Residential Patterns and Public Goods in Urban Brazil

NBER Summer Institute
Real Estate / Urban Economics Workshop

July 2024

Nina Harari
The Wharton School, University of Pennsylvania

How do developing country cities accommodate the poor?

Rio

São Paulo

How do developing country cities accommodate the poor?

- Urban poor value proximity to high-access / high-amenity neighborhoods (Barnhardt et al., 2017, Rojas Ampuero and Carrea, 2023)

- … but policy makers worry about negative externalities and opportunity costs from poor neighborhoods in high land-value areas (Henderson et al., 2021)

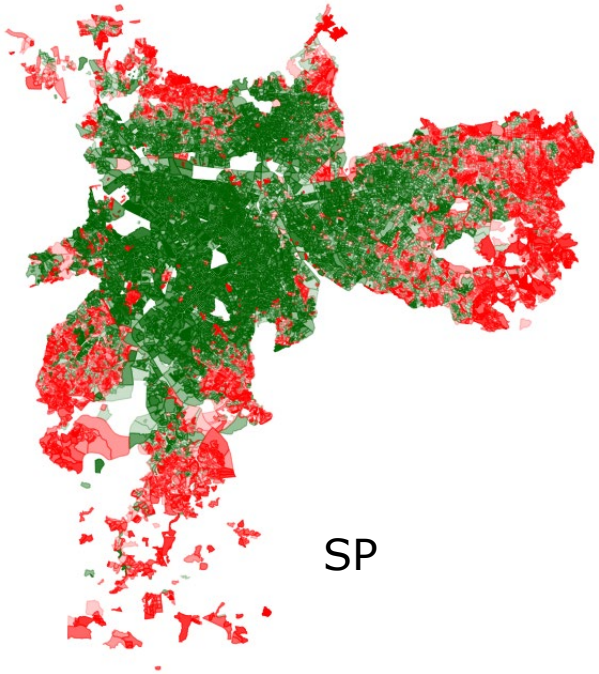
- Tradeoff at the core of the debate around many urban policies related to public goods:
  - Slum upgrading on site (Harari and Wong, 2024)
  - Infrastructure (Tsivanidis, 2023; Khanna et al., 2024)
  - Sanitation (Xu, 2023; Kresh et al., 2020; Feler and Henderson, 2011)

- Little systematic evidence on the geography of income and public goods access within developing country cities

This project:
spatial distribution of rich and poor neighborhoods in Brazil

- 200 mm urban residents + world's 9th most unequal country

Rio



SP

Key contributions:

(i) measurement: "distance segregation"

(ii) identification: residential patterns → public goods

Share of households below min. wage by Census block, 2010

[0, 0.1) [0.1, 0.2) [0.2, 0.3) [0.3, 0.4) [0.4, 0.5)
[0.5, 0.6) [0.6, 0.7) [0.7, 0.8) [0.8, 0.9) [0.9, 1.00)

Outline and preview of findings

- Measurement & stylized facts:
  - New city-level metrics of "distance segregation" by income / race / formality
  - Descriptives
- Causal evidence: residential patterns → public goods
  - IV: spatial distribution of geographic features within cities predicts segregation
  - Key finding: Lower level of public goods access in segregated cities
  - Mechanisms considered: redistributive preferences, engineering costs , externalities
- Other determinants of distance segregation

Related literature

- Residential segregation in developing countries
  - Asher et al. (2024): lower public goods access in minority neighborhoods in India
    - ➢ This paper: consider space + IV

- Local public goods provision in Brazil:
  - Feler and Henderson (2011): strategically withhold public services to discourage in-migration of poor migrants and crowding
  - Xu (2023) on "externalities-correcting public goods" in SP
    - ➢ This paper: IV

Outline and preview of findings

- Measurement & stylized facts
- Causal evidence: residential patterns → public goods
- Historical determinants of distance segregation

Data

- Sample of ~ 600 cities:
  - Municipalities with >50k residents and >50% urban residents
  - Admin boundaries overestimate urban areas: procedure to trim low-density blocks

Data

- Sample of ~ 600 cities:
  - Municipalities with >50k residents and >50% urban residents
  - Admin boundaries overestimate urban areas: procedure to trim low-density blocks
- 2010 Census, block-level (*setor*) data + maps
  - Median block: ~200 households and 0.09 sq km
  - Avg. income, racial composition, slum dummy (*aglomerado subnormal*)
    - illegal occupation of land AND at least one among: narrow and irregular roads, irregular buildings, precarious basic public services.

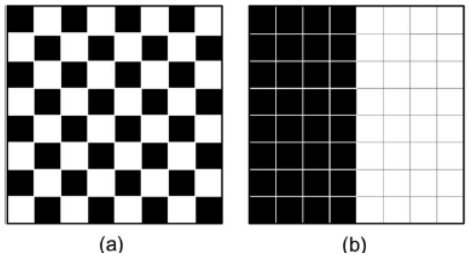
Data

- Sample of ~ 600 cities:
  - Municipalities with >50k residents and >50% urban residents
  - Admin boundaries overestimate urban areas: procedure to trim low-density blocks
- 2010 Census, block-level (*setor*) data + maps
  - Median block: ~200 households and 0.09 sq km
  - Avg. income, racial composition, slum dummy (*aglomerado subnormal*)
- Municipal public goods by block:
  - Positional public goods from Census:
    - % of residents with access to public sewerage / public water
    - Neighborhood public goods index: paved streets, sidewalks, no street garbage, no open sewer, addresses, lighting, curbs, manholes, ramps, greenery
  - Public amenities from Open Street Map: presence of fire stations, police stations, post offices, parks within 3km of block

Measuring "distance segregation"

- Physical distance matters for
  - access to jobs and amenities
  - public goods delivered along spatial networks
  - spatial decay in externalities

- Standard measure of segregation (dissimilarity index) are a-spatial: only internal composition of own neighborhood matters
  - "checkerboard paradox" :

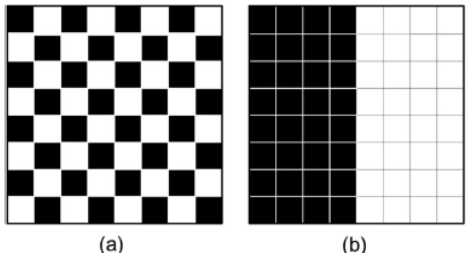


→ same dissimilarity index

Measuring "distance segregation"

- Physical distance matters for
  - access to jobs and amenities
  - public goods delivered along spatial networks
  - spatial decay in externalities

- Standard measure of segregation (dissimilarity index) are a-spatial:
  only internal composition of own neighborhood matters
  - "checkerboard paradox" :



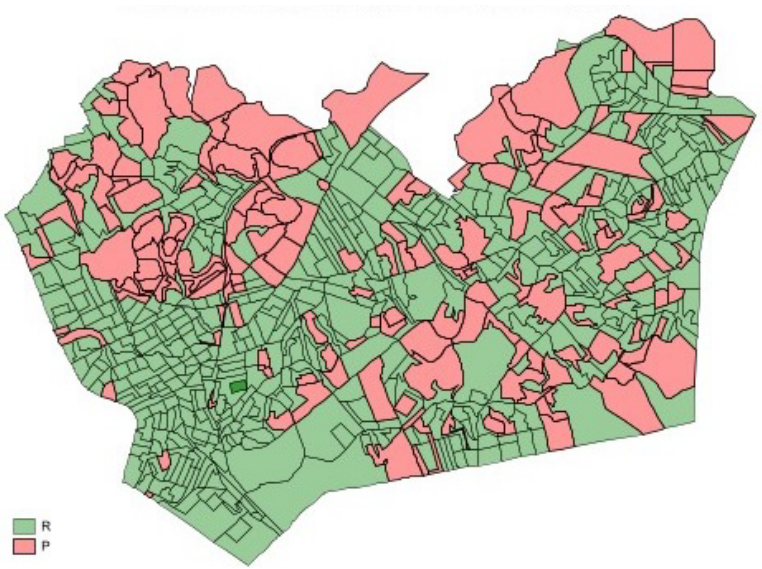
→ same dissimilarity index

- This paper: "distance segregation"
  = average distance between P and R neighborhoods

Examples: segregated vs. integrated cities

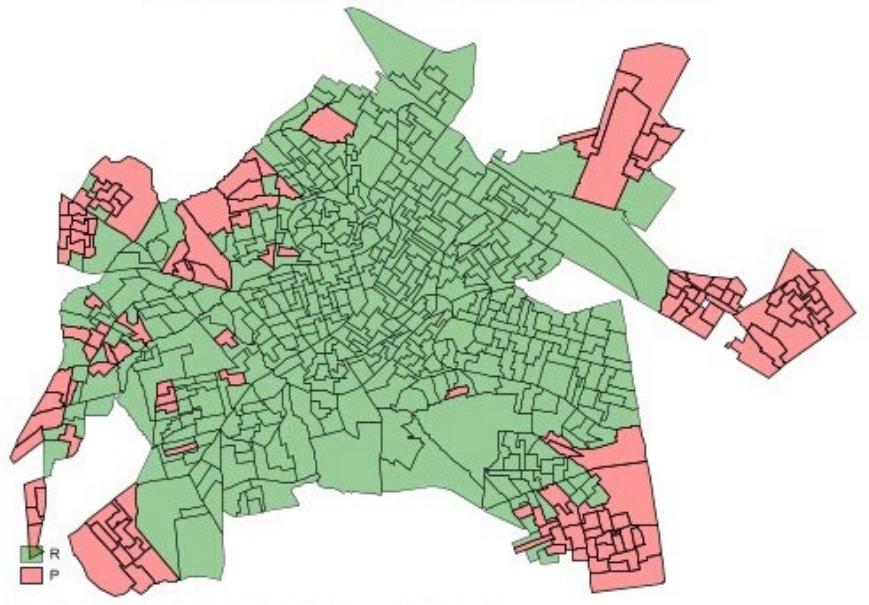
$$D_C^{PR}, \text{norm.} = \frac{\text{avg. distance P to R blocks}}{\text{avg. distance between any block}}$$

Belford Roxo, RJ



$$D_C^{PR}, \text{norm.} = \frac{4{,}860 \text{ m}}{4{,}780 \text{ m}} = 1.02$$

Uberlândia, MG



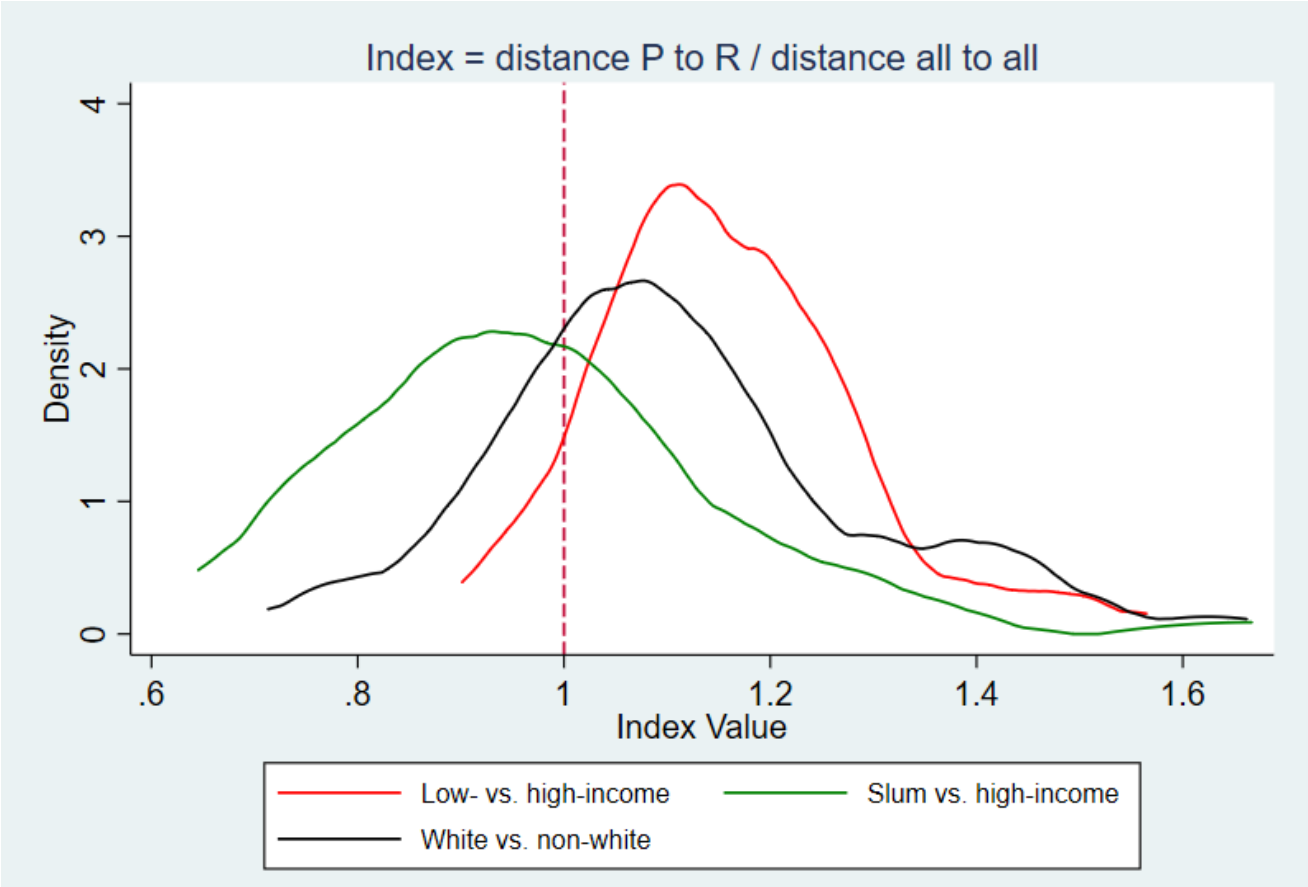
$$D_C^{PR}, \text{norm.} = \frac{7{,}647 \text{ m}}{6{,}189 \text{ m}} = 1.24$$

P = avg. income is in bottom 25% of municipality

Other ex.

