Slum Upgrading and Long-run Urban Development: Evidence from Indonesia

December 2021

Mariaflavia (Nina) Harari

The Wharton School,
University of Pennsylvania

Maisy Wong

The Wharton School,
University of Pennsylvania,
and NBER

Slums and urbanization

How do cities grow out of informality

- Massive urbanization in developing countries
  - 2.5 billion people by 2050, >1 billion in slums today
  - Weak property rights, land scarcity

- A popular policy: slum upgrading = provide public goods on site

- Jakarta's Kampung Improvement Program (KIP)
  - 5 million beneficiaries, 25% of Jakarta's area, 1969-1984
  - Basic upgrades + 15-year verbal non-eviction guarantee
    - Eg. Roads, drains, sanitation, health centers, schools

- Dynamic inefficiency?
  - Upgrades improve well-being of many residents (World Bank, 1995)
  - Preserving slums at the expense of formal developments can generate opportunity costs

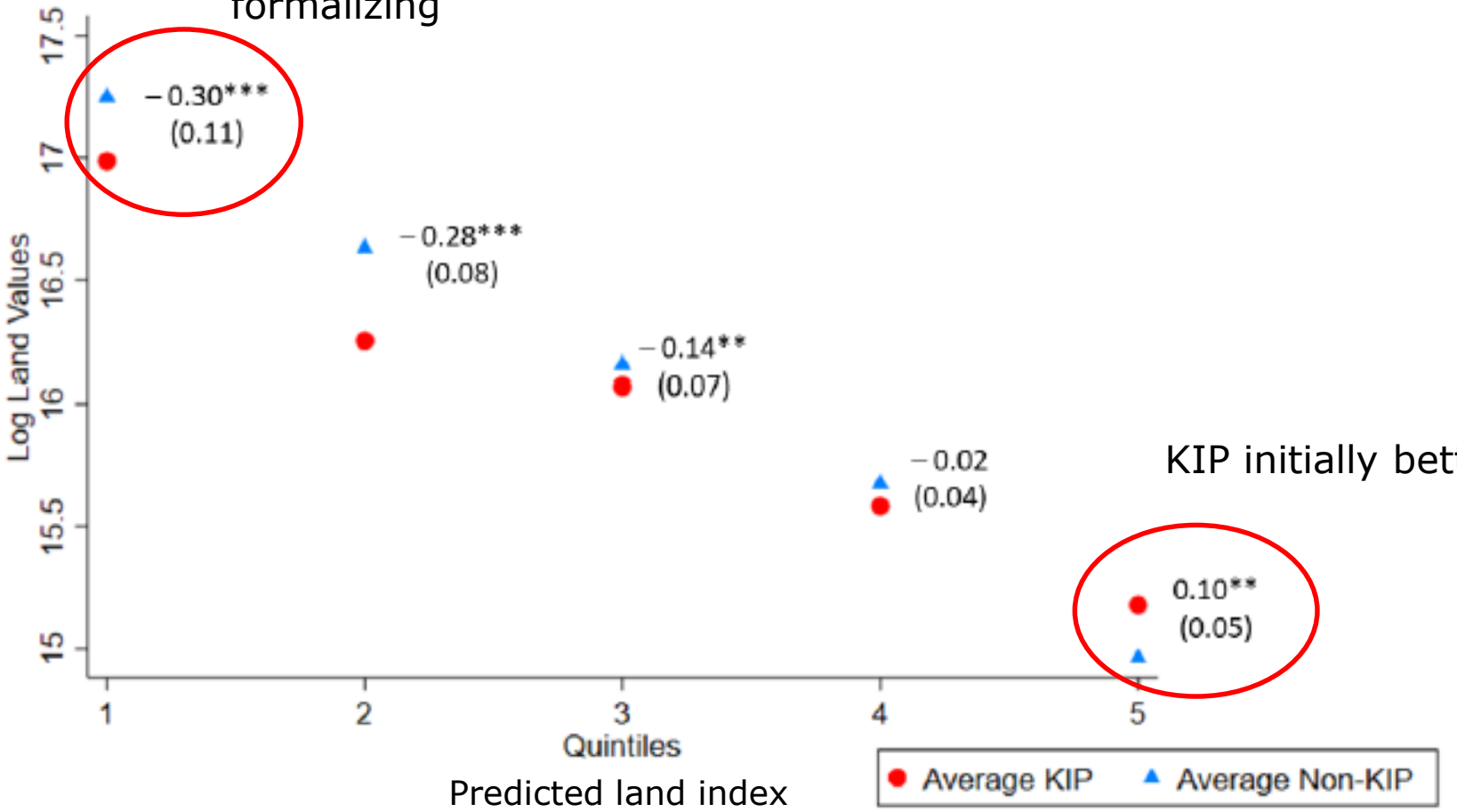
Causal evidence on world's largest slum upgrading program

- **Setting:** Jakarta, a mega-city growing out of informality
- **ID threats:** Measurement and coverage issues, selection bias
- **Research design**: KIP vs. non-KIP
  - Historical kampungs + neighborhood FE's
  - Boundary analysis (200m)
  - Staggered rollout to assess program selection bias
- Relative to non-KIP historical kampungs, KIP areas today:
  - 15% lower land values, 50% fewer tall buildings (>3 fl.) on average
  - Are more informal: quality index based on photos (+0.3 sd), share of unregistered land parcels (+3%)
  - Selection bias exists but goes away with granular controls

Cross-sectional heterogeneity by real-estate market potential

-Neighborhoods at early vs. late stages of urban development

Reversal as non-KIP starts formalizing



KIP initially better…

Should slums be upgraded or formalized?

Informing the debate

- Lower land values in KIP ≠ inefficiency
  - By law, slum residents without titles do not get compensated
  - Formalization can be privately profitable for developers but socially inefficient
- Surplus comparison: KIP vs non-KIP counterfactual
  - Granular data + treatment effects to quantify key trade-offs
  - Gains from formalizing: higher formal land values and heights
  - Losses from displacing slum residents: horizontal coverage in slums
- Where is inefficiency the greatest?
  - Concentrated losses: 90% in Q1 and Q2 (half of KIP areas)
  - KIP delivers sizeable surplus in Q3-Q5 (3 million ppl)
- Case studies to illustrate equity considerations

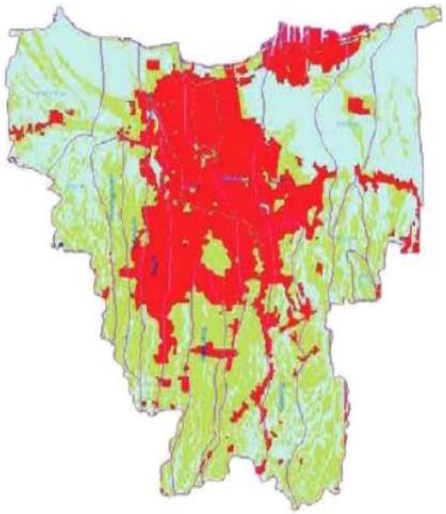
Related literature

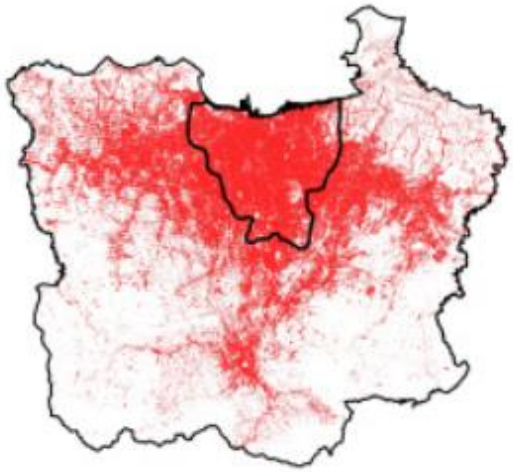
- **Urban development with informality:**
  - Slums and opportunity cost of land use (Henderson et al., 2021; Gechter and Tsivanidis, 2020)
  - Bleakley and Lin (2012), Libecap and Lueck (2013), Brooks and Lutz (2016), Hornbeck and Keniston (2017) …
- **Shelter provision and slum policy:**
  - Sites and services (Michaels et al., 2021)
  - Public housing (Picarelli, 2019; Barnhardt et al., 2017; Franklin (2019, 2020))
- **Urban renewal and place-based policies:**
  - Kline and Moretti (2014)
- **Our contribution**
  1. Novel causal estimates of long-term impacts of slum upgrading
  2. Data on formal and informal areas
  3. Quantify trade-offs associated with preserving vs redeveloping slums

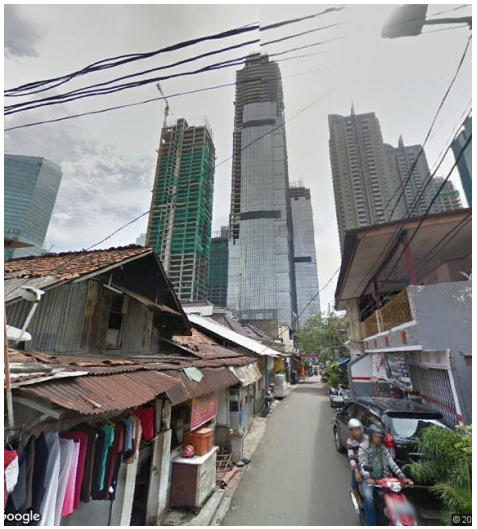
Outline

- Introduction
- **Background**
- Data
- Empirical strategy and results
- Surplus calculations
- Conclusions

Jakarta: a mega-city growing out of informality







1970s

- 10m people in city
- 30m in metro area

Today

KIP
(1969 – 1984)

The KIP program in Jakarta

- 3 waves: I (1969-1974), II (1974-1979), III (1979-1984)
- Goal: improve neighborhood conditions
  - Basic physical upgrades (estimated useful life ~ 15 years)
  - + verbal non-eviction guarantee for 15 years
- KIP components:
  - Road paving and widening
  - Drainage canals, sanitation (flooding concerns)
  - Health clinics and schools
- Selection criteria: scoring rule for neighborhood conditions, age, population density, income + even distribution across 5 districts
- 1995 World Bank evaluation:
  - positive impacts on neighborhood quality and human capital
  - KIP "crucial to establishing the permanence of the kampungs"

Kampungs in Jakarta, before and after KIP

KIP and kampung redevelopment

- WB report: policy should take into account "when and how the transformation of kampungs into modern real estate is likely to take place"
  - Effectively started in mid 2000s
- Redeveloping kampungs is complex: high land assembly costs (disputes, high fragmentation and density, political costs…)
- Slum upgrading programs can delay formalization:
  - Higher land values from upgrades
  - Strengthened perceptions of occupancy rights
  - Slums more attractive ➔ people stay ➔ greater population density and land fragmentation

Outline

- Introduction
- Background
- **Data**
- Empirical strategy and results
- Surplus calculations
- Conclusions

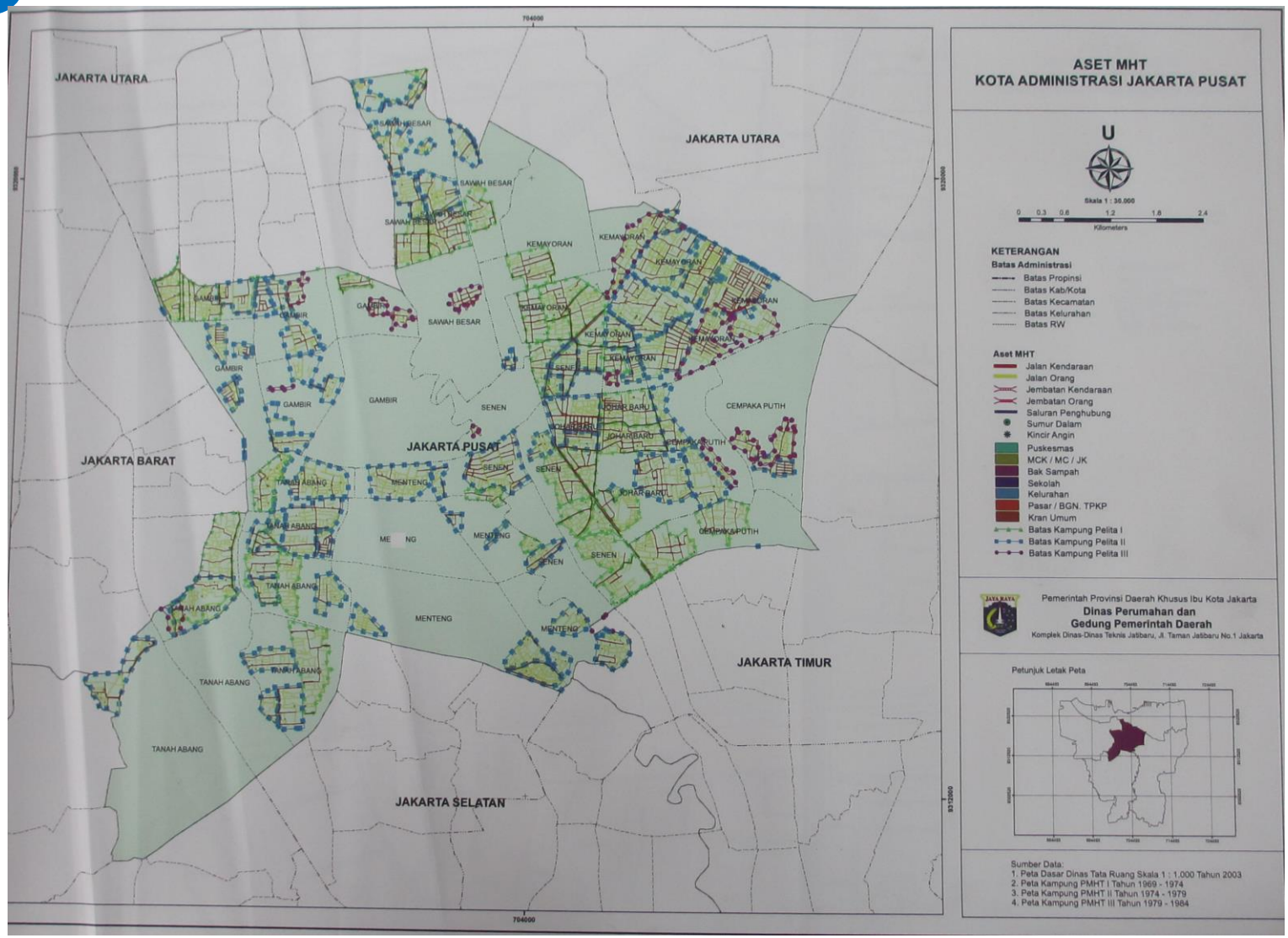
Core datasets:

1. Maps: KIP, historical slums
2. 2015 assessed land values
3. Photos: heights, informality

Auxiliary:

