Sewers and Urban Development

Sean McCulloch (Brown University)
Matthew Schaelling (Brown University)
Matthew Turner (Brown University and NBER)

March 14, 2024

How does sewer construction affect the development of cities? I

Why is this an important/interesting question?

- Rank the most important things a local government does for you, e.g.; water, law and order, sewer, fire, other public health, transportation, schools.
- We've studied the effects of transportation a lot, but sewers only a little. This doesn't seem to line up with their importance.
- If spatial equilibrium reflects the balancing of agglomeration with congestion forces, understanding the importance of sewers for mitigating congestion is fundamental to understanding equilibrium.
- Are the effects of sewers on cities the same everywhere?

How does sewer construction affect the development of cities? II

- The absence of sewers is one of the defining features of slums

  ... slums are areas where households lack durable housing, sufficient living space, secure tenure, or easy access to safe water or adequate sanitation facilities. (WB website Feb 10, 2024)

- How does the provision of modern sanitation change slums?
- What is the incidence of this change?

Research design

- Measure extent of sewer network in a sample of cities using geocoded census data.
- Measure outcomes, primarily population density and income, using the same census data.
- Estimate treatment effects by comparing census tracts on opposite sides of drainage basin divides.

This overcomes two of the main impediments to studying the effects of sewers; observing them and understanding the assignment process.

It is also cheap. All we need is a geocoded census at about the tract level, and digital elevation maps.

Literature

- Water and sewer infrastructure has large, well documented effects on health in the developing world (Ashraf et al. [2017], Bhalotra et al. [2021], Galiani et al. [2005]) and in developed world cities in the late 19th/early 20th centuries (Alsan and Goldin [2019], Anderson et al. [2018], Ferrie and Troesken [2008]).
- Sewers have no effect on infant mortality in Brazil 1990-2010 [Gamper-Rabindran et al., 2010].
- Sewers have a large effect on land prices in late 19th century Chicago Coury et al. [2022].
- There does not seem to be a lot of sorting in response to sewers Alsan and Goldin [2019].

Longrightarrow weak prior that sewers lead to large increase in density and not much change in demographics. No prior over heterogeneity of effects.

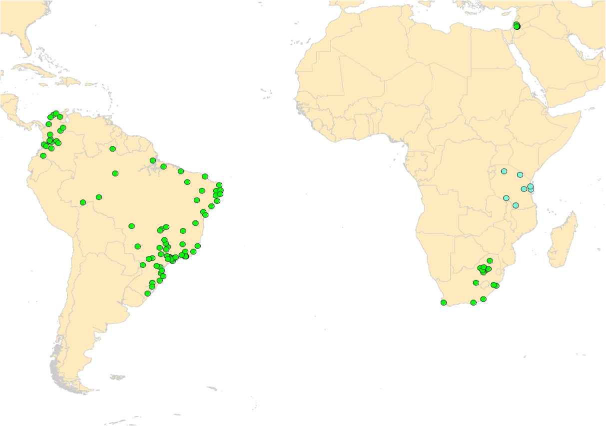
Data I

Cities

- The UN Cities data is a census of all cities that had a population 300,000 or more in 2014. These data report the location of the center of each city.
- We focus attention on areas (1) near the boundary of the drainage basin containing the city center, and (2) within 75km of the city center.
- We estimate treatment effects using all Cities in the UN Cities data in; Brazil, Colombia, South Africa, Jordan. We can evaluate counterfactuals in these cities and those in Tanzania and India (maybe, TBD).

Data II

Cities



Locations of cities in our sample. Tanzania is only included in counterfactual exercise.

Data

Sewers

Sewer data all comes from census questions;

- Brazil: Brazilian Institute of Geography and Statistics [2012], 'Is the bathroom or toilet drain connected to the public sewer system?'
- Colombia: National Administrative Department of Statistics [2018], 'Does your house have sewage service?'
- South Africa: National Administrative Department of Statistics [2018], 'Is the main type of toilet facility used by this household a flush toilet connected to sewerage system'
- Tanzania: National Bureau of Statistics (Tanzania), Office of the Chief Government Statistician [2012], 'Does your house have a flush toilet connected to a piped sewer system?'
- Jordan: Department of Statistics (Jordan) [2015], 'Does your house have sanitation connected to a public network?'

We calculate the share of households in a 'tract' with sewer access and map the extent of tracts with sewers.

Data

Population density, other outcomes

Population density, income measures, and other outcomes all come from the same censuses.

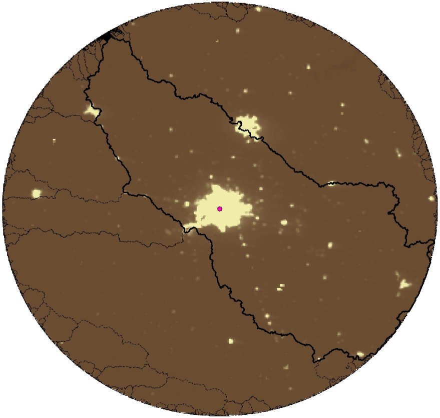
For all countries but Jordan, population density is the full count of people divided by tract area. For Jordan, it is the full count of households divided by tract area.

Demographic and neighborhood outcomes also come from these censuses and vary from country to country.

Data

Drainage basins

- We construct drainage basins from digital elevation maps using tools for this purpose in ARCGIS.
- We use two; the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) DEM and the Shuttle Radar Topography Mission (SRTM) DEM.
- ASTER is derived from stereoscopic imagery that is thought to be less prone than SRTM to confuse trees and rooftops with the ground. We rely primarily on the ASTER DEM, but consider SRTM for robustness checks.
- A comparison with LIDAR data shows that average error of ASTER is about 4m in four small study areas. SRTM is about the same. [Uuemaa et al., 2020].