Distance segregation by income / race / formality

More segregation by income than race, slums tend to be integrated.



More

Note: high-income = block is in top quartile by avg income; low-income = block in in the bottom quartile. (Non-)white = majority (non-)white residents.

Correlates of distance-segregated neighborhoods

Worse access to public goods

	(1) % Sewerage	(2) % Water	(3) Neighborhood public goods index	(4) OSM amenities
Log distance to CBD	-0.0229***	-0.00514	-0.0380***	-0.0332***
	(0.00586)	(0.00506)	(0.00660)	(0.00733)
Log distance to R blocks	-0.134***	-0.0596***	-0.176***	-0.234***
	(0.0149)	(0.0157)	(0.0143)	(0.0181)
Log distance to R x P	-0.00754***	-0.00221***	-0.0245***	-0.00401***
	(0.000857)	(0.000444)	(0.00153)	(0.000414)
Constant	1.975***	1.591***	1.659***	2.920***
	(0.151)	(0.119)	(0.145)	(0.140)
Observations	161,460	161,460	152,097	161,422
R-squared	0.576	0.440	0.571	0.579
Mean dep. var.	0.718	0.930	0	0.696

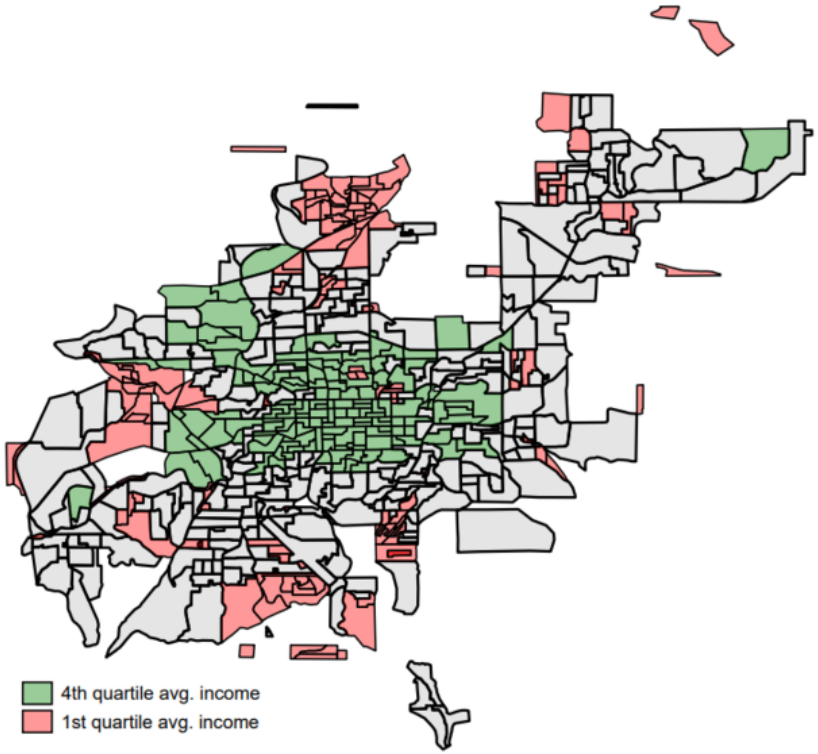
Notes: Each observation is a census block. All specifications include city fixed effects and geography controls (distance to the shoreline, distance to rivers or streams, distance from lakes, the share of a block covered by water bodies, average elevation, average slope). Additional controls include share white residents, share black residents, and a slum dummy. Standard errors clustered at the city level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

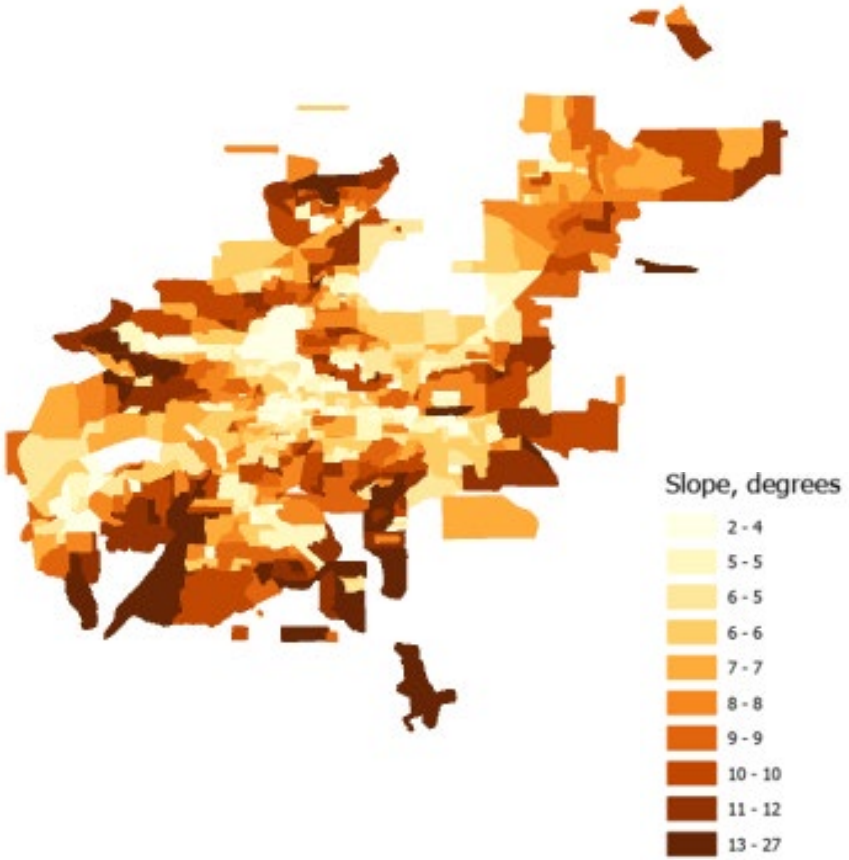
Outline

- Measurement & stylized facts
- Causal evidence: residential patterns → public goods
- Historical determinants of distance segregation

Determinants of residential patterns: natural advantage

Lee and Lin (2018): natural amenities anchor rich neighborhoods in the US





Caxias do Sul, RS

Determinants of residential patterns: natural advantage

- "0-th stage" : within cities, geography predicts location of P and R blocks

	(1)	(2)	(3)	(4)
	Avg. income bottom quartile	Avg. income top quartile	Slum	>50% non-white residents
Slope	0.0240***	-0.0110***	0.0161***	0.0150***
	(0.00274)	(0.00293)	(0.00210)	(0.00293)
Distance to rivers	-0.0299**	0.0525***	-0.0307***	-0.0433***
	(0.0139)	(0.0198)	(0.00586)	(0.0160)
Distance to lakes	0.0335***	-0.0390***	0.00704***	0.0288***
	(0.00474)	(0.00669)	(0.00182)	(0.00697)
Observations	162,464	162,464	162,064	162,064
R-squared	0.043	0.037	0.147	0.500

Notes: Each observation is a census block, from 605 cities. All specifications include city fixed effects. Standard errors clustered at the city level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Determinants of residential patterns: natural advantage

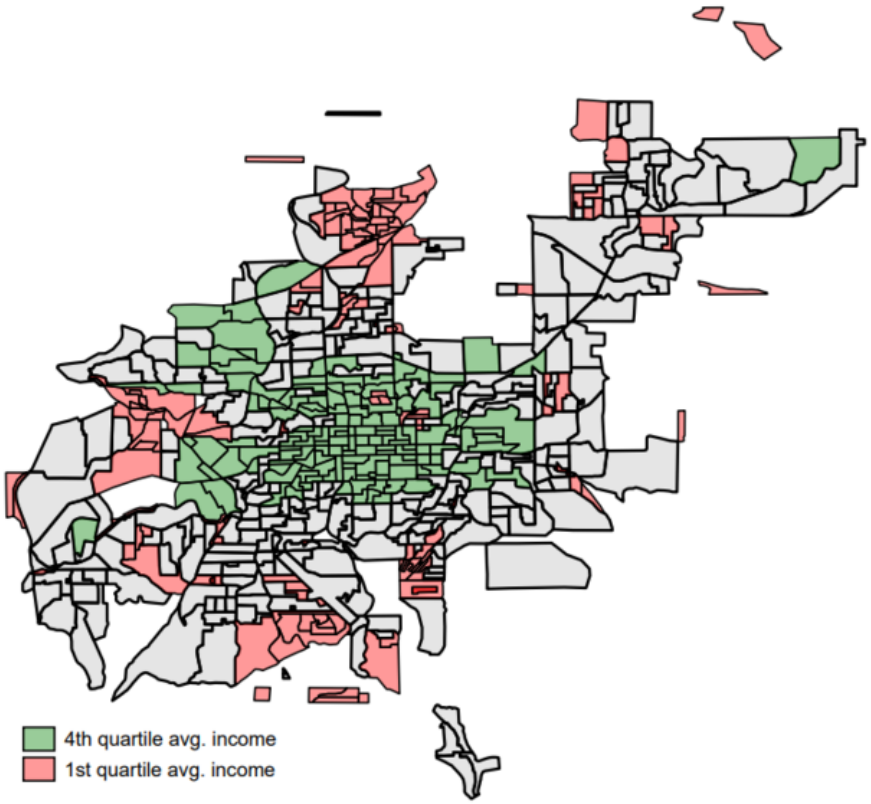
	(1) Avg. income bottom quartile	(2) Avg. income top quartile	(3) >50% non-white residents	(4) Slum
Slope	-0.0349***	0.0442***	-0.0460***	-0.0325***
	(0.0107)	(0.0103)	(0.00962)	(0.00791)
Slope^2	0.00627***	-0.00601***	0.00640***	0.00452***
	(0.00109)	(0.000982)	(0.00112)	(0.000817)
Slope^3	-0.000165***	0.000160***	-0.000164***	-9.95e-05***
	(3.07e-05)	(2.84e-05)	(3.45e-05)	(1.81e-05)
Distance to rivers	-0.0763***	0.0873**	-0.0538*	-0.0612***
	(0.0252)	(0.0353)	(0.0300)	(0.00969)
Distance^2 to rivers	0.0151***	-0.0125**	0.00505	0.00904***
	(0.00503)	(0.00622)	(0.00557)	(0.00179)
Distance to lakes	0.0506***	-0.0633***	0.0473***	0.00991***
	(0.00656)	(0.00995)	(0.0109)	(0.00328)
Distance^2 to lakes	-0.00149***	0.00212***	-0.00163***	-0.000261
	(0.000423)	(0.000688)	(0.000525)	(0.000260)
Distance to shore	0.0176***	-0.0372***	0.0311***	-0.00265
	(0.00611)	(0.00729)	(0.00688)	(0.00262)
Distance^2 to shore	-0.000298***	0.000495***	-0.000460***	-4.13e-06
	(7.13e-05)	(8.80e-05)	(7.69e-05)	(3.07e-05)
Constant	0.107**	0.440***	0.355***	0.152***
	(0.0496)	(0.0409)	(0.0299)	(0.0199)
Observations	162,464	162,464	162,064	162,064
R-squared	0.058	0.063	0.516	0.158
F statistic	1202	377	1282	165

Notes: Each observation is a census block, from 605 cities. All specifications include city fixed effects. Standard errors clustered at the city level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

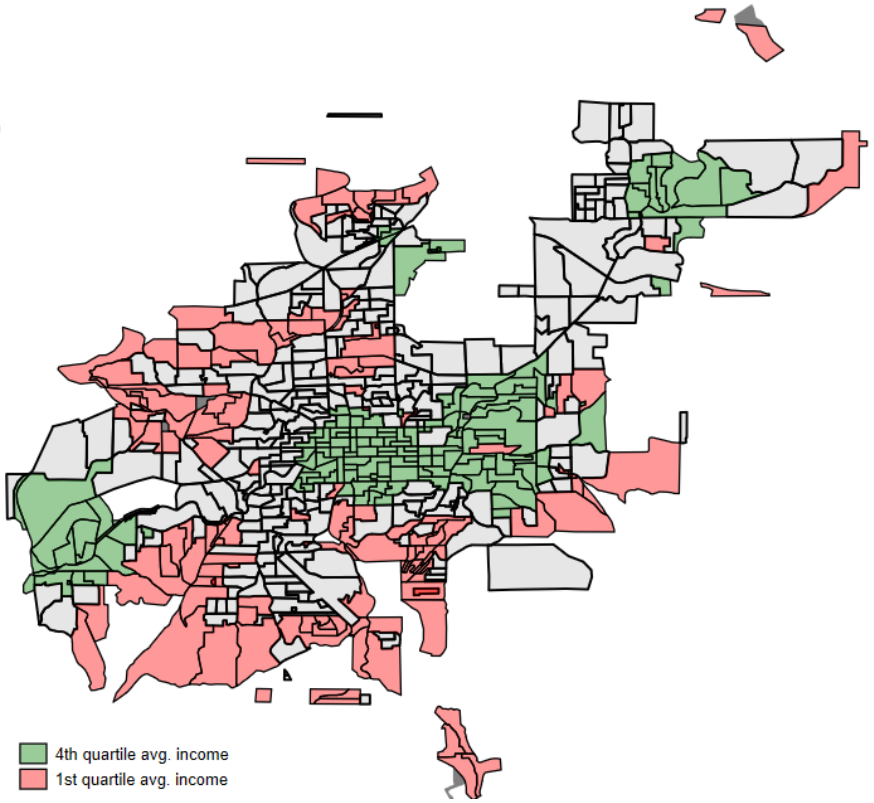
"0-th stage" refined

Predicting distance segregation: from 0 to 1st stage

Actual



Predicted



"Geography segregation" = city-level distance segregation between $\hat{P}$ and $\hat{R}$

