- raw pre-period building density means show (1) differences between project areas and spillover areas and (2) differences between constructed and unconstructed projects
- we are going to use two strategies:
    1. our high-res spatial data provides a unique opportunity to control for spatially precise fixed effects (2001 census subplaces) which erase mean differences between constructed and unconstructed projects in the pre-period and;
    2. after controlling for 2001 census subplace fixed effects, constructed and unconstructed projects have the same pattern in the preperiod: both have higher informal building density in project areas, which means it is important to include unconstructed projects as counterfactuals in a DDD setting

    - by project type, greenfield and in-situ have perfect pre-trends, while unlabeled projects have a few differences
- raw building changes over time suggest some crowd-in close to project boundaries, especially for in-situ upgrading projects[1]
- empirical strategy is a DDD with each observation linked to its closest project and controlling for 2001 census subplace-by-year fixed effects
    - there are around 1000 sub-places in 2001
    - using 2001 subplaces is a nice way to tie our hands on geography, and subplaces are likely to be drawn in more thoughtful ways than simply imposing arbitrary 3km grid cells
    - the fixed effects results have some differences with the results without fixed effects (in none_250.pdf), which are consistent with where the identification is coming from (ie. subplaces on the edge of projects so that they are exposed to two treatment conditions, which is probably important to discuss)
    - specifications define spillover areas as 0-250m distance away from projects
        * why 250m? because the average area of the 250m buffers is just about equal to the average area of the projects themselves (which is 1.2 km2), which is consistent with some rough assumption that spatial effects might be proportional to size of footprint; also, the effects mostly dissipate at further distances
        * one concern is that there is error in drawing the project boundaries so what we are calling spillover may actually just be mislabeled project areas: three responses

[1] I don't want to show "demeaned" changes over time because that is not exactly what the regression is picking up, in fact the regression results are more consistent with the raw changes than the demeaned changes

- the raw data show seriously sharp breaks in densities at boundaries
- we find opposite informal density substitution patterns: we find substitution toward bkyd shacks inside projects, but away from bkyd shacks just outside projects
- casual observation suggests that boundaries are often drawn generously (ie. in many cases, houses do not fill the total shapes)

- Standard errors are clustered at the project/treatment level (ie. about 300 projects by inside, spillover, and control areas for about 900 groups total). This approach follows from recommendations from Imbens and Athey's recent paper. I also tried spatially auto-correlated Conley errors, which are actually a lot smaller than the clustered errors (I don't think you can do these errors on top of clustering), but might be worth reporting also

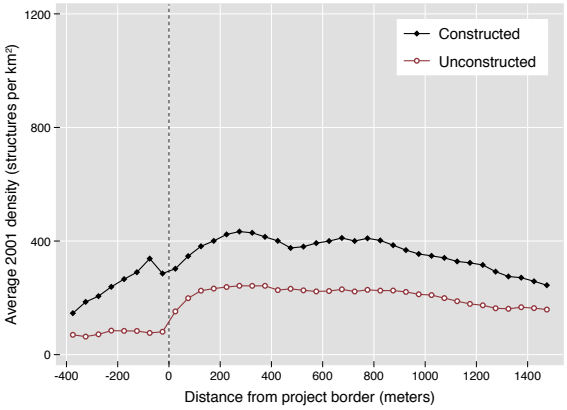
- **My headline takeaways:**
  - Greenfield projects are characterized by increases in formal housing (with improved services) and backyard informal housing, and small decreases in informal non-bkyd housing. These new houses slightly crowd-out new development around their footprints. New and improved greenfield houses act as a positive amenity, noisily increasing house prices
  - In-situ projects knock down a lot of informal non-bkyd housing (which is only partially offset by new bkyd housing) and replace this informal housing with new formal housing (with improved services). This big shift in housing quality crowds-in new development of both formal and informal housing (with no strong net effect on services) nearby. This increase in local housing supply offsets the gain to local amenities and drives down local formal housing prices