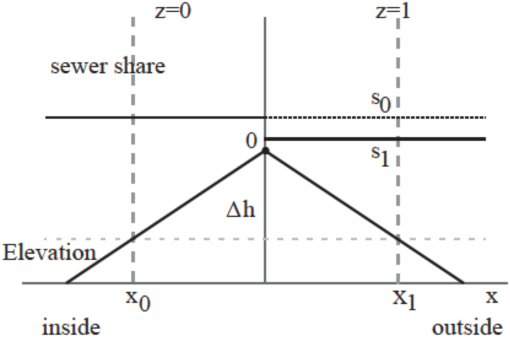
Drainage basin boundaries in a 75km disk centered on Uberlandia, Brazil. Background is lights at night, grey is all drainage basin boundaries calculated from the ASTER DEM. Black is the boundary of the drainage basin containing the center city.

Identification I

- Sewers work on gravity. Moving sewage on a grade of less than 1:200 is hard. Uphill is harder.
- Sewer networks generally serve a (part of a) single drainage basin.
- Two census tracts on opposite sides of a basin divide should be similar (on average), but one may require moving sewage uphill to get to an existing sewer network.

How can we use this intuition to think about the effect of sewer service on urban development?

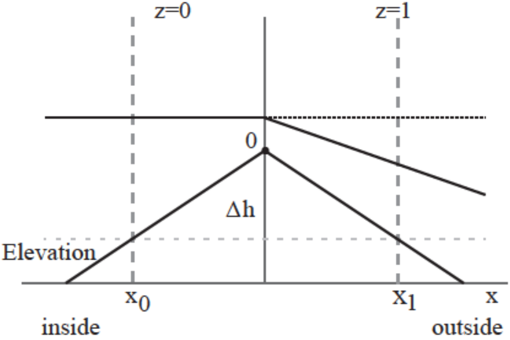
Identification with a discontinuity in sewer share



Plot of elevation and sewer share in a neighborhood of a basin divide. 'inside' is uphill from existing network. x is distance to the basin divide. Elevation is relative to basin divide.

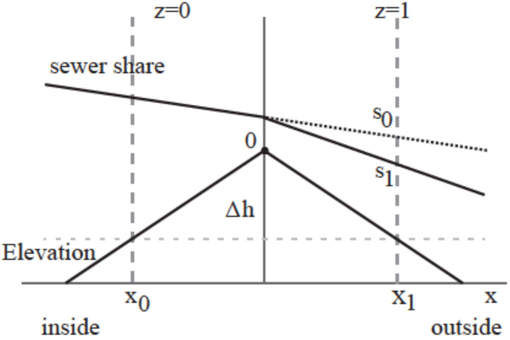
Treatment is $s_0 - s_1$. We want to compare outcomes for census tracts at x_0 and x_1.

Identification with a kink in sewer share



No strong prior over whether crossing a basin divide will lead to a step or a kink in sewer share. It depends on the scale over which costs increase.

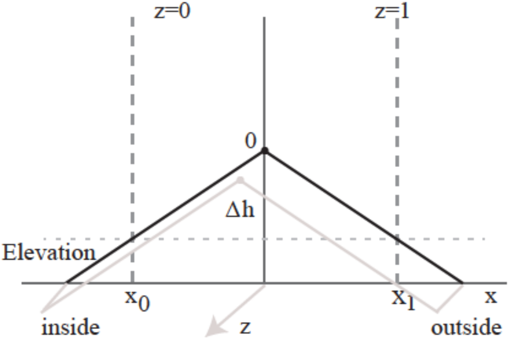
Identification with a kink in sewer share and trend in x



Distance to the basin divide may have an independent effect on sewer share. We need to look for a kink (or step) in sewer share net of the effect of x displacement.

Is x best thought of as (log) displacement perpendicular to the basin divide, or as (log) radial distance to the CBD?

Identification, elevation vs displacement



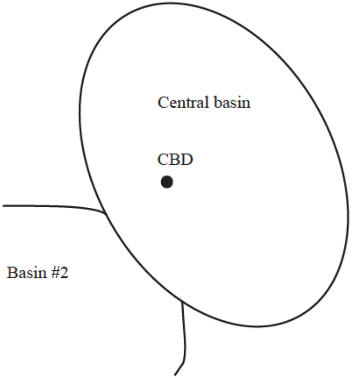
Distance to the basin divide and climbing to the elevation divide should both matter. With elevation, we can exploit variation independent of x.

Identification, further issues

- Taking the presence of gravity as given, crossing a basin divide must increase the cost of sewer access.
- To turn this intuition into a research design, we need some empirical analog to the figures.
  - We can already draw basin divides.
  - How do we define 'inside'?
  - What is the best analog for x?
  - What window width should we consider? (within 4km of divide).
  - The pictures describe a line. How to translate to a plane? How wide a strip should we use?
  - What is the shape of the sewer share response, kink vs. step vs. both?

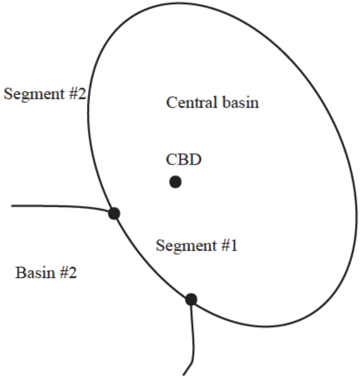
Geography: Central basins

DESA Population Division [2018] gives CBD coordinates for cities with pop. $\geq$ 300,000 in 2014. All have sewer networks. Focus on drainage basins containing CBDs and $<$ 75km from CBD.



