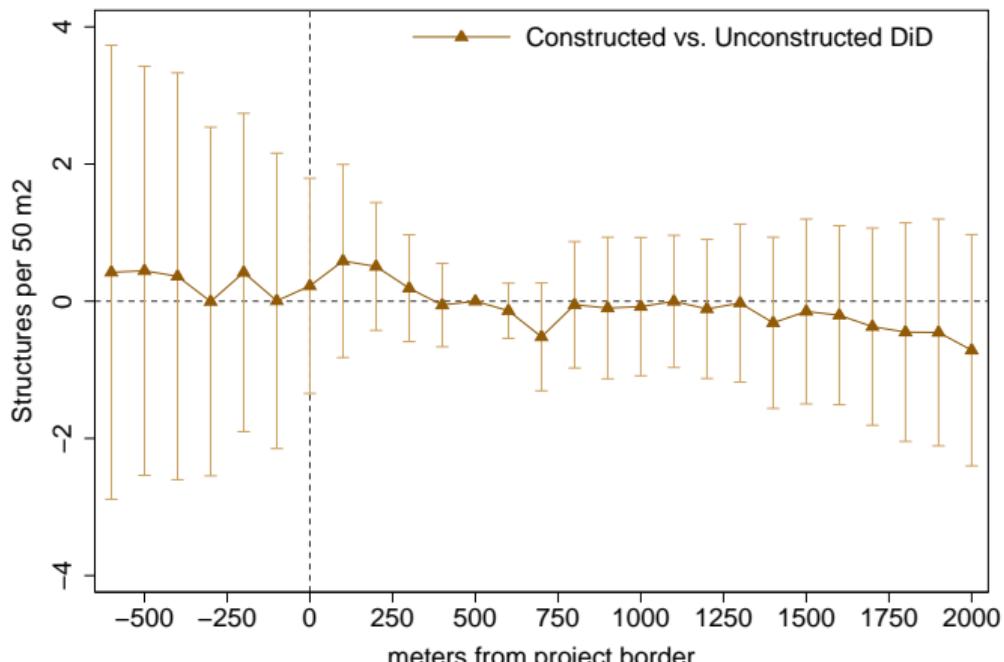
# Empirical Specification

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

## Density Outcomes:

- ▶ total buildings.
- ▶ formal buildings.
- ▶ informal buildings.
- ▶ backyard/non-backyard informal.

