

# Loss in the Time of Cholera

Ambrus, A., Field, E., & Gonzalez, R. (2020). Loss in the time of cholera: Long-run impact of a disease epidemic on the urban landscape. *American Economic Review*, 110(2), 475-525.

- Loss in the Time of Cholera
  - Authors
    - Attila Ambrus
    - Erica Field
    - Robert M. Gonzalez
  - Research Question
    - Common sense
  - Literature Review & Main Contribution
  - Background
    - Two main reasons for focusing on this particular event
  - Empirical Strategy
    - A. Catchment Area Boundary
    - B. RD Specification
    - C. Validity of RD Design
  - Results
    - 1. Impacts on Cholera Exposure
    - 2. Short-run Impacts on property prices
    - 3. Short-run Impacts on other attributes
      - Migration
      - Residential Crowding
    - 4. Long-run Impacts on property prices & socioeconomic status
    - 5. Robustness Checks
      - BSP Boundary Definition
      - Falsification Tests
    - 6. Theoretical Framework
      - Some features to exclude confounding channels
      - Three key elements of "Tipping"
  - Reading Notes

## Authors

## Attila Ambrus

- Professor of Economics, Duke University
- **Research Interest:** Microeconomic Theory; Game Theory; Experimental Economics; Development Economics; Industrial Organization
- **Publication:** Top 5 (Except ECMA); Many JPubE; JET; ...

## Erica Field

- Professor of Economics and Global Health, Duke University
- **Research Interest:** Development Economics; Labor Economics; Health Economics

## Robert M. Gonzalez

- Assistant Professor, Department of Economics, University of South Carolina
- **Research Interest:** The role of information-communication technologies (ICTs), the use of novel GIS data and methods to study conflict, crime, and public health...

## Research Question

Can disease exert a permanent effect on the geography of urban poverty?

## Common sense

- Residential migration will preserve the spatial distribution of income in the long run, erasing such shocks from the map over time.

Yet, in reality, **spatial discontinuities in urban land values are frequently observed and do not always appear related to discrete changes in local amenities.**

## Literature Review & Main Contribution

1. *A large literature in economic geography on the long-run persistence of income differences across space:*

- There are a handful of papers that show evidence of persistent income differences across cities or towns even long after specific sources of economic advantage have become obsolete.

- Channels: changes in population growth that accompanied economic development or economic shocks; the path dependence that demographic trends created via economies of scale in industry; the lens of institutions ...
- Not applicable within the parish setting, which constitutes a single economic and institutional environment.
- **This paper's contribution:** provide an alternative means through which temporary and localized economic shocks can lead to long-term changes in the **spatial distribution of poverty**.  
⇒ also broaden the channels of **residential sorting** through which we might observe persistent effects of historic shocks on the long-run economic growth of communities than has been considered in the literature.

## 2. *A large literature documenting and modeling neighborhood externalities in real estate values:*

- Our findings also illustrate the potential economic cost of spatially correlated shocks when there are significant externalities from neighbors' socioeconomic status;

## 3. *Our theoretical model is closely related to spatial models of location choice and segregation, but different:*

- **First**, while agents in the models above follow simple behavioral rules, in our model they are fully forward-looking utility-maximizers;
- **Second**, in our model rent-setting landlords coordinate the movement of tenants in and out of the block.
- **Third**, instead of a self-contained city, our model features a block situated in an open world where tenants can move in and out.
- **The focus of our paper is also different:** most of our focus is on characterizing the initial conditions under which the block converges to poverty.

## 4. Technically, our paper is related to asynchronous-move dynamic games, and the role of asynchronicity of moves in coordination problems.

# Background

We examine this question in the context of a cholera epidemic that hit a single urban parish of London in 1854.

Over the course of one month, 660 residents living in the 0.5-mile radius of St. James Parish died from cholera, implying that roughly 5 percent of families were suddenly impoverished because of the loss of a potential wage earner.

The outbreak was eventually attributed to contaminated groundwater that leached into one of the 13 wells serving the parish. As a result, the impact was concentrated in one particular neighborhood wherein a much higher fraction of families, 13 percent, experienced a loss.

- In this paper, we test empirically whether the cholera outbreak had a causal effect on neighborhood poverty, as measured by real estate prices, soon and long after the epidemic ended.

## **Two main reasons for focusing on this particular event**

1. Unique natural experiment: sharp changes in death rates at the boundary of the catchment area for the contaminated well, where other property attributes such as access to public goods can be assumed to change smoothly.
2. Very detailed microdata on deaths and property characteristics are available from this particular neighborhood at the time of the epidemic.