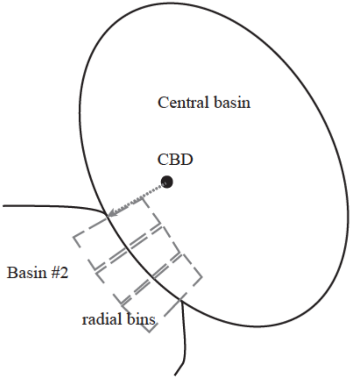
Geography: Segments

Partition central basin divide into a ‘segment’ for each opposite
basin



Geography: Radial bins

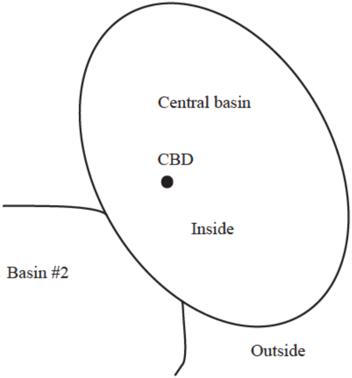
Define ‘radial bins’, 2km wide, and 2×4km deep. Radial bin zero starts at the point on the basin divide closest to the CBD.



Tract elevation is relative to highest tract centroid in the same radial bin $\leq$ 2km from the basin divide. NB: larger elevations are lower.

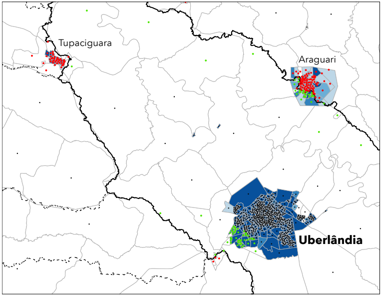
‘Basin inside’

Say a census tract is ‘inside’ or ‘basin inside’ if its centroid lies in the central basin. If Basin #2 is also a CBD basin, then ‘basin inside’ is not defined for any tract in a radial bin on the common divide boundary.



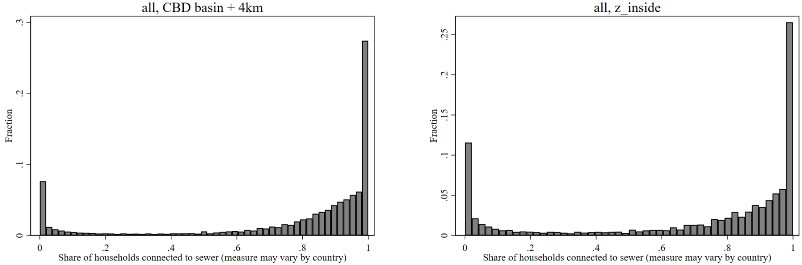
‘Basin inside’ near Uberlandia, Brazil

Say a census tract is ‘inside’ or ‘basin inside’ if its centroid lies in the central basin.



- Tracts in the central basin are ‘inside’.
- Drop tracts with centroids more than 4km from the basin divide.

Sewer share, Universe



- Left: all tracts in central basin or within 4km ('Universe')
- Right: all tracts within 4km of central basin divide ('Estimation sample')
- Only about 40% of tracts have sewer share 0 or 1; treatment is continuous.
- Estimation sample is slightly less sewered than the Universe.

Description, Universe

	CBD Basin + 4km	$\pm$ 4km Basin Divide
Num cities	86	85
Mean area cbd basin (kmsq)	6,613	1,494
Num segments	1,103	582
Num radial-bins	9,245	3,288
Num tracts	396,967	76,995
Share inside	0.88	0.48
Mean tract area (kmsq)	1.08	0.90
Mean dist to CBD (km)	22.45	13.45
Mean log dist to CBD (m)	9.44	8.83
Mean dist to basin divide (km)	12.21	1.67
Sewer share	0.67	0.69
Pop density (persons per kmsq)	26,019	21,798
Income (pay per month)	1,582	1,513

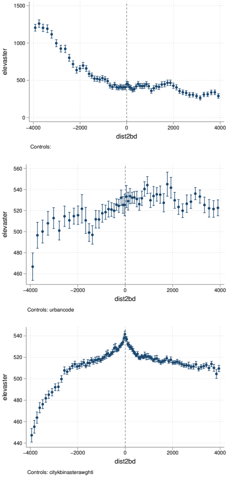
- About 25 tracts per bin.
- Tracts are about 1km square. Central basins average about 80km square.
- Population density is high. Quarter acre lots with 4 people per household is about 4000 per km^2.

Estimation sample; Brazil, Colombia, South Africa, Jordan

Country	cities	π-bins	tracts	share inside	Tract area km^2	People/km^2	Sewer Share
Brazil	62	2164	40,117	0.45	1.27	15,603	0.68
Colombia	18	1049	36,706	0.51	0.47	28,667	0.70
Jordan	6	75	172	0.46	5.32	871	0.26
South Africa	15	499	7,927	0.52	1.95	8,226	0.76
Tanzania	7	387	3918	0.68	2.09	8,677	0.01

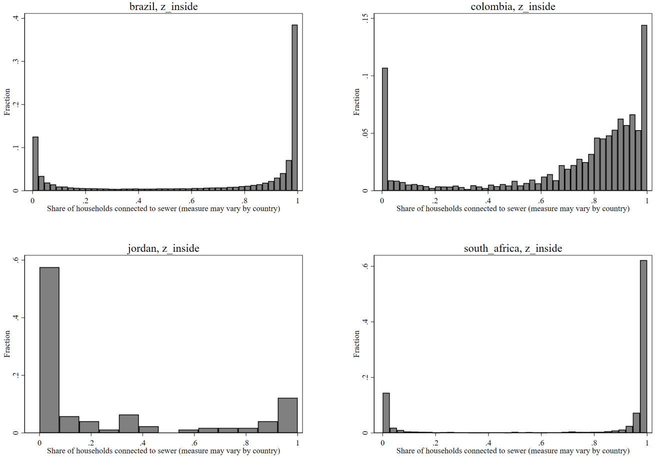
- Jordan is households, not people.
- The economic geography of these places is really different.
- Estimations will be at tract/bin level, so most of the weight will come from Colombia and Brazil.
- Tanzania has no sewers.