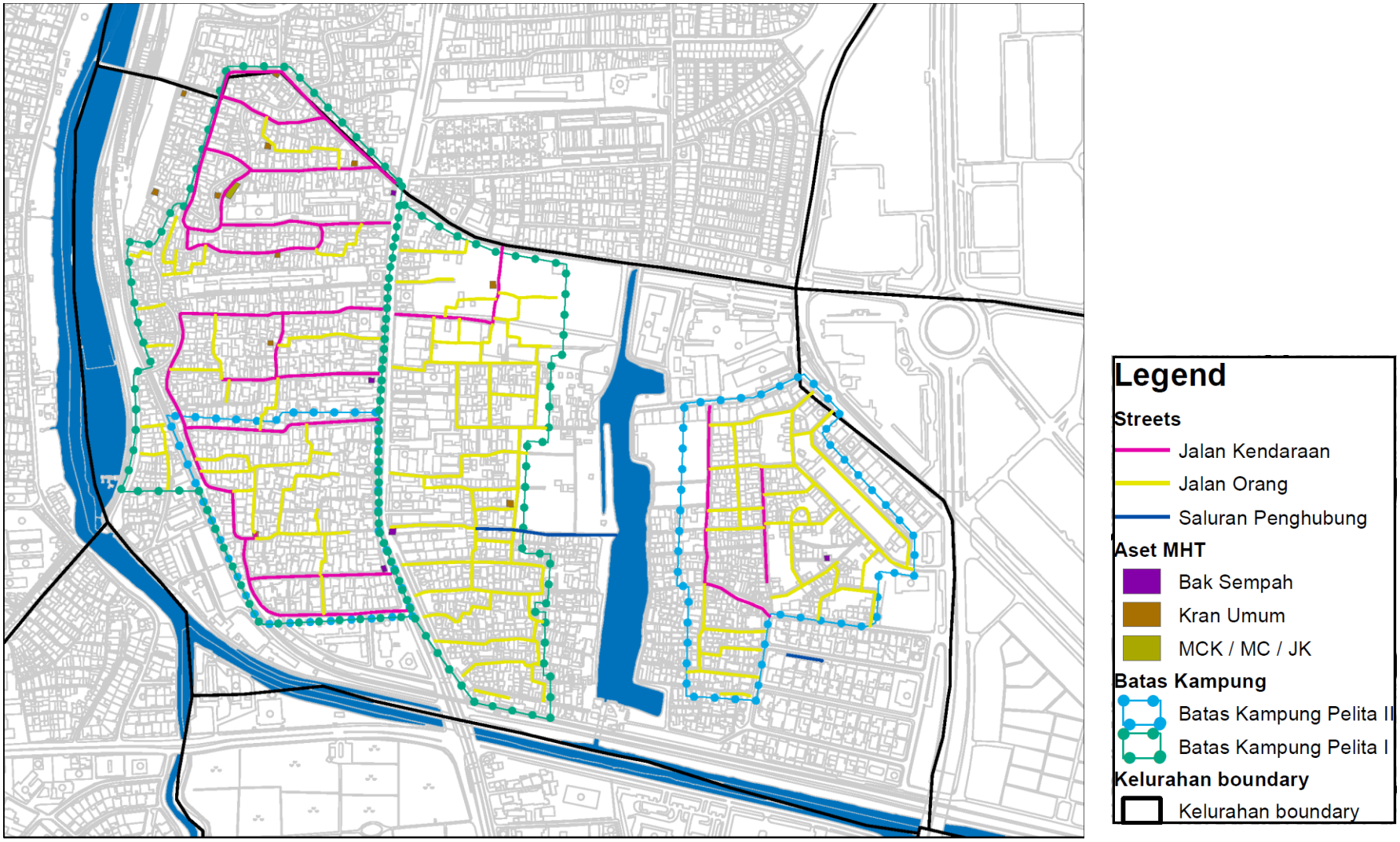
4. 2011 land parcels
5. 2010 Population Census
6. Land use, amenities …
7. Registered land titles
8. Geographic, landmarks, distance controls

Policy maps: KIP coverage

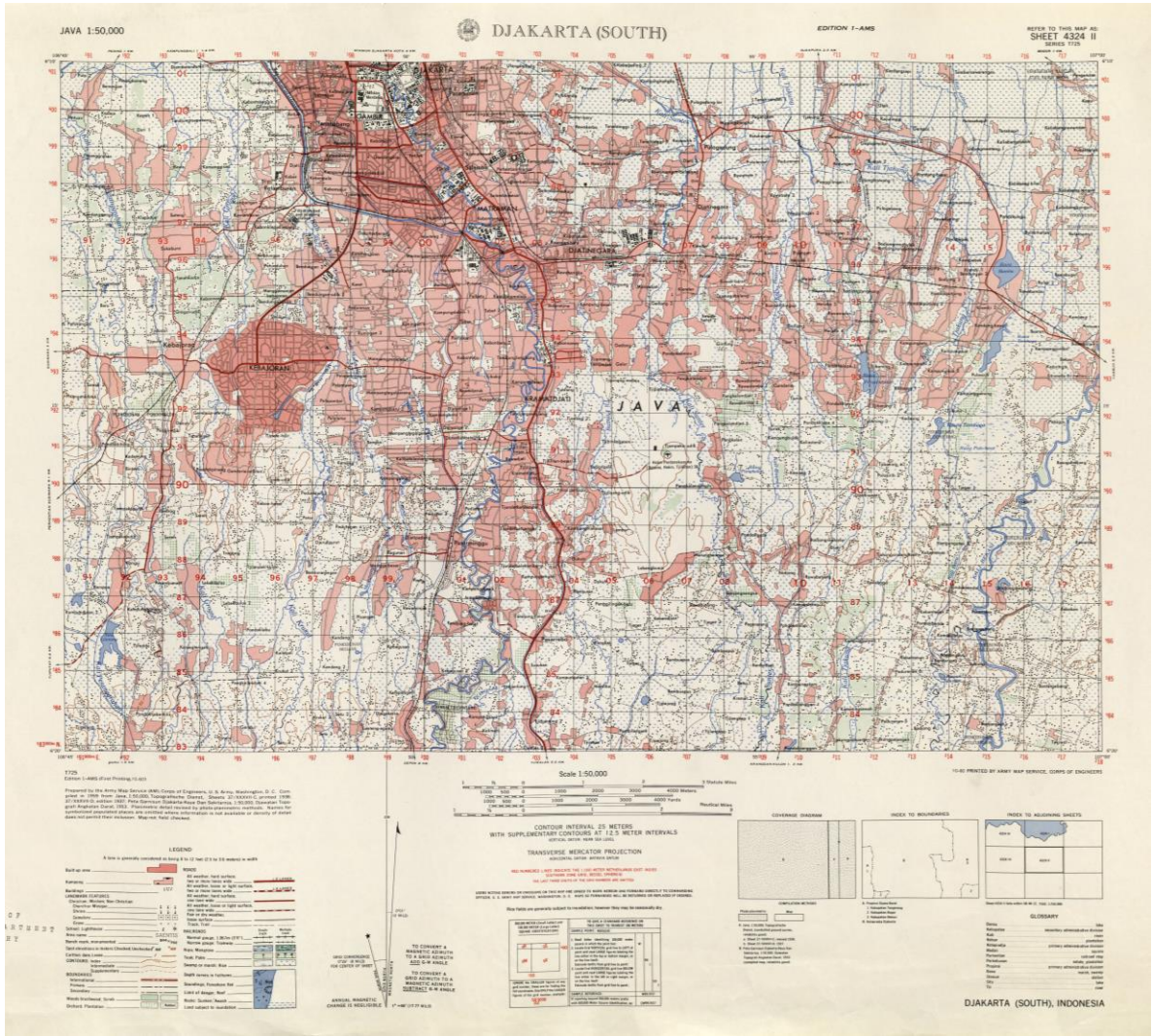


Jakarta Department of Housing , 2011

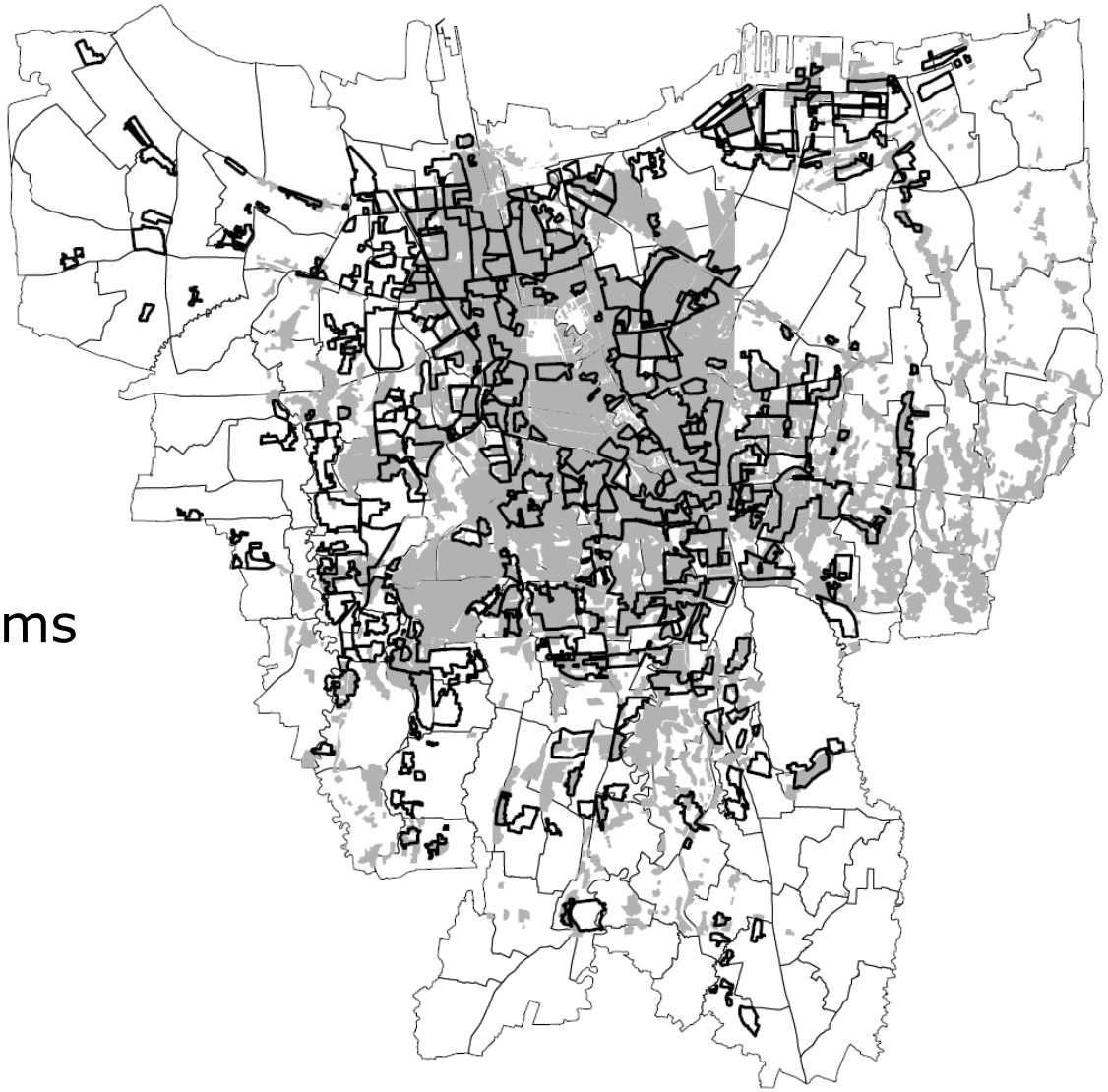
Policy maps: KIP boundaries and assets



Historical maps: Kampungs in 1959



Assembled data: treated and control slums



KIP areas

Historical slums (from 1937, 1959 maps)

Assessed land values, 2015

N = 19,848 sub-blocks

- Market-based assessment:
  - Goal: Property taxes
  - Start from broker data / listings, other sources
  - Adjustments (hedonic, field visits)
  - Subtract cost of structure based on engineering cost
- Validation check using transacted prices

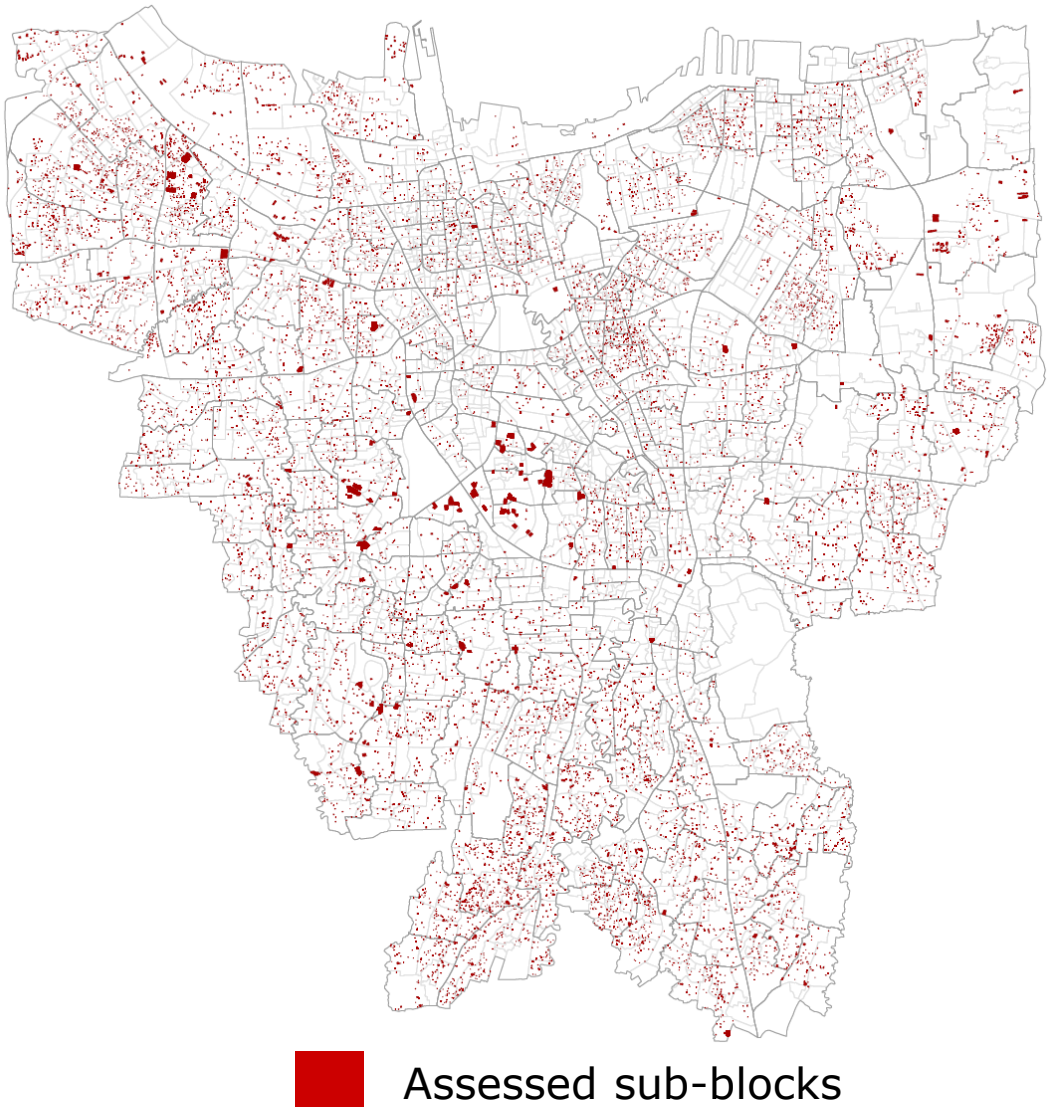


Photo survey for building height + informality

Real quantities/quality measure

- Representative sample: 19,518 pixels (75m x 75m)
- Google StreetView + field photos
  - For each pixel: take 4 photos (4 angles) from centroid
  - Photos from the field to overcome coverage bias in Google:
    - 19% photos: for slums
    - 5% photos: private gated developments
- Outcome: 1(tallest building in pixel > 3 floors)

- For a subset of ~ 28,000 photos in our key estimation samples:

  nr of floors and informality indices (more on this later)

Outline

- Introduction
- Background
- Data
- **Empirical strategy and results**
- Surplus calculations
- Conclusions

Empirical strategy

$$Y_{ij} = \alpha + \beta\, 1(KIP_{ij}) + X_{ij} + \varepsilon_{ij}$$

- Y: land values, building heights
- i = sub-block (land values) or pixel (for heights)
- j = geographic unit

- Estimation samples:
  - Historical kampungs: KIP vs. non-KIP within localities
  - Boundary discontinuity design: 200m