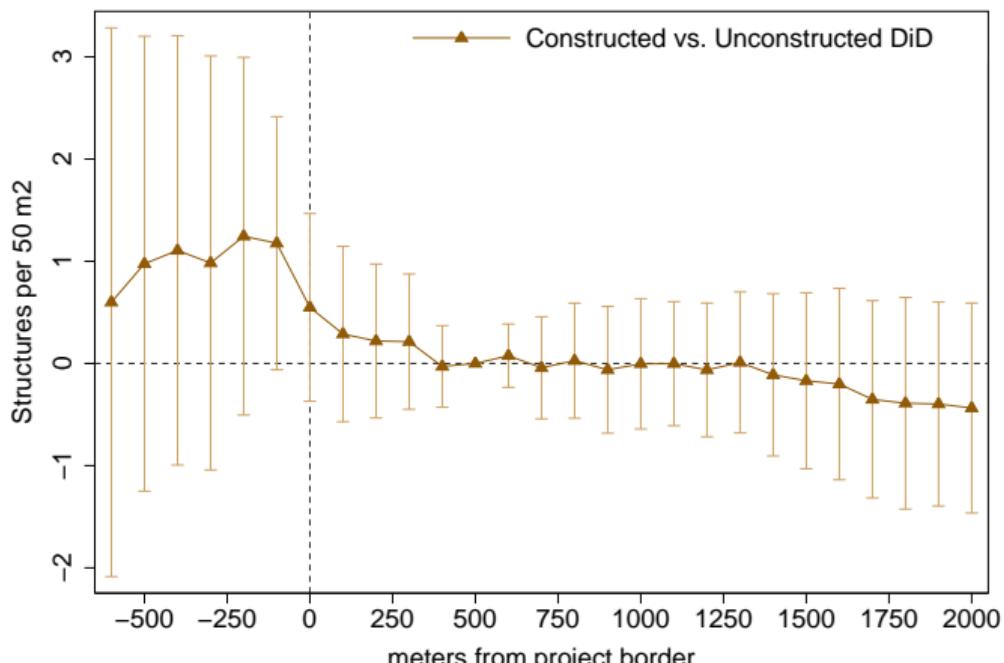
# Building Density: Estimation Results



Mean Structures per 50 m<sup>2</sup>: 5.77

Figure: total structures

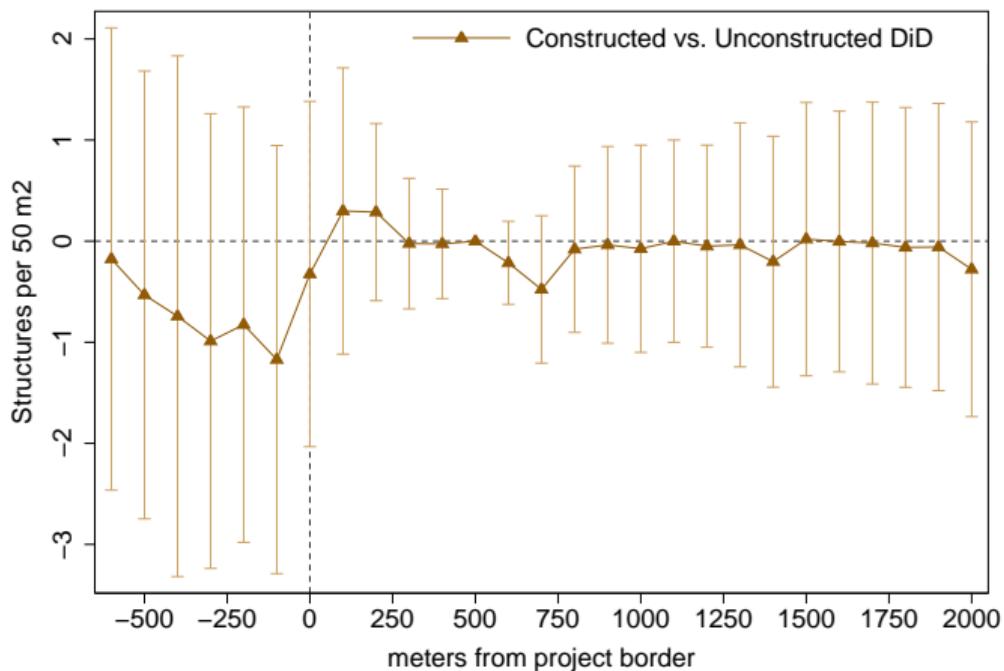
# Building Density: Estimation Results



Mean Structures per 50 m<sup>2</sup>: 2.75

Figure: formal structures

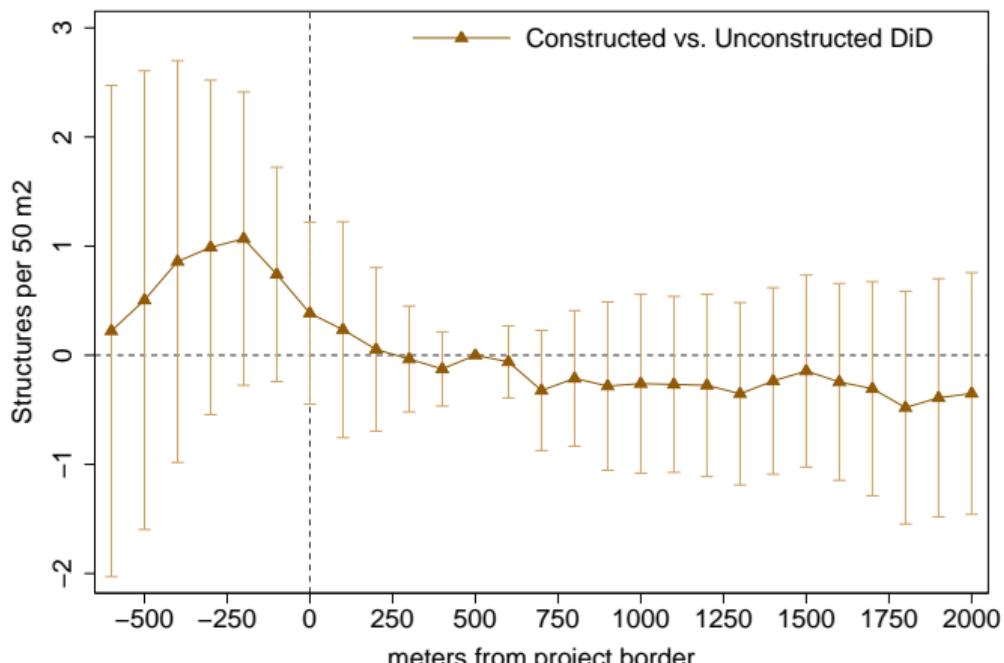
# Building Density: Estimation Results



Mean Structures per 50 m<sup>2</sup>: 3.02

Figure: informal structures

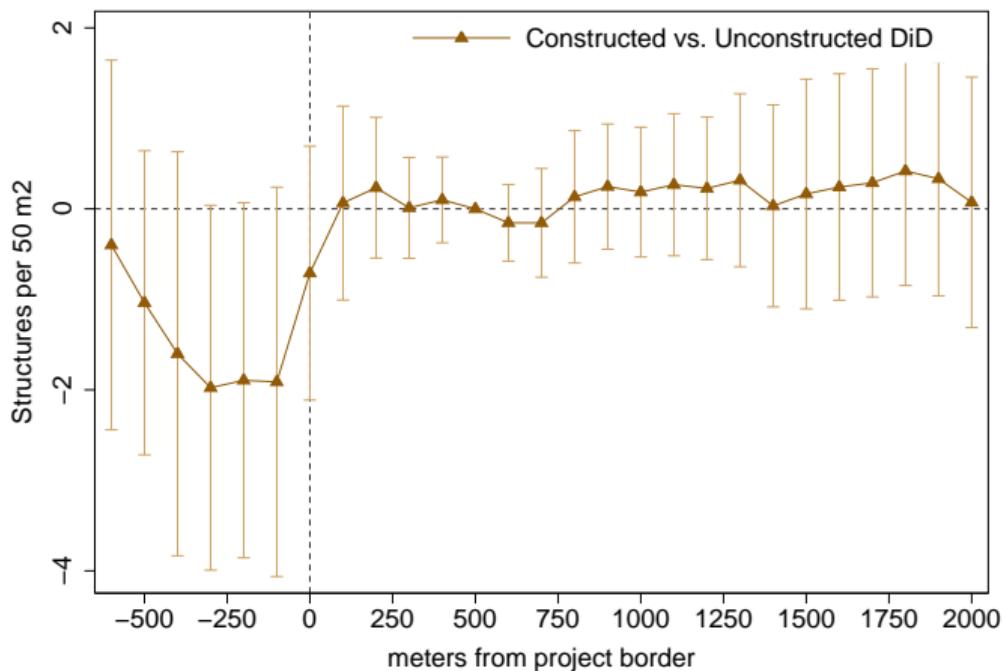
# Building Density: Estimation Results



Mean Structures per 50 m<sup>2</sup>: 1.46

Figure: informal backyard structures

# Building Density: Estimation Results



Mean Structures per 50 m<sup>2</sup>: 1.55

Figure: informal non-backyard structures

## Triple Differences Tests

	total	formal	informal	backyrd	non-bckyrd
inside	0.185 (1.045)	1.125* (0.658)	-0.940 (1.022)	1.076 (0.726)	-2.016** (0.839)
within 500m	0.264 (0.423)	0.195 (0.358)	0.068 (0.475)	0.247 (0.389)	-0.179 (0.412)
<i>N</i>	399,406	399,406	399,406	399,406	399,406
<i>R</i> <sup>2</sup>	0.818	0.828	0.785	0.741	0.774