Table 1. Housing Project Areas Description

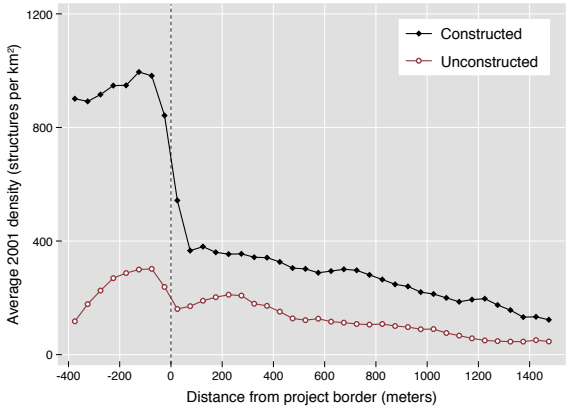
	All		**Greenfield**		**In-Situ**	
	Const.	Unconst.	Const.	Unconst.	Const.	Unconst.
Number of Projects	164	137	41	19	25	29
Area (km2)	1.21	1.19	1.75	2.53	1.60	0.88
Median Construction Yr.	2006	2006	2006	2005	2004	2006
Delivered Houses	302	0	420	0	557	0
House Price in 1 km (R[†])	189,304	218,635	194,214	186,841	179,596	208,571
Distance to CBD[‡] (km)	32.4	28.0	40.6	40.5	32.3	30.6

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.

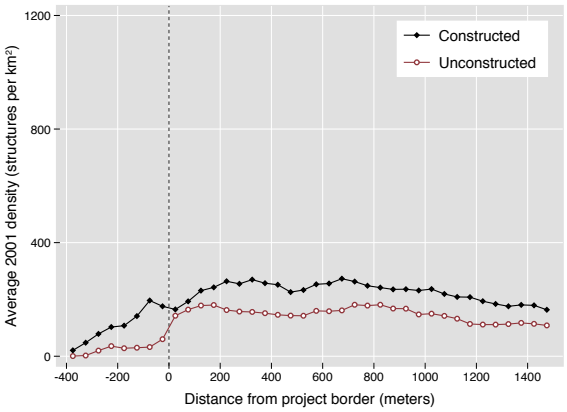
(a) All Projects pre-period formal raw data



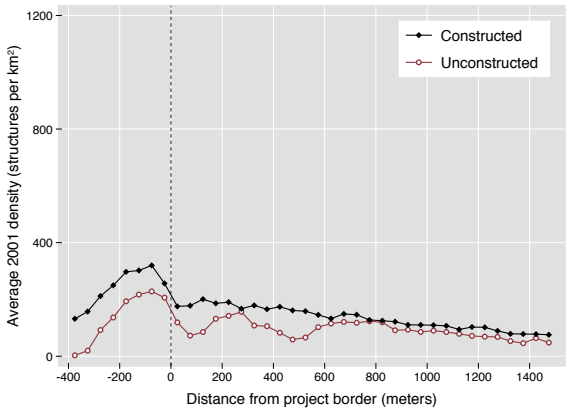
(b) All Projects pre-period informal raw data



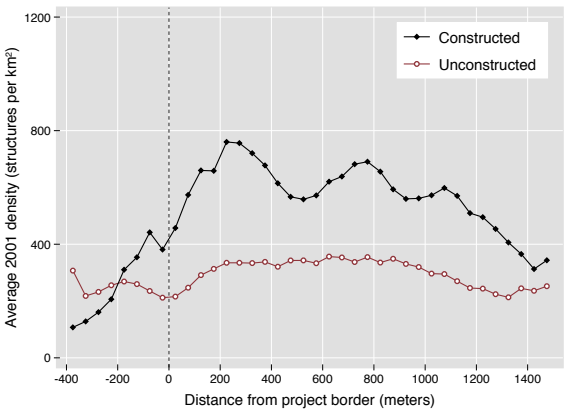
(c) Greenfield pre-period formal raw data