Balance test

- KIP vs. non-KIP differences cannot explain main results

Unit of analysis:	Sub-block level			Pixel level		
Sample:	Full sample (1)	Historical kampung (2)	BDD 200m (3)	Full sample (4)	Historical kampung (5)	BDD 200m (6)
Panel A: Landmark controls						
Log Distance to Monument	-0.37***	-0.02	0.002	-0.31***	0.001	0.004
	[0.00]	[0.12]	[0.80]	[0.00]	[0.85]	[0.31]
Log Distance to Tanjung Priok Harbor	-0.26***	-0.004	-0.004*	-0.19***	-0.004	-0.002
	[0.00]	[0.67]	[0.08]	[0.00]	[0.31]	[0.16]
Log Distance to Old Batavia	-0.31***	-0.02**	-0.005	-0.28***	0.005	0.08
	[0.00]	[0.01]	[0.66]	[0.00]	[0.64]	[0.33]
Log Distance to Concert Hall	-0.36***	-0.02	0.01	-0.31***	0.003	0.01
	[0.00]	[0.13]	[0.29]	[0.00]	[0.64]	[0.27]
Log Distance to Hotel Des Indes	-0.37***	-0.02*	-0.01	-0.34***	-0.001	-0.00006
	[0.00]	[0.09]	[0.33]	[0.00]	[0.88]	[0.99]
Log Distance to Bioscoop Metropool	-0.41***	-0.01	-0.01	-0.32***	0.01	0.002
	[0.00]	[0.55]	[0.20]	[0.00]	[0.42]	[0.58]
Log Distance to Akademi Nasional	-0.06	-0.03	0.01	-0.08**	-0.01	0.01
	[0.18]	[0.55]	[0.73]	[0.04]	[0.70]	[0.55]
Log Distance to Ragunan Zoo	0.15**	0.01	0.0004	0.07**	0.01	-0.001
	[0.03]	[0.45]	[0.93]	[0.04]	[0.30]	[0.53]
N	19848	3144	1291	88861	11002	3835
Geography FE	District	Locality	KIP Boundary	District	Locality	KIP Boundary

Summary statistics

Balance test

- KIP vs. non-KIP differences cannot explain main results

Unit of analysis:	Sub-block level			Pixel level		
Sample:	Full sample (1)	Historical kampung (2)	BDD 200m (3)	Full sample (4)	Historical kampung (5)	BDD 200m (6)
Panel B: Infrastructure controls						
Log Distance to Historical Main Road	-0.33***	-0.05	-0.01	-0.31***	0.002	0.03
	[0.00]	[0.18]	[0.76]	[0.00]	[0.95]	[0.14]
Presence of Wells or Pipes within 1000m	0.08***	0.01	-0.003	0.09***	-0.002	0.002
	[0.00]	[0.24]	[0.44]	[0.00]	[0.89]	[0.47]
Log Average Distance to Railway Stations	-0.64***	-0.02	0.01	-0.52***	0.01	-0.01
	[0.00]	[0.20]	[0.53]	[0.00]	[0.41]	[0.56]
Log Average Distance to Tram Stations	-0.52***	-0.02	-0.002	-0.45***	0.004	0.003
	[0.00]	[0.19]	[0.84]	[0.00]	[0.74]	[0.63]
Panel C: Topography controls						
Elevation, m	-4.90***	-0.58	0.14	-3.90***	-0.25	0.09
	[0.00]	[0.49]	[0.79]	[0.00]	[0.37]	[0.79]
Slope, Degrees	-0.09	-0.20	-0.24	-0.001	-0.06	-0.17
	[0.71]	[0.62]	[0.36]	[0.99]	[0.64]	[0.25]
Log Average Distance to 1959 Waterways	-0.15***	0.002	0.0005	-0.12***	-0.002	-0.0002
	[0.00]	[0.77]	[0.84]	[0.00]	[0.53]	[0.89]
Flow Accumulation	0.12	0.92	0.49	-0.11*	0.18	-0.002
	[0.59]	[0.13]	[0.35]	[0.06]	[0.38]	[0.99]
Log Distance to Coast	-0.22***	-0.005	-0.001	-0.17***	-0.01	0.0005
	[0.00]	[0.75]	[0.83]	[0.00]	[0.32]	[0.92]
Log Distance to Surface Water Occurrence	-0.08	-0.01	-0.004	-0.12*	-0.03	0.01
	[0.38]	[0.88]	[0.79]	[0.10]	[0.33]	[0.49]
N	19848	3144	1291	88861	11002	3835
Geography FE	District	Locality	KIP Boundary	District	Locality	KIP Boundary

Results: a roadmap

- Main results: land values , heights
- Threats to identification: program selection bias
- Heterogeneity by market potential
- Channels

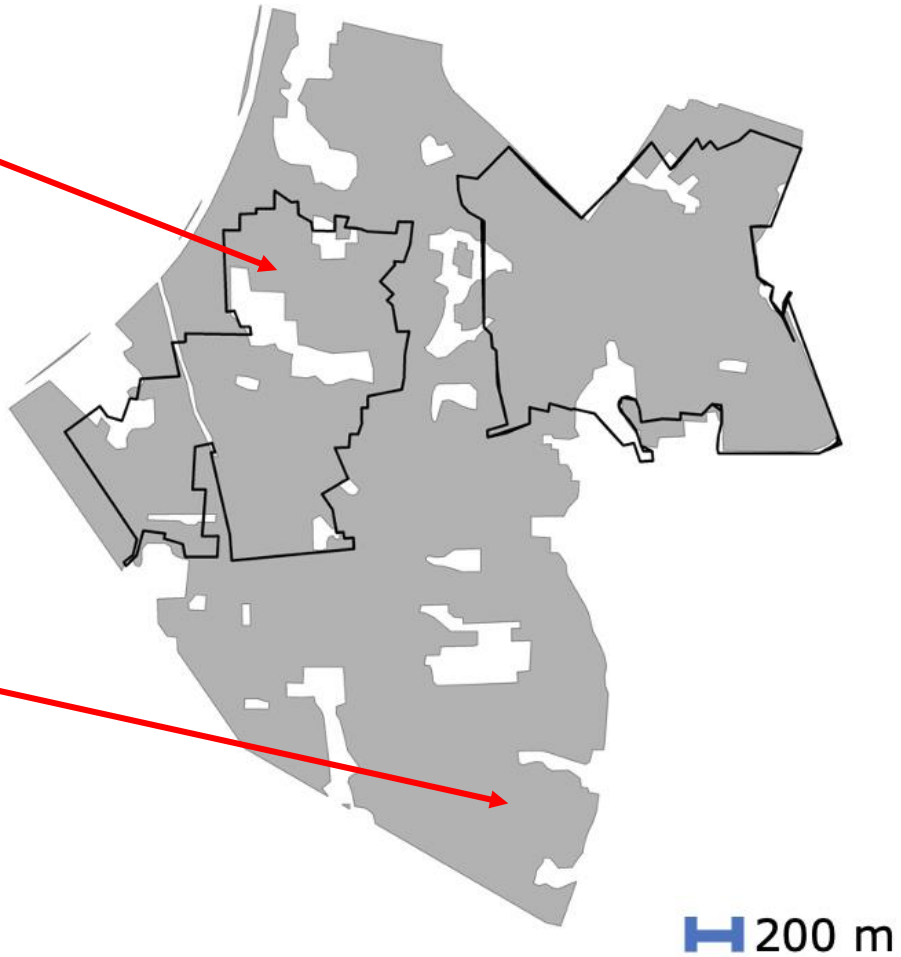
Land values: -15% in KIP

- 196 locality FE's, 123 boundary FE's

Dependent variable:	**Log land values**
Sample:	Historical kampung
	(1)
KIP	-0.14***
	(0.05)
N	3144
R-Squared	0.73
Distance	Y
Topography	Y
Landmarks	Y
Distance to KIP boundary	N
Geography FE	Locality

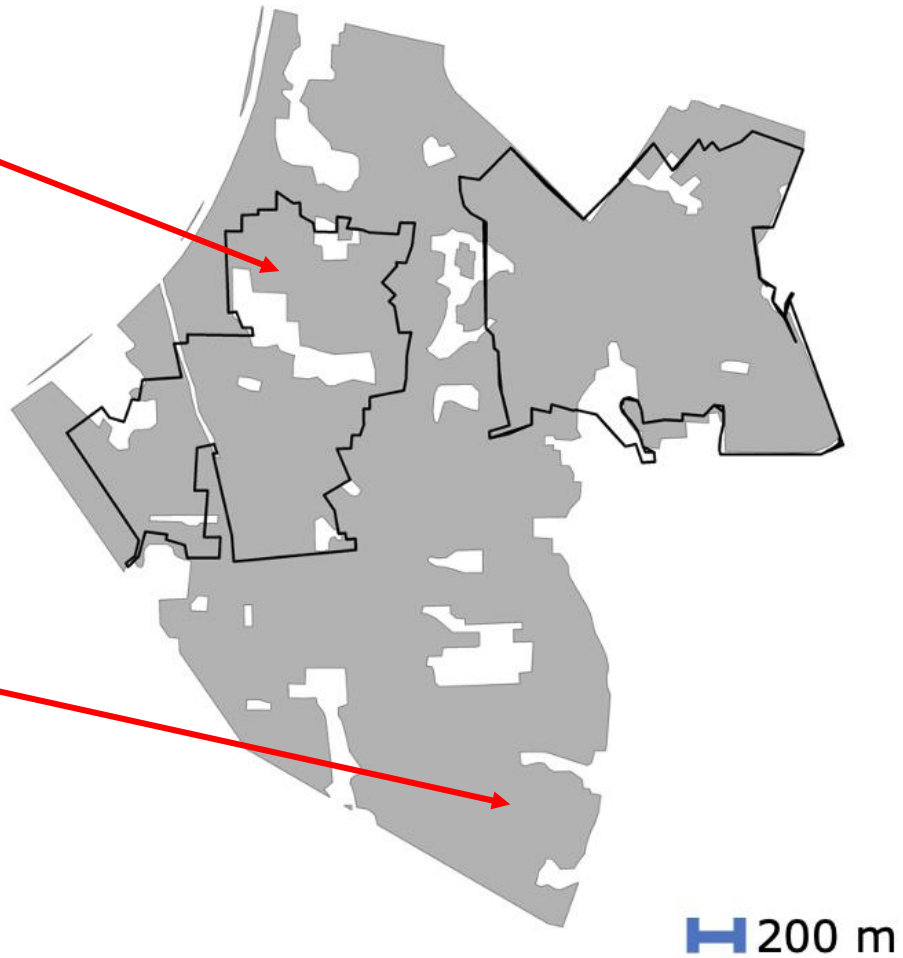
Compare historical slums with and without KIP

Case study: Setia Budi



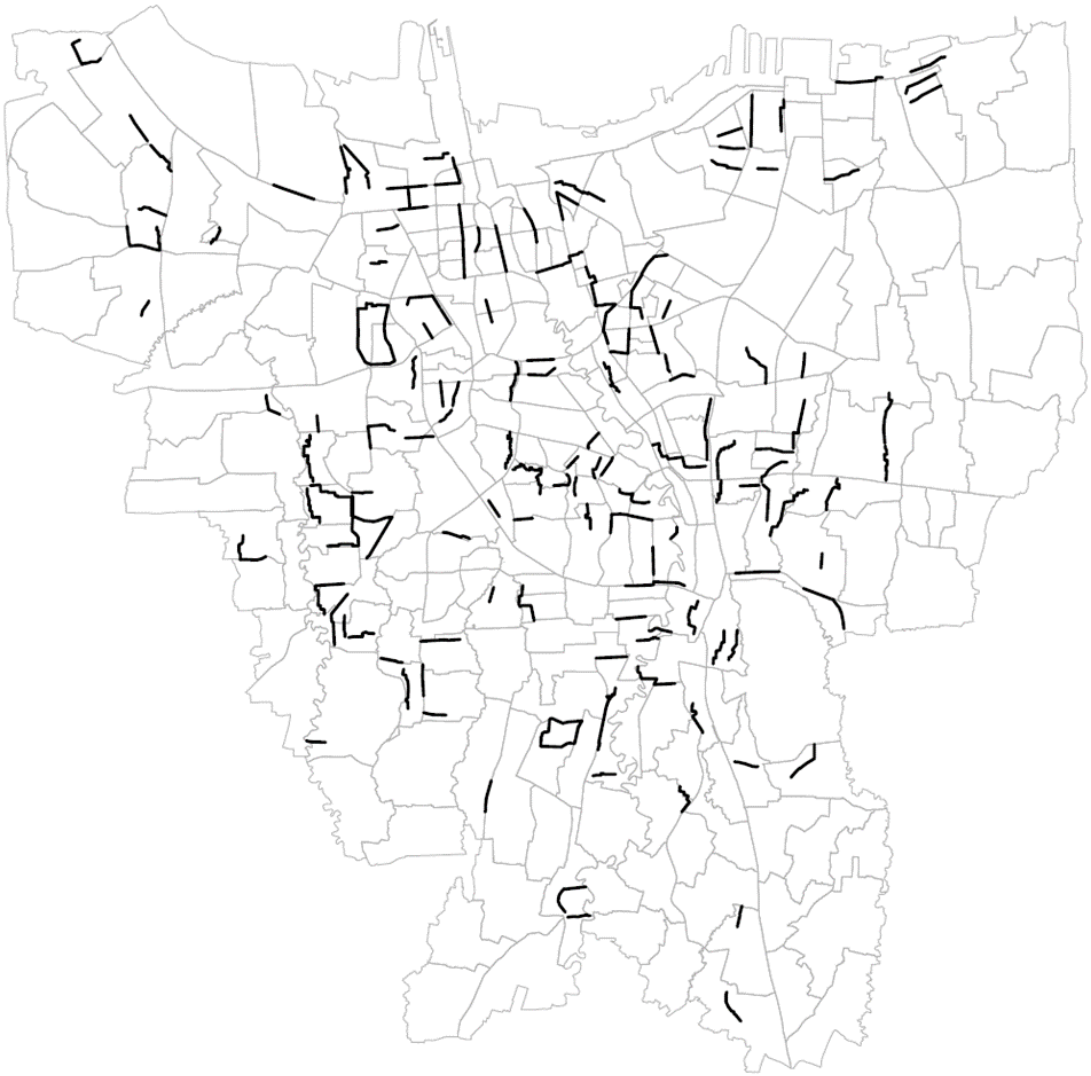
Compare historical slums with and without KIP

Case study: Setia Budi



Empirical strategy

- Boundary discontinuity design



Land values: -15% in KIP

- 196 locality FE's, 123 boundary FE's

Dependent variable:	Log land values	
Sample:	Historical kampung	BDD 200m
	(1)	(2)
KIP	-0.14***	-0.17***
	(0.05)	(0.06)
N	3144	1291
R-Squared	0.73	0.81
Distance	Y	Y
Topography	Y	Y
Landmarks	Y	Y
Distance to KIP boundary	N	Y
Geography FE	Locality	KIP Boundary

KIP: half as many tall buildings (1 if > 3 floors)

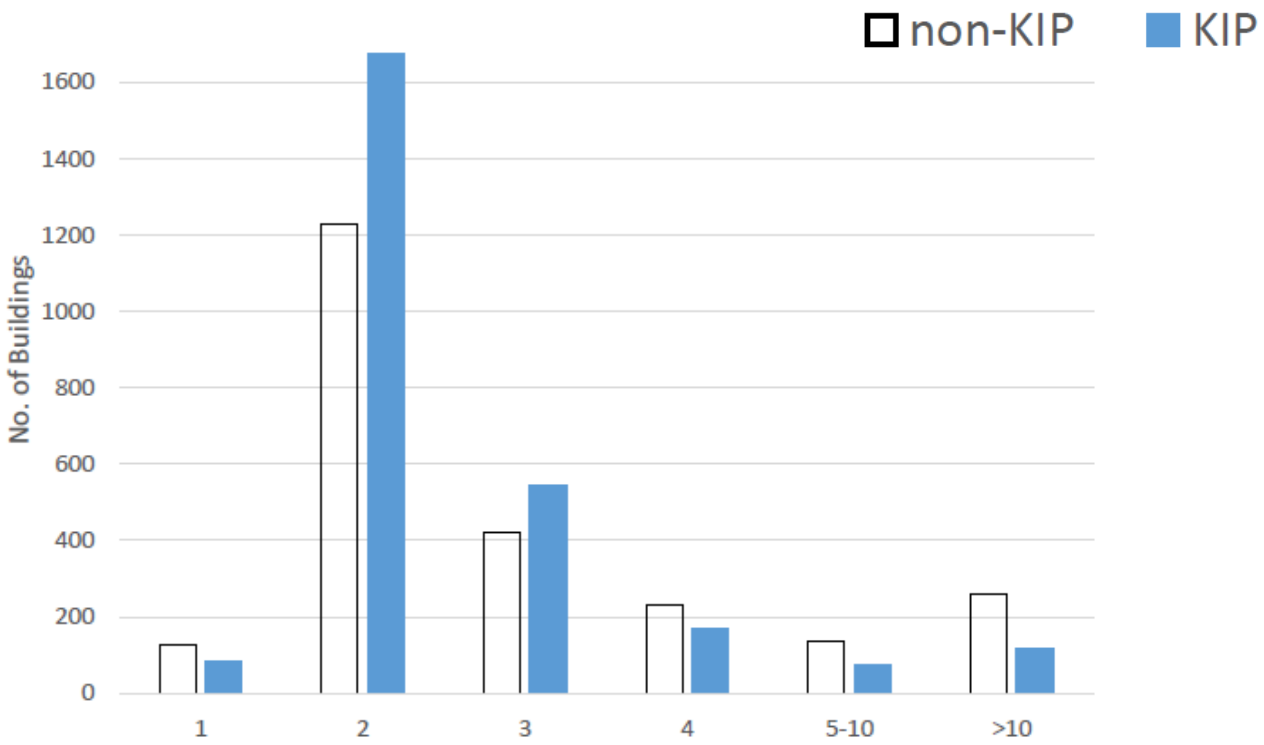
-12pp (relative to control group mean of 0.24)

