Subsidized Housing with Slum Externalities: Evidence from South Africa

Stefano Polloni
joint with Ben Bradlow and Will Violette

February 2017

Public Housing and Development

- Public Housing $\rightarrow$ primary government response to slums
- Positive effects on direct recipients
  (Cateneo et al. [2009], Franklin et al. [2016], Galiani et al. [2017])
- **Question:** What are the spillover effects of public housing in developing countries?
  - *Positive:* Amenity value
  - *Negative:* Crowd-in slums
- **Setting:** 150+ projects in South Africa; GPS price and slum data
- **Findings:** Home prices drop by 16% within 3 yrs and 400m of a project
  - Home quality improves within project footprint but declines nearby (slum crowd-in)

Public Housing in South Africa

- Over 4.3 million houses since 1994 (13% of pop.)
  - 50 to 500 houses per project
  - Fully serviced (roads, water, sanitation, electricity)
  - Greenfield projects on undeveloped land near slums
  - In-Situ upgrading replacing existing slums
- Who gets a house?
  - Official Policy:
    - National/provincial waiting list; no resale within 7 years
    - Must be eligible: Citizens, new homeowners, married or dependents, inc/month $<$R3,500
  - In Practice:
    - Waiting lists/eligibility weakly enforced
    - Only 82% of houses occupied by initial owners within 5 yrs

Measuring Public Housing and Spillovers

- Focus on Gauteng Province (includes Johannesburg and Pretoria)

1. **Property Transactions** 500,000 deeds records (bottom 20% of formal housing market)
   - Buyer/seller name, GPS, price, date from 2002-2011
2. **Building Census:** GPS for over 4 mil. buildings in 2001 and 2011
3. **Population Census:** in 2001 and 2011
4. **Admin. Project Records:** location, dates, costs
   - Includes planned but unconstructed projects

Identifying Housing Projects

Completed projects:

- Use sales from government sellers on previously empty land plots
- Cluster sales into projects based on geographic proximity
- Include projects where over 50% of sales occurred in the same year

Identifying Housing Projects

1. **Seller Identity:** match government names and housing authorities in seller-names from transactions

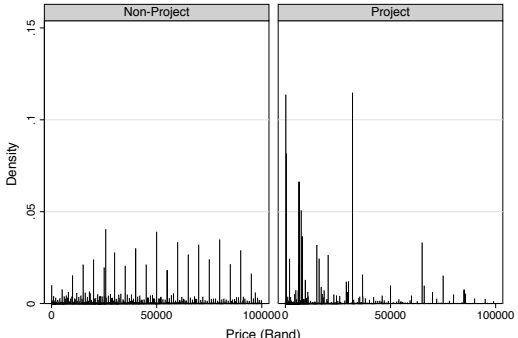
Figure: Top 5 Seller Names

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. **Seller Identity:** match government names and housing authorities in seller-names from transactions
2. **Subsidy Value:** exclude purchase prices R50,000 above subsidy value (<4% of remaining transactions)

Figure: Purchase Price Densities

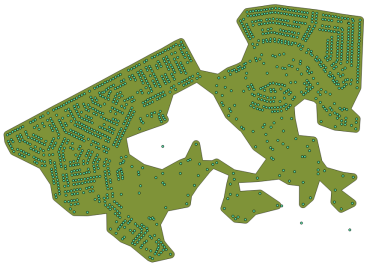


Identifying Housing Projects

1. **Seller Identity:** match government names and housing authorities in seller-names from transactions
2. **Subsidy Value:** exclude purchase prices R50,000 above subsidy value ($<$4% of remaining transactions)
3. **Pre-Existing Formal Dwellings:** exclude land plots with formal structures in 2001 building census (31% of remaining transactions)

Identifying Housing Projects

1. **Seller Identity:** match government names and housing authorities in seller-names from transactions
2. **Subsidy Value:** exclude purchase prices R50,000 above subsidy value ($<$4% of remaining transactions)
3. **Pre-Existing Formal Dwellings:** exclude land plots with formal structures in 2001 building census (31% of remaining transactions)
4. **Spatial Clustering:** collect nearby houses into projects with density-based clustering algorithm