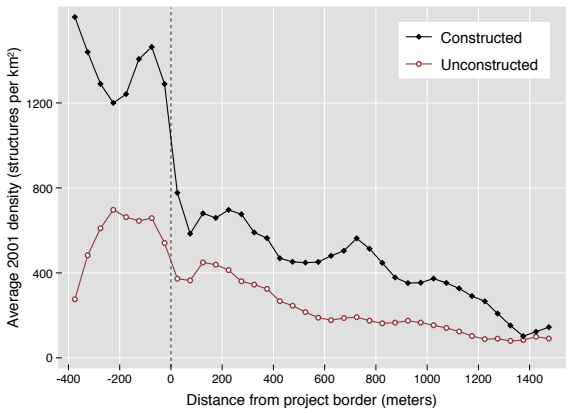
(d) Greenfield pre-period informal raw data



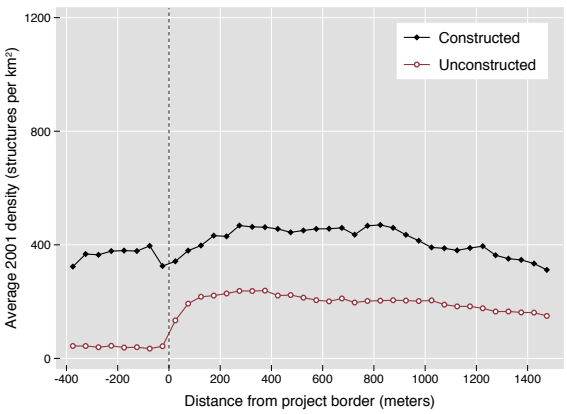
(e) In-Situ pre-period formal raw data



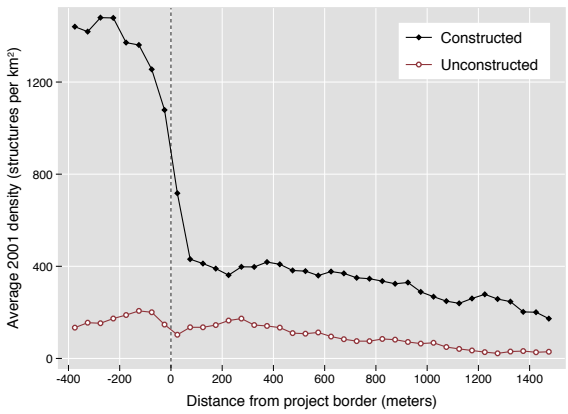
(f) In-Situ pre-period informal raw data