Dependent variable:	**Log land values**		**1(Height>3)**	
Sample:	Historical kampung	BDD 200m	Historical kampung	BDD 200m
	(1)	(2)	(3)	(4)
KIP	-0.14***	-0.17***	-0.12***	-0.08**
	(0.05)	(0.06)	(0.02)	(0.03)
N	3144	1291	5277	1036
R-Squared	0.73	0.81	0.29	0.38
Distance	Y	Y	Y	Y
Topography	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y
Distance to KIP boundary	N	Y	N	Y
Geography FE	Locality	KIP Boundary	Locality	KIP Boundary

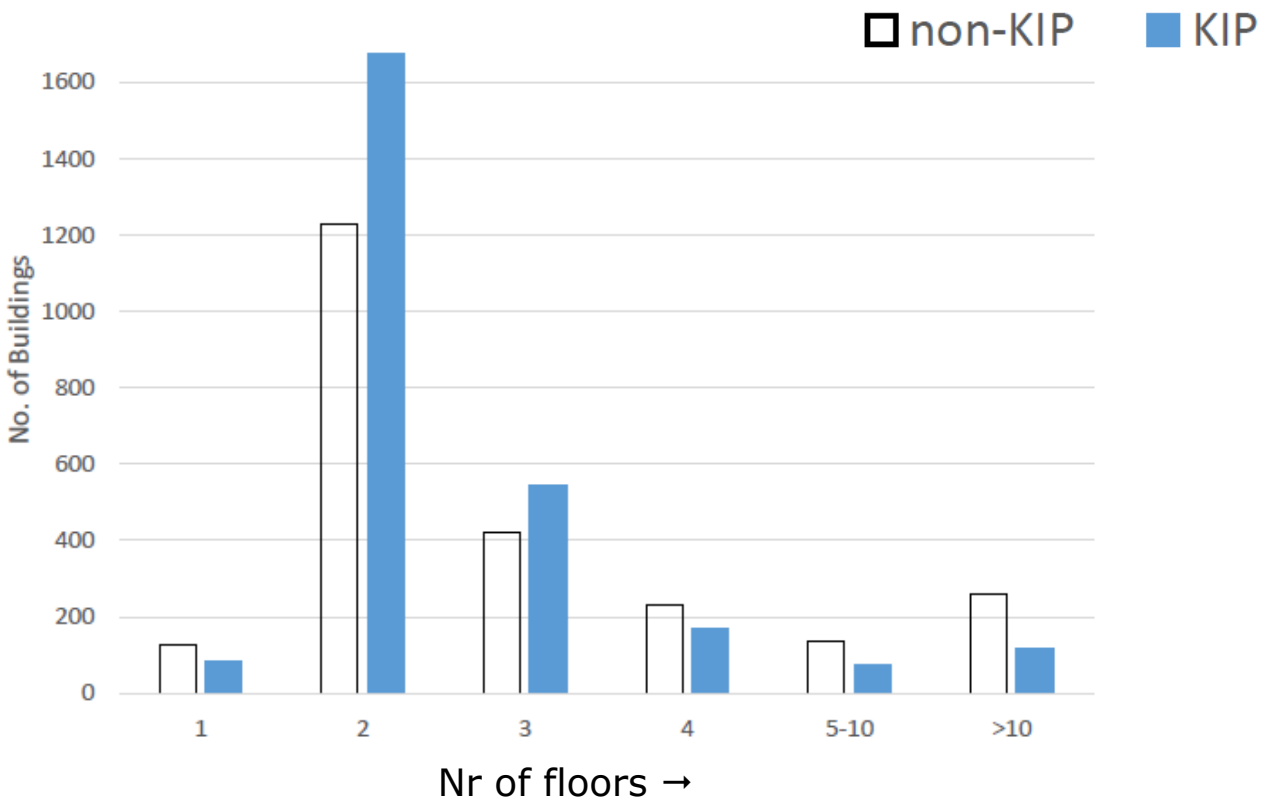
Robustness

KIP areas have half as many tall buildings

+ bunching at 2 floors



Height results validate land values result



Translating height effects to land values

Given estimated price premium of high-rises:
(from hedonic reg. in non-KIP areas)

➔ Missing high-rises in KIP explain 90% of difference in land values

Results: a roadmap

- Main results:

  - 15% land values , 50% fewer tall buildings in KIP

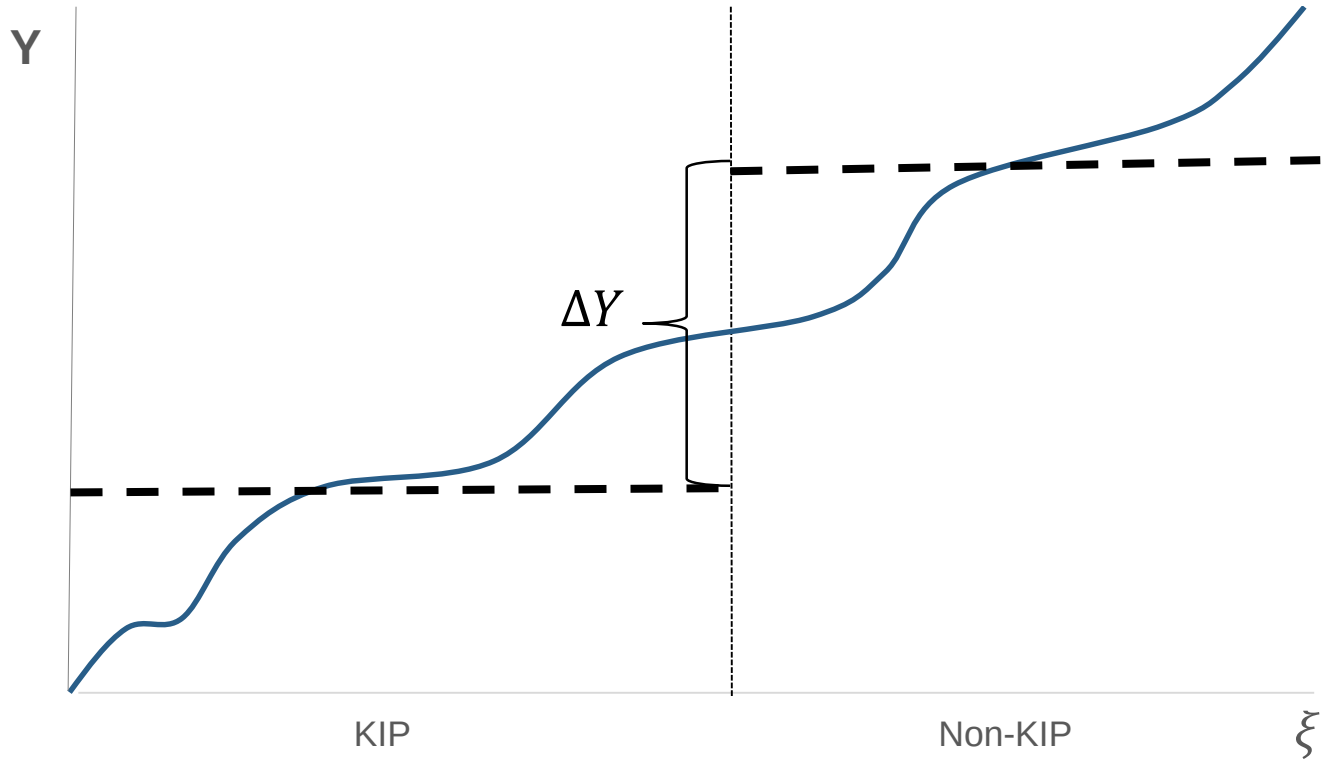
- **Threats to identification: program selection bias**

- Heterogeneity by market potential

- Channels

Using staggered roll-out to assess selection bias

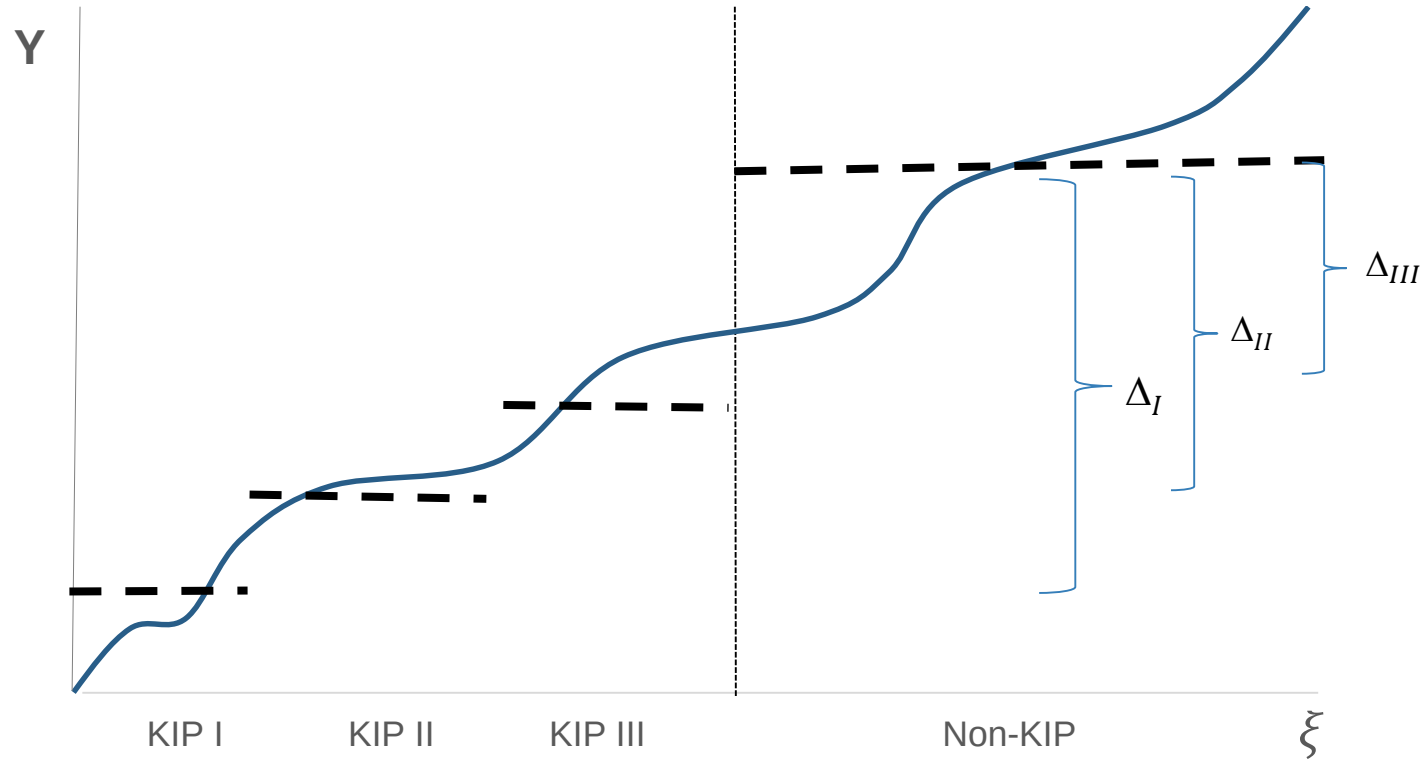
$$Y_{ij} = \alpha + \beta KIP_{ij} + \xi_j + \varepsilon_{ij}$$



$$\Delta Y = \beta + E(\xi_j | KIP) - E(\xi_j | nonKIP)$$

Selection bias

Scoring rule implies: $\Delta_I < \Delta_{II} < \Delta_{III}$



Monotonic pattern consistent with scoring rule

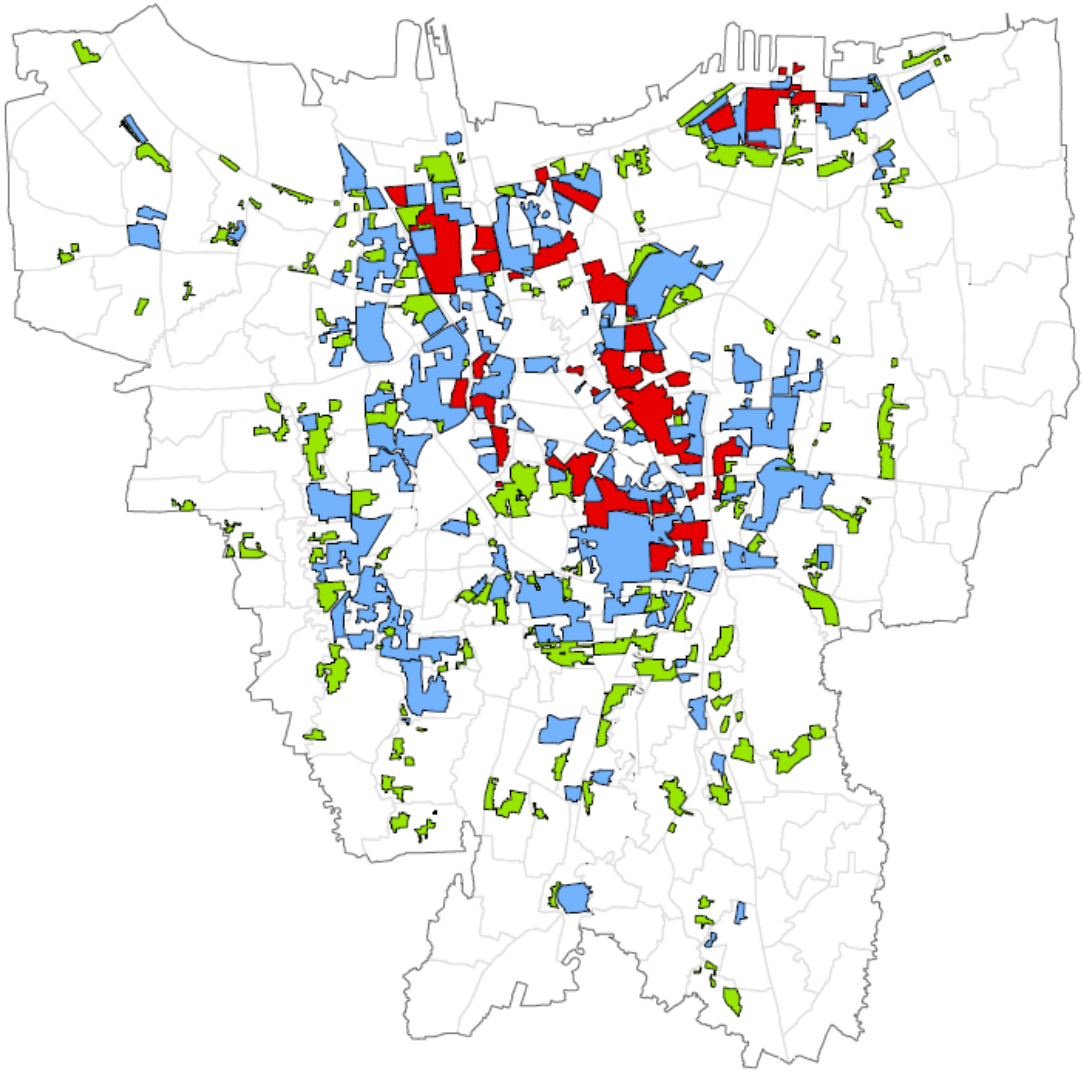
Dependent variable:	**Log land values**
Sample:	Full sample
	(1)
KIP I (1969-1974)	-0.40***
	(0.07)
KIP II (1974-1979)	-0.29***
	(0.07)
KIP III (1979-1984)	-0.17**
	(0.08)
N	19848
R-Squared	0.57
p-val ($H_0 : \|\beta_I\| \leq \|\beta_{II}\|$)	0.108
p-val ($H_0 : \|\beta_{II}\| \leq \|\beta_{III}\|$)	0.119
Distance	Y
Topography	Y
Landmarks	Y
KIP investments	N
Distance tercile	N
Geography FE	District