Identifying Housing Projects

1. **Seller Identity:** match government names and housing authorities in seller-names from transactions
2. **Subsidy Value:** exclude purchase prices R50,000 above subsidy value ($<$4% of remaining transactions)
3. **Pre-Existing Formal Dwellings:** exclude land plots with formal structures in 2001 building census (31% of remaining transactions)
4. **Spatial Clustering:** collect nearby houses into projects with density-based clustering algorithm
5. **Temporal Clustering:** include clusters with $>$50% of transactions during modal year (%50 of clusters)

- Overlaps well with completed projects from admin. data

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

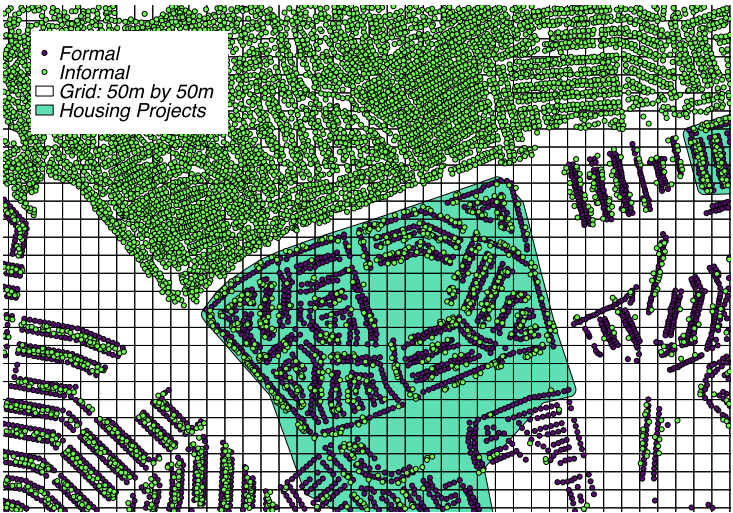
Housing Projects

	Completed	Uncompleted
Formal Homes per km2: 2001	340.6	344.9
Formal Homes per km2: 2011	1,783.1	861.8
Informal Homes per km2: 2001	443.0	1,450.0
Informal Homes per km2: 2011	1,064.6	2,038.4
Median Year (est.)	2005	2006
Distance to CBD (km)	28.9	26.5
Total Projects	56	101

Home counts are from building census.

How do projects affect housing growth?

- Count structures within 50 by 50 meter grids



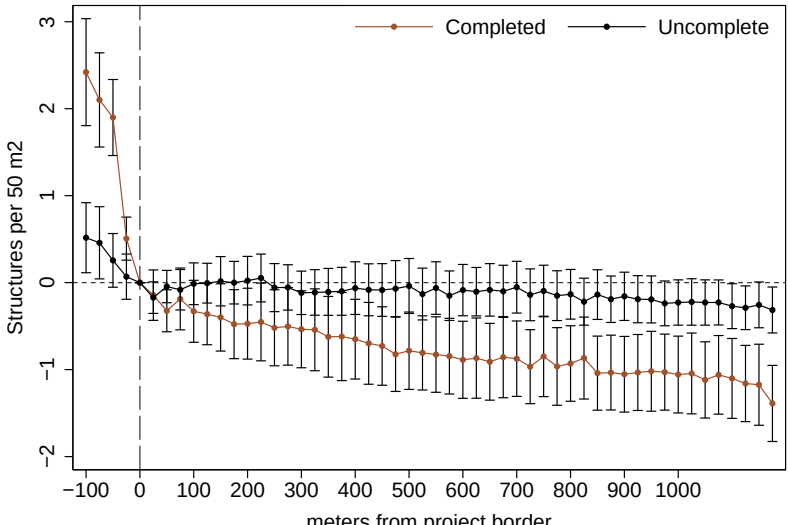
Estimating Differences-in-Differences

$$H_{gtp} = \sum_{d=1}^{D} \alpha_d \mathbb{1}[dist = d] Post_{tp} C_p + \sum_{d=1}^{D} \beta_d \mathbb{1}[dist = d] Post_{tp} U_p + \lambda_g + \varepsilon_{gtp}$$

- g: grid cell, t: year (2001, 2011), p: project
- H_{gtp}: Houses per grid cell
- $dist$: distance from boundary (< 0 within project)
- $Post_{tp}$: After project
- C_{bp}: Completed, U_{bp}: Uncompleted
- λ_g: grid cell fixed effect