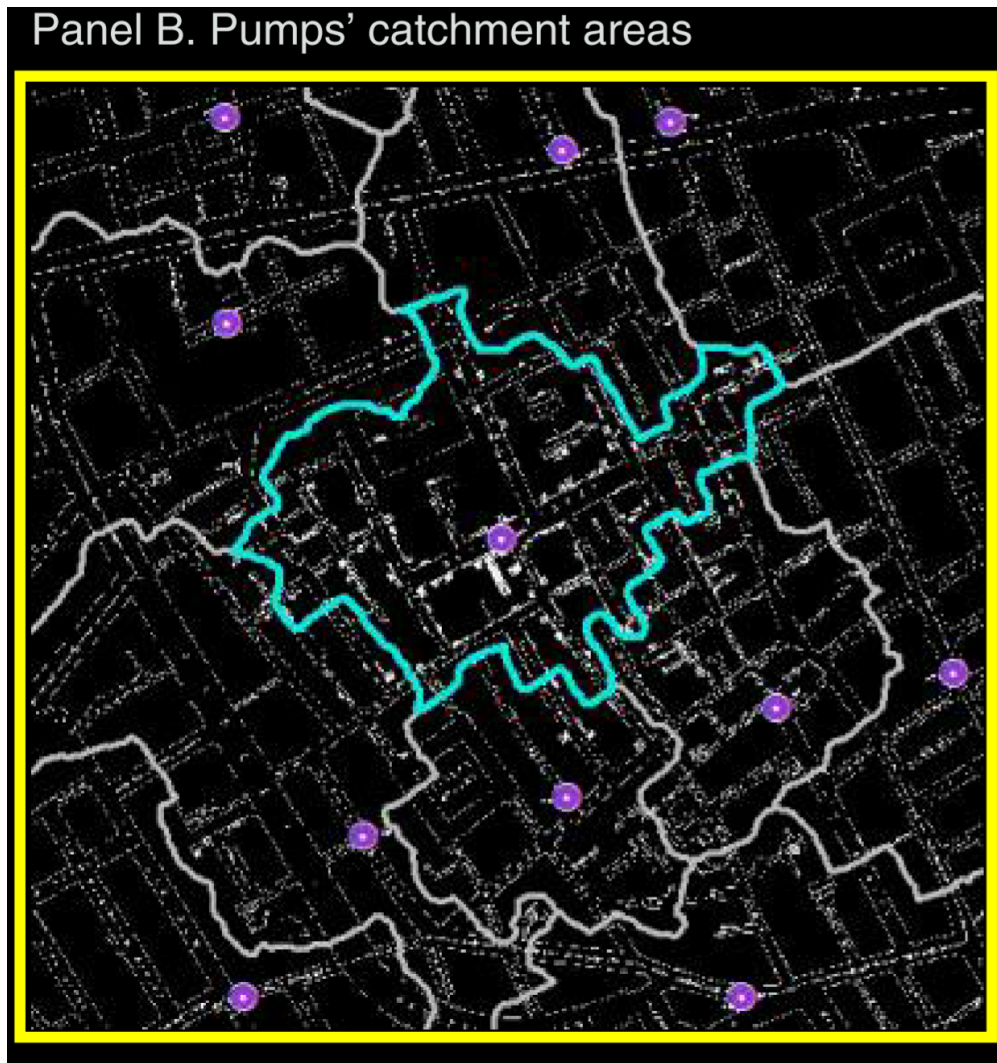
## **Empirical Strategy**

### **A. Catchment Area Boundary**

We define the catchment area according to a network Voronoi diagram of St. James, in which each of the 13 water pumps defines a point and the cells are determined according to the walking distance by road to each point.

⇒

- The resulting catchment area has a total area of 0.07 square kilometers and a boundary of about 1,161 meters.
- 76 percent of cholera deaths occurred within the catchment area.



## B. RD Specification

Specifying a **one-dimensional** forcing variable: distance

$$y_{it} = \alpha + \gamma BSP_i + f(X_i) + \mathbf{W}'_{it}\beta + \epsilon_{it}, \quad X_i < h,$$

where  $y$  is a measure of property value,  $BSP$  is dummy,  $X_i$  is distance.

- Two types of  $f(\cdot)$ :
  - $f(X_i) = \delta_0 X_i + \delta_1 BSP_i * X_i$
  - $f(X_i) = \sum_{k=1}^K \delta_k X_i^k$
- Covariates: balance test

## C. Validity of RD Design

Examining the similarity across the boundary of neighborhood features, including rental price and neighborhood amenities.