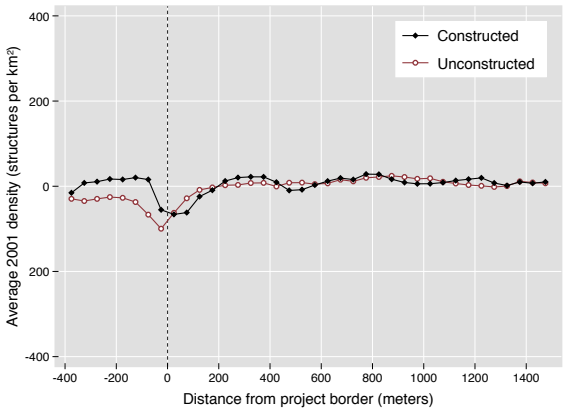
(g) Other pre-period formal raw data



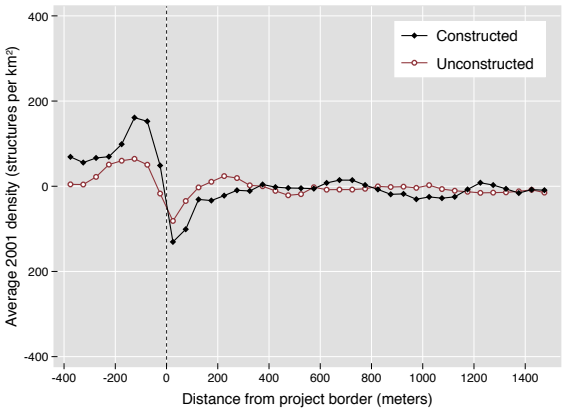
(h) Other pre-period informal raw data



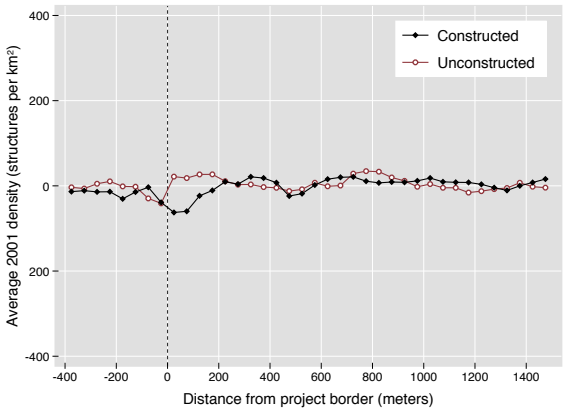
(a) All Projects pre-period formal fe



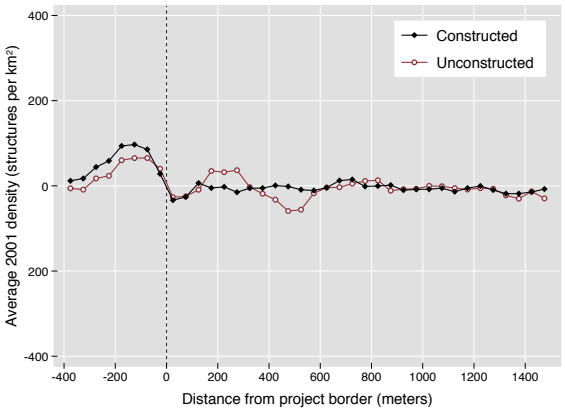
(b) All Projects pre-period informal fe



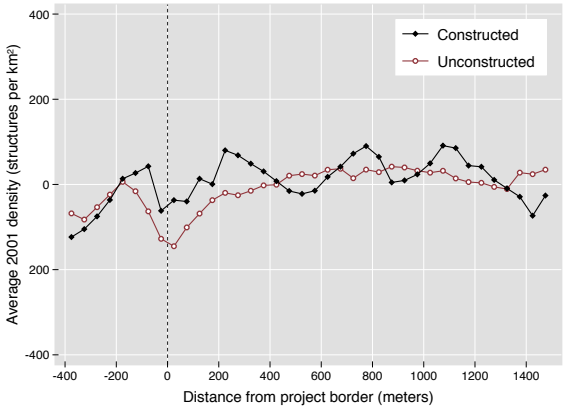
(c) Greenfield pre-period formal fe



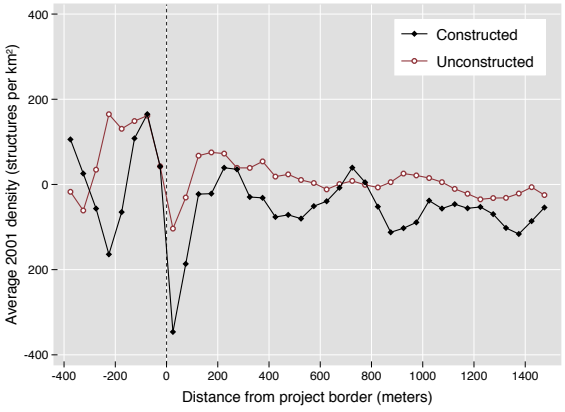