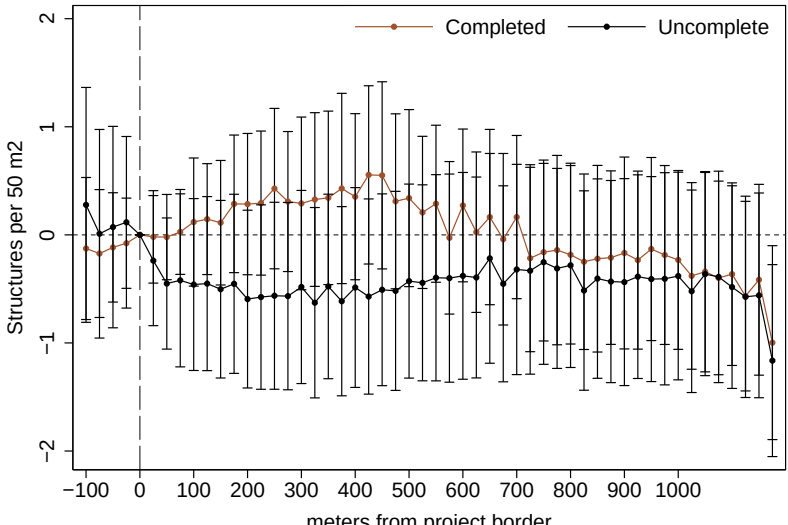
- **Identification**: Counterfactual outcomes for completed projects would have changed in the same way as uncompleted projects.

Formal Houses



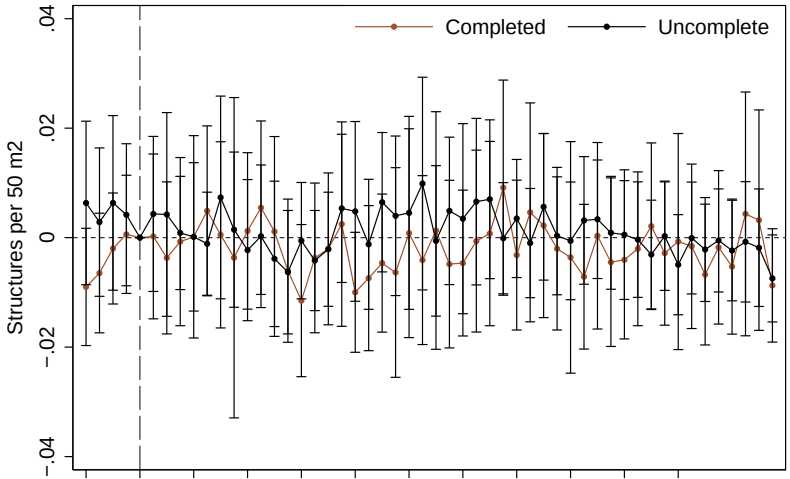
Mean Structures per 50 m2: 2.21

Informal Houses (Slums)



Mean Structures per 50 m2: 2.03

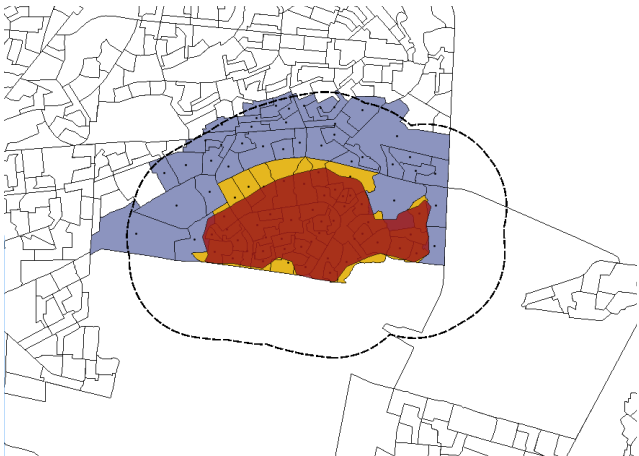
Commercial Buildings



Mean Structures per 50 m2: .05

How do projects affect census demographics?