Instrumenting for segregation: first stage

Dependent variable: distance segregation, km

	(1)	(2)	(3)
	Income	Race	Slum
Geography distance-segregation, km	0.610***	0.549***	0.394***
	(0.0463)	(0.0717)	(0.0763)
Equivalent area radius, km	0.512***	0.515***	0.677***
	(0.0735)	(0.0903)	(0.0953)
Average elevation, m	0.000241	9.86e-06	-0.000162
	(0.000183)	(0.000299)	(0.000378)
Ruggedness, m	6.34e-05	0.000938	-0.000565
	(0.000771)	(0.000793)	(0.00119)
Average slope, degrees	-0.0257	0.0930	0.0428
	(0.0477)	(0.127)	(0.111)
Water bodies within 30 km, sqkm	-0.000342	-0.001000	0.00137**
	(0.000349)	(0.000653)	(0.000539)
% land available within 30km	-0.651*	0.144	1.078
	(0.372)	(0.608)	(0.692)
Slope adjustment factor for distance	-2.592	-6.371**	2.498
	(2.227)	(3.073)	(5.882)
Distance to state capital, km	-0.000378	-0.000616	-0.000241
	(0.000280)	(0.000518)	(0.000619)
Distance to Atlantic, km	-0.000405	-0.00103***	0.000932
	(0.000255)	(0.000385)	(0.000570)
Observations	597	447	223
R-squared	0.903	0.828	0.863
F statistic	174	59	27

Notes: Each observation is a city. All specifications include as additional controls latitude, longitude, precipitation, sunshine, soil type dummies, % low-fertility soil, landslide risk. Standard errors clustered at the meso-region level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Discussion of instrument

- Geography-driven variation in cities' susceptibility to spatial segregation (similar to Ananat, 2011)
- Variation from the relative positioning of slopes/water within the city(Harari, 2020)
- Key threat to identification: direct effects of geography on public goods
  - Would tend to bias against results
  - Control for city-wide geography – Balance
  - Robustness by sample cuts: exclude mountainous, coastal, etc.
  - Robust to residualizing public goods measures by local geography

Main results

Lower levels of local public goods access in segregated cities

	(1)	(2)	(3)	(4)
Distance-segregation by income, km	% sewerage	% water	Nbhd pub goods index	OSM amenities
IV	-0.0220***	-0.0170***	-0.0505***	-0.0207***
	(0.00832)	(0.00532)	(0.0129)	(0.00427)
OLS	-0.015**	-0.016***	-0.037***	-0.020***
	(0.006)	(0.004)	(0.009)	(0.003)
Observations	597	597	597	597
R-squared	0.582	0.460	0.607	0.182
IV F statistic	174	174	174	174
Mean dep. var.	0.571	0.899	0.0104	0.630

Notes: Each observation is a city. All specifications include geographic controls. Standard errors clustered at the meso-region level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

1 st. dev. distance segregation ~ 1.7 km

Similar results for segregation by race and slum status

Other measures of public goods

Discussion: naïve OLS vs. IV

Which confounders does IV address?

1. Reverse causality:

   Low public goods provision → rich move away from poor ($\beta_{OLS} \downarrow$)

2. Omitted variables:

   Local institutions / state capacity:

   in "integrated" cities government may be worse at enforcing property rights (Henderson et al. 2020) AND at providing public goods ($\beta_{OLS} \uparrow$)

Other robustness

- Specification / functional form:
  - Control for city shape
  - Log distance
  - "Market access" version with exponential distance
  - Results stronger with index weighted by P and R counts

- Local PF:
  - Results stronger in cities with municipal company
  - Control for state FEs

Mechanisms

Why different levels of public goods access in segregated cities?

1) Engineering costs of expanding services network

2) Differences in overall redistributive preferences (Trounstine, 2016)

3) Different spatial targeting of far-away vs. close-by poor

Mechanisms

Why different levels of public goods access in segregated cities?

1) Engineering costs of expanding services network

   - Poor neighborhoods next to rich ones are along the connecting path for network public goods (Troesken, 2002)

2) Differences in overall redistributive preferences (Trounstine, 2016)

3) Different spatial targeting of far-away vs. close-by poor

Mechanisms

Why different levels of public goods access in segregated cities?

1) Engineering costs of expanding services network

   - Poor neighborhoods next to rich ones are along the connecting path for network public goods (Troesken, 2002)

2) Differences in overall redistributive preferences (Trounstine, 2016)

3) Different spatial targeting of far-away vs. close-by poor

   Two competing forces:

   - Correcting externalities:

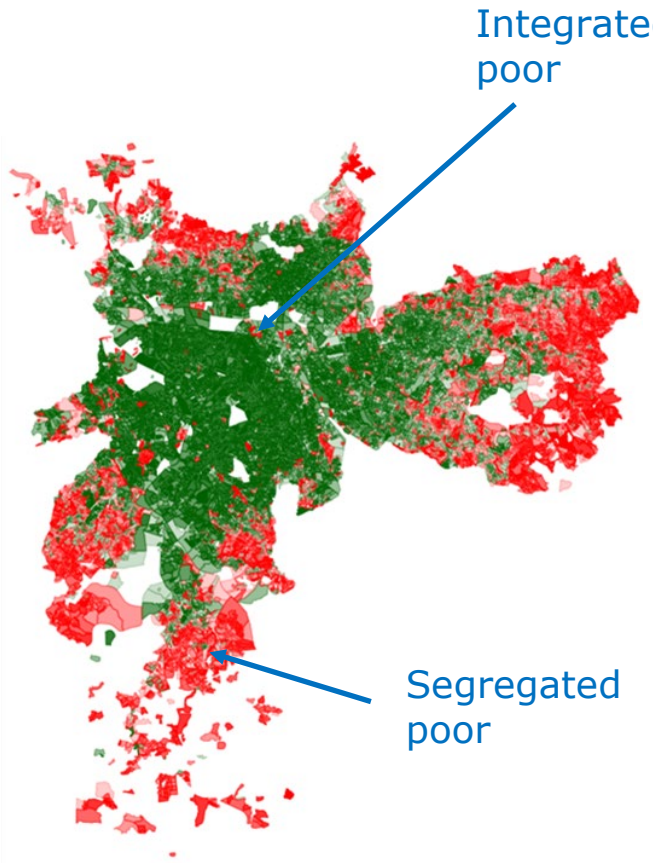
     under-provide to the *far-away* poor: too far to exert negative externalities (Xu, 2023)

   - Deterrence:

     strategically under-provide to the *close-by* poor to discourage in-migration and crowding (Feler and Henderson, 2011)

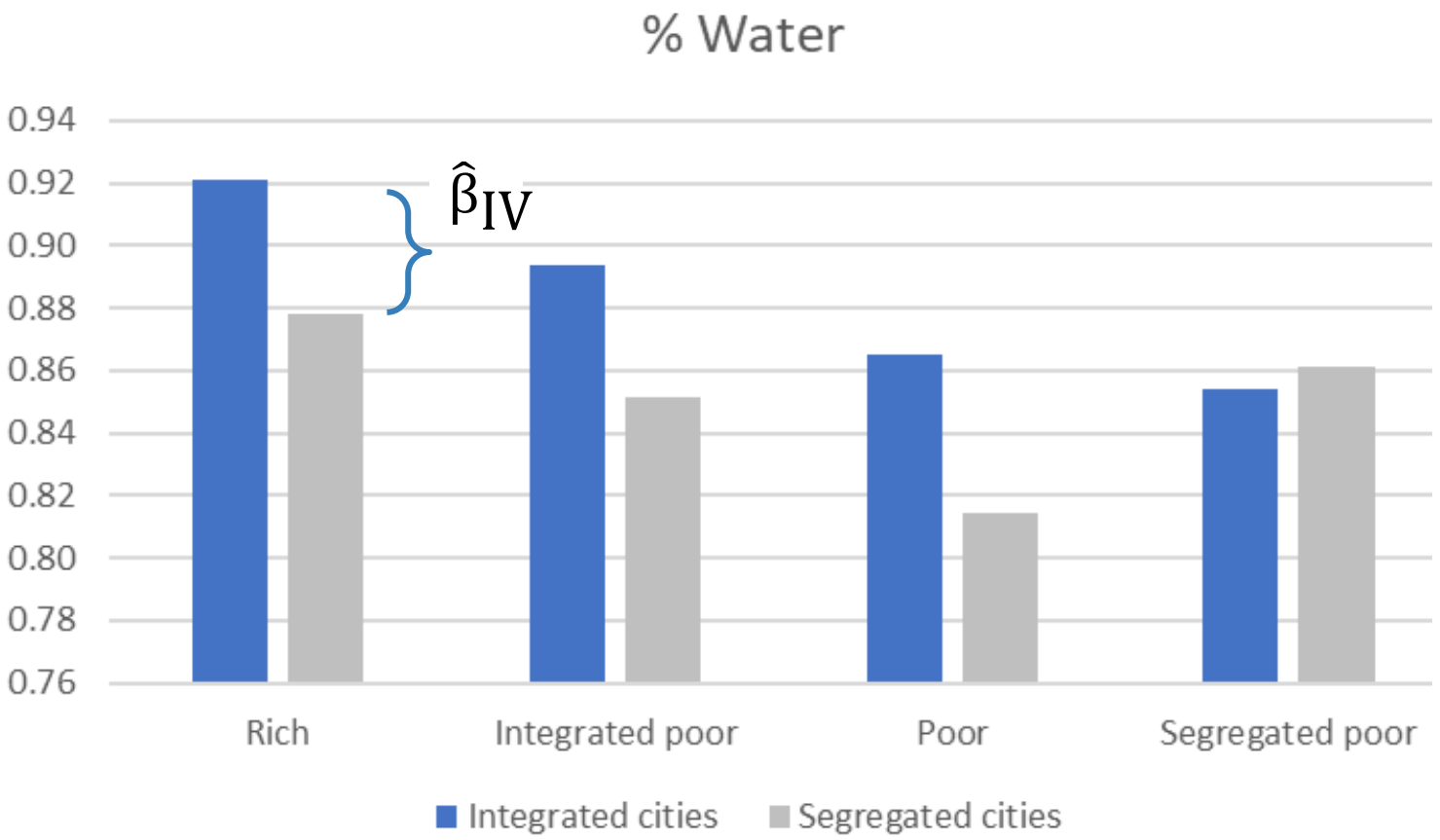
Where are public goods being over/under provided?