(d) Greenfield pre-period informal fe



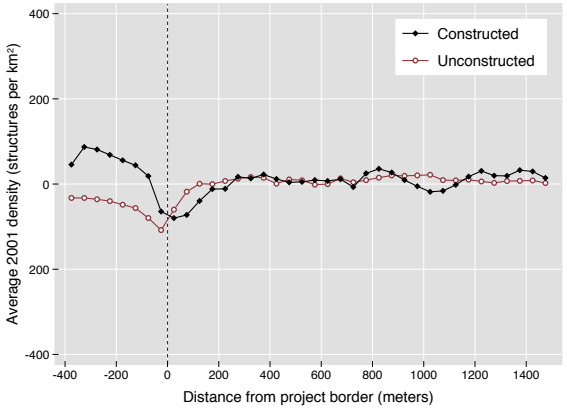
(e) In-Situ pre-period formal fe



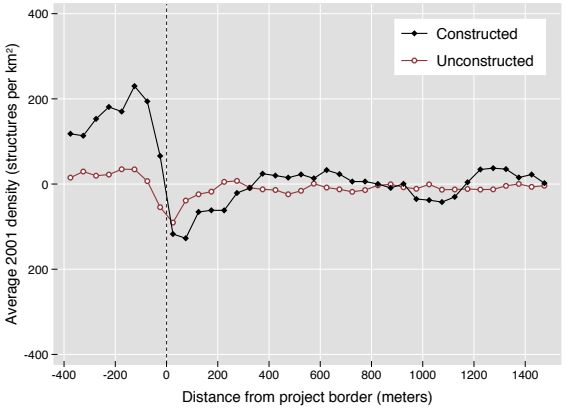
(f) In-Situ pre-period informal fe



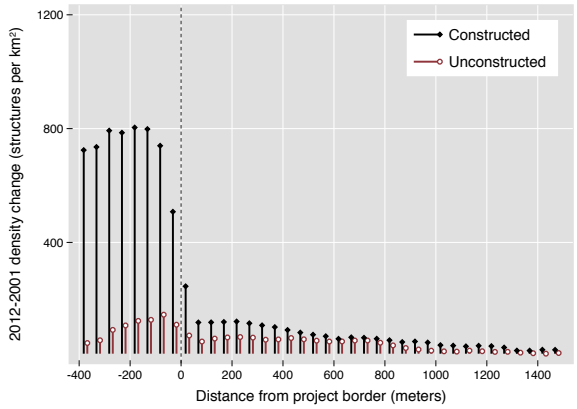
(g) Other pre-period formal fe



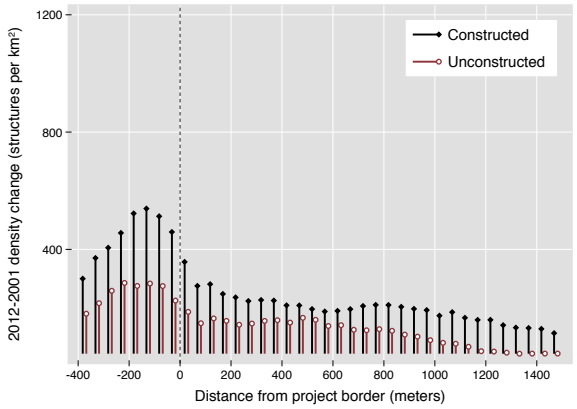
(h) Other pre-period informal fe



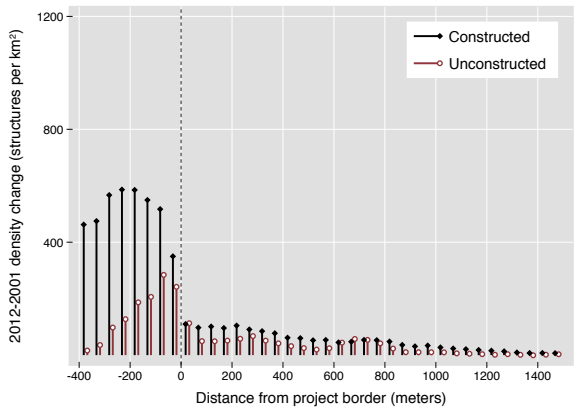
(a) All Projects changes formal raw data