- Project Blocks: >30% overlap (yellow)
- Spillover Blocks: <30% overlap, centroids within 1.2 km (blue)



Census Descriptives at Baseline (2001)

- Uncompleted project areas have worse outcomes
- Spillover areas are comparable

	Within Project (>30% Overlap)		Spillover (<30% Overlap)	
	Completed	Uncompleted	Completed	Uncompleted
Flush Toilet	0.56	0.26	0.77	0.78
Piped Water	0.21	0.11	0.41	0.37
Owner	0.57	0.43	0.47	0.51
Elec. Cooking	0.58	0.24	0.68	0.63
Elec. Light	0.79	0.36	0.74	0.78
Single House	0.51	0.45	0.52	0.57
Number of Rooms	2.93	3.05	3.11	3.28
Household Size	3.59	3.54	3.27	3.50
Census Blocks	883	967	2,370	2,463
Households	59,460	75,768	213,061	212,005

Census Difference-in-Differences

$$Y_{hbtp} = \alpha_1 \, Post_{tp} \, C_{bp} \, Project_{bp} + \alpha_2 \, Post_{tp} \, C_{bp} \, Spillover_{bp}$$

$$+ \theta_1 \, Post_{tp} \, Project_{bp} + \theta_2 \, Post_{tp} \, Spillover_{bp}$$

$$+ \theta_3 \, C_{bp} \, Spillover_{bp} + \theta_4 \, Spillover_{bp} + \lambda_p + \varepsilon_{hbtp}$$

- h: household, b: census block, t: year (2001, 2011), p: project
- $Post_{tp}$: After project
- C_{bp}: Completed
- $Project_{bp}$: $>$%30 overlap
- $Spillover_{bp}$: $\leq$%30 overlap
- λ_p: Project fixed effect

- **Identification**: Counterfactual outcomes for completed projects would have changed in the same way as uncompleted projects.

Census Differences-in-Differences Estimates

VARIABLES	(1) Flush Toilet	(2) Piped Water Inside	(3) Electric Cooking	(4) Electric Lighting
Project X Post X Complete	0.210**	0.202***	0.0679	-0.0482
	(0.0824)	(0.0540)	(0.0849)	(0.0998)
Spillover X Post X Complete	-0.0723*	-0.0464	-0.130***	-0.0461
	(0.0411)	(0.0302)	(0.0449)	(0.0385)
Project X Post	0.115**	0.148***	0.353***	0.270***
	(0.0448)	(0.0298)	(0.0760)	(0.0822)
Spillover X Post	0.136***	0.182***	0.268***	0.138***
	(0.0313)	(0.0205)	(0.0347)	(0.0285)
Spillover X Complete	-0.205*	-0.0732	-0.262***	
	(0.122)	(0.106)	(0.0885)	(0.0935)
Spillover	0.328***	0.185**	0.373***	
	(0.0873)	(0.0719)	(0.0788)	(0.0831)
Constant	0.498***	0.217***	0.389***	0.537***
	(0.0514)	(0.0425)	(0.0379)	(0.0415)
Observations	1,544,285	1,544,285	1,544,285	1,544,285
R-squared	0.360	0.243	0.301	0.306
Project FE	YES	YES	YES	YES

Robust standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Standard errors are clustered at the project level.

Census Differences-in-Differences Estimates

VARIABLES	(5) Single House	(6) Owns House	(7) No. Rooms	(8) Household Size
Project X Post X Complete	0.183***	-0.0523	0.286*	0.0992
	(0.0451)	(0.0645)	(0.158)	(0.0915)
Spillover X Post X Complete	-0.0957**	0.00820	-0.102	-0.00151
	(0.0374)	(0.0501)	(0.0923)	(0.0462)
Project X Post	0.0535**	0.223***	0.353***	-0.242***
	(0.0238)	(0.0498)	(0.106)	(0.0771)
Spillover X Post	0.129***	0.289***	0.390***	-0.236***
	(0.0312)	(0.0245)	(0.0670)	(0.0351)
Spillover X Complete	-0.0540	-0.217***	-0.148	-0.202*
	(0.0731)	(0.0724)	(0.222)	(0.122)
Spillover	0.147***	0.128**	0.474***	0.150
	(0.0560)	(0.0624)	(0.180)	(0.103)
Constant	0.439***	0.468***	2.746***	3.334***
	(0.0301)	(0.0292)	(0.0927)	(0.0515)
Observations	1,479,342	1,496,636	1,459,677	1,532,866
R-squared	0.195	0.147	0.174	0.057
Project FE	YES	YES	YES	YES

How do projects affect local housing prices?