Pattern disappears: historical kampungs + locality FE's

Dependent variable:	Log land values	
Sample:	Full sample	Historical kampung
	(1)	(2)
KIP I (1969-1974)	-0.40***	-0.13
	(0.07)	(0.09)
KIP II (1974-1979)	-0.29***	-0.11*
	(0.07)	(0.06)
KIP III (1979-1984)	-0.17**	-0.14*
	(0.08)	(0.08)
N	19848	3144
R-Squared	0.57	0.73
p-val ($H_0 : \|\beta_I\| \leq \|\beta_{II}\|$)	0.108	0.357
p-val ($H_0 : \|\beta_{II}\| \leq \|\beta_{III}\|$)	0.119	0.609
Distance	Y	Y
Topography	Y	Y
Landmarks	Y	Y
KIP investments	N	N
Distance tercile	N	N
Geography FE	District	Locality

KIP waves

Red (wave I), blue (wave II), green (wave III)



Robust to controlling for heterog. treatment by waves

Dependent variable:	Log land values		
Sample:	Full sample	Historical kampung	Historical kampung
	(1)	(2)	(3)
KIP I (1969-1974)	-0.40***	-0.13	-0.09
	(0.07)	(0.09)	(0.10)
KIP II (1974-1979)	-0.29***	-0.11*	-0.07
	(0.07)	(0.06)	(0.06)
KIP III (1979-1984)	-0.17**	-0.14*	-0.09
	(0.08)	(0.08)	(0.08)
N	19848	3144	3144
R-Squared	0.57	0.73	0.74
p-val ($H_0 : \lvert\beta_I\rvert \leq \lvert\beta_{II}\rvert$)	0.108	0.357	0.378
p-val ($H_0 : \lvert\beta_{II}\rvert \leq \lvert\beta_{III}\rvert$)	0.119	0.609	0.581
Distance	Y	Y	Y
Topography	Y	Y	Y
Landmarks	Y	Y	Y
KIP investments	N	N	Y
Distance tercile	N	N	Y
Geography FE	District	Locality	Locality

Similar patterns for heights

Dependent variable:	Log land values			1(Height>3)		
Sample:	Full sample	Historical kampung	Historical kampung	Photo sample	Historical kampung	Historical kampung
	(1)	(2)	(3)	(4)	(5)	(6)
KIP I (1969-1974)	-0.40***	-0.13	-0.09	-0.13***	-0.10***	-0.08***
	(0.07)	(0.09)	(0.10)	(0.03)	(0.03)	(0.03)
KIP II (1974-1979)	-0.29***	-0.11*	-0.07	-0.10***	-0.10***	-0.08***
	(0.07)	(0.06)	(0.06)	(0.01)	(0.02)	(0.02)
KIP III (1979-1984)	-0.17**	-0.14*	-0.09	-0.04**	-0.07**	-0.05
	(0.08)	(0.08)	(0.08)	(0.02)	(0.03)	(0.03)
N	19848	3144	3144	19518	5277	5277
R-Squared	0.57	0.73	0.74	0.15	0.29	0.29
p-val ($H_0 : \|\beta_I\| \leq \|\beta_{II}\|$)	0.108	0.357	0.378	0.133	0.525	0.418
p-val ($H_0 : \|\beta_{II}\| \leq \|\beta_{III}\|$)	0.119	0.609	0.581	0.010	0.199	0.245
Distance	Y	Y	Y	Y	Y	Y
Topography	Y	Y	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y	Y	Y
KIP investments	N	N	Y	N	N	Y
Distance tercile	N	N	Y	N	N	Y
Geography FE	District	Locality	Locality	District	Locality	Locality

Other threats to identification

- Generic persistence of slums
- Spatial spillovers
- BDD robustness: bandwidth, overlapping boundaries

Results: a roadmap

- Main results:

  - 15% land values , 50% fewer tall buildings in KIP

- Threats to identification

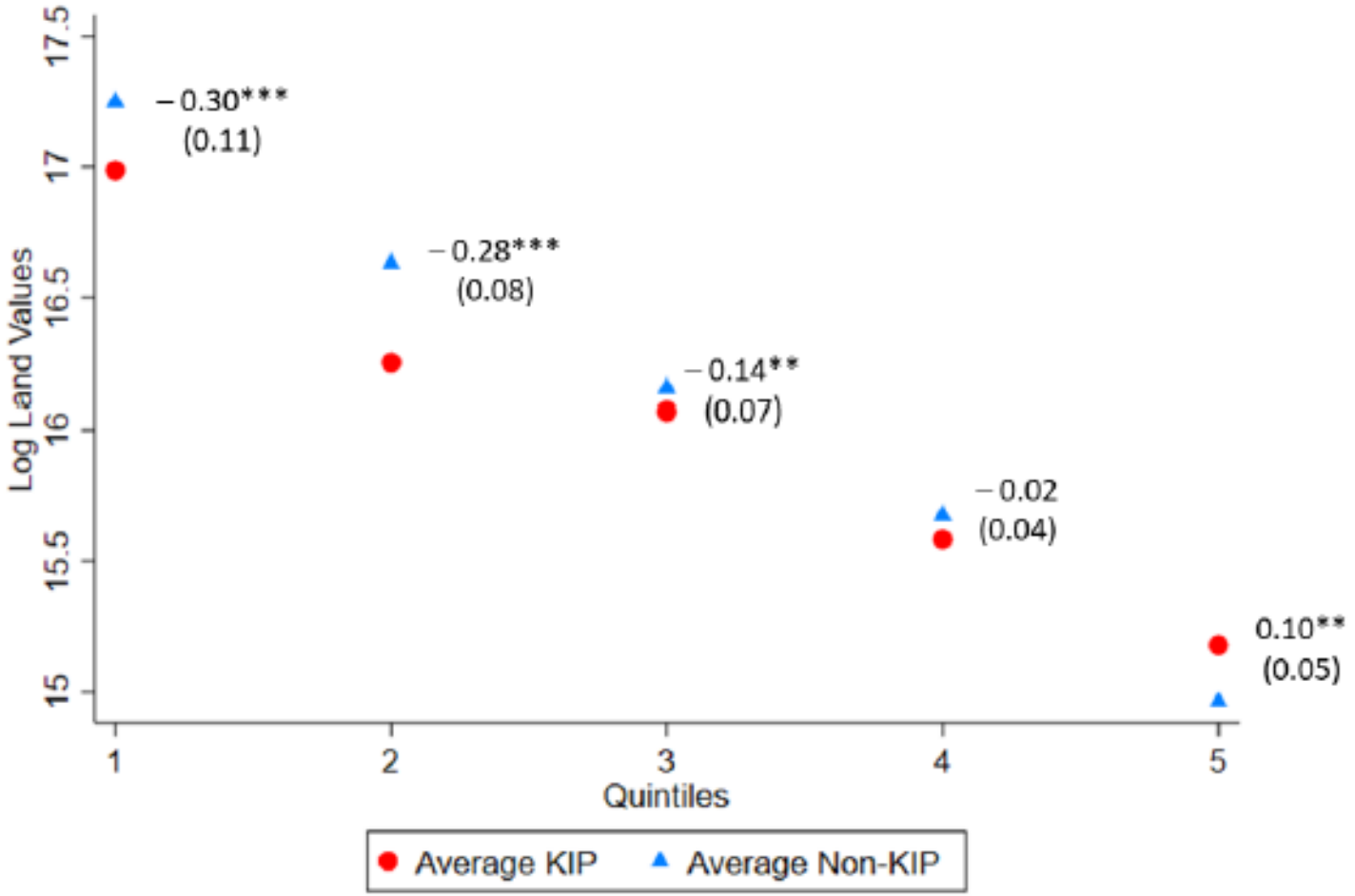
  Historical sample + granular FEs + controls address most selection

- **Heterogeneity by market potential**

- Channels

Heterogeneity by market potential

Effect of KIP on log land values by quintiles of non-KIP land values



Potential displacement

Results: a roadmap

- Main results:

  - 15% land values , 50% fewer tall buildings in KIP

- Threats to identification

  Historical sample + granular FEs + controls address most selection

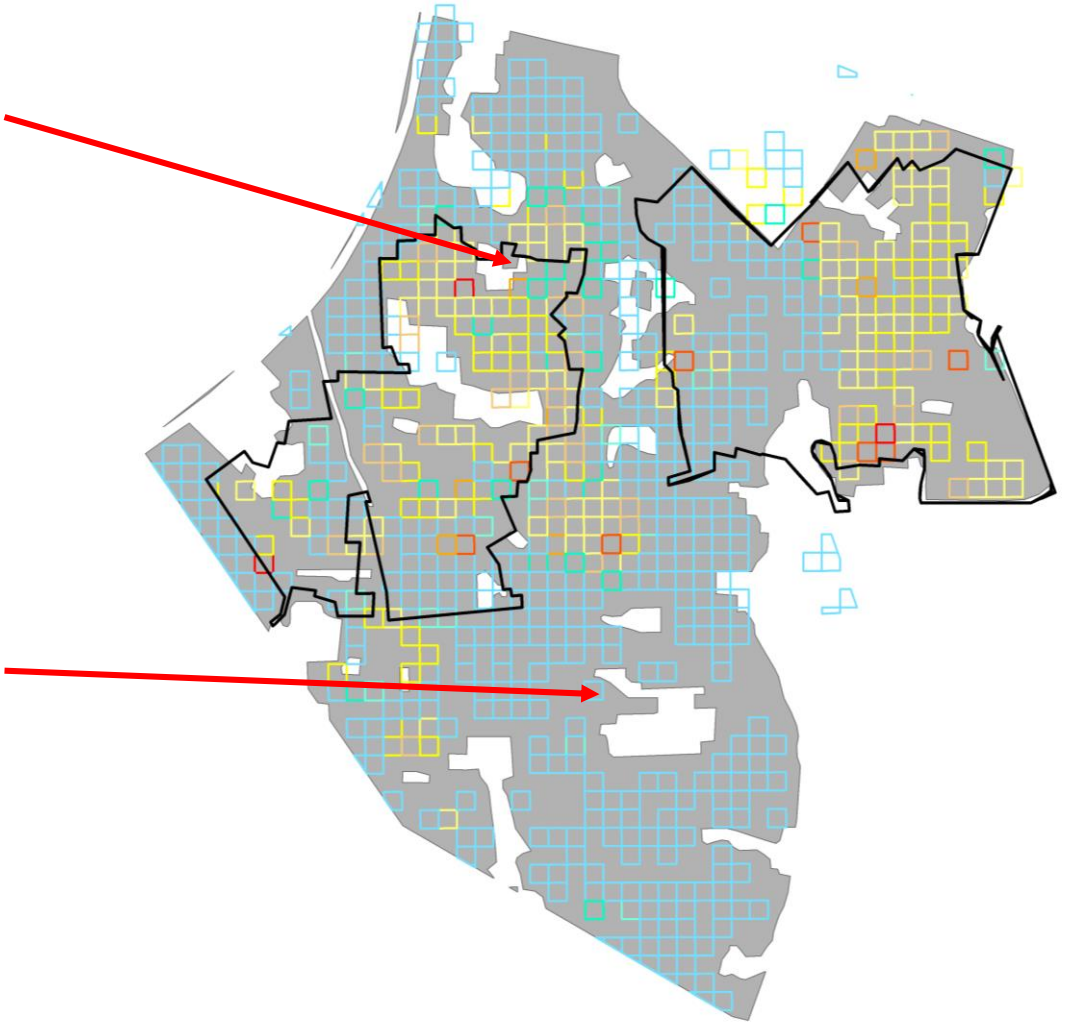
- Heterogeneity by market potential

  Positive KIP effects in Q5, negative in Q1

- **Channels:** Why do upgraded areas have low land values and heights?

Why do upgraded areas have low land values and heights?

Non-KIP neighborhoods formalize, KIP stays informal



blue = very formal
red = very informal

Informality index from photo sample

N=7,101 pixels (~ 28,000 photos) from historical kampung + BDD samples

Rank-based index

- Subjective ranking by 2 RA's
- Averaged (robust to RA FE's)

0 = very formal

1

2

3

4 = very informal

Attribute-based index

- Manually code 15 attributes and average z-scores:
  - Vehicular access (e.g. paved roads)
  - Structures (e.g. permanent wall)
  - Appearance (e.g. exposed wires)

Rank index and attributes

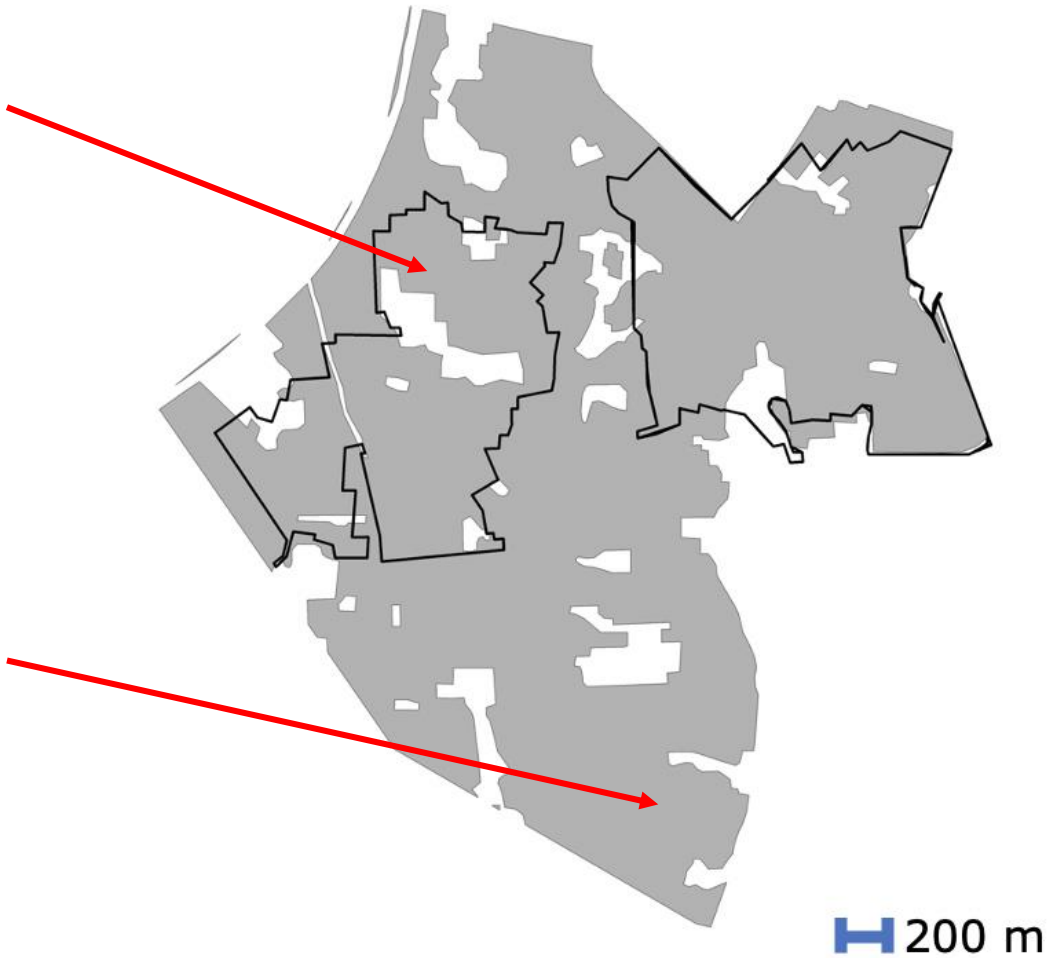
KIP areas more likely to be kampungs

+ 0.29 sd + 0.05 sd

Dependent variable:	Rank-based index		Attribute-based index		Unregistered parcels (shares)	
Sample:	Historical kampung	BDD 200m	Historical kampung	BDD 200m	Historical kampung	BDD 200m
	(1)	(2)	(3)	(4)	(5)	(6)
KIP	0.29***	0.38***	0.05**	0.06	0.03**	0.04*
	(0.05)	(0.13)	(0.02)	(0.05)	(0.01)	(0.02)
N	5277	1036	5277	1036	5277	1036
R-Squared	0.26	0.39	0.17	0.26	0.35	0.40
Distance	Y	Y	Y	Y	Y	Y
Topography	Y	Y	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y	Y	Y
Distance to KIP boundary	N	Y	N	Y	N	Y
Geography FE	Locality	KIP boundary	Locality	KIP boundary	Locality	KIP boundary

Cadastral maps of parcels, 2011

- Parcel count as proxy of land assembly costs



Land fragmentation and population density

- 9 more parcels and 11 more households per pixel in KIP

Dependent variable:	Parcel density		Log population density
Sample:	Historical kampung (1)	BDD 200m (2)	Historical kampung (3)
KIP	8.55***	13.84***	0.33***
	(1.07)	(1.29)	(0.07)
N	11002	3835	1184
R-Squared	0.51	0.44	0.56
Distance	Y	Y	Y
Topography	Y	Y	Y
Landmarks	Y	Y	Y
Distance to KIP boundary	N	Y	N
Geography FE	Locality	KIP boundary	Locality

- Pre-KIP population density is a confounder but not large enough to explain the effects

Heterogeneous effects on fragmentation Congestion Mortality and fertility Migration

Results: a roadmap

- Main results:

  - 15% land values , 50% fewer tall buildings in KIP

- Threats to identification

- Heterogeneity by market potential

- **Channels**

  - Consistent with delayed formalization:

    KIP more informal today, greater parcels and population density

Results: a roadmap

- Main results:

  - 15% land values , 50% fewer tall buildings in KIP

- Threats to identification

- Heterogeneity by market potential

- **Channels**

  - Consistent with delayed formalization:

    KIP more informal today, greater parcels and population density

  - Other channels we consider:

    - KIP-provided amenities
    - Current amenities
    - Human capital

- Additional robustness checks

Outline

- Introduction
- Background
- Data
- Empirical strategy and results

- 15% land values , 50% fewer tall buildings in KIP

- **Surplus calculations**

where is it inefficient to preserve slums?