Integrated city



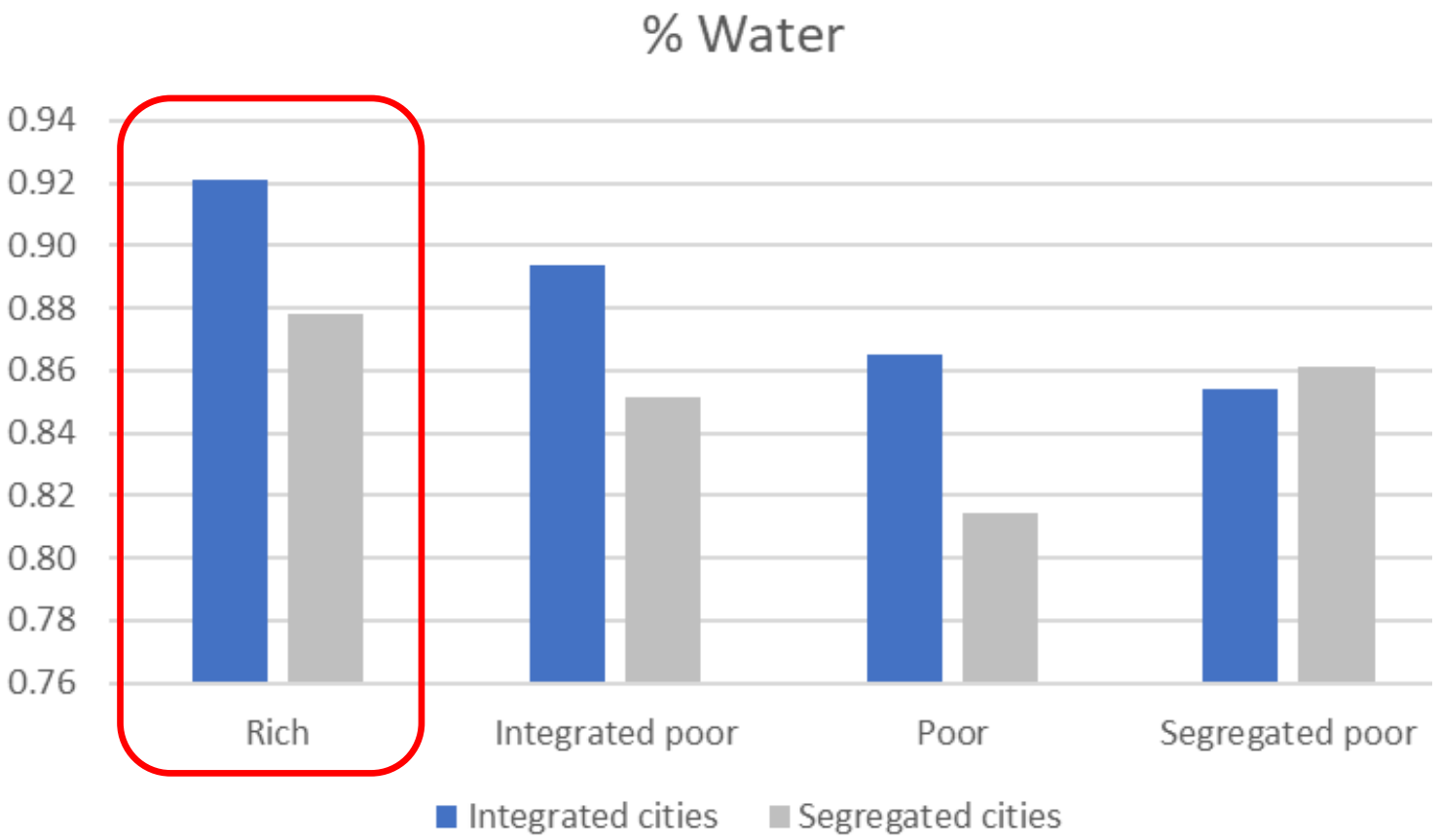
Segregated city

Public goods access across neighborhoods



Regression tables

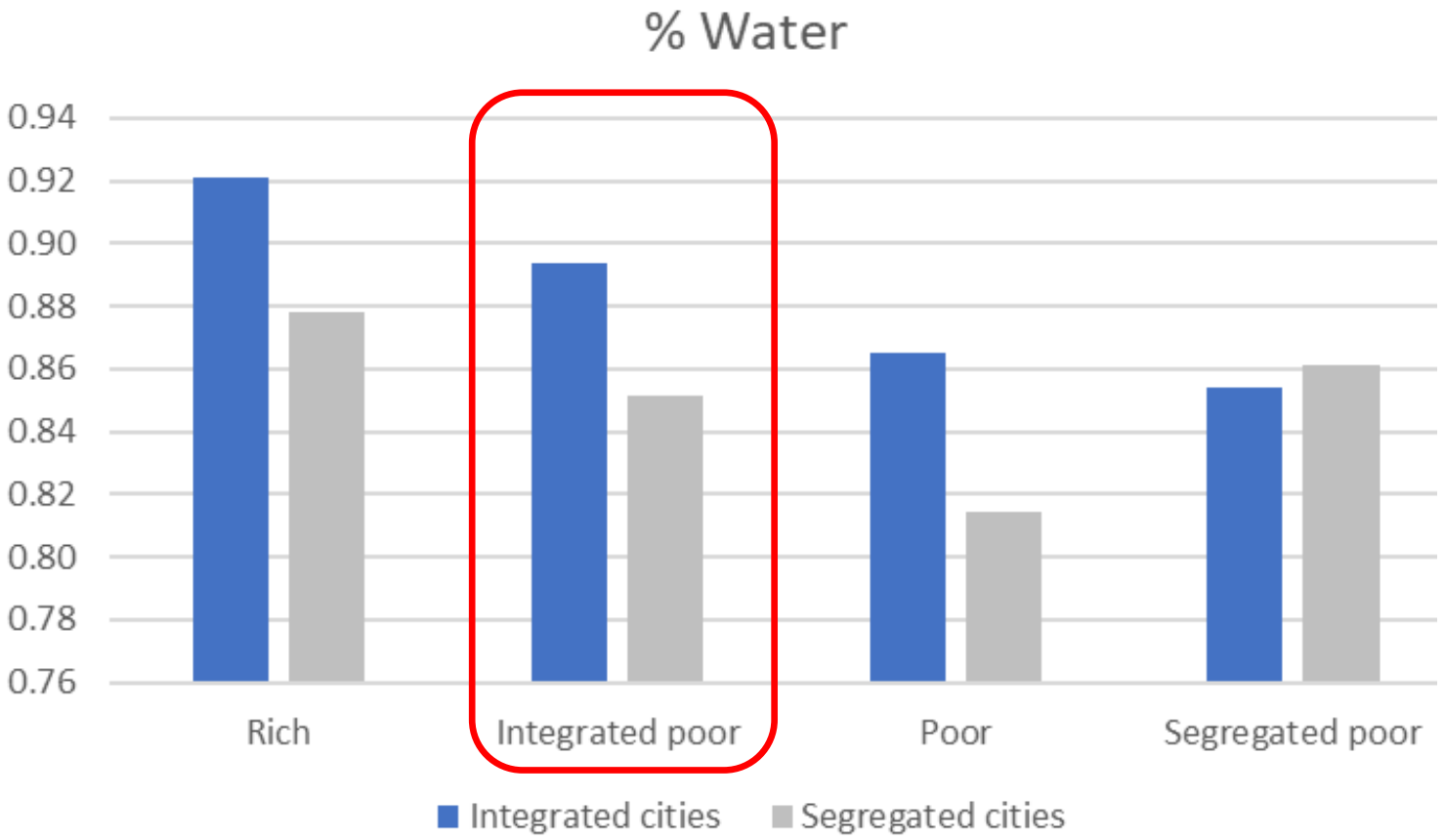
In segregated cities, lower provision even among the rich



Consistent with weaker preferences for redistribution in segregated cities

- Likely a combination of sorting + direct effects (e.g. from lack of exposure)

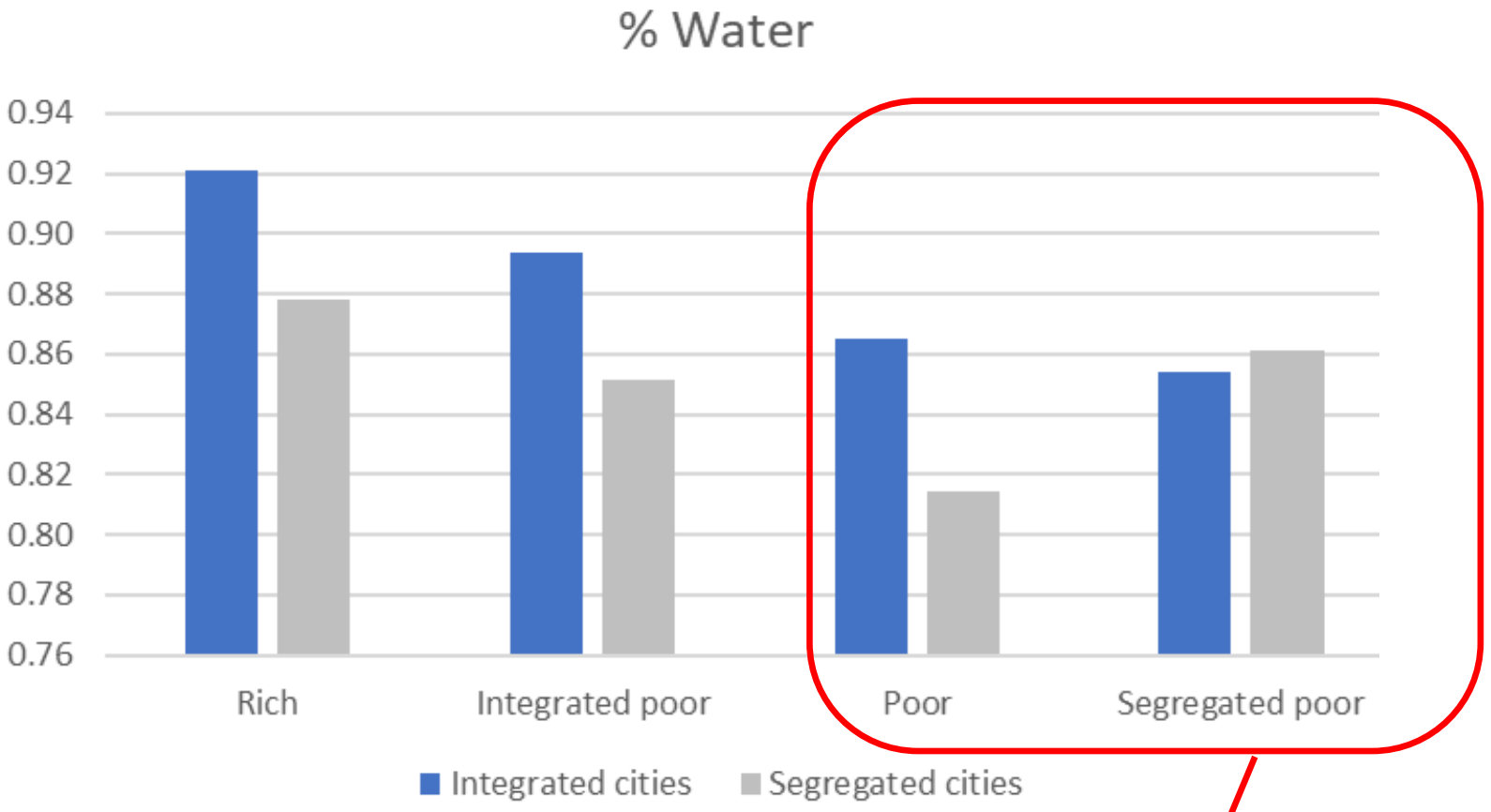
In segregated cities, lower provision even among the close-by poor



Within cities, close-by poor are better provided than far-away poor (consistent with engineering argument)…

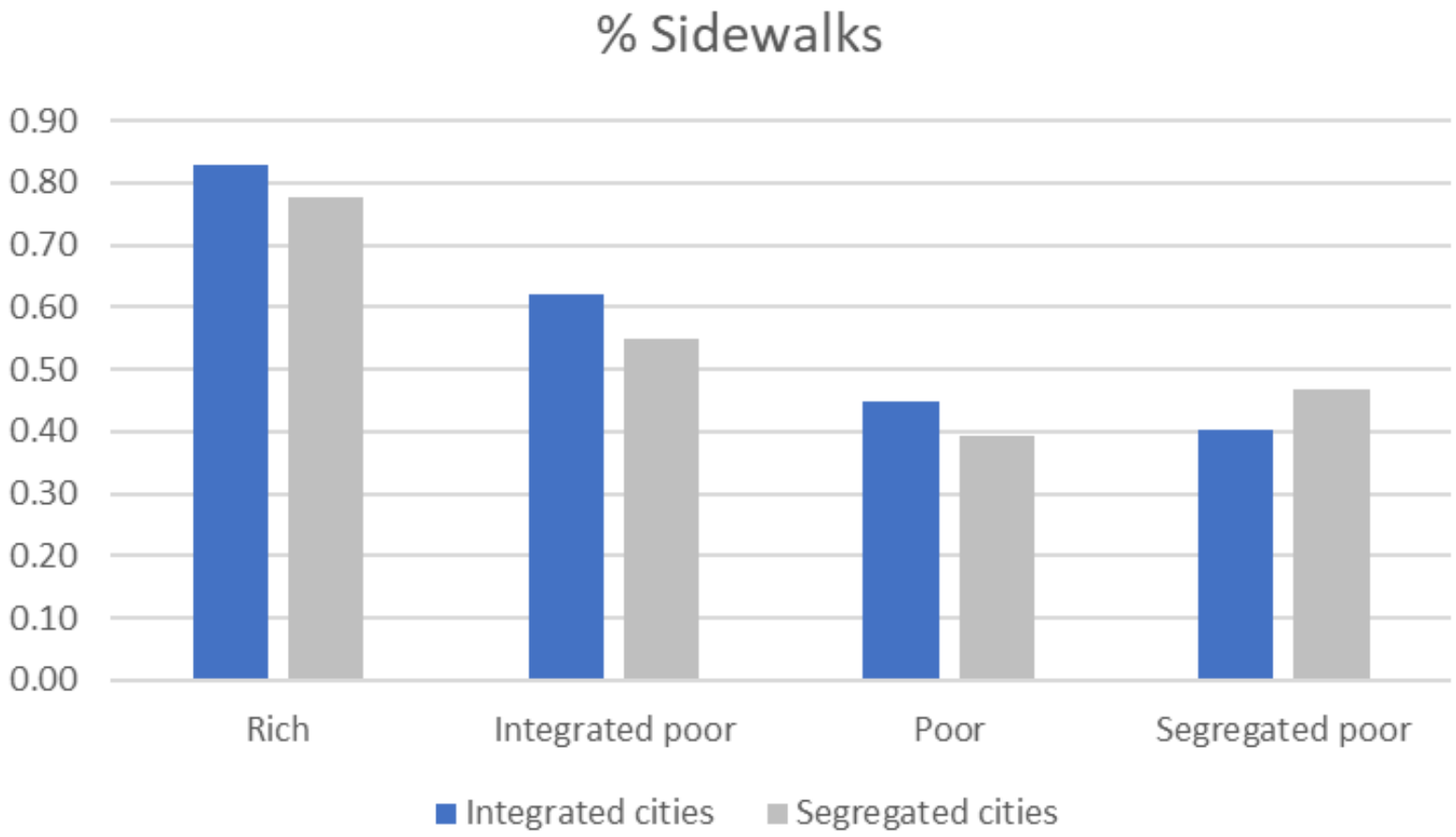
…but engineering argument alone cannot explain differences across cities.