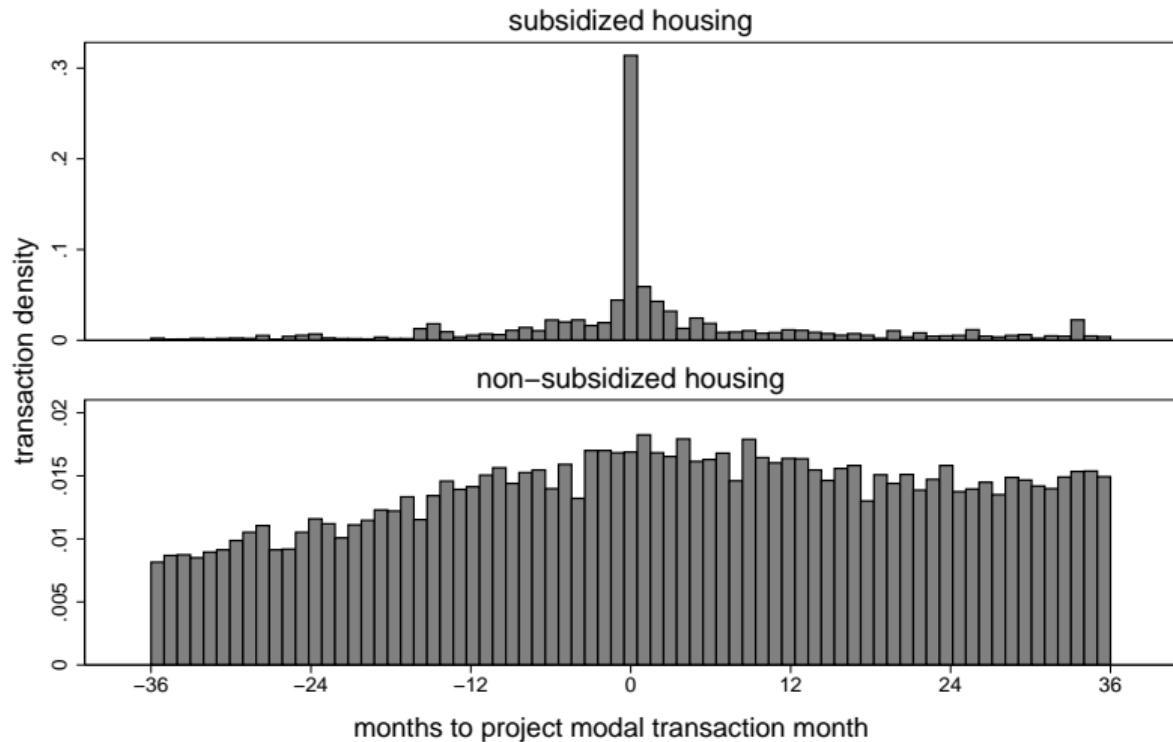
All specifications include project Fixed Effects. Standard errors clustered at the project level.

## Spillovers on Housing Prices

## Empirical Set-up

- ▶ Sample composed of non-subsidized formal housing transactions outside of project borders.
- ▶ Exact transaction date is observed.
- ▶ Assign event date to **constructed** projects.
  - ▶ modal transaction month for subsidized transactions.
- ▶ Assign "event date" to **unconstructed** projects.
  - ▶ apply average delay between project announcement and delivery.

# Transaction Densities



# Empirical Specifications

DID model:

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

with:

- ▶  $P_{itp}$ : log-price of property  $i$  sold at time  $t$ , in vicinity of project  $p$ .
- ▶  $D_{tp} = 1$  if date  $t$  is after modal construction month.
- ▶  $T_{ip} = 1$  if property  $i$  within 700m of project border.
- ▶  $X_i$ : quadratic in lot size of property  $i$ .
- ▶  $\lambda_p$ : project fixed-effect.
- ▶  $\eta_t$ : time (year  $\times$  month) fixed-effect.
- ▶  $\varepsilon_{itp}$ : error term