- Conclusions

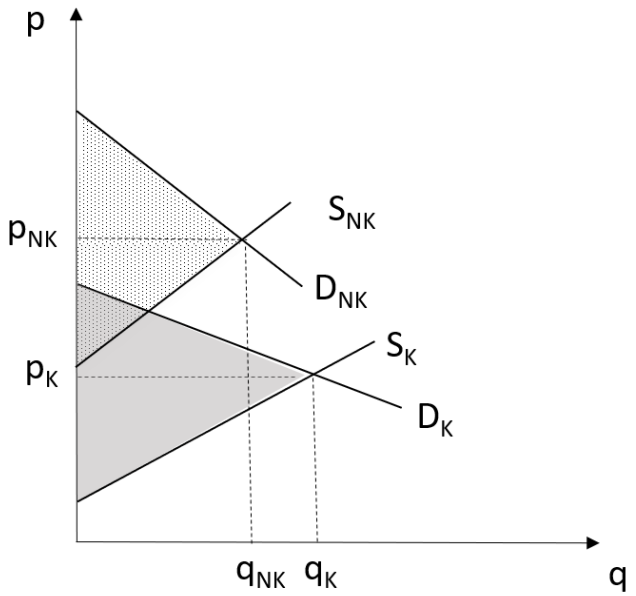
Where is KIP inefficient?

- Lower land values in KIP do not imply inefficiency:

Need to compare gains from formalization with loss in informal surplus

- Back-of-the-envelope calculation neighborhood by neighborhood:

$\Delta Surplus_{KIP\text{-}NKIP} = CS_{KIP} + PS_{KIP} - (CS_{NKIP} + PS_{NKIP})$



Caveats:

- focus on individual neighborhoods abstracting from city-level effects (e.g. externalities, KIP-induced displacement, open-city migration…)
- exercise quantifies opportunity costs on preserving slums today (not overall effect of KIP program on people)

Surplus calculations

- Observe p_K, q_K in data and use treatment effects to pin down p_{NK}, q_{NK}

	(1)	(2)	(3)	(4)	(5)	(6)
	KIP treatment effects		*Value of built-up stock*		*ΔValue* $_{K-NK}$	*ΔSurplus* $_{K-NK}$
	Log land values	*Log heights*	*KIP*	*Non-KIP*		
(1) Q1 (highest)	-0.30***	-0.25	$1,873	$3,098	-$1,225	-$2,369
	(0.11)	(0.16)				
(2) Q2	-0.28***	-0.01	$1,112	$1,716	-$603	-$1,044
	(0.08)	(0.05)				
(3) Q3	-0.14**	-0.15**	$972	$1,317	-$345	-$382
	(0.07)	(0.07)				
(4) Q4	-0.02	0.06	$717	$738	-$22	$398
	(0.04)	(0.04)				
(5) Q5 (lowest)	0.10**	-0.03	$489	$478	$11	$347
	(0.05)	(0.08)				
(6) Overall			$1,113	$1,626	-$513	-$781

Land index

Q1 = high real estate market potential

Surplus calculations

Value of built-up stock = value of land + value of structures

				$p_K q_K$	$p_{NK} q_{NK}$		
		(1)	(2)	(3)	(4)	(5)	(6)
		KIP treatment effects		*Value of built-up stock*		*ΔValue K - NK*	*ΔSurplus K - NK*
		Log land values	*Log heights*	***KIP***	***Non-KIP***		
(1)	Q1 (highest)	-0.30*** (0.11)	-0.25 (0.16)	$1,873	$3,098	-$1,225	-$2,369
(2)	Q2	-0.28*** (0.08)	-0.01 (0.05)	$1,112	$1,716	-$603	-$1,044
(3)	Q3	-0.14** (0.07)	-0.15** (0.07)	$972	$1,317	-$345	-$382
(4)	Q4	-0.02 (0.04)	0.06 (0.04)	$717	$738	-$22	$398
(5)	Q5 (lowest)	0.10** (0.05)	-0.03 (0.08)	$489	$478	$11	$347
(6)	Overall			$1,113	$1,626	-$513	-$781

= land value x land area + constr. cost x volume

= land value x land area + constr. cost x building height x horiz. coverage

(35% KIP, 18% non KIP)

Figures are 2015 $ per squared meter. Avg nKIP value psqm: $1100.

Surplus calculations

- Real estate value $_{KIP}$ – Real estate value $_{non\text{-}KIP}$

	(1)	(2)	(3)	(4)	(5)	(6)
	KIP treatment effects		*Value of built-up stock*		*ΔValue* $_{K-NK}$	*ΔSurplus* $_{K-NK}$
	Log land values	*Log heights*	*KIP*	*Non-KIP*		
(1) Q1 (highest)	-0.30*** (0.11)	-0.25 (0.16)	$1,873	$3,098	-$1,225	-$2,369
(2) Q2	-0.28*** (0.08)	-0.01 (0.05)	$1,112	$1,716	-$603	-$1,044
(3) Q3	-0.14** (0.07)	-0.15** (0.07)	$972	$1,317	-$345	-$382
(4) Q4	-0.02 (0.04)	0.06 (0.04)	$717	$738	-$22	$398
(5) Q5 (lowest)	0.10** (0.05)	-0.03 (0.08)	$489	$478	$11	$347
(6) Overall			$1,113	$1,626	-$513	-$781

Trade-off: formal areas are more valuable and taller, but have lower horizontal coverage (35% in KIP, 18% in non-KIP)

Surplus calculations

From value to surplus

	(1)	(2)	(3)	(4)	(5)	(6)
	KIP treatment effects		*Value of built-up stock*		*ΔValue* $_{K-NK}$	*ΔSurplus* $_{K-NK}$
	Log land values	*Log heights*	*KIP*	*Non-KIP*		
(1) Q1 (highest)	-0.30*** (0.11)	-0.25 (0.16)	$1,873	$3,098	-$1,225	-$2,369
(2) Q2	-0.28*** (0.08)	-0.01 (0.05)	$1,112	$1,716	-$603	-$1,044
(3) Q3	-0.14** (0.07)	-0.15** (0.07)	$972	$1,317	-$345	-$382
(4) Q4	-0.02 (0.04)	0.06 (0.04)	$717	$738	-$22	$398
(5) Q5 (lowest)	0.10** (0.05)	-0.03 (0.08)	$489	$478	$11	$347
(6) Overall			$1,113	$1,626	-$513	-$781

- Functional form assumptions:
  - linear demand approximation (validated)
  - Cobb-Douglas supply
- Literature elasticity estimates for K and nKIP (informal more inelastic)

Where is KIP inefficient?

Heterogeneity and concentrated losses

	(1)	(2)	(3)	(4)	(5)	(6)
	KIP treatment effects		*Value of built-up stock*		*ΔValue* $_{K-NK}$	*ΔSurplus* $_{K-NK}$
	Log land values	*Log heights*	*KIP*	*Non-KIP*		
(1) Q1 (highest)	-0.30*** (0.11)	-0.25 (0.16)	$1,873	$3,098	-$1,225	-$2,369
(2) Q2	-0.28*** (0.08)	-0.01 (0.05)	$1,112	$1,716	-$603	-$1,044
(3) Q3	-0.14** (0.07)	-0.15** (0.07)	$972	$1,317	-$345	-$382
(4) Q4	-0.02 (0.04)	0.06 (0.04)	$717	$738	-$22	$398
(5) Q5 (lowest)	0.10** (0.05)	-0.03 (0.08)	$489	$478	$11	$347
(6) Overall			$1,113	$1,626	-$513	-$781

- Q1, Q2: 24% of KIP coverage, 90% of losses (26 times rental value)
- Q3, Q4, Q5: KIP delivers sizable surplus to 3 million people

Gains and losses from formalizing slums

KIP vs non KIP as a lower bound for slum vs formal

Case studies from recent kampung redevelopments:

	(1)	(2)	(3)	(4)	(5)	(6)
	KIP treatment effects		***Value of built-up stock***		$\Delta Value_{K-NK}$	$\Delta Surplus_{K-NK}$
	Log land values	***Log heights***	***KIP***	***Non-KIP***		
(1) Kalijodo, (Q2)	-0.28***	-0.01	$1,224	$1,894	-$670	-$910
	(0.08)	(0.05)				
(2) Kali Pessangrahan, (Q3)	-0.14**	-0.15**	$838	$1,171	-$332	-$307
	(0.07)	(0.07)				
(3) Bukit Duri, (Q5)	0.10**	-0.03	$764	$727	$38	$572
	(0.05)	(0.08)				

Gains and losses from formalizing slums

Equity considerations: sharing the gains is challenging

Case studies from recent kampung redevelopments:

		(6) $\Delta Surplus_{K-NK}$
(1) Kalijodo, (Q2)	Residents offered apartments 24 km away (54% surplus loss)	-$910
(2) Kali Pessangrahan, (Q3)	Land sale negotiations stalling since 2015, despite titles	-$307
(3) Bukit Duri, (Q5)	Residents not paid, despite successful class action	$572

Key policy lessons:

- gains and losses from formalizing slums are heterogeneous across locations
- some redevelopments that deliver higher market values are not socially efficient
- difficult to share surplus with residents under current land mkt institutions

Outline

- Introduction
- Background
- Data
- Empirical strategy and results
- Surplus calculations
- **Conclusions**

Conclusion
-Slum upgrading in city growing out of informality

- Novel causal estimates of the long-term impacts of a large-scale slum upgrading program using granular data
  - 15% lower land values, half as many tall buildings
  - Heterogeneity across neighborhoods at different stages
  - Delayed formalization in KIP, greater parcel and population density
- Policy lessons:
  - Where / when to do slum upgrading
  - Opportunity costs concentrated in high market potential areas
- Ongoing work:
  - Mapping property titles
  - Methodology: using photos for poverty mapping

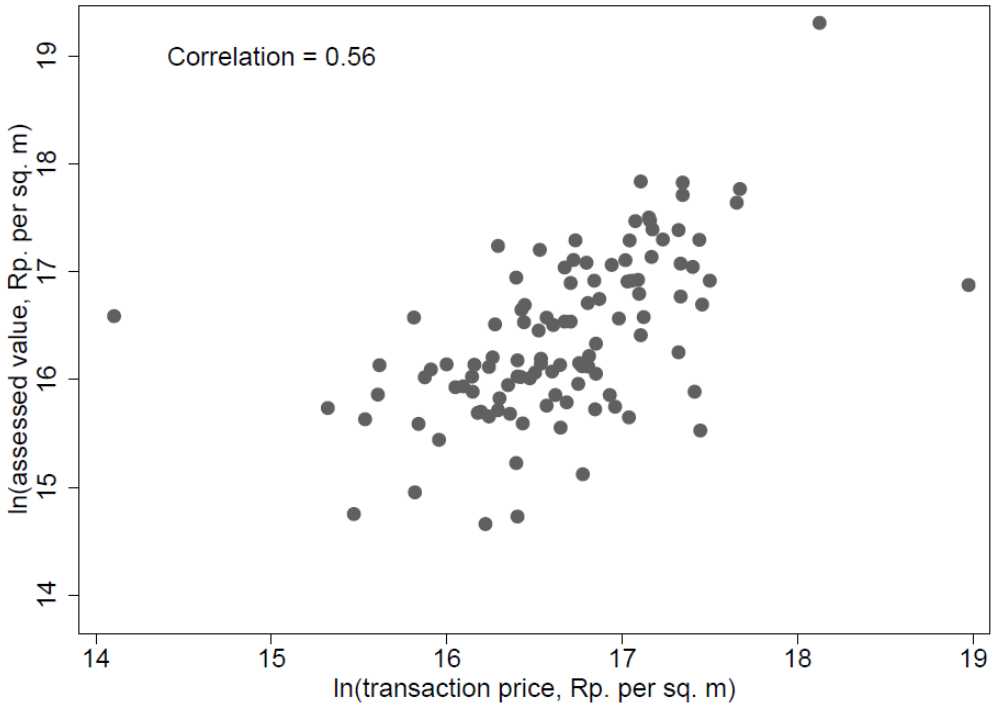
Appendix

Assessed land values, 2015

- Correlation with property transaction prices

Validation check:

compare with 4000 manually geo-referenced property transaction prices from Brickz website



Summary statistics

Variable name	N	Unit	Mean	SD
Panel A: Outcomes				
Assessed Land Values, Thousand Rupiahs per sqm	19848	sub-block	12388	14690
1(Height>3)	19518	pixel	0.17	0.37
1(Rank-Based Informality Index > 1)	7101	pixel	0.47	0.50
Rank-Based Informality Index	7101	pixel	1.11	1.12
Attribute-Based Informality Index	7101	pixel	0.00	0.42
Parcel Count	88861	pixel	15.86	16.19
Retail Density	88861	pixel	0.02	0.10
Office Density	88861	pixel	0.04	0.16
Population density	2533	hamlet	25698	25918
Panel B: Controls				
Distance to Monument, m	88861	pixel	10707.52	4666.42
Distance to Historical Main Road, m	88861	pixel	6983.48	4448.38
Presence of Wells or Pipes within 1000m	88861	pixel	0.11	0.31
Average Distance to Railway Stations, m	88861	pixel	7391.57	4063.21
Average Distance to Tram Stations, m	88861	pixel	8334.91	4173.37
Distance to Tanjung Priok Harbor, m	88861	pixel	15920.54	6200.10
Distance to Old Batavia, m	88861	pixel	12560.57	5879.22
Distance to Schouwburg Weltevreden Concert Hall, m	88861	pixel	11113.72	4774.49
Distance to Hotel Des Indes, m	88861	pixel	11283.16	5105.56
Distance to Bioscoop Metropool, m	88861	pixel	10096.05	4108.94
Distance to Akademi Nasional, m	88861	pixel	11528.63	5708.23
Distance to Ragunan Zoo, m	88861	pixel	13580.41	6654.04
Elevation, m	88861	pixel	21.91	14.78
Slope, Degrees	88861	pixel	4.86	3.38
Average Distance to 1959 Waterways, m	88861	pixel	2754.79	1212.98
Flow Accumulation	88861	pixel	2.91	7.24
Distance to Coast, m	88861	pixel	11501.98	6926.17
Distance to Surface Water Occurrence, m	88861	pixel	2427.23	1509.44