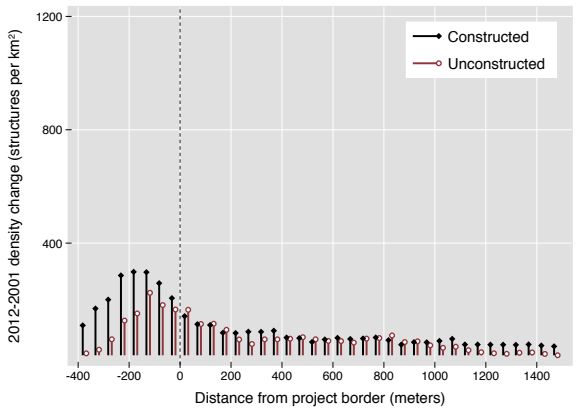
(b) All Projects changes informal raw data



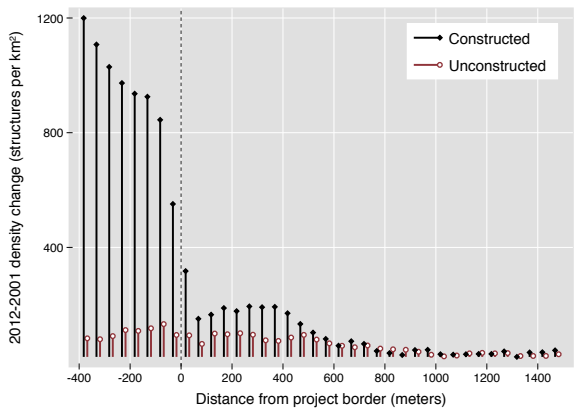
(c) Greenfield changes formal raw data



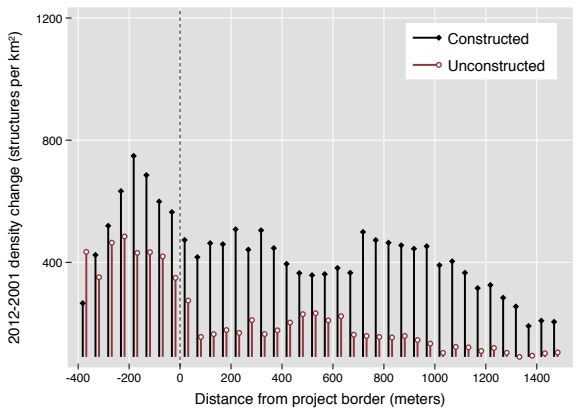
(d) Greenfield changes informal raw data