Close-by poor vs. far-away poor: correcting externalities vs. deterrence

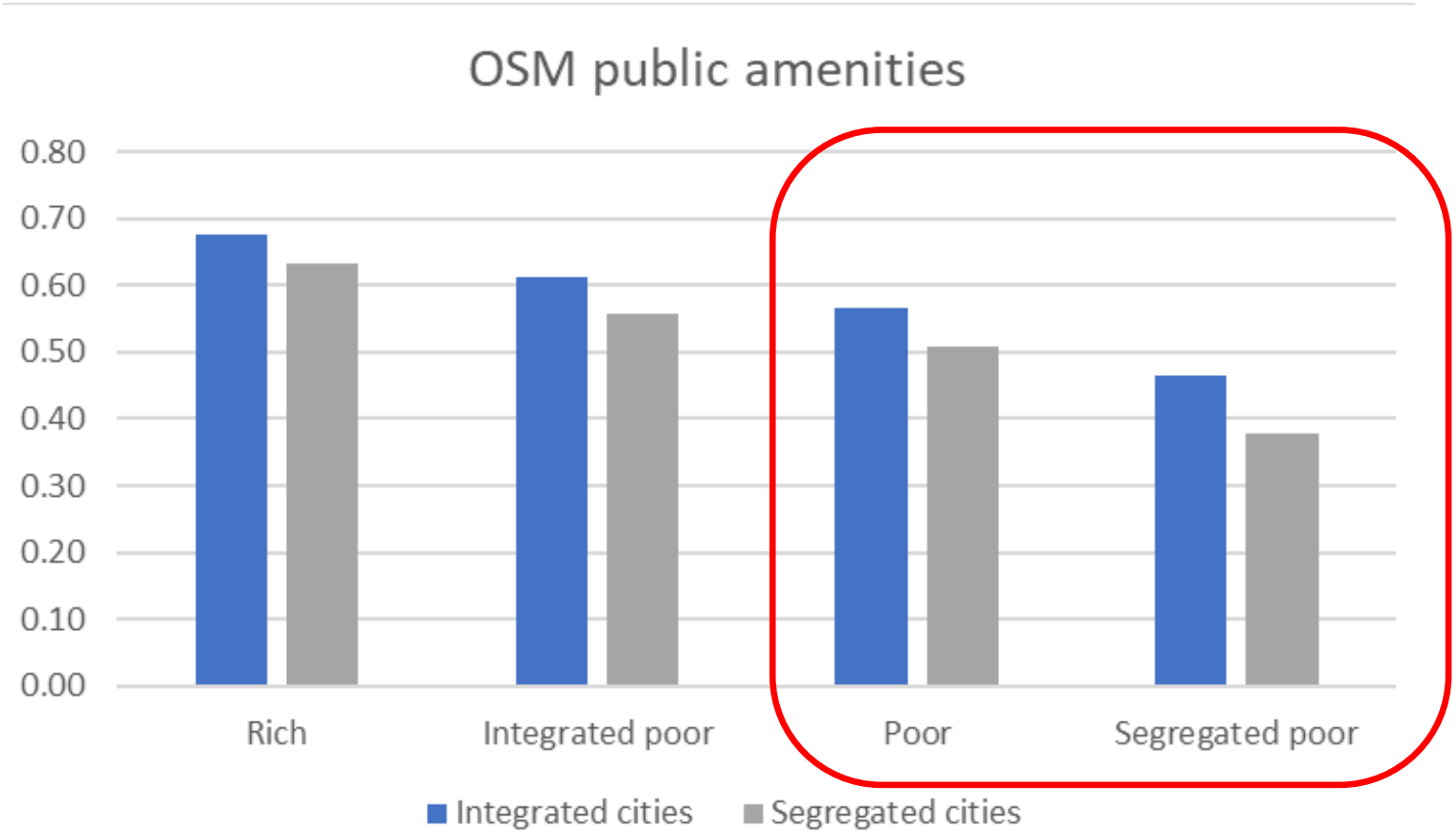


In segregated cities, deterrence mechanism is stronger

Similar patterns for other positional public goods



Deterrence argument less applicable to non-positional public goods



Outline

- Measurement & stylized facts

- Causal evidence: residential patterns → public goods

- Other determinants of distance segregation:
  - Beyond "first nature": externalities, path dependence
  - Historical factors: e.g. disease outbreaks in 19th century prompting slum clearance in the center

Conclusion

- Characterize within-city residential patterns in a developing country context, using new metrics of distance-segregation
- Document inequality in access to public goods within cities

- Identification: instrument for distance segregation based on geography

- Key finding: more segregated cities → lower public goods access
  - Mechanism: redistributive preferences, externalities, deterrence

Conclusion

- Characterize within-city residential patterns in a developing country context, using new metrics of distance-segregation
- Document inequality in access to public goods within cities

- Identification: instrument for distance segregation based on geography

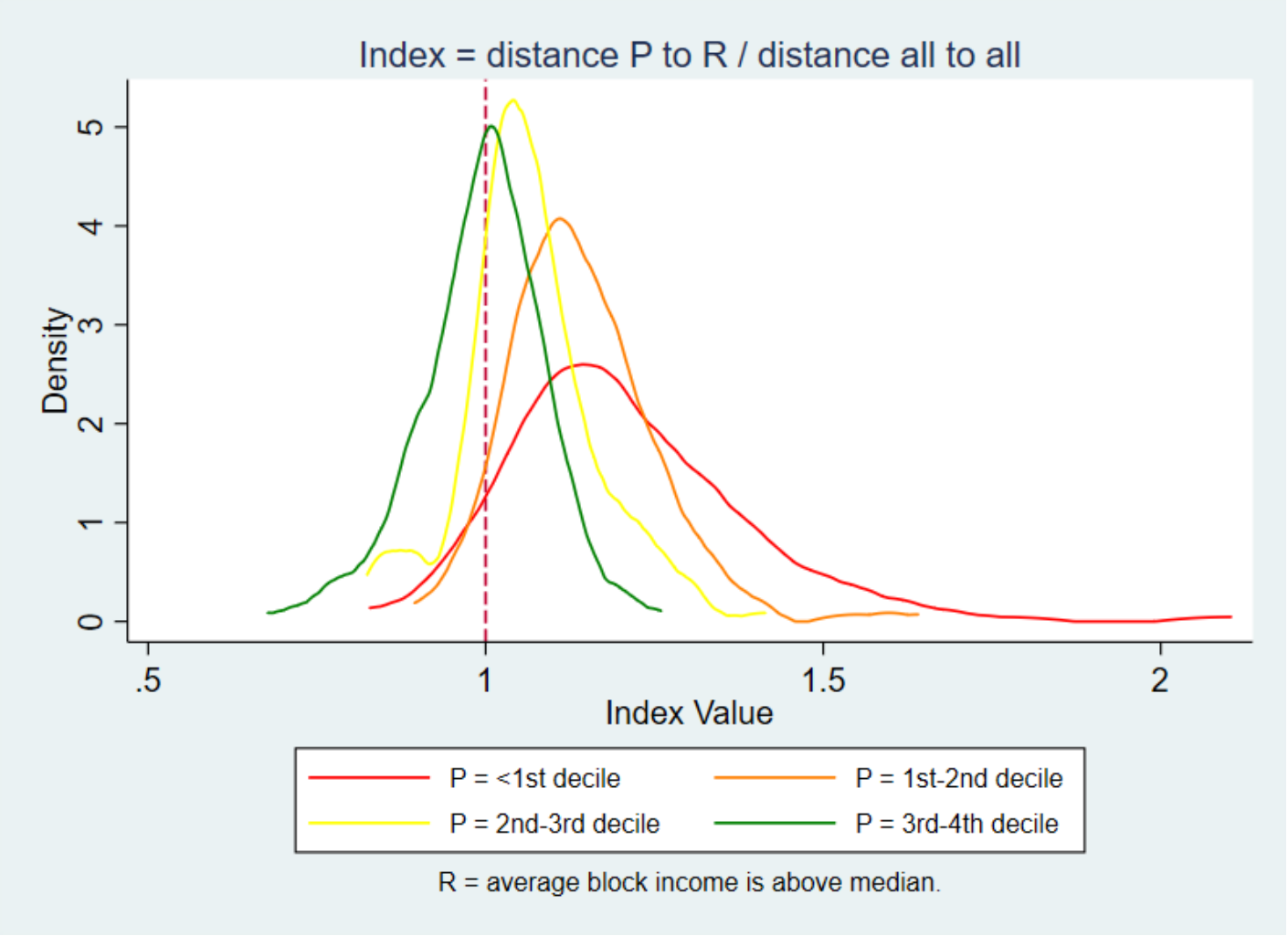
- Key finding: more segregated cities → lower public goods access
  - Mechanism: redistributive preferences, externalities, deterrence

- Inform policies affecting where poor and rich live: affordable housing, slum upgrading, infrastructure…

- Beyond developing countries: insight from adding a spatial angle to the measurement of segregation
  - Ongoing project on racial segregation and inequality in access in US