# Housing Prices: Estimation Results

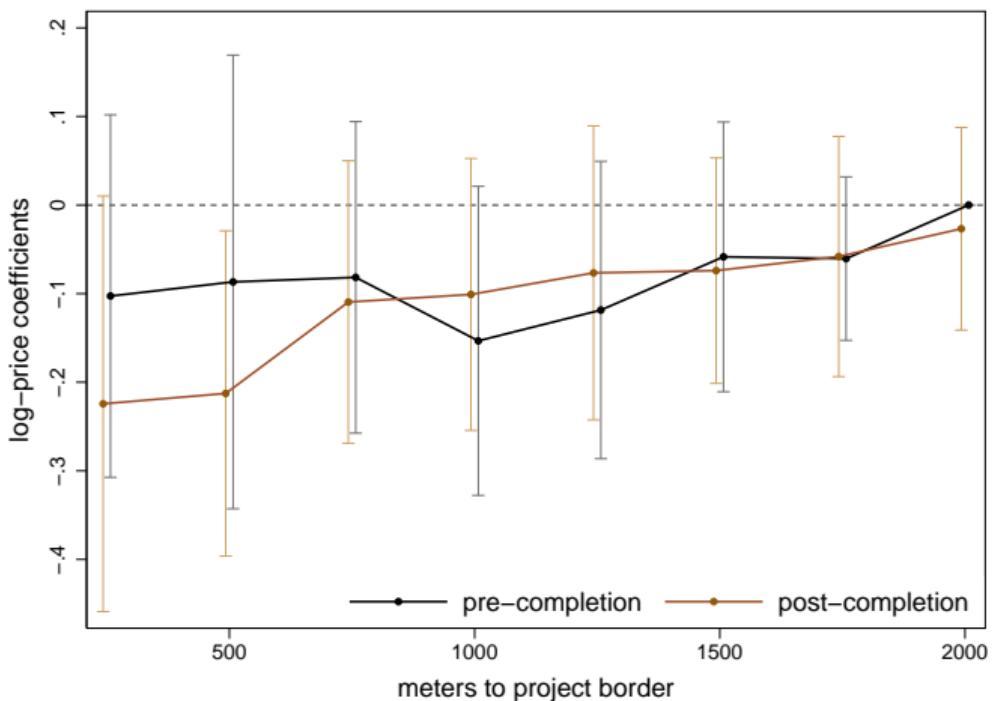


Figure: proximity effects: constructed projects

# Housing Prices: Estimation Results

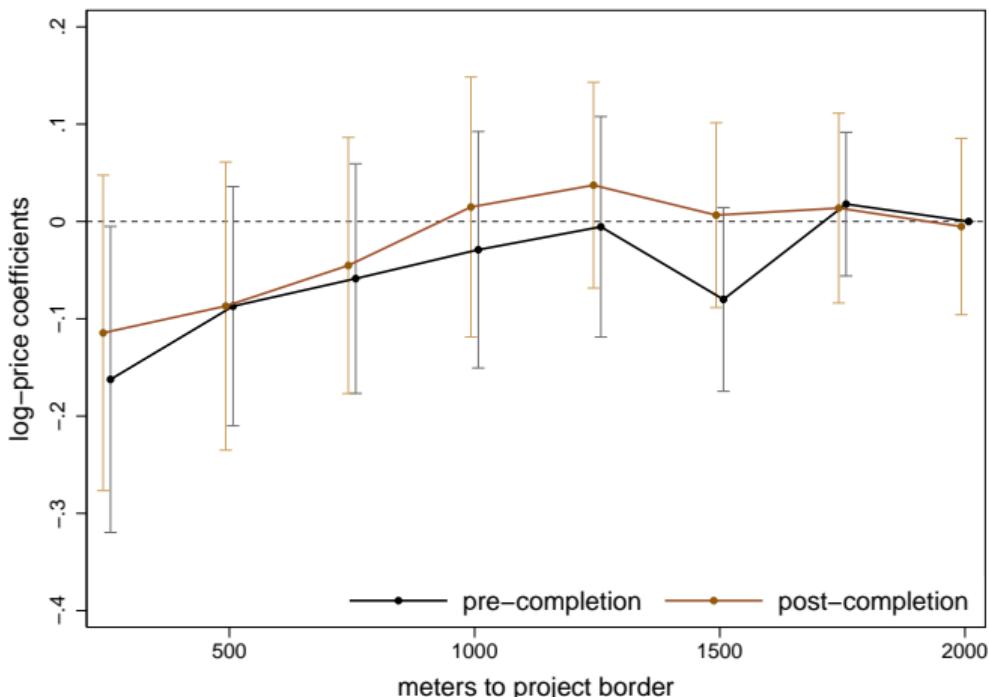


Figure: proximity effects: non-constructed projects