Why radial-bin controls?



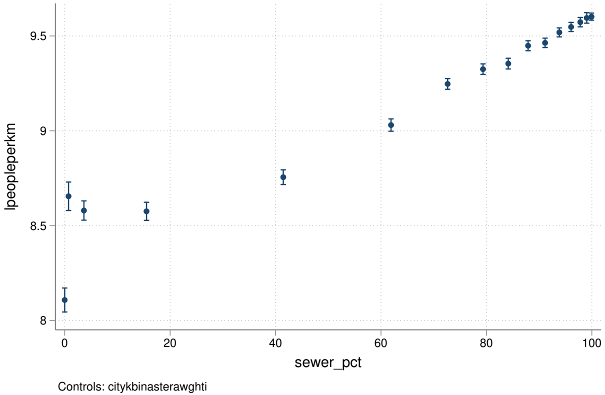
Mean elevation by distance to basin divide; raw data (top), net of segment mean (middle), and net of radial bin mean (bottom).

Sewer share, estimation sample, by country



- Brazil, South Africa do better with the ‘last mile problem’
- Share of tracts with 0 or 100% is about 25% (Colombia), 55%(Brazil), 70% (Jordan, South Africa).

Population density vs sewer %



Mean log population by tract sewer percentage. All tracts within 4km of a basin divide, conditional on radial bin. 100% increase $\approx$ factor of three for pop. density, i.e. 3% elasticity.

Estimation/Identification (1)

$$
\begin{aligned}
i &\sim \text{tract, } k \sim \text{radial bin} \\
s_{ik} &\sim \text{sewer \% (0,100)} \\
y_{ik} &\sim \ln(\text{pop/km}^2) \text{ or } \ln(\text{mean income}) \\
x_{ik} &\sim \text{meters from tract centroid to basin divide} \\
\Delta E_{ik} < 0 &\sim \text{tract centroid meters below basin divide} \\
\mathbb{1}(\text{Outside})_{ik} &\sim \text{indicator for `outside'}
\end{aligned}
$$

Compare tracts on opposite sides of basin divide:

$$
\begin{aligned}
s_{ik} &= \mathbb{1}(k)_{ik} + \mathbb{1}(k)x_{ik} + \alpha\mathbb{1}(\text{Outside})_{ik} + \eta_{ik} \text{ (1st stage)} \\
y_{ik} &= \mathbb{1}(k)_{ik} + \mathbb{1}(k)x_{ik} + \gamma\mathbb{1}(\text{Outside})_{ik} + \eta_{ik} \text{ (Reduced form)} \\
y_{ik} &= \mathbb{1}(k)_{ik} + \mathbb{1}(k)x_{ik} + \beta_{ik}s_{ik} + \eta_{ik} \text{ (OLS/TSLS)}
\end{aligned}
$$

Why bin fixed effect? Why interaction with distance? NB: only parameters pooled across bins are α, β, γ

Estimation/Identification (2)

$$\begin{aligned} i &\sim \text{tract, } k \sim \text{radial bin} \\ s_{ik} &\sim \text{sewer \% (0,100)} \\ y_{ik} &\sim \ln(\text{pop/km}^2) \text{ or } \ln(\text{mean income}) \\ x_{ik} &\sim \text{meters from tract centroid to basin divide} \\ \Delta E_{ik} < 0 &\sim \text{tract centroid meters below basin divide} \\ \mathbb{1}(\text{Outside})_{ik} &\sim \text{indicator for `outside'} \end{aligned}$$

Compare tracts on opposite sides of basin divide:

$$s_{ik} = \mathbb{1}(k)_{ik} + \mathbb{1}(k)x_{ik} + \delta\Delta E_{ik} + \alpha\mathbb{1}(\text{Outside})_{ik}\Delta E_{ik} + \eta_{ik} \text{ (1st stage)}$$
$$y_{ik} = \mathbb{1}(k)_{ik} + \mathbb{1}(k)x_{ik} + \delta\Delta E_{ik} + \gamma\mathbb{1}(\text{Outside})_{ik}\Delta E_{ik} + \eta_{ik} \text{ (Reduced form)}$$
$$y_{ik} = \mathbb{1}(k)_{ik} + \mathbb{1}(k)x_{ik} + \delta\Delta E_{ik} + \beta_{ik}s_{ik} + \eta_{ik} \text{ (OLS/TSLS)}$$

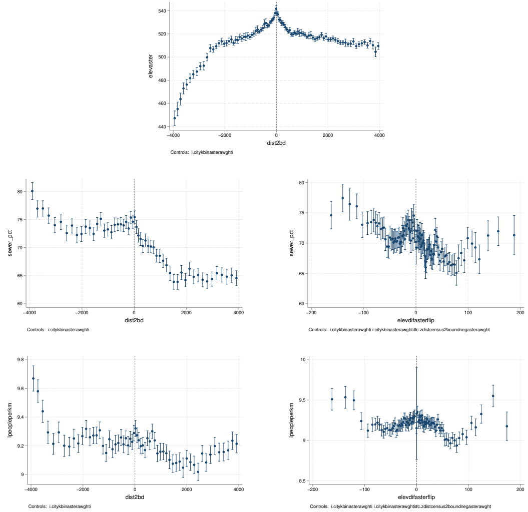
Why bin fixed effect? Why interaction with distance? NB: only parameters pooled across bins are α, β, γ, δ

Estimation/Identification

We want to visualize two sorts of variation in the data

- Change in sewer % and log(pop/km^2) and distance to basin divide
  $\Longrightarrow$ binscatter of each against x net of π-bin mean.
- Change in sewer % and log(pop/km^2) and elevation to basin divide
  $\Longrightarrow$ binscatter of each against meters to divide, net of π-bin mean and π-bin slope in x.