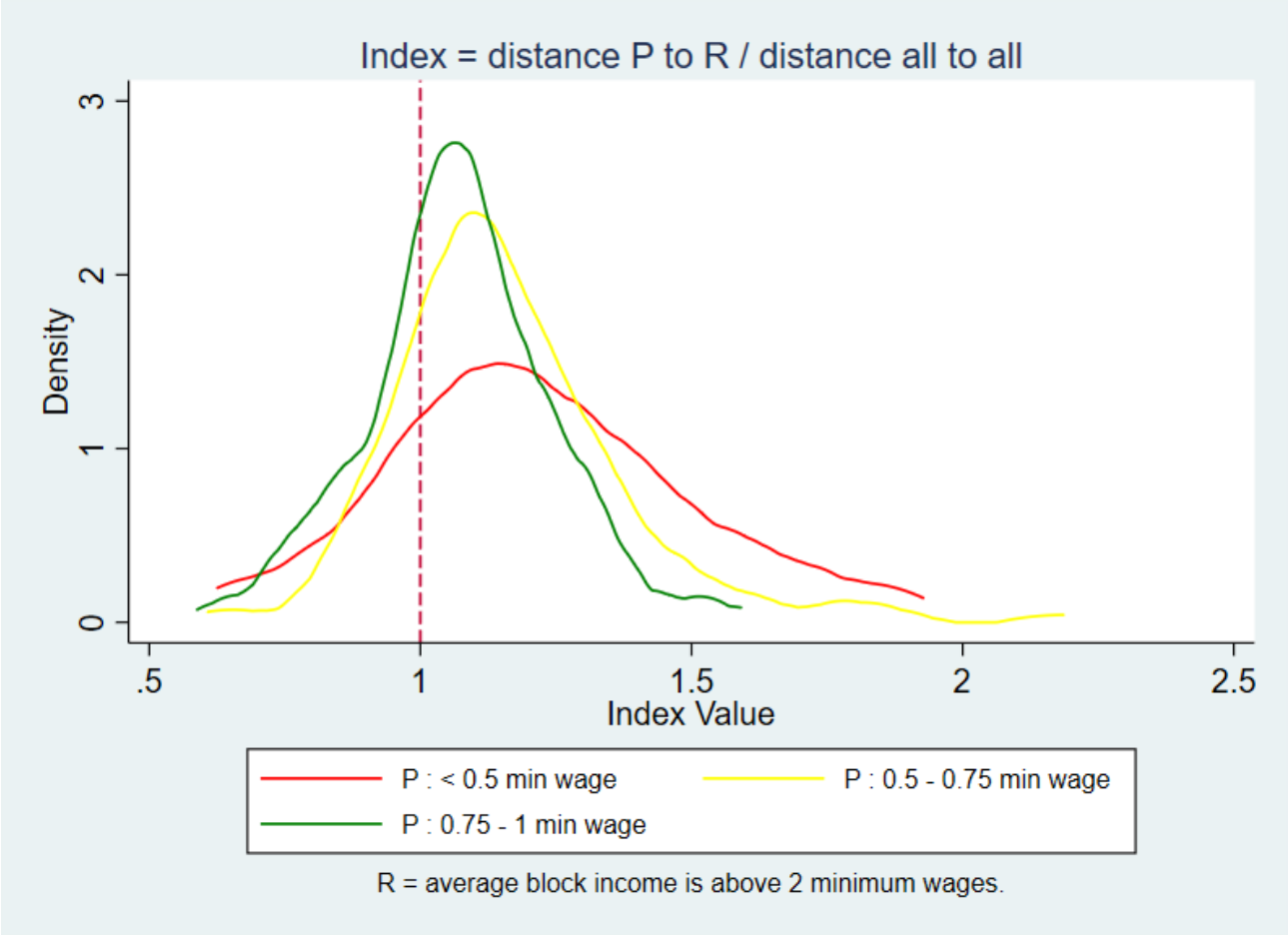
Distance segregation by relative income

More income inequality = more distance-segregation



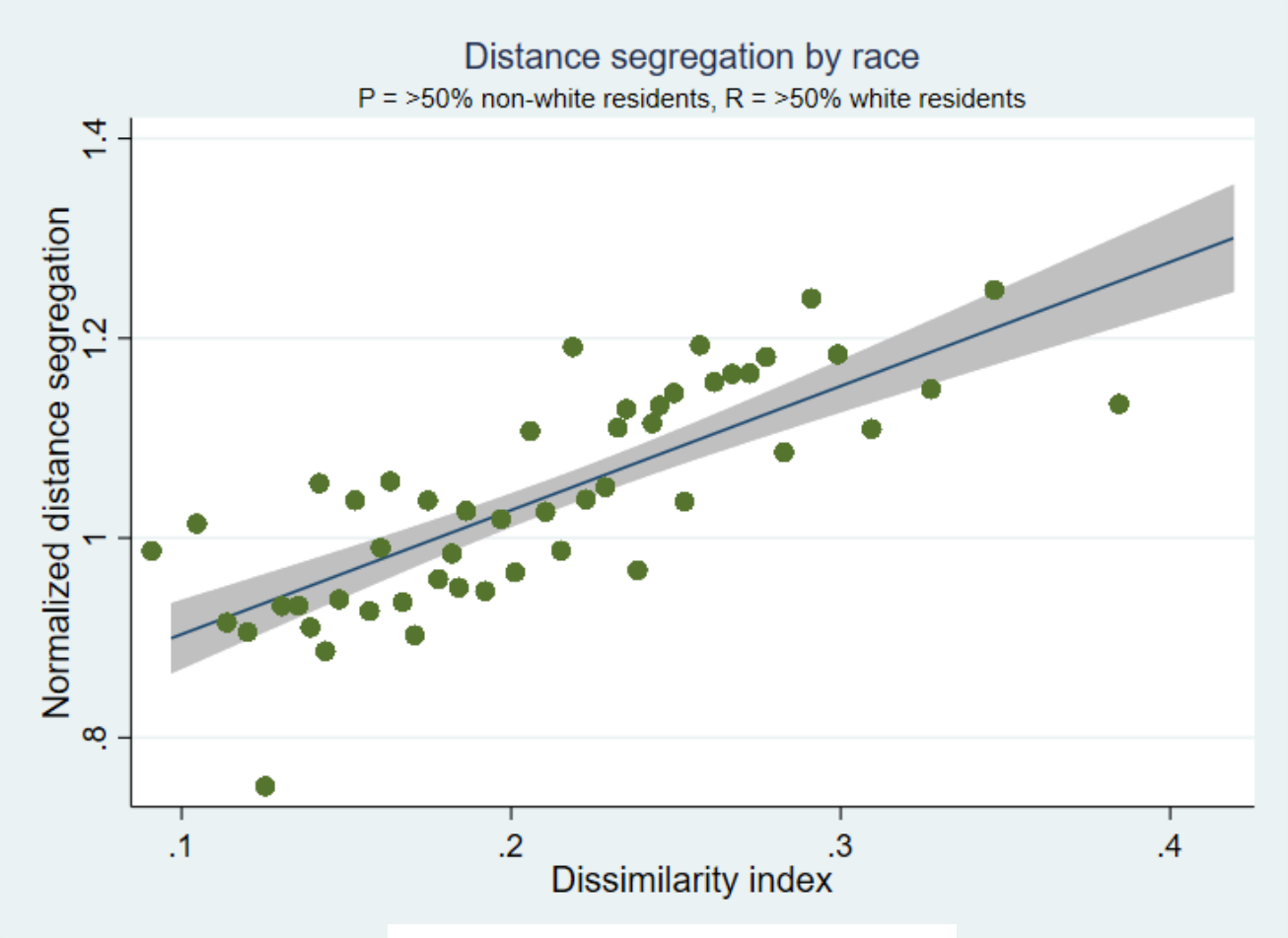
Distance segregation: absolute income

The poorer the neighborhood, the more distance-integrated



Distance segregation vs. dissimilarity

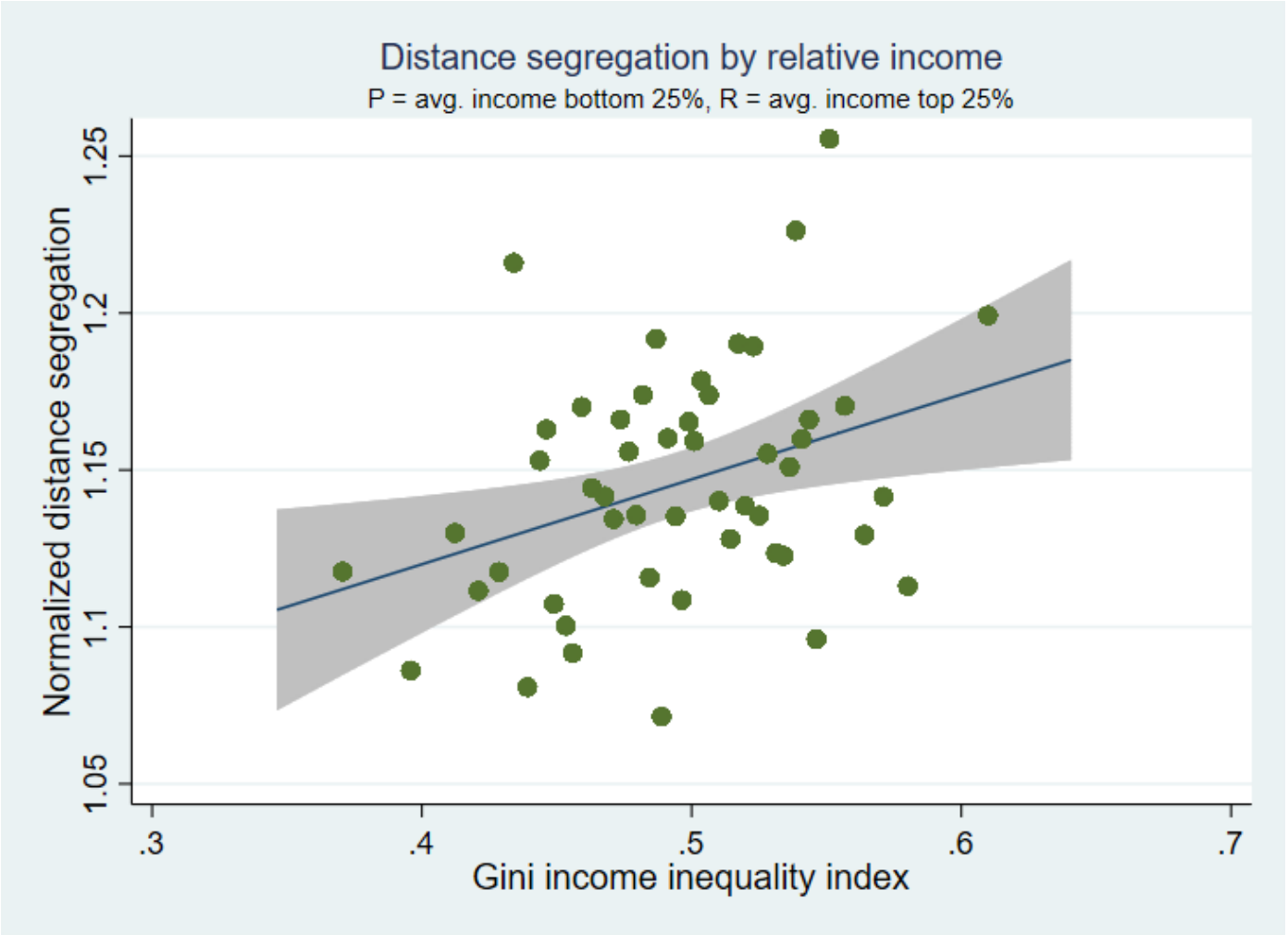
Conventionally segregated cities are also distance segregated



$$\text{Index of dissimilarity} = \frac{1}{2}\sum_{i=1}^{N}\left|\frac{black_i}{black_{total}} - \frac{nonblack_i}{nonblack_{total}}\right|$$

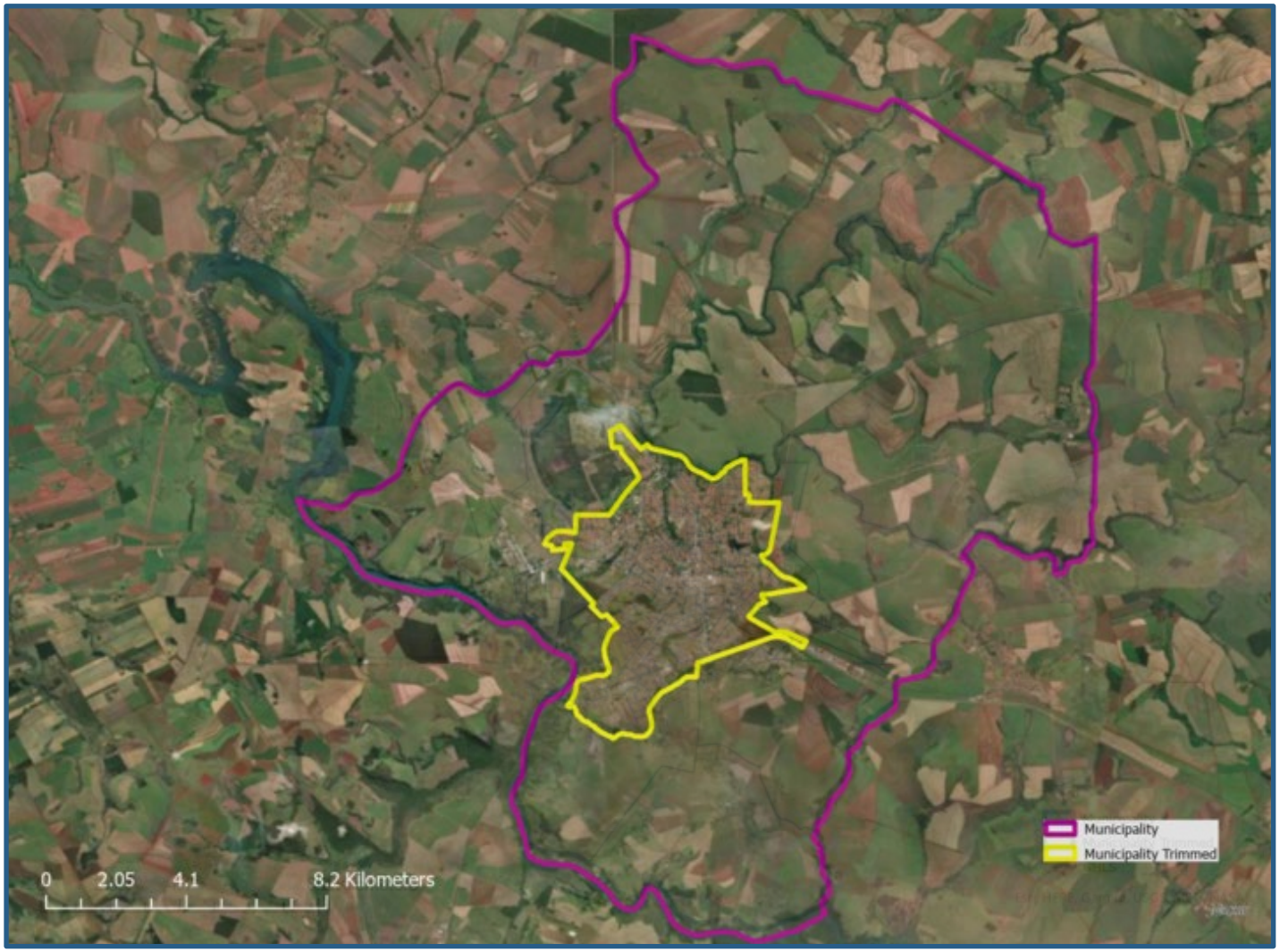
Distance segregation vs. income inequality