Robustness checks for building heights

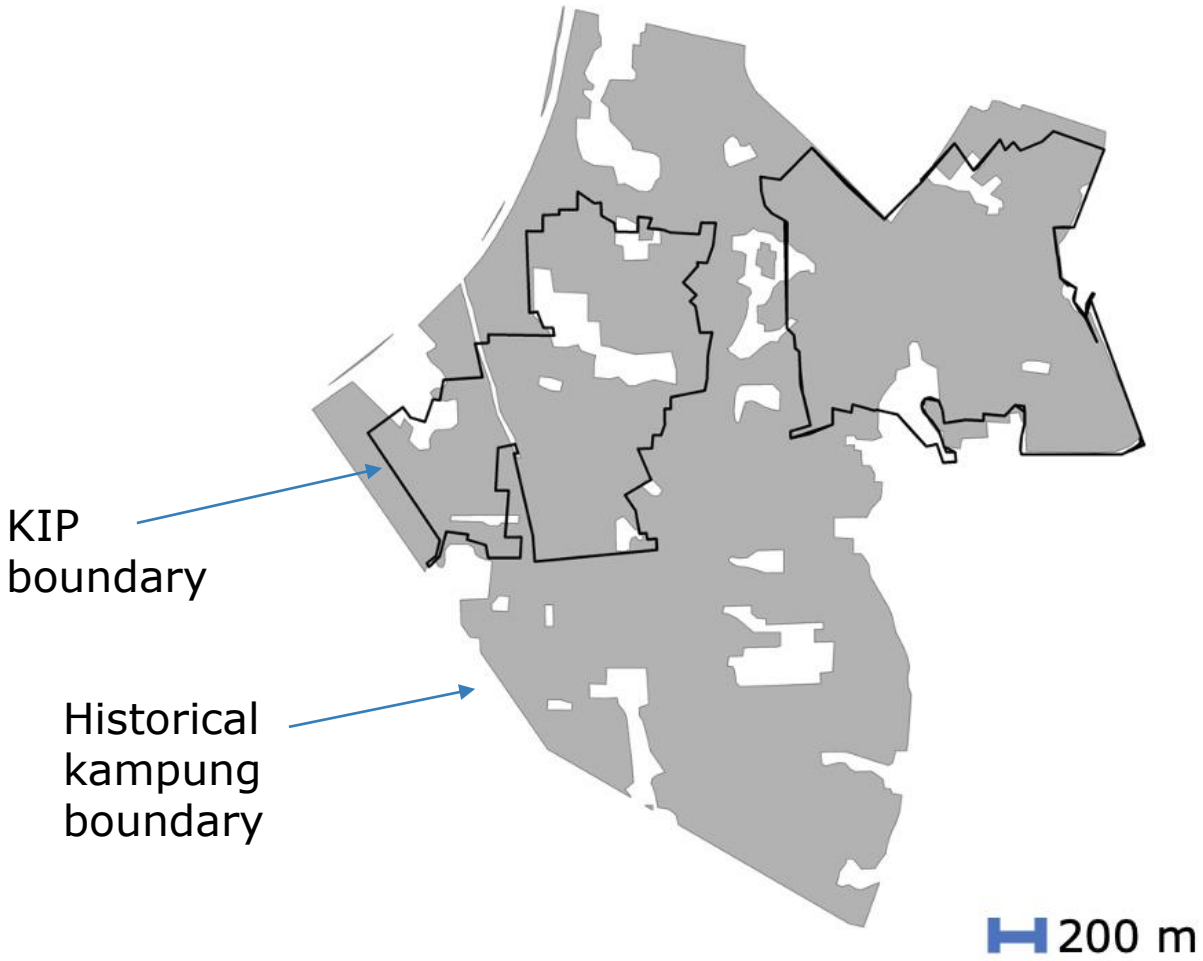
Dependent variable:	Building Heights		Log(height)	
Sample:	Historical kampung	BDD 200m	Historical kampung	BDD 200m
	(1)	(2)	(3)	(4)
KIP	-1.61***	-1.28**	-0.19***	-0.11**
	(0.37)	(0.58)	(0.04)	(0.05)
N	5277	1036	5061	1008
R-Squared	0.32	0.54	0.37	0.56
Distance	Y	Y	Y	Y
Topography	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y
Distance to KIP boundary	N	Y	N	Y
Exclude photos outside pixel	N	N	N	N
Replace photos outside pixel	N	N	N	N
Geography FE	Locality	KIP boundary	Locality	KIP boundary

Selection for building heights

Dependent variable:	1(Height>3)					
Sample:	Historical kampung	BDD 200m	Historical kampung	BDD 200m	Historical kampung	BDD 200m
	(1)	(2)	(3)	(4)	(5)	(6)
KIP	-0.12***	-0.08**	-0.13***	-0.08	-0.13***	-0.11**
	(0.02)	(0.03)	(0.04)	(0.15)	(0.02)	(0.04)
N	5081	1011	856	180	3617	691
R-Squared	0.29	0.38	0.44	0.63	0.31	0.40
Distance	Y	Y	Y	Y	Y	Y
Topography	Y	Y	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y	Y	Y
Distance to KIP Boundary	N	Y	N	Y	N	Y
Exclude no building pixels	Y	Y	N	N	N	N
Only pixels zoned for services	N	N	Y	Y	N	N
Only pixels near predetermined roads	N	N	N	N	Y	Y
Geography FE	Locality	KIP Boundary	Locality	KIP Boundary	Locality	KIP Boundary

KIP effect vs. generic persistence of slums

Falsification test: placebo boundaries



KIP effect vs. generic persistence of slums

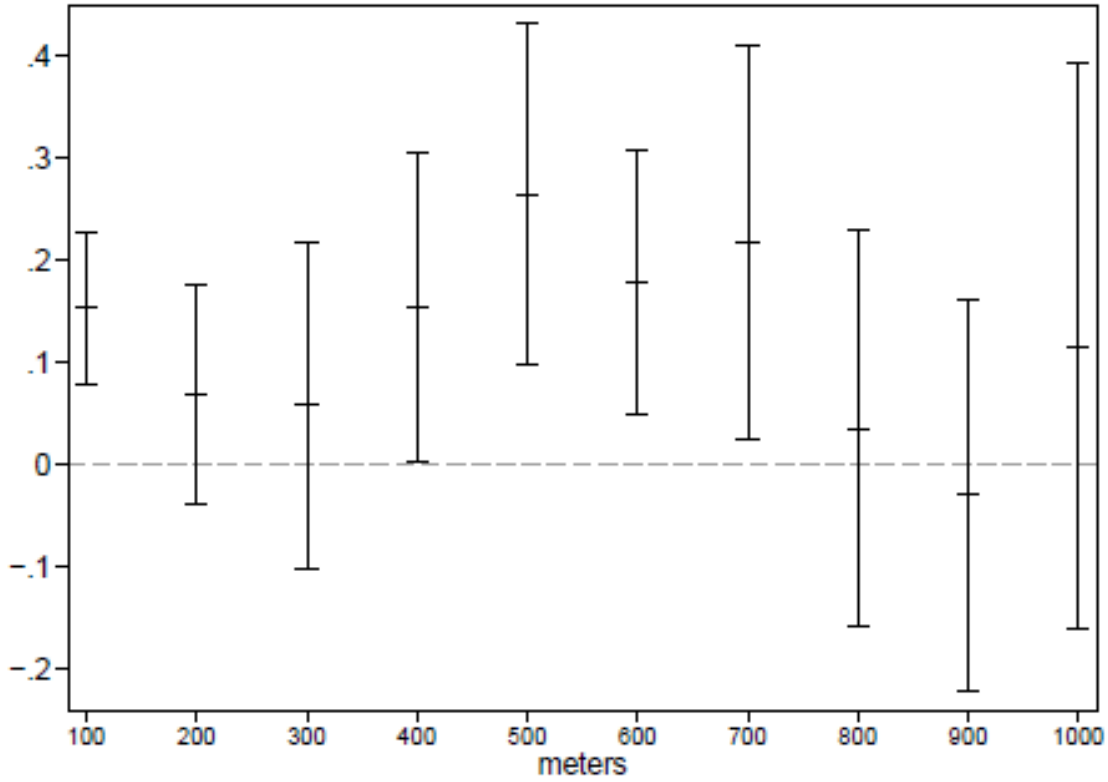
Falsification test: placebo boundaries

Dependent variable:	**Log land values**	
Sample:	Placebo Boundaries	
	BDD 200m	BDD 500m
	(1)	(2)
Ever Kampung	-0.003	0.001
	(0.04)	(0.05)
N	1793	2631
R-Squared	0.50	0.50
Distance	Y	Y
Topography	Y	Y
Landmarks	Y	Y
Distance to boundary	Y	Y
Geography FE	Boundary	Boundary

Non-KIP historical slum boundaries

Spatial spillovers from KIP onto controls (Turner et al., 2014)

- Decay away from KIP boundaries. Not large relative to 12% effect



Each point corresponds to a coefficient and 95 percent confidence interval for coefficients on distance bins, historical sample with locality fixed effects.

KIP-provided amenities: likely depreciated

- Heterogeneous treatment effects by KIP component

Dependent variable:	Log land values	
Sample:	Historical kampung	Full sample
	(1)	(2)
KIP	-0.09*	-0.11***
	(0.05)	(0.04)
Length of Vehicular Roads (in km)	-0.03	-0.02
	(0.03)	(0.02)
Length of Pedestrian Roads (in km)	-0.01	0.01
	(0.02)	(0.02)
Number of Sanitation Facilities	0.005	0.003
	(0.008)	(0.008)
Number of Public Buildings	0.014	-0.000
	(0.03)	(0.02)
KIP X Length of Vehicular Roads	-0.001	0.004
	(0.03)	(0.02)
KIP X Length of Pedestrian Roads	-0.005	-0.005
	(0.02)	(0.02)
KIP X Number of Sanitation Facilities	0.002	-0.004
	(0.007)	(0.008)
KIP X Number of Public Buildings	-0.02	0.03
	(0.03)	(0.02)
N	3144	19848
R-Squared	0.73	0.85
Distance	Y	Y
Topography	Y	Y
Landmarks	Y	Y
Geography FE	Locality	Hamlet

- Intensity of KIP investments within 500 m of each obs.
- No differential effects

Rank-based informality index and attributes

Dependent variable:	Rank-based informality index	
Sample:	Historical kampung	BDD 200m
	(1)	(2)
Panel A: Access		
Road accessible by car (1=no)	0.81***	0.78***
	(0.05)	(0.09)
Paved road (1=no)	0.34	0.39
	(0.26)	(0.85)
Unpaved road (1= yes)	0.09	0.54*
	(0.13)	(0.28)
Damaged road pavement (1=yes)	0.10	-0.17
	(0.08)	(0.16)
Garden (1=no)	0.32***	0.20**
	(0.05)	(0.08)
Panel B: Neighborhood appearance		
Exposed wires (1=yes)	0.41***	0.37***
	(0.04)	(0.07)
Drainage canals (1=no)	0.25***	0.31**
	(0.06)	(0.14)
Trash (1=yes)	0.26***	0.48***
	(0.05)	(0.11)
N	5277	1036
R-Squared	0.60	0.66
Distance	Y	Y
Topography	Y	Y
Landmarks	Y	Y
Distance to KIP boundary	N	Y
Geography FE	Locality	KIP boundary

Dependent variable:	Rank-based informality index	
Sample:	Historical kampung	BDD 200m
	(1)	(2)
Panel C: Permanence of structures		
Unfinished buildings (1=yes)	-0.23***	-0.10
	(0.08)	(0.13)
Permanent wall (1=no)	0.88**	-1.04**
	(0.35)	(0.46)
Unfinished wall (1=yes)	0.50***	0.52***
	(0.03)	(0.09)
Non-permanent wall (1=yes)	0.36***	0.33***
	(0.04)	(0.07)
Damaged wall (1=yes)	0.23***	0.15**
	(0.04)	(0.07)
Permanent fence (1=no)	0.09*	-0.04
	(0.05)	(0.08)
Rust (1=yes)	0.22***	0.20
	(0.05)	(0.14)
N	5277	1036
R-Squared	0.60	0.66
Distance	Y	Y
Topography	Y	Y
Landmarks	Y	Y
Distance to KIP boundary	N	Y
Geography FE	Locality	KIP boundary

Other measures of informality

Dependent variable:	**Rank-based index, pooling**		**Lack of access**		**Poor neighborhood appearance**		**Non-permanent structures**	
Sample:	Historical kampung	BDD 200m	Historical kampung	BDD 200m	Historical kampung	BDD 200m	Historical kampung	BDD 200m
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
KIP	0.29***	0.38***	0.03	0.05	0.05	0.19**	0.06**	-0.00
	(0.05)	(0.12)	(0.03)	(0.05)	(0.04)	(0.08)	(0.02)	(0.04)
N	10554	2072	5277	1036	5277	1036	5277	1036
R-Squared	0.23	0.35	0.13	0.21	0.14	0.21	0.13	0.23
Distance	Y	Y	Y	Y	Y	Y	Y	Y
Topography	Y	Y	Y	Y	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y	Y	Y	Y	Y
Distance to KIP boundary	N	Y	N	Y	N	Y	N	Y
Geography FE	Locality	KIP boundary	Locality	KIP boundary	Locality	KIP boundary	Locality	KIP boundary

Displacement across market potential areas

Non-KIP areas in Q2 are more developed than KIP areas in Q1

Dependent variable:	Log land values	1(Height>3)
	(1)	(2)
KIP X Quintile 1	-0.59***	-0.13***
	(0.11)	(0.03)
KIP X Quintile 2	-0.46***	-0.11***
	(0.06)	(0.02)
KIP X Quintile 3	-0.20***	-0.10***
	(0.05)	(0.02)
KIP X Quintile 4	-0.09*	-0.05***
	(0.05)	(0.01)
KIP X Quintile 5	0.16***	-0.06***
	(0.06)	(0.01)
Quintile 1	1.60***	0.10***
	(0.07)	(0.04)
Quintile 2	1.24***	0.06**
	(0.05)	(0.02)
Quintile 3	0.89***	0.03
	(0.04)	(0.02)
Quintile 4	0.47***	-0.01
	(0.04)	(0.02)
N	19848	19518
R-Squared	0.80	0.22
p-val ($H_0 : \beta_{Quintile1KIP} \geq \beta_{Quintile2Control}$)	0.01	0.01
Distance	Y	Y
Topography	Y	Y
Landmarks	Y	Y
Geography FE	Locality	Locality

Heterogeneous analysis for parcel and pop density

- Is KIP more fragmented just because it is more informal?
- Restrict to places that are informal or periphery

Dependent variable:	**Parcel density**		**Log population density**	
Sample:	Q4 & Q5 sample	Informal sample	Q4 & Q5 sample	Informal sample
	(1)	(2)	(3)	(4)
KIP	9.50***	9.53***	0.46***	0.28***
	(0.90)	(2.63)	(0.10)	(0.07)
N	28175	2307	811	762
R-Squared	0.52	0.74	0.52	0.55
Distance	Y	Y	Y	Y
Topography	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y
Geography FE	Hamlet	Hamlet	Locality	Locality

KIP grid roads

vs.