Identification, Universe



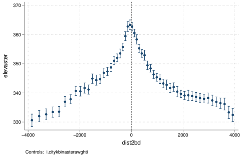
Note elevation range around divide. On average the divide is not a dramatic feature.

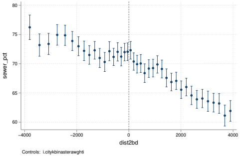
Sewers and population density, Universe

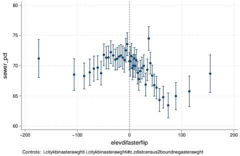
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. OLS									
sewer %	0.0127***	0.0127***	0.0127***	0.0125***	0.0125***	0.0125***	0.0112***	0.0112***	0.0112***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
2. First stage									
outside	-5.4384***		-3.4005***	-2.9845***		-1.5576***	-2.8406***		-2.1131***
	(0.2577)		(0.3390)	(0.3172)		(0.3870)	(0.3544)		(0.4293)
ΔElev*Outside		-0.0668***	-0.0422***		-0.0503***	-0.0393***		-0.0379***	-0.0197***
		(0.0037)	(0.0048)		(0.0054)	(0.0065)		(0.0061)	(0.0074)
3. Reduced form pop density									
outside	-0.0612***		0.0265	-0.0432***		0.0465**	-0.0395**		0.0111
	(0.0127)		(0.0182)	(0.0160)		(0.0215)	(0.0174)		(0.0236)
ΔElev*Outside		-0.0016***	-0.0018***		-0.0021***	-0.0025***		-0.0013***	-0.0014***
		(0.0002)	(0.0003)		(0.0003)	(0.0004)		(0.0003)	(0.0004)
4. IV									
sewer %	0.0113***	0.0243***	0.0178***	0.0145***	0.0425***	0.0337***	0.0139**	0.0337***	0.0208***
	(0.0022)	(0.0027)	(0.0021)	(0.0051)	(0.0055)	(0.0046)	(0.0058)	(0.0081)	(0.0057)
F	516.1	519.8	326.1	100.3	146.2	83.29	82.22	62.75	46.94
N	80897	80897	80897	80897	80897	80897	80897	80897	80897
π bins	Y	Y	Y	Y	Y	Y	Y	Y	Y
seg×x	.	.	.	Y	Y	Y	.	.	.
πbin×x	.	.	.	.	.	.	Y	Y	Y

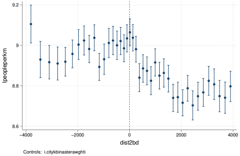
Note: Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Focus on (7), (8), (9). I think (8) is best $\Longrightarrow$ 1%=OLS$\leq$ IV $\approx$ 3%.

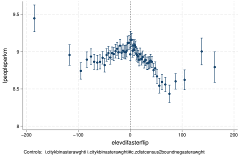
Identification, Brazil









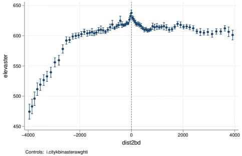


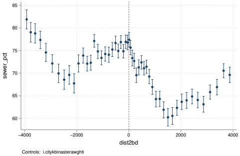
Sewers and population density, Brazil

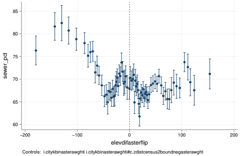
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. OLS									
sewer %	0.0126***	0.0126***	0.0126***	0.0126***	0.0126***	0.0126***	0.0115***	0.0115***	0.0115***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0004)
2. First stage									
outside	-4.2966***		-5.1857***	-2.4269***		-2.4486***	-2.3699***		-2.0924***
	(0.3870)		(0.4948)	(0.5034)		(0.6010)	(0.5398)		(0.6499)
ΔElev*Outside		-0.0270***	0.0200***		-0.0221***	0.0007		-0.0294***	-0.0089
		(0.0046)	(0.0058)		(0.0083)	(0.0099)		(0.0088)	(0.0106)
3. Reduced form pop density									
outside	-0.1318***		-0.1480***	-0.1066***		-0.0430	-0.0854***		-0.0132
	(0.0175)		(0.0233)	(0.0240)		(0.0293)	(0.0250)		(0.0304)
ΔElev*Outside		-0.0010***	0.0004		-0.0025***	-0.0021***		-0.0024***	-0.0023***
		(0.0003)	(0.0003)		(0.0005)	(0.0006)		(0.0005)	(0.0006)
4. IV									
sewer %	0.0307***	0.0362***	0.0299***	0.0439***	0.1127***	0.0433***	0.0360***	0.0830***	0.0428***
	(0.0041)	(0.0092)	(0.0040)	(0.0108)	(0.0390)	(0.0107)	(0.0108)	(0.0240)	(0.0113)
F	142.2	27.28	75.66	25.71	6.951	12.86	22.18	11.00	11.45
N	39137	39137	39137	39137	39137	39137	39137	39137	39137
π-bins	Y	Y	Y	Y	Y	Y	Y	Y	Y
seg×x	.	.	.	Y	Y	Y	.	.	.
π-bin×x	.	.	.	.	.	.	Y	Y	Y

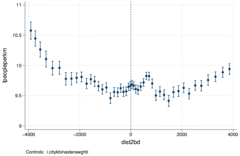
Note: Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Focus on (7), (8), (9). I think (8) is best $\implies$ 1%=OLS$\leq$ IV $\approx$ 4%.