# Housing Prices: Estimation Results

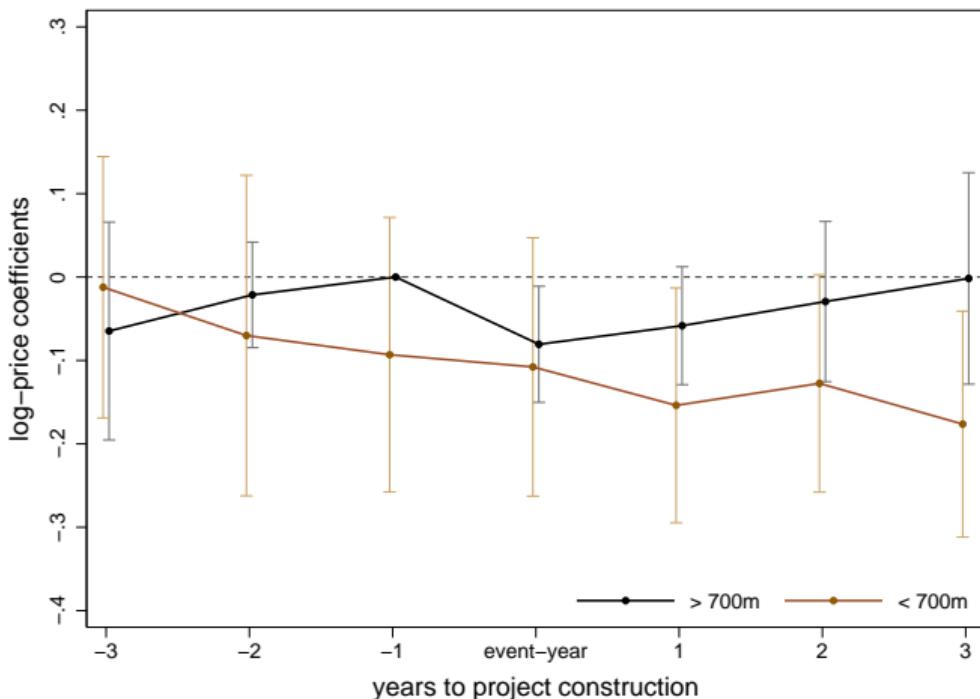


Figure: timing effects: constructed projects

# Housing Prices: Estimation Results

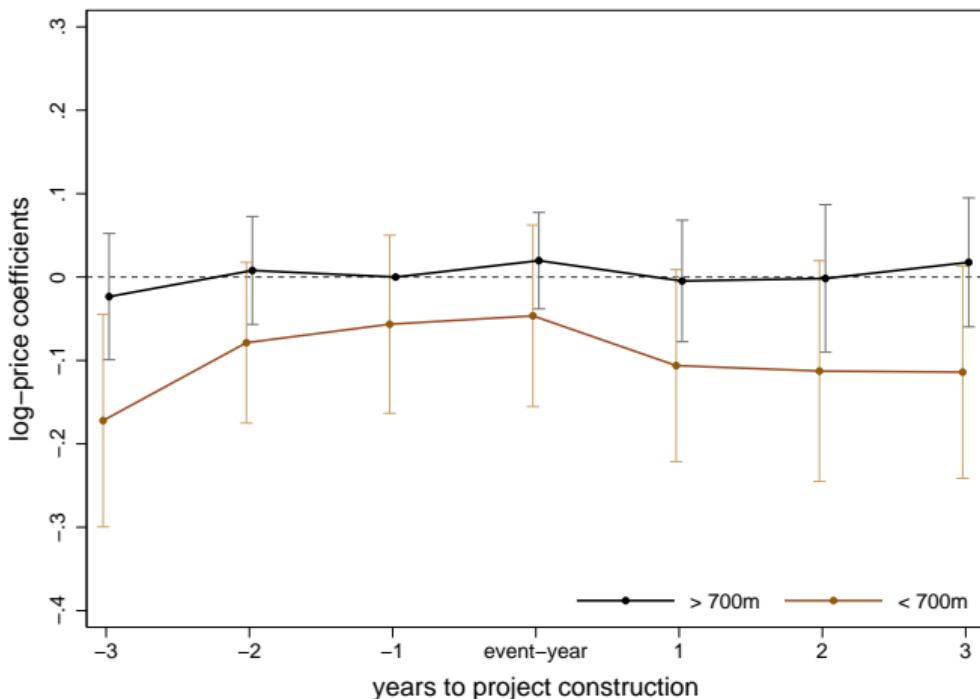


Figure: timing effects: non-constructed projects