Unequal cities are also distance segregated



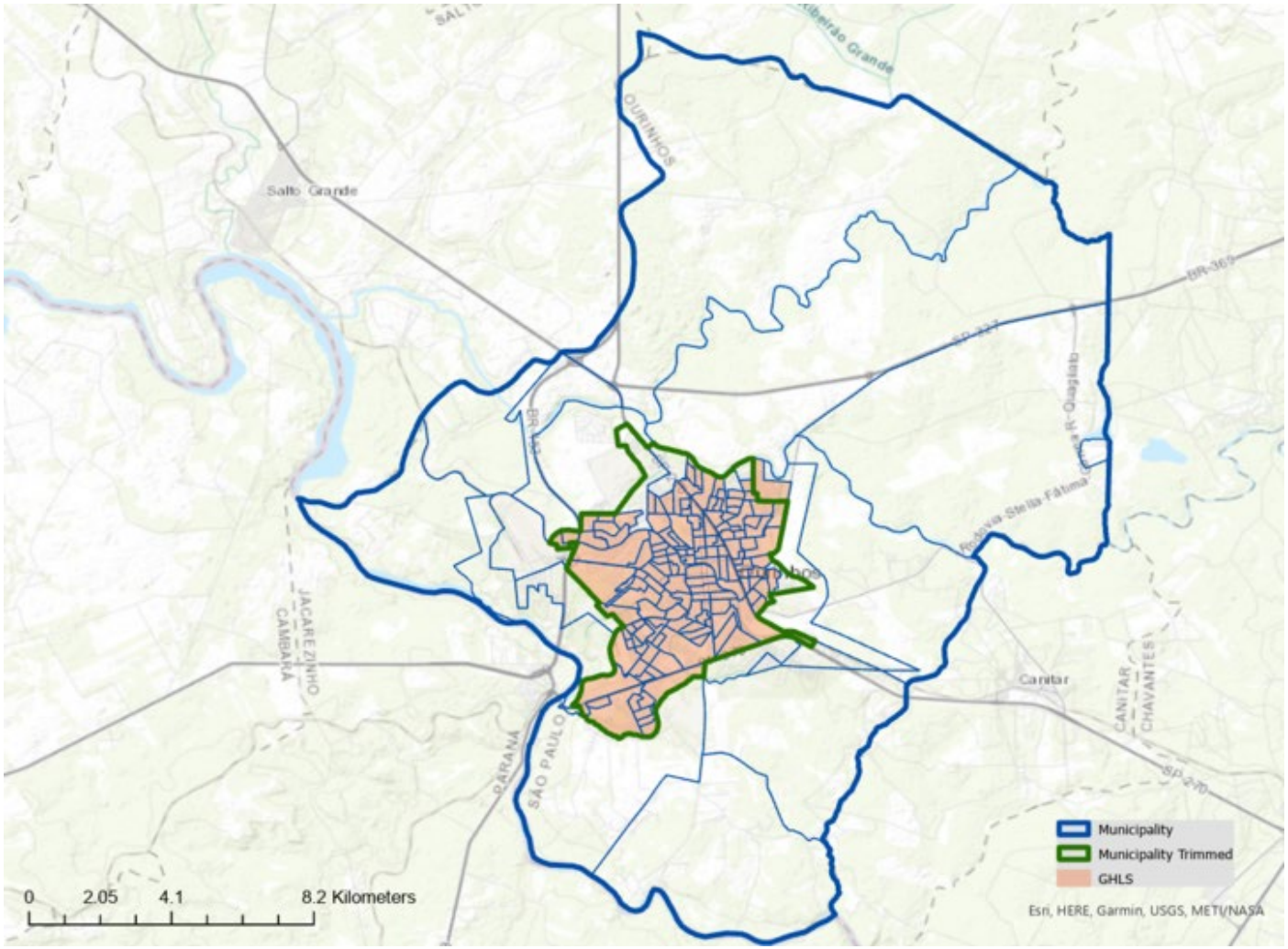
Data: defining cities

Ourinhos (SP)



Data: defining cities

Ourinhos (SP)



Data: defining cities

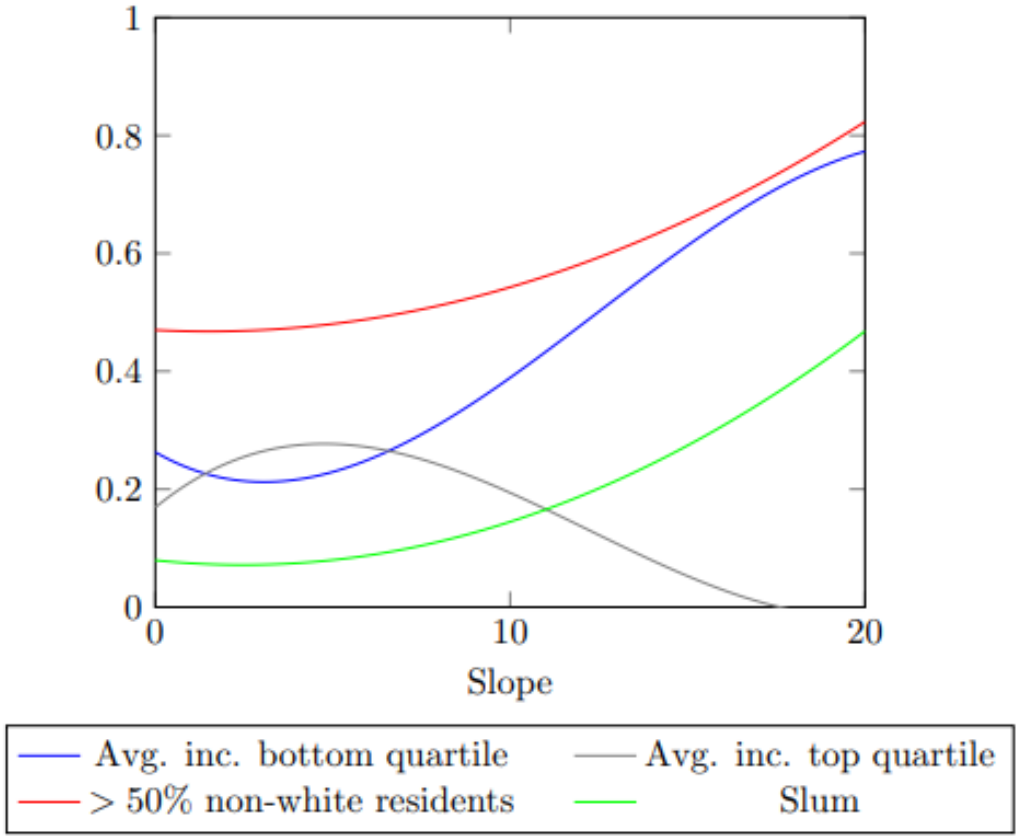
Delineating urban areas (esp. in developing countries) notoriously complicated (Duranton, 2021)

- Procedure:
  1. Consider municipalities with >50,000 residents and 50% residents classified as "urban" in Census
  2. Exclude blocks in the bottom 25% of the country for population density
  3. Manually exclude blocks that survive this procedure but appear very disconnected from urban core
     - Visually inspect against satellite imagery / OSM / Google maps
  4. Control for cities that are not contiguous and underwent manual cleaning above

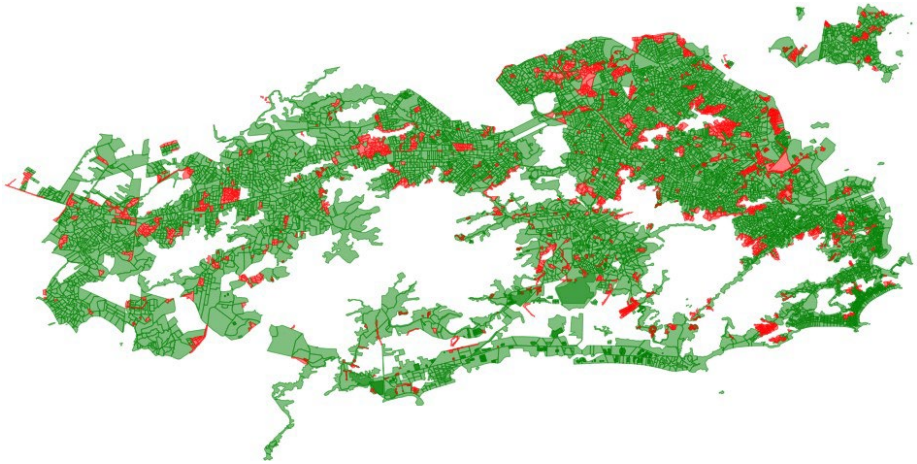
- Also considered alternative candidate units:
  - *arranjos populacionais* (Chauvin, 2018); aggregation based on commuting flows (Dingel et al. (2021): aggregates of municipalities
  - Global Human Settlement Layer (GHLS) aggregations

Determinants of residential patterns: natural advantage

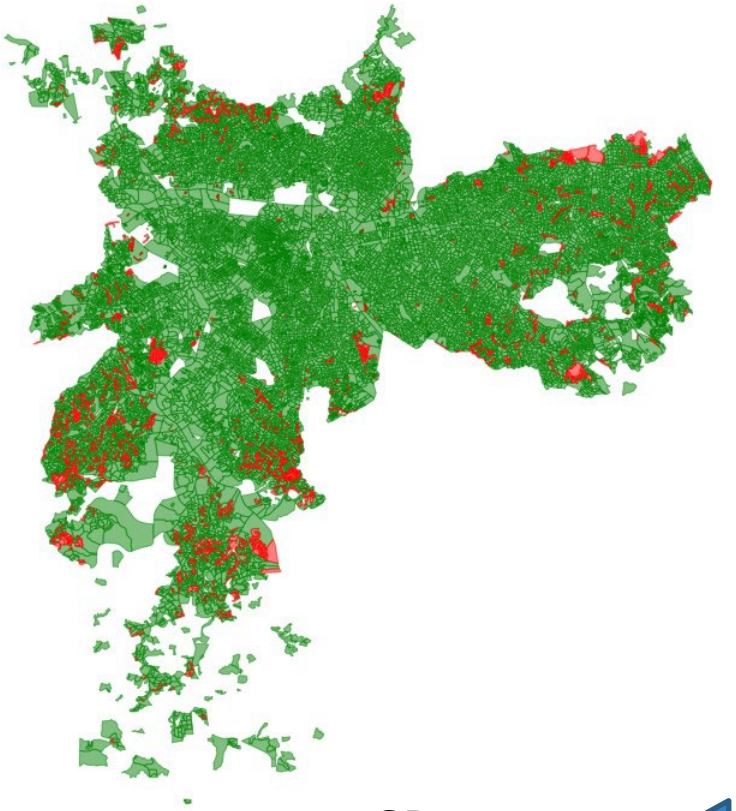
"0-th stage" refined



Geography of poor and rich neighborhoods in Brazil



Rio



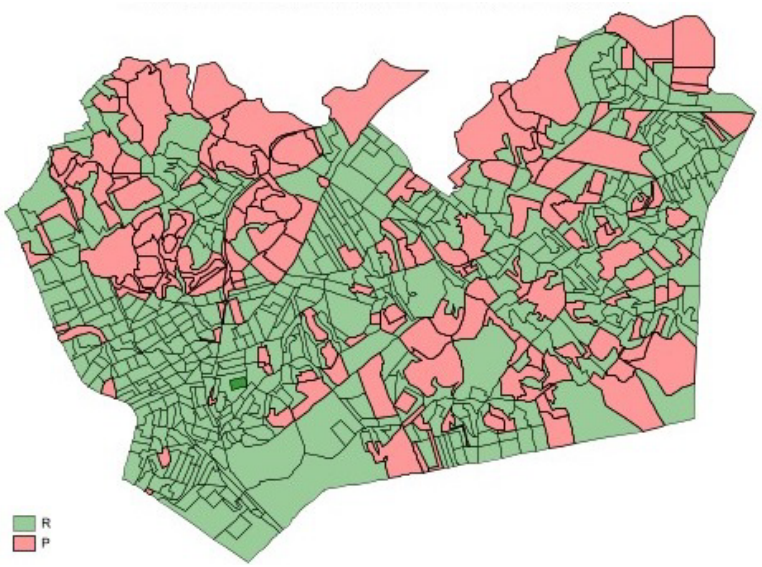
SP

Slum

Examples: segregated vs. integrated cities

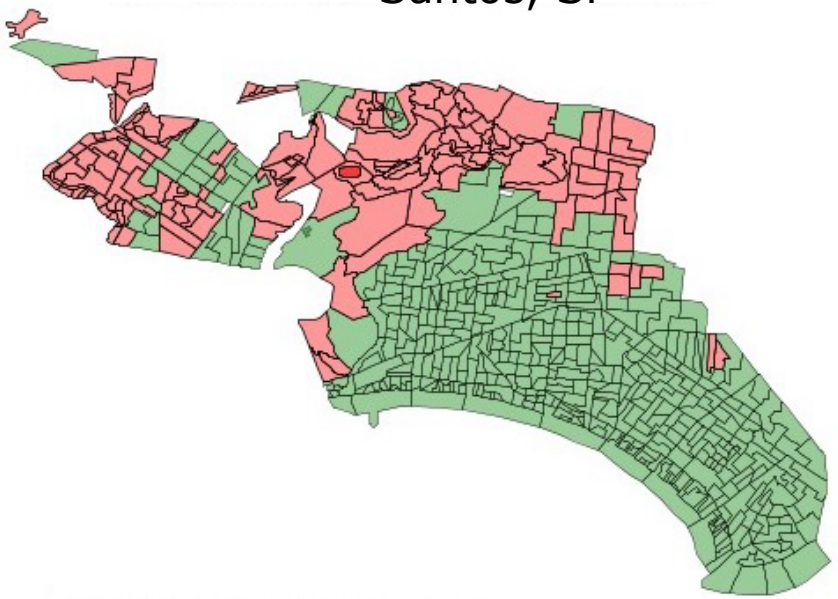
$$D_C^{PR}, \text{norm.} = \frac{\text{avg. distance P to R}}{\text{avg. distance}}$$

Belford Roxo, RJ



$$D_C^{PR}, \text{norm.} = \frac{4{,}860 \text{ m}}{4{,}780 \text{ m}} = 1.02$$

Santos, SP



$$D_C^{PR}, \text{norm.} = \frac{5{,}245 \text{ m}}{3{,}870 \text{ m}} = 1.36$$

P = avg. income is in bottom 25% of city

Instrument: balance test

Correlations small or bias against + addressed in robustness checks

	(1)	(2)
	OLS	Sample mean
Average elevation, m	-12.449**	414.3
	(4.831)	
Ruggedness, m	2.632	93.98
	(2.661)	
Average slope, degrees	0.010	5.006
	(0.078)	
Water bodies within 30 km, sqkm	-1.440	56.07
	(1.818)	
Low-fertility soil within 30 km, sqkm	-29.759**	1517
	(14.451)	
% land available within 30km	0.004	0.845
	(0.003)	
Landslide risk	0.002	1.455
	(0.010)	
Slope adjustment factor for distance	0.000	1.190
	(0.000)	
Distance to state capital, km	-9.936***	181.1
	(3.312)	
Distance to Atlantic, km	-12.230**	216.4
	(5.383)	
Latitude, degrees	0.048	-16.74
	(0.174)	
Longitude, degrees	0.193*	-45.84
	(0.098)	
Precipitation, annual avg., mm/day	1.372**	124.4
	(0.654)	
Sunshine, annual avg., wh/m2.day	-2.463	4935
	(10.176)	