KIP paved roads

KIP footpaths

Presence of grid roads reverse direct KIP effect

Dependent variable:	**Parcel count**	
Sample:	Historical kampung (5)	Full sample (6)
KIP	8.76***	11.70***
	(1.05)	(0.65)
Grid roads	-7.59**	2.33
	(3.33)	(2.11)
Grid roads x	-8.23**	-9.47***
Vehicular	(4.13)	(2.71)
N	11002	88859
R-Squared	0.51	0.33
Distance	Y	Y
Topography	Y	Y
Landmarks	Y	Y
Geography FE	Locality	Locality

- Regularity and coordination of plots (Libecap and Lueck, 2011; Fuller and Romer, 2014; Baruah, Henderson, and Peng, 2017)

Current public amenities: likely converged

- Negligible differences in access

Dependent variable:	Log distance to				Density	
	School	Hospital	Police	Bus Stop	Retail	Office
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Historical kampung						
KIP	0.05	0.07*	0.00	0.31***	-0.01***	-0.04**
	(0.06)	(0.04)	(0.04)	(0.05)	(0.00)	(0.02)
N	11002	11002	11002	11002	11002	11002
R-Squared	0.25	0.60	0.71	0.56	0.15	0.26
Panel B: Full Sample						
KIP	-0.02	0.03**	0.01	0.08***	-0.01***	-0.02***
	(0.03)	(0.01)	(0.02)	(0.02)	(0.00)	(0.01)
N	88861	88861	88861	88861	88861	88861
R-Squared	0.51	0.78	0.87	0.85	0.25	0.40
Distance	Y	Y	Y	Y	Y	Y
Topography	Y	Y	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y	Y	Y

Historical population density

Not a confounder once we include controls + FEs

Dependent variable:	**Parcel count**	**Parcel count**	**Parcel count**	**Parcel count**
	(1)	(2)	(3)	(4)
KIP	9.00***	14.87***	7.78***	8.97***
	(1.86)	(2.05)	(2.16)	(1.32)
N	4214	3925	4214	3925
R-Squared	0.19	0.19	0.45	0.48
Distance	N	N	Y	Y
Topography	N	N	Y	Y
Landmarks	N	N	Y	Y
Geography FE	District	District	Locality	Locality
Density	Low	High	Low	High

Results pass Oster (2019) selection test

KIP boundaries same as administrative boundaries? No

Sample:	BDD 500m			
Dependent variable:	Log land values		1(Height>3)	
	(1)	(2)	(3)	(4)
KIP	-0.15***	-0.14**	-0.09***	-0.09***
	(0.05)	(0.06)	(0.02)	(0.03)
N	2781	2715	3196	3196
R-Squared	0.81	0.86	0.27	0.34
Distance	Y	Y	Y	Y
Topography	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y
Distance to KIP boundary	Y	Y	Y	Y
KIP Boundary FE	Y	Y	Y	Y
Locality FE	N	Y	N	Y

Drop overlapping boundaries

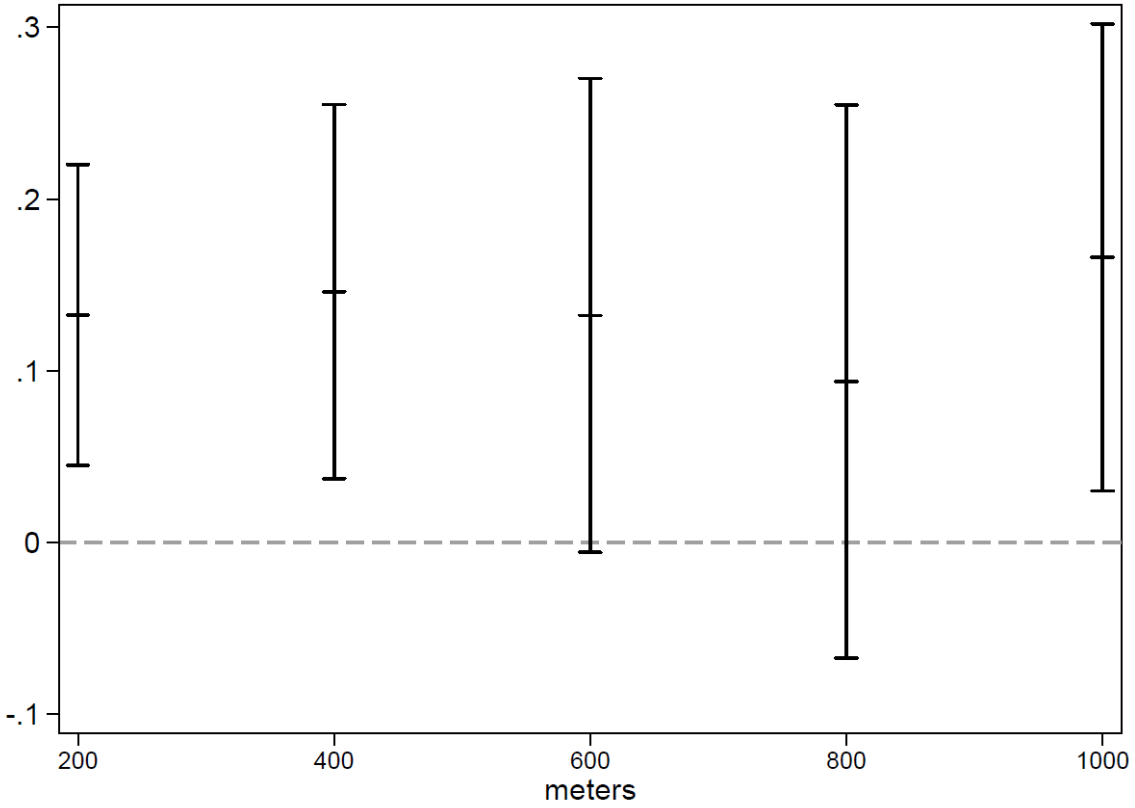
Dependent variable:	**Log land values**			**1(Height>3)**		
	BDD 200m	BDD 200m	BDD 200m	BDD 200m	BDD 200m	BDD 200m
	(1)	(2)	(3)	(4)	(5)	(6)
KIP	-0.17**	-0.12*	-0.13*	-0.08**	-0.12***	-0.11**
	(0.06)	(0.07)	(0.07)	(0.03)	(0.04)	(0.04)
N	1191	888	829	971	610	572
R-Squared	0.80	0.84	0.84	0.39	0.43	0.45
Distance	Y	Y	Y	Y	Y	Y
Topography	Y	Y	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y	Y	Y
Distance to KIP boundary	Y	Y	Y	Y	Y	Y
Drop Boundaries	Railways	Waterways	Both	Railways	Waterways	Both

BDD robustness by distance band

Dependent variable:	Log land values			1(Height>3)		
	BDD 150m	BDD 300m	BDD 500m	BDD 150m	BDD 300m	BDD 500m
	(1)	(2)	(3)	(4)	(5)	(6)
KIP	-0.19***	-0.18***	-0.15***	-0.06	-0.08***	-0.09***
	(0.07)	(0.06)	(0.05)	(0.04)	(0.03)	(0.02)
N	929	1991	2781	602	2295	3196
R-Squared	0.81	0.81	0.81	0.43	0.30	0.27
Distance	Y	Y	Y	Y	Y	Y
Topography	Y	Y	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y	Y	Y
Distance to KIP boundary	Y	Y	Y	Y	Y	Y

Is congestion reducing land values?

- As we move away from high-density non-KIP hamlets, cannot detect large enough decay in land values to explain -12% effect



Effect on land values of being at different distance bins to 45 non-KIP hamlets with population density above median

KIP and population density

- Where do the extra people in KIP come from?
- Not detected in proxies of fertility nor mortality

Dependent variable:	**Household size**	**Number of children deaths per 1000 live births**	**Number of children**
	(1)	(2)	(3)
KIP	0.01	0.13	0.02
	(0.03)	(0.68)	(0.01)
N	2533	2012188	2012188
R-Squared	0.34	0.03	0.24
Distance	Y	Y	Y
Topography	Y	Y	Y
Landmarks	Y	Y	Y
Age FE	Y	Y	Y
Geography FE	Locality	Locality	Locality

Robustness to full sample analysis

- Effects do not cancel out in full sample

Dependent variable:	Log land values	1(Height>3)	Rank index	Attribute index	Unregistered parcels	Parcel count	Log population density
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
KIP	-0.11***	-0.07***	0.28***	0.06***	0.03***	10.13***	0.48***
	(0.03)	(0.02)	(0.05)	(0.02)	(0.01)	(0.55)	(0.06)
N	19848	19518	7101	7101	88861	88861	2533
R-Squared	0.85	0.36	0.47	0.38	0.39	0.52	0.46
Distance	Y	Y	Y	Y	Y	Y	Y
Topography	Y	Y	Y	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y	Y	Y	Y
Geography FE	Hamlet	Hamlet	Hamlet	Hamlet	Hamlet	Hamlet	Locality

Selection into land values dataset

Dependent variable	**1(Has assessed values)**	
Sample	Full sample	Historical kampung
	(1)	(2)
KIP	0.03***	0.03***
	(0.00)	(0.01)
N	97563	11537
R-Squared	0.09	0.08
Distance	Y	Y
Topography	Y	Y
Landmarks	Y	Y
Geography FE	Hamlet	Locality

Robustness to excluding Dutch areas

Dependent variable:	**Log land values**		**1(Height>3)**	
Sample:	Historical kampung	BDD 200m	Historical kampung	BDD 200m
	(1)	(2)	(3)	(4)
KIP	-0.14***	-0.17**	-0.12***	-0.08**
	(0.05)	(0.08)	(0.02)	(0.03)
N	1885	730	5240	1010
R-Squared	0.72	0.83	0.29	0.38
Distance	Y	Y	Y	Y
Topography	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y
Distance to KIP boundary	N	Y	N	Y
Exclude hamlets with Dutch settlements	Y	Y	Y	Y
Geography FE	Locality	KIP Boundary	Locality	KIP Boundary

Standard errors robustness

	P-values of ATE				
Dependent variable	Cluster: locality (1)	Cluster: sub-district (2)	Conley, 700m cutoff (3)	Conley, 900m cutoff (4)	Conley, 1200m cutoff (5)
Log Land Values	0.00	0.00	0.00	0.00	0.00
1(Height>3)	0.00	0.00	0.00	0.00	0.00

Sorting: KIP residents are slightly more educated

- Biased against lower land values
- Universe of current residents age >25, matched to hamlets

Dependent variable:	**Junior Secondary**	**High School**	**College**	**Years of Schooling**
	(1)	(2)	(3)	(4)
KIP	0.01**	0.02**	-0.005	0.07
	(0.01)	(0.01)	(0.01)	(0.07)
N	4924774	4924774	4924774	4924774
R-Squared	0.11	0.10	0.06	0.13
Distance	Y	Y	Y	Y
Topography	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y
Gender FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Geography FE	Locality	Locality	Locality	Locality

Stayers in KIP have slightly more schooling

- Restrict to those born in the district

Dependent variable:	Junior Secondary	High School	College	Years of Schooling
	(1)	(2)	(3)	(4)
KIP	0.01**	0.02**	-0.003	0.08
	(0.01)	(0.01)	(0.01)	(0.07)
N	2136737	2136737	2136737	2136737
R-Squared	0.22	0.20	0.08	0.25
Distance	Y	Y	Y	Y
Topography	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y
Gender FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Born in the same district	Y	Y	Y	Y
Geography FE	Locality	Locality	Locality	Locality

Sorting:
KIP has fewer migrants, and more educated

Dependent variable:	Migrant by birthplace	5-year migrant	Years of schooling	Years of schooling
	(1)	(2)	(3)	(4)
KIP	-0.02***	-0.01***	0.03	0.11
	(0.005)	(0.003)	(0.08)	(0.09)
N	8621849	7861339	2788037	339213
R-Squared	0.14	0.06	0.10	0.11
Distance	Y	Y	Y	Y
Topography	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y
Gender FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Migrant by birthplace	N	N	Y	N
5-year migrant	N	N	N	Y
Geography FE	Locality	Locality	Locality	Locality

Consistent with high share of long-term stayers in 1995 WB report and own 2016 hh survey (>30 years)

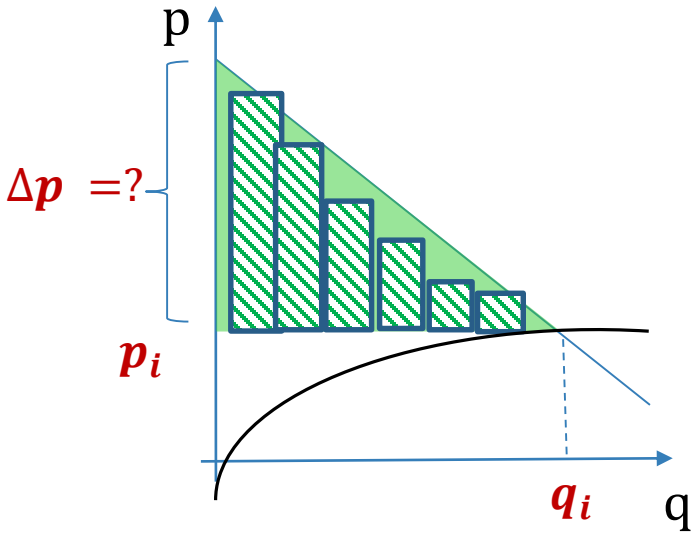
Sorting:
KIP has fewer migrants

Dependent variable:	Migrant by birthplace	5-year migrant	Years of schooling	Years of schooling
	(1)	(2)	(3)	(4)
KIP	-0.02***	-0.01***	0.03	0.11
	(0.005)	(0.003)	(0.08)	(0.09)
N	8621849	7861339	2788037	339213
R-Squared	0.14	0.06	0.10	0.11
Distance	Y	Y	Y	Y
Topography	Y	Y	Y	Y
Landmarks	Y	Y	Y	Y
Gender FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Migrant by birthplace	N	N	Y	N
5-year migrant	N	N	N	Y
Geography FE	Locality	Locality	Locality	Locality

Consistent with high share of long-term stayers in 1995 WB report and own 2016 hh survey (>30 years)

Surplus calculation: functional form assumptions

Consumer surplus: linear approximation (validated by reduced form test)



- CS_i = base x height / 2 = $(\Delta p_i \cdot q_i)/2$.
- Def of demand elasticity in (p_i, q_i) : $\beta = \frac{Dq}{Dp} \cdot \frac{p_i}{q_i}$
- Dp is unknown, $Dq = q_i \Rightarrow Dp = \frac{Dq}{\beta} \cdot \frac{p_i}{q_i} = \frac{q_i}{\beta} \cdot \frac{p_i}{q_i} = \frac{p_i}{\beta}$
- CS = $\frac{q_i \cdot p_i}{2\beta}$

Producer surplus: Cobb Douglas supply (Combes, Duranton, Gobillon, 2021)

$$PS = p^* q^* - \int_0^{q^*} (Aq)^{\frac{1}{\delta}} dq = \frac{p^* q^*}{1+\delta}.$$

Tot surplus = real estate value (p*q*) rescaled by elasticities

Surplus calculation: calculating real estate value

- h = building height in nr of floors — observed in data
- v = land value per squared meter of land — observed in data
- L = area of plot
- l = building footprint area
- c = construction costs per square meter of built-up space — from industry reports
- φ = horizontal coverage = l/ L — observed in data

Total real estate value on plot of area L:

pq = value of land + value of structure = vL+chl

per squared meter of land: pq / L = v + ch φ

Empirical implementation in KIP and non KIP:

$$\widehat{p_K q_K} = \left(\overline{v_K} + c_K \cdot \overline{h_K} \cdot \phi_K\right) \qquad \widehat{p_{NK} q_{NK}} = \left(\overline{v_K} \cdot e^{-\widehat{\beta_v}} + c_{NK} \cdot \overline{h_K} \cdot e^{-\widehat{\beta_h}} \cdot \phi_{NK}\right)$$

Surplus calculation: key parameters

- Construction costs
  - c_{NK}= \$1016 psqm in Q1,Q2; \$738 in Q3; \$422 in Q4, Q5 from industry report
  - c_K = 30% of c_{NK} (implied by difference in supply elasticities) = \$650
- Horizontal coverage:
  - φ_K = 35%, φ_{NK} = 18% from cadastral maps (in line with Henderson, Regan and Venables (2020))
- Supply elasticity:
  - δ_{NK} = 1.4 , δ_K = 1.3 from formal and informal elasticities in Henderson, Regan and Venables (2020)
- Demand elasticity:
  - ε_{NK} = 0.2 from Malpezzi and Mayo (1987)
  - ε_K = 0.16 (applying difference in housing budget share)

Surplus calculation: linear demand validation

- From $p_K q_K = v_K + c_K h_K \varphi_K$: calculate Dp/Dq as a function of Dv/Dq
- Calculate Dv/Dq = (Dv/DKIP)/(Dq/DKIP) at different distances from center / quintiles
- Obtain negative values comparable in magnitude

$\Rightarrow$ Conclude that slope of demand is constant for different q's.