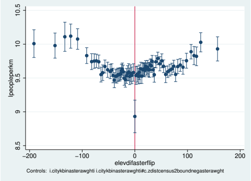
Identification, Colombia









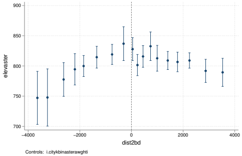


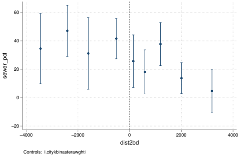
Sewers and population density, Colombia

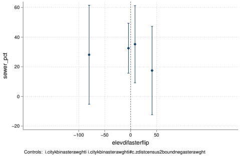
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. OLS									
sewer %	0.0152***	0.0152***	0.0152***	0.0147***	0.0147***	0.0147***	0.0128***	0.0128***	0.0128***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
2. First stage									
outside	-5.7843***		-1.8809***	-2.9270***		-0.4675	-2.9053***		-2.1532***
	(0.3648)		(0.4814)	(0.4256)		(0.5409)	(0.5009)		(0.6186)
ΔElev*Outside		-0.0851***	-0.0734***		-0.0585***	-0.0555***		-0.0345***	-0.0169
		(0.0057)	(0.0072)		(0.0070)	(0.0087)		(0.0085)	(0.0105)
3. Reduced form pop density									
outside	0.0315		0.1806***	0.0373		0.1489***	0.0306		0.0539
	(0.0198)		(0.0293)	(0.0229)		(0.0336)	(0.0255)		(0.0389)
ΔElev*Outside		-0.0017***	-0.0028***		-0.0016***	-0.0025***		-0.0001	-0.0005
		(0.0003)	(0.0004)		(0.0004)	(0.0005)		(0.0005)	(0.0007)
4. IV									
sewer %	-0.0054	0.0198***	0.0155***	-0.0127	0.0272***	0.0249***	-0.0105	0.0025	-0.0061
	(0.0036)	(0.0033)	(0.0031)	(0.0086)	(0.0060)	(0.0058)	(0.0095)	(0.0134)	(0.0093)
F	292.2	545.7	283.4	55.05	155.7	78.38	47.87	36.74	26.75
N	36072	36072	36072	36072	36072	36072	36072	36072	36072
π-bins	Y	Y	Y	Y	Y	Y	Y	Y	Y
seg$\times$x	.	.	.	Y	Y	Y	.	.	.
π-bin$\times$x	.	.	.	.	.	.	Y	Y	Y

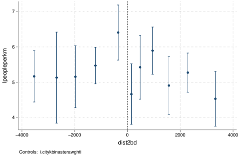
Note: Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Focus on (7), (8), (9). I think (8) is best $\Longrightarrow$ 1%=OLS$\geq$ IV $\approx$ 0%. Compare to Brazil. NB: First stage looks close.

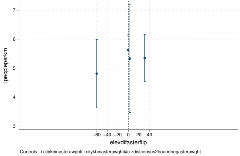
Identification, Jordan









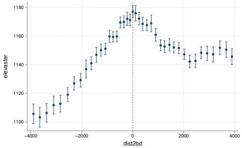


Sewers and population density, Jordan

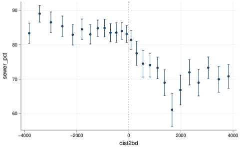
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. OLS									
sewer %	0.0217***	0.0217***	0.0217***	0.0221***	0.0221***	0.0221***	0.0233***	0.0233***	0.0233***
	(0.0043)	(0.0043)	(0.0043)	(0.0040)	(0.0040)	(0.0040)	(0.0056)	(0.0056)	(0.0056)
2. First stage									
outside	-20.0780***		-22.1156***	-17.7914**		-21.1823***	-15.6821		-16.1444
	(7.4818)		(7.6816)	(7.7331)		(7.9664)	(13.5692)		(14.2676)
ΔElev*Outside		-0.1628	0.1275		-0.0688	0.2411		-0.3605	0.0300
		(0.1718)	(0.1537)		(0.1936)	(0.1774)		(0.4895)	(0.4524)
3. Reduced form pop density									
outside	-0.4995		-0.3943	-0.4584		-0.3967	-0.1572		-0.1224
	(0.3132)		(0.3151)	(0.3091)		(0.2991)	(0.5409)		(0.6267)
ΔElev*Outside		-0.0118	-0.0066		-0.0102	-0.0044		-0.0052	-0.0023
		(0.0098)	(0.0098)		(0.0119)	(0.0119)		(0.0157)	(0.0166)
4. IV									
ewer %	0.0249	0.0722	0.0220	0.0258	0.1482	0.0204	0.0100	0.0145	0.0099
	(0.0152)	(0.0830)	(0.0136)	(0.0175)	(0.4370)	(0.0127)	(0.0311)	(0.0399)	(0.0311)
F	11.37	0.820	5.871	8.302	0.109	4.742	3.634	1.164	1.782
N	144	144	144	144	144	144	144	144	144
π-bins	Y	Y	Y	Y	Y	Y	Y	Y	Y
seg×x	.	.	.	Y	Y	Y	.	.	.
π-bin×x	.	.	.	.	.	.	Y	Y	Y

Note: Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
OLS=2%, Only (4) has reasonable first stage F and IV=OLS. More than COlumbia, less than Brazil?