Instrument: robustness

IV estimates on distance segregation

	(1)	(2)	(3)	(4)		
	% sewerage	% water	Nbhd pub goods index	OSM amenities	Obs.	IV F stat.
Panel A: Sample cuts						
Exclude state capitals	-0.0208**	-0.0183***	-0.0538***	-0.0202***	570	135
	(0.00945)	(0.00549)	(0.0153)	(0.00446)		
Exclude near state capitals	-0.0124*	-0.0164**	-0.0423***	-0.0172***	452	94
	(0.00711)	(0.00640)	(0.0131)	(0.00489)		
Exclude elevated	-0.0210**	-0.0177***	-0.0464***	-0.0224***	542	135
	(0.00919)	(0.00586)	(0.0133)	(0.00463)		
Exclude coastal	-0.0278***	-0.00383	-0.0446**	-0.0223***	507	104
	(0.00912)	(0.00780)	(0.0204)	(0.00735)		
Exclude top largest	-0.0207**	-0.0180***	-0.0494***	-0.0219***	566	152
	(0.00905)	(0.00571)	(0.0146)	(0.00450)		
Exclude bottom largest	-0.0234***	-0.0214***	-0.0546***	-0.0222***	566	370
	(0.00896)	(0.00394)	(0.0136)	(0.00407)		

Less local public goods provision in more segregated cities

	(1)	(2)	(3)	(4)
	% sewerage		% water	
Residualized by	topography	topography + distance	topography	topography + distance
Distance-segregation by income, km				
IV	-0.0209**	-0.0188**	-0.0167***	-0.0168***
	(0.00825)	(0.00806)	(0.00539)	(0.00539)
Observations	597	597	597	597
R-squared	0.127	0.139	0.129	0.128
IV F statistic	174	174	174	174
Mean dep. var	-0.0703	-0.0750	-0.0202	-0.0201

Notes: Each observation is a city. All specifications include the controls in Table 3. Standard errors clustered at the meso-region level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Robustness: control for city shape

IV estimates on distance	(1)	(2)	(3)	(4)		
	% Sewerage	% Water	Nbhd public goods	OSM amenities	Obs.	IV F stat.
Disconnected city	-0.0206**	-0.0175***	-0.0517***	-0.0223***	597	133
	(0.00837)	(0.00556)	(0.0129)	(0.00429)		
Nr water basins	-0.0245***	-0.0184***	-0.0543***	-0.0222***	597	161
	(0.00867)	(0.00586)	(0.0139)	(0.00463)		
Nr polygons, available land within 30km	-0.0232***	-0.0170***	-0.0537***	-0.0212***	597	173
	(0.00838)	(0.00544)	(0.0132)	(0.00427)		
Perimeter/area ratio, available land within 30km	-0.0222***	-0.0169***	-0.0524***	-0.0218***	597	171
	(0.00835)	(0.00551)	(0.0133)	(0.00437)		

Notes: this table reports IV coefficients of distance segregation by income on the four primary outcomes. Specifications are analogous to those in Table 3.

Robustness: specification / functional form

IV estimates on distance	(1)	(2)	(3)	(4)		
	% Sewerage	% Water	Nbhd public goods	OSM amenities	Obs.	IV F stat.
Include distance to CBD in prediction	-0.0232***	-0.0134**	-0.0597***	-0.0234***	597	113
	(0.00802)	(0.00659)	(0.0147)	(0.00409)		
Population density-weighted distance segregation	-0.0324***	-0.0251***	-0.0745***	-0.0305***	597	311
	(0.0106)	(0.00921)	(0.0171)	(0.00727)		
Index based on log distance	-0.164***	-0.105**	-0.233*	-0.0962	597	187
	(0.0634)	(0.0468)	(0.139)	(0.0596)		
Exposure Index	0.067***	0.030***	0.169***	0.025***	597	841
	(0.018)	(0.0101)	(0.040)	(0.007)		

Notes: this table reports IV coefficients of distance segregation by income on the four primary outcomes. Specifications are analogous to those in Table 3. Exposure index coefficients are standardized to the effect of one standard deviation of the index.

Robustness: specification / functional form

- Population density weighted index:

$$D_{Wc}^{PR} = \frac{\sum_i D_{Wi}^R \left(\frac{L_i^P}{a_i}\right)}{\sum_i \left(\frac{L_i^P}{a_i}\right)} \text{ where } D_{Wi}^R = \frac{\sum_j d_{ij} \left(\frac{L_j^R}{a_j}\right)}{\sum_j \left(\frac{L_j^R}{a_j}\right)}$$

- Exposure index:

$$E_c^{PR} = \frac{\sum_i E_i^R \left(\frac{L_i^P}{a_i}\right)}{\sum_i \left(\frac{L_i^P}{a_i}\right)} \text{ where } E_i^R = \frac{\sum_j \frac{1}{\tilde{d}_{ij}} \left(\frac{L_j^R}{a_j}\right)}{\sum_j \frac{1}{\tilde{d}_{ij}}}$$

$$\tilde{d}_{ij} = \exp\left(k \frac{d_{ij}}{s}\right) \text{ with } k = 0.013 \text{ (Tsivanidis, 2023) and } s = 30 \text{ km/h}$$

Robustness: local PF

IV	(1) Share residents with public sewerage	(2) Share residents with public water	(3) Neighborhood public goods index	(4) OSM amenities within 3km
Distance-segregation	-0.0178**	-0.0172***	-0.0468***	-0.0216***
	(0.00796)	(0.00417)	(0.0127)	(0.00410)
Distance-segregation	-0.0171**	-0.0117**	-0.0177	-0.00322
x municipal company	(0.00747)	(0.00520)	(0.0186)	(0.00702)
Observations	582	582	582	582
R-squared	0.595	0.491	0.621	0.178
IV F statistic	227	181	227	181
	181	227	181	227
Mean dep. var.	0.580	0.904	0.0294	0.633

Notes: Each observation is a city. All columns report IV estimates. All specifications include the controls in Table 3. Standard errors clustered at the meso-region level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Less local public goods provision in more segregated cities

Panel A: distance-segregation by race

	(1)	(2)	(3)	(4)
	% sewerage	% water	Nbhd pub goods index	OSM amenities
IV	-0.0356***	-0.0136***	-0.0762***	-0.0262***
	(0.00962)	(0.00461)	(0.0179)	(0.00745)
OLS	-0.017***	-0.015***	-0.037***	-0.017***
	(0.006)	(0.005)	(0.012)	(0.005)
Observations	447	447	447	447
R-squared	0.494	0.399	0.478	0.117
IV F statistic	59	59	59	59
Mean dep. var	0.647	0.928	0.151	0.645

Less local public goods provision in more segregated cities

Panel B: distance-segregation by slum status

	(1)	(2)	(3)	(4)
	% sewerage	% water	Nbhd pub goods index	OSM amenities
IV	-0.0568***	-0.0228*	-0.0898**	-0.0139
	(0.0157)	(0.0117)	(0.0349)	(0.0111)
OLS	-0.025***	-0.014*	-0.026	-0.010
	(0.008)	(0.008)	(0.020)	(0.008)
Observations	223	223	223	223
R-squared	0.569	0.413	0.475	0.216
IV F statistic	27	27	27	27
Mean dep. var	0.602	0.900	0.0543	0.662

Notes: Each observation is a city. All specifications include the controls in Table 3. Standard errors clustered at the meso-region level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Other measures of public goods

Distance-segregation by income, km	(1) Share municipal spending on health	(2) Nr hospital beds per 1000 ppl	(3) Park	(4) Post office	(5) Police station	(6) Fire station	(7) Paved streets, % residents	(8) Sidewalks, % residents
IV	-0.00738***	-0.190***	142.4***	580.9***	414.9***	613.6***	-0.0134***	-0.0240***
	(0.00200)	(0.0453)	(34.03)	(95.63)	(90.74)	(145.5)	(0.00399)	(0.00564)
Observations	593	595	593	419	496	332	597	597
R-squared	0.156	0.098	0.336	0.311	0.155	0.357	0.387	0.492
Mean dep. var	0.239	2.362	619.2	2510	2031	3111	0.771	0.645

Notes: Each observation is a city. All specifications include the controls in Table 3. The dependent variable in columns 3 though 6 is distance in meter to the nearest OSM amenity. Standard errors clustered at the meso-region level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Human capital and demographics

	(1)	(2)	(3)	(4)	(5)
Distance-segregation by income, km	% Literacy rate	% High school degree or higher	% Employment in service sector	% non-White	% Prime age males
IV	-0.00275***	-0.00600***	-0.00648***	0.00502	-0.000106
	(0.00100)	(0.00145)	(0.00196)	-0.00318	(0.000335)
Observations	597	597	597	597	597
R-squared	0.757	0.468	0.547	0.782	0.617
F statistic	174	174	174	174	174
Mean dep. var.	0.908	0.412	0.366	0.516	0.208

Notes: Each observation is a city. All specifications include state fixed effects and the controls in Table 3. Standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Spatial targeting of integrated vs. segregated neighborhoods

IV estimates for distance-segregation by income, km

	(1)	(2)	(3)	(4)	(5)
	% sewerage	% water	% paved streets	% sidewalks	OSM amenities
A. Poor	-0.0216**	-0.0199***	-0.0124**	-0.0217***	-0.0222***
	(0.00923)	(0.00531)	(0.00617)	(0.00732)	(0.00416)
	[0.479]	[0.867]	[0.610]	[0.439]	[0.556]
B. Rich	-0.0193***	-0.0169**	-0.0143***	-0.0212***	-0.0167***
	(0.00692)	(0.00738)	(0.00347)	(0.00461)	(0.00533)
	[0.667]	[0.922]	[0.909]	[0.831]	[0.696]
C. Poor / Rich	-0.0244	-9.14e-05	-0.00231	-0.0154**	-0.0146**
	(0.0244)	(0.0117)	(0.00567)	(0.00705)	(0.00584)
	[0.759]	[0.945]	[0.652]	[0.503]	[0.828]
D. Poor close to Rich	-0.0210***	-0.0182**	-0.0153***	-0.0245***	-0.0174***
	(0.00769)	(0.00749)	(0.00424)	(0.00550)	(0.00568)
	[0.629]	[0.920]	[0.845]	[0.735]	[0.698]
E. Poor far from Rich / Poor	0.0258	0.0279***	0.0496**	0.113***	-0.0319***
	(0.0390)	(0.0108)	(0.0230)	(0.0403)	(0.00933)
	[0.885]	[0.991]	[0.982]	[1.054]	[0.776]
Observations	597	597	597	597	597

Notes: Each panel shows results for different aggregations of the outcome variable at the city level. In Panel A the dependent variables are outcomes averaged among the poor neighborhoods in the city. The table reports IV coefficients, standard errors in parentheses, and sample averages of each dependent variable in square brackets.

Historical determinants: disease externalities

Cities exposed to outbreaks are more segregated today

Dependent variable: distance-segregation index, normalized

	(1)	(2)	(3)	(4)
	Income	Income	Income	Race
Distance segregation 1900 radius	0.318***			
	(0.0505)			
Mosquito suitability		0.00103		
		(0.0909)		
Mosquito suitability		0.0396**		
x Established after outbreak in state		(0.0178)		
Avenidas			0.0388***	
			(0.0127)	
% Slaves in 1872				0.291**
				(0.134)
Established after slavery abolished				0.0638
				(0.0390)
% Slaves in 1872				-0.435***
x Established after slavery abolished				(0.138)
Observations	516	600	600	451
R-squared	0.233	0.169	0.076	0.367

Notes: Each observation is a city. Average distance between P and R neighborhoods is normalized by average distance between any two neighborhoods. All specifications include the geography controls from Table 3. Columns 2 and 4 additionally controls for state fixed effects and dummies for period of settlement. Standard errors clustered at the meso-region level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Yellow fever outbreaks prompted "hygienist" interventions to clear tenements from the center and segregate poor away

Historical determinants

Dependent variable: distance-segregation index, normalized

	(1)	(2)	(3)	(4)
	Income	Income	Income	Race
Distance segregation 1900 radius	0.318***			
	(0.0505)			
Mosquito suitability		0.00103		
		(0.0909)		
Mosquito suitability		0.0396**		
x Established after outbreak in state		(0.0178)		
Avenidas			0.0388***	
			(0.0127)	
% Slaves in 1872				0.291**
				(0.134)
Established after slavery abolished				0.0638
				(0.0390)
% Slaves in 1872				-0.435***
x Established after slavery abolished				(0.138)
Observations	516	600	600	451
R-squared	0.233	0.169	0.076	0.367

Notes: Each observation is a city. Average distance between P and R neighborhoods is normalized by average distance between any two neighborhoods. All specifications include the geography controls from Table 3. Columns 2 and 4 additionally controls for state fixed effects and dummies for period of settlement. Standard errors clustered at the meso-region level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Beyond geography: historical determinants of distance segregation

- Cities that experienced disease outbreaks in 19th century are more income segregated today
  - Key urban externality in 19th century: infectious disease
  - Proxy using mosquito suitability
  - Outbreaks prompted interventions to clear tenements from the center and segregate poor away

- Cities with a history of slavery are more race segregated today

- Segregation in the historical portion of the city

Potential avenue for future work!