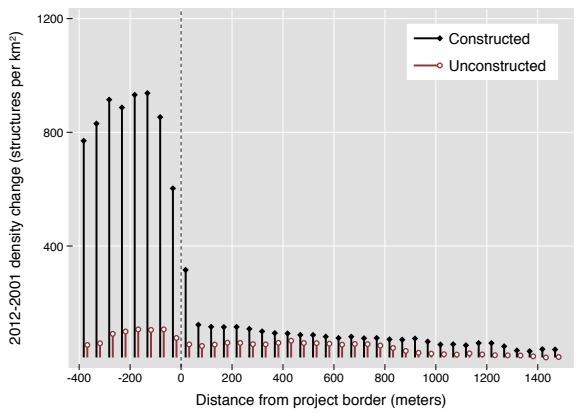
(e) In-Situ changes formal raw data



(f) In-Situ changes informal raw data



(g) Other changes formal raw data



(h) Other changes informal raw data

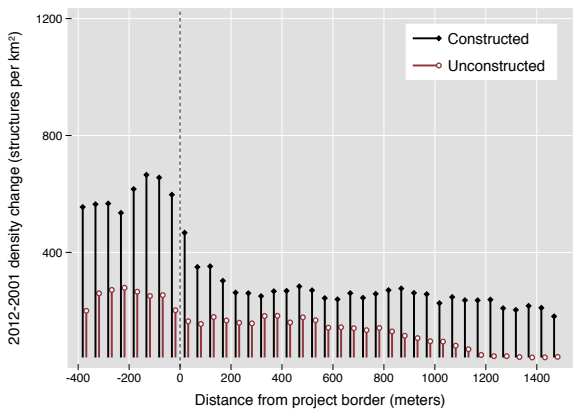


Table 2. Building Density

	(1) Total	(2) Formal	(3) Informal	(4) Informal Bkyd.	(5) Informal Non-Bkyd.
All Projects					
inside project	458.657[a]	437.088[a]	21.569	257.528[a]	-235.958[a]
	(119.427)	(64.029)	(84.154)	(71.621)	(65.963)
0-250m outside project	30.928	4.747	26.181	-12.779	38.960
	(41.744)	(20.850)	(32.003)	(26.733)	(25.231)
R^2	0.442	0.409	0.405	0.387	0.346
Greenfield					
inside project	290.271[c]	227.318[b]	62.953	100.890	-37.937
	(155.997)	(97.511)	(75.374)	(71.544)	(37.696)
0-250m outside project	-22.059	-13.467	-8.592	-34.453	25.861
	(54.721)	(29.279)	(36.258)	(25.044)	(28.150)
In-Situ Upgrading					
inside project	253.124	573.329[a]	-320.205	86.406	-406.611
	(468.589)	(128.696)	(400.741)	(308.662)	(294.362)
0-250m outside project	107.252	53.251	54.001	-34.073	88.074
	(166.688)	(75.018)	(129.139)	(116.068)	(96.378)
Other					
inside project	733.024[a]	604.727[a]	128.297	476.514[a]	-348.217[a]
	(126.059)	(65.941)	(93.433)	(91.385)	(80.071)
0-250m outside project	34.494	18.061	16.432	-9.068	25.501
	(56.279)	(26.526)	(44.439)	(36.456)	(34.928)
Mean Outcome 2001	526.22	261.56	264.66	96.43	168.23
Mean Outcome 2011	838.62	385.14	453.48	286.79	166.69
R^2	0.443	0.411	0.406	0.388	0.347
N	1,705,534	1,705,534	1,705,534	1,705,534	1,705,534

Table 3. Effect of Housing Projects on Socio-demographics

	(1) Age	(2) P.O.B. not Gauteng	(3) Unemployed	(4) Years of Education	(5) Monthly Income
All Projects					
inside project	-0.135	0.002	-0.041[b]	0.199	755.808[c]
	(0.316)	(0.021)	(0.019)	(0.150)	(457.417)
0-250m outside project	0.137	-0.010	-0.022	0.162	200.144
	(0.230)	(0.016)	(0.016)	(0.103)	(316.010)
R^2	0.702	0.771	0.529	0.692	0.666
Greenfield					
inside project	-0.193	-0.023	0.012	0.516[c]	1203.244
	(0.589)	(0.056)	(0.046)	(0.285)	(1088.395)
0-250m outside project	0.147	0.021	0.014	0.005	532.664
	(0.514)	(0.026)	(0.043)	(0.225)	(657.908)
In-Situ Upgrading					
inside project	0.744	0.066[c]	-0.050[b]	-0.027	142.708
	(0.536)	(0.036)	(0.025)	(0.326)	(675.644)
0-250m outside project	0.184	-0.003	-0.016	0.131	145.960
	(0.314)	(0.029)	(0.031)	(0.236)	(640.210)
Other					
inside project	-0.333	-0.020	-0.047[c]	0.232	896.161
	(0.420)	(0.031)	(0.026)	(0.197)	(647.119)
0-250m outside project	0.148	-0.013	-0.029	0.171	-28.946
	(0.330)	(0.022)	(0.022)	(0.130)	(421.127)
Mean Outcome 2001	27.30	0.37	0.47	8.27	2,477.01
Mean Outcome 2011	28.30	0.43	0.33	9.68	4,486.48
R^2	0.704	0.774	0.529	0.694	0.669
N	12,734	12,727	12,724	12,728	12,724