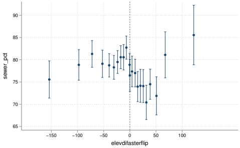
Identification, South Africa



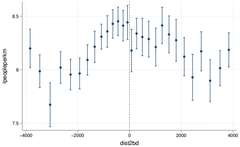
Controls: i.citykbinasterawghti



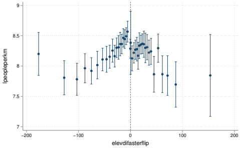
Controls: i.citykbinasterawghti



Controls: i.citykbinasterawghti i.citykbinasterawghti#c.zdistcensus2boundnegasterawght



Controls: i.citykbinasterawghti



Controls: i.citykbinasterawghti i.citykbinasterawghti#c.zdistcensus2boundnegasterawght

Sewers and population density, South Africa

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. OLS									
sewer %	-0.0009	-0.0009	-0.0009	-0.0008	-0.0008	-0.0008	-0.0004	-0.0004	-0.0004
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0010)	(0.0010)	(0.0010)
2. First stage									
outside	-11.0076***		-9.6987***	-7.2291***		-8.2255***	-5.1822***		-6.3499***
	(0.9269)		(1.0645)	(1.1801)		(1.2194)	(1.1769)		(1.3017)
ΔElev*Outside		-0.1298***	-0.0318*		-0.0138	0.0469*		0.0027	0.0557*
		(0.0162)	(0.0180)		(0.0235)	(0.0240)		(0.0271)	(0.0295)
3. Reduced form pop density									
outside	-0.0675		-0.2640***	-0.2087***		-0.2579***	-0.2083***		-0.3193***
	(0.0511)		(0.0719)	(0.0626)		(0.0733)	(0.0720)		(0.0848)
ΔElev*Outside		0.0021**	0.0048***		0.0004	0.0023		0.0026	0.0053**
		(0.0011)	(0.0014)		(0.0016)	(0.0018)		(0.0019)	(0.0022)
4. IV									
sewer %	0.0061	-0.0162*	0.0031	0.0289***	-0.0297	0.0308***	0.0402**	0.9684	0.0508***
	(0.0047)	(0.0084)	(0.0048)	(0.0101)	(0.1223)	(0.0099)	(0.0171)	(9.7682)	(0.0174)
F	140.5	78.68	71.68	34.90	0.365	19.24	16.46	0.0113	10.21
N	5544	5544	5544	5544	5544	5544	5544	5544	5544
π-bins	Y	Y	Y	Y	Y	Y	Y	Y	Y
seg×x	.	.	.	Y	Y	Y	.	.	.
π-bin×x	.	.	.	.	.	.	Y	Y	Y

Note: Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Identification comes from step in x. OLS=0. IV is suspicious and sensitive to specification.

Summary

- A 1% increase in sewer share gives about a 3% increase in population density in the Universe. This is about the same for both instruments, and double or triple the OLS effect. This conclusion has been robust to different sampling rules
- There is weak evidence for heterogeneity across countries. The treatment effect is smaller for Colombia, is around 4% for Brazil and South Africa, and Jordan is too small to allow us to say much.
- Caveat: Distance to CBD and distance to boundary. How to fix?

Sewers and mean income, Brazil

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
2. First stage									
outside	-4.2966***		-5.1857***	-2.4269***		-2.4486***	-2.3699***		-2.0924***
	(0.3870)		(0.4948)	(0.5034)		(0.6010)	(0.5398)		(0.6499)
ΔElev*Outside		-0.0270***	0.0200***		-0.0221***	0.0007		-0.0294***	-0.0089
		(0.0046)	(0.0058)		(0.0083)	(0.0099)		(0.0088)	(0.0106)
3. Reduced form pop income									
outside	-0.0611***		-0.1002***	-0.0059		-0.0066	-0.0136		-0.0098
	(0.0070)		(0.0089)	(0.0085)		(0.0102)	(0.0094)		(0.0116)
ΔElev*Outside		-0.0000	0.0009***		-0.0000	0.0000		-0.0002	-0.0001
		(0.0001)	(0.0002)		(0.0002)	(0.0002)		(0.0002)	(0.0002)
4. IV									
sewer %	0.0143***	0.0011	0.0161***	0.0025	0.0017	0.0025	0.0059	0.0075	0.0061
	(0.0019)	(0.0047)	(0.0019)	(0.0035)	(0.0076)	(0.0035)	(0.0040)	(0.0066)	(0.0040)
F	142.2	27.28	75.66	25.71	6.951	12.86	22.18	11.00	11.45
N	39131	39131	39131	39131	39131	39131	39131	39131	39131
π-bins	Y	Y	Y	Y	Y	Y	Y	Y	Y
seg×x	.	.	.	Y	Y	Y	.	.	.
π-bin×x	.	.	.	.	.	.	Y	Y	Y

Note: Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Who benefits in Brazil? A one percent increase in sewer share gives no increase in average income. This suggest that income sorting on sewer access is not important.

Sewers and mean income, South Africa

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
2. First stage									
outside	-4.2966***		-5.1857***	-2.4269***		-2.4486***	-2.3699***		-2.0924***
	(0.3870)		(0.4948)	(0.5034)		(0.6010)	(0.5398)		(0.6499)
ΔElev*Outside		-0.0270***	0.0200***		-0.0221***	0.0007		-0.0294***	-0.0089
		(0.0046)	(0.0058)		(0.0083)	(0.0099)		(0.0088)	(0.0106)
3. Reduced form pop income									
outside	-0.3312***		-0.1133***	-0.0654*		-0.0429	-0.0736*		-0.0183
	(0.0281)		(0.0366)	(0.0338)		(0.0360)	(0.0377)		(0.0428)
ΔElev*Outside		-0.0065***	-0.0053***		-0.0014*	-0.0011		-0.0028***	-0.0026**
		(0.0006)	(0.0007)		(0.0008)	(0.0008)		(0.0010)	(0.0012)
4. IV									
sewer %	0.0318***	0.0532***	0.0345***	0.0088**	0.0591	0.0066*	0.0136**	0.1834	0.0062
	(0.0029)	(0.0055)	(0.0030)	(0.0041)	(0.0492)	(0.0040)	(0.0063)	(0.2728)	(0.0062)
N	5520	5520	5520	5520	5520	5520	5520	5520	5520
F	132.8	72.60	67.54	39.43	1.092	20.95	19.10	0.379	10.45
kbins	Y	Y	Y	Y	Y	Y	Y	Y	Y
segment	.	.	.	.	.	.	.	.	.
segXx	.	.	.	Y	Y	Y	.	.	.
kbinXx	.	.	.	.	.	.	Y	Y	Y

Note: Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

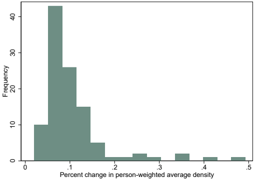
Who benefits in South Africa?