- Focus on 1.2 km buffers around housing projects

Housing Price Descriptives

	In 1.2 km buffer		
	Completed	Uncompleted	Other
Purchase Price (Rand)	248,181.0	230,410.1	243,484.9
	[1440258.5]	[754,489.7]	[304,045.0]
Plot Size (m3)	819.2	865.2	1,888.5
	[34,138.9]	[4,243.3]	[55,518.5]
Sold At Least Once	0.326	0.350	0.331
Median Purchase Year	2006	2006	2006
Observations	28,943	20,700	167,578

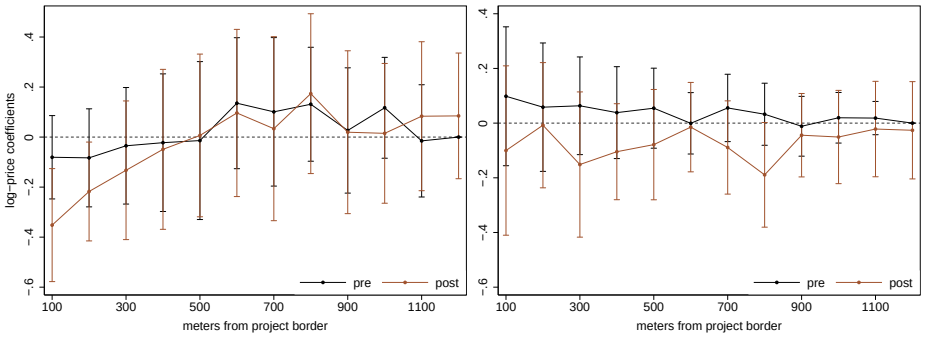
Estimating Differences-in-Differences

$$P_{itp} = \sum_{d=1}^{D} \alpha_d \mathbb{1}[dist = d] Post_{tp} + \sum_{d=1}^{D} \alpha_d \mathbb{1}[dist = d] Pre_{tp} + \gamma_t + \lambda_p + \theta X_i + \varepsilon_{itp}$$

$$P_{itp} = \sum_{e=1}^{E} \alpha_j \mathbb{1}[time = e] Near_{tp} + \sum_{e=1}^{E} \alpha_j \mathbb{1}[time = e] Far_{tp} + \gamma_t + \lambda_p + \theta X_i + \varepsilon_{btp}$$

- i: transaction, t: year-month, p: project
- P_{gtp}: Purchase price (formal houses)
- $dist$: dist to project, $Near_{tp}$: <400m, Far_{tp}: ≥400m & <1200
- $time$: months to project, $Post_{tp}$: 36 months after, Pre_{tp} 36 before
- λ_p: project fixed effect, γ_t: calendar month fixed effect

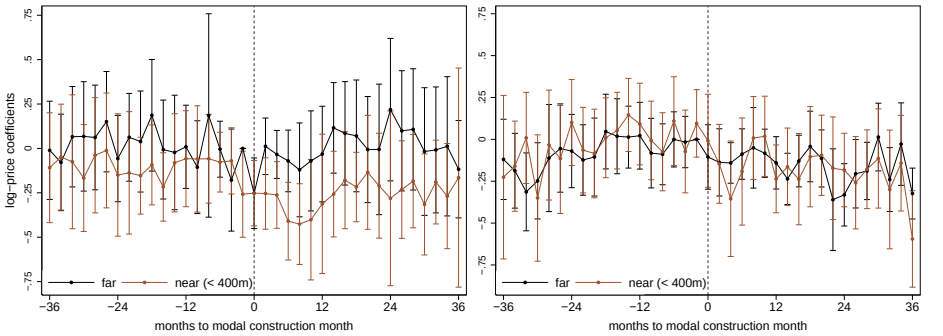
Figure: Price Estimates over Distance



Completed Projects

Uncompleted Projects

Figure: Price Estimates over Time



Completed Projects

Uncompleted Projects