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

Table 4. Census Household-level Estimates

	(1) Flush Toilet	(2) Water Indoors	(3) Electricity Cooking	(4) Electricity Heating	(5) Electricity Lighting	(6) Number of Rooms	(7) Household Size	(8) Population Density
All Projects								
inside project	0.084[c]	0.156[a]	0.124[b]	0.099[c]	0.083	0.215	0.075	-567.957
	(0.050)	(0.044)	(0.057)	(0.058)	(0.063)	(0.163)	(0.074)	(686.903)
0-250m outside project	0.011	0.081[b]	-0.007	-0.012	-0.011	0.074	-0.032	-391.311
	(0.032)	(0.039)	(0.035)	(0.034)	(0.036)	(0.108)	(0.054)	(746.383)
R^2	0.615	0.593	0.668	0.633	0.638	0.676	0.691	0.606
Greenfield								
inside project	0.153	0.145	0.203[c]	0.074	0.170	0.710[c]	0.027	1600.773
	(0.131)	(0.124)	(0.112)	(0.135)	(0.119)	(0.373)	(0.186)	(1393.881)
0-250m outside project	0.044	0.125	0.014	-0.012	-0.020	0.337	0.150	-531.120
	(0.058)	(0.094)	(0.055)	(0.058)	(0.065)	(0.259)	(0.110)	(1416.472)
In-Situ Upgrading								
inside project	0.026	0.069	-0.007	0.040	-0.106	-0.060	-0.061	-2324.992[c]
	(0.098)	(0.076)	(0.106)	(0.092)	(0.117)	(0.247)	(0.123)	(1312.602)
0-250m outside project	0.042	0.009	-0.059	-0.075	-0.062	-0.127	-0.063	-1668.991
	(0.070)	(0.068)	(0.076)	(0.069)	(0.080)	(0.179)	(0.100)	(1323.043)
Other								
inside project	0.079	0.187[a]	0.153[c]	0.117	0.150[c]	0.197	0.160	-208.542
	(0.068)	(0.057)	(0.082)	(0.084)	(0.086)	(0.240)	(0.103)	(970.275)
0-250m outside project	-0.039	0.093[c]	-0.020	-0.005	-0.016	0.039	-0.081	68.527
	(0.042)	(0.053)	(0.044)	(0.045)	(0.046)	(0.152)	(0.075)	(1074.844)
Mean Outcome 2001	0.79	0.35	0.66	0.62	0.77	3.30	3.51	8,566.83
Mean Outcome 2011	0.83	0.54	0.81	0.72	0.82	3.56	3.18	9,823.82
R^2	0.625	0.602	0.674	0.638	0.644	0.680	0.693	0.608
N	12,732	12,732	12,732	12,732	12,732	12,709	12,730	12,734

All regressions include 3km grid Fixed-Effects. Standard errors clustered at the project level in parenthesis. [c] $p<0.10$,[b] $p<0.05$,[a] $p<0.01$

Table 5. Triple Difference Estimates on Log-Prices

	(1)	(2)
All Projects		
inside project	-0.152	-0.139
	(0.220)	(0.216)
0-250m outside project	-0.016	-0.010
	(0.061)	(0.060)
Lot Size Controls		✓
r2	0.52	0.52
N	67,751	67,751
Greenfield		
inside project	0.118	-0.010
	(0.167)	(0.172)
0-250m outside project	0.171	0.185
	(0.164)	(0.162)
In-Situ Upgrading		
inside project	0.082	0.142
	(0.308)	(0.286)
0-250m outside project	-0.179[c]	-0.177[b]
	(0.091)	(0.089)
Other		
inside project	-0.321	-0.280
	(0.274)	(0.267)
0-250m outside project	0.045	0.047
	(0.079)	(0.079)
Lot Size Controls		✓
r2	0.52	0.52
N	67,751	67,751

Standard errors clustered at the project level in parenthesis.
[c] p<0.10,[b] p<0.05,[a] p<0.01