How important are sewers? I

- Add sewer connections for 1% of people to a city.
- Start with the densest census tract first, and work down to less dense tracts.
- Assume each 1% increase in sewer connections increases tract population by 3%.
- This gives a 3% increase in population
- Compare: Baum-Snow et al. [2017] find that radial highways in China have no impact on population.

How important are sewers? II

The effect on sewers on city average density (person weighted) is

much larger, 

- 1% of sewer connection often results in a 10% increases in person weighted density.
- Baum-Snow [2007] finds that each radial interstate highway decreased the density of US central cities by 9%.
- With a 5% agglomeration effect, this is 0.5% increase in city average wage.
- ...plus whatever wage increase is experienced by the 3% of new residents.

Conclusion I

- We've estimated the effects of sewer access on population density in a sample of developing world cities.
  - At 3%, the average effect seems large, both absolutely, and in comparison with (nearly) comparable estimates for highways.
  - We have two distinct identification strategies.
  - There is no evidence for resorting in response to sewer access (preliminary).
  - There is a lot of cross-country heterogeneity (preliminary).
  - TBD. Work out econometrics that allow for heterogenous treatment effects.

Bibliography I

Nava Ashraf, Edward Glaeser, Abraham Holland, and Bryce Millett Steinberg. Water, health and wealth. Technical report, National Bureau of Economic Research, 2017.

Sonia R Bhalotra, Alberto Diaz-Cayeros, Grant Miller, Alfonso Miranda, and Atheendar S Venkataramani. Urban water disinfection and mortality decline in lower-income countries. *American Economic Journal: Economic Policy*, 13(4):490–520, 2021.

Sebastian Galiani, Paul Gertler, and Ernesto Schargrodsky. Water for life: The impact of the privatization of water services on child mortality. *Journal of Political Economy*, 113(1):83–120, 2005.

Marcella Alsan and Claudia Goldin. Watersheds in child mortality: The role of effective water and sewerage infrastructure, 1880–1920. *Journal of Political Economy*, 127(2):586–638, 2019.

Bibliography II

D Mark Anderson, Kerwin Kofi Charles, and Daniel I Rees. Public health efforts and the decline in urban mortality. Technical report, National Bureau of Economic Research, 2018.

Joseph P Ferrie and Werner Troesken. Water and chicago's mortality transition, 1850–1925. *Explorations in Economic History*, 45(1):1–16, 2008.

Shanti Gamper-Rabindran, Shakeeb Khan, and Christopher Timmins. The impact of piped water provision on infant mortality in brazil: A quantile panel data approach. *Journal of Development Economics*, 92(2):188–200, 2010.

Michael Coury, Toru Kitagawa, Allison Shertzer, and Matthew Turner. The value of piped water and sewers: Evidence from 19th century chicago. Technical report, National Bureau of Economic Research, 2022.

Bibliography III

Brazilian Institute of Geography and Statistics. Brazil demographic census 2010. Technical report, 2012. Rio de Janeiro, Brazil, 2012.

National Administrative Department of Statistics. Colombia population and housing census 2018. Technical report, 2018. Bogotá, Colombia: National Administrative Department of Statistics.

National Bureau of Statistics (Tanzania), Office of the Chief Government Statistician. Tanzania population and housing census 2012. Technical report, 2012. National Bureau of Statistics (Tanzania), Office of the Chief Government Statistician (Zanzibar).

Department of Statistics (Jordan). Jordan population and housing census 2015. Technical report, 2015. Department of Statistics, Aman Jordan.

Bibliography IV

Evelyn Uuemaa, Sander Ahi, Bruno Montibeller, Merle Muru, and Alexander Kmoch. Vertical accuracy of freely available global digital elevation models (aster, aw3d30, merit, tandem-x, srtm, and nasadem). *Remote Sensing*, 12(21):3482, 2020.

UN DESA Population Division. World urbanization prospects: the 2018 revision, 2018.

Nathaniel Baum-Snow, Loren Brandt, J Vernon Henderson, Matthew A Turner, and Qinghua Zhang. Roads, railroads, and decentralization of chinese cities. *Review of Economics and Statistics*, 99(3):435–448, 2017.

Nathaniel Baum-Snow. Did highways cause suburbanization? *The quarterly journal of economics*, 122(2):775–805, 2007.

Abstract: We investigate the effect of sewer access on neighborhood density and demographic composition in a sample of large cities in Brazil, Colombia, South Africa and Jordan. Sanitary sewers rely on gravity to operate, and it is difficult and expensive to move sewage uphill. This means that otherwise similar census tracts, close to, but on opposite sides of a drainage basin divide may have different costs of accessing a sewer network. We compare such tracts to learn about the effects of sewer access on neighborhood density and demographics. We find that a 1% increase in the share of households with sewer access causes about a 2% increase in population density. We find preliminary evidence suggesting economically important heterogeneity in the size of this effect across countries, and no evidence to suggest that people resort in response to changes in sewer access. The effect of sewers on density appears to be economically large, and to be large relative to the effects of highways.