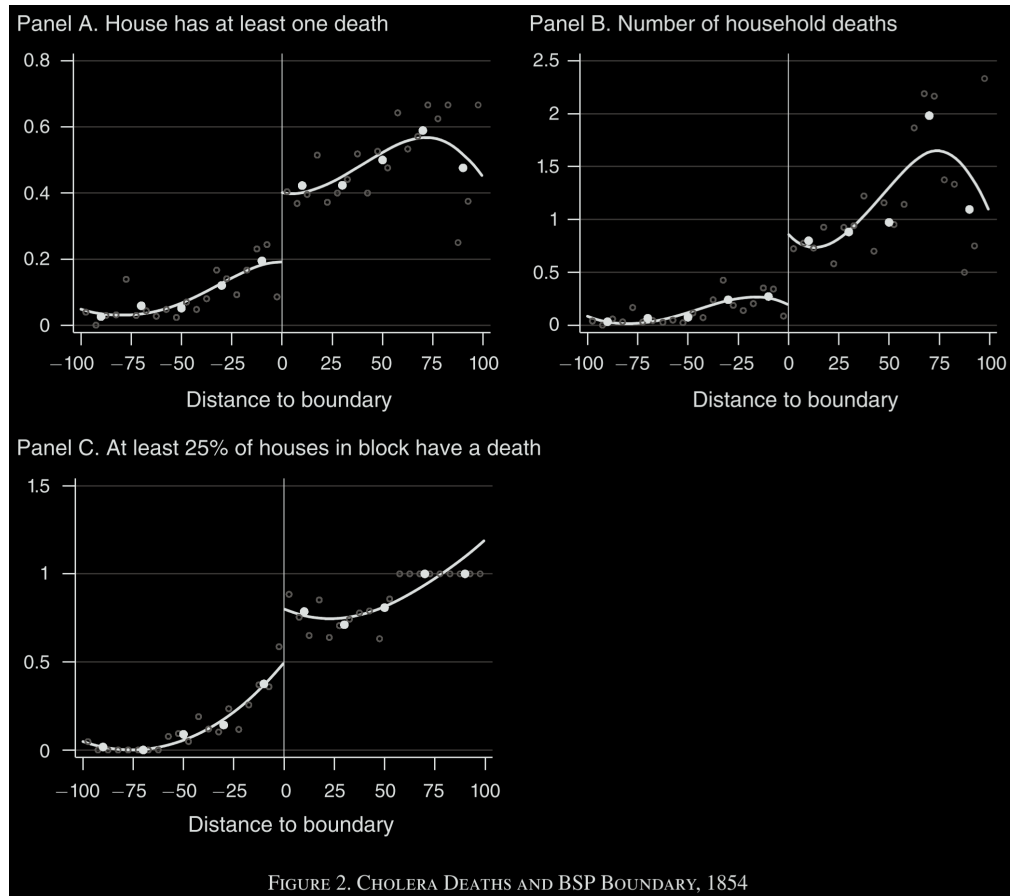
⇒ Differences between catchment area and noncatchment area properties decrease significantly in magnitude, as expected.

TABLE 1—HOUSE CHARACTERISTICS, 1853								
	Full sample			Within 100 m			RD estimates	
	Inside (1)	Outside (2)	SE (3)	Inside (4)	Outside (5)	SE (6)	RD coefficient (7)	SE (8)
Rental characteristics:								
Rental price (in logs)	3.713	3.775	(0.057)	3.709	3.740	(0.066)	0.062	(0.147)
Tax assessed (in logs)	0.446	0.515	(0.057)	0.442	0.495	(0.064)	0.044	(0.152)
Tax exonerated (yes = 1)	0.0648	0.231	(0.040)	0.0653	0.222	(0.052)	−0.034	(0.056)
Sewer access:								
Old/existing	0.472	0.565	(0.085)	0.467	0.593	(0.091)	−0.067	(0.189)
New sewer	0.401	0.275	(0.082)	0.404	0.259	(0.084)	0.248	(0.151)
No access	0.128	0.160	(0.054)	0.129	0.148	(0.061)	−0.181	(0.124)
Distance (m/100) to:								
Closest pump	1.045	0.958	(0.079)	1.051	1.065	(0.094)	0.034	(0.098)
Soho centroid	1.319	2.472	(0.118)	1.323	2.180	(0.130)	−0.152	(0.275)
Presumed plague pit	2.359	3.137	(0.224)	2.358	2.630	(0.216)	0.211	(0.401)
Public square	2.586	2.715	(0.137)	2.583	2.709	(0.141)	−0.190	(0.356)
Church	1.311	1.717	(0.130)	1.311	1.609	(0.142)	−0.184	(0.275)
Police station	4.376	5.412	(0.264)	4.379	4.792	(0.225)	0.264	(0.552)
Fire station	3.601	2.665	(0.185)	3.596	2.854	(0.224)	0.389	(0.458)
Theater	4.020	5.303	(0.236)	4.027	4.680	(0.225)	0.043	(0.512)
Pub	0.286	0.407	(0.037)	0.287	0.405	(0.046)	−0.158	(0.121)
Urinal	0.874	1.123	(0.087)	0.870	1.020	(0.088)	−0.282	(0.207)
Sewer vent	0.426	0.556	(0.047)	0.427	0.563	(0.050)	−0.166	(0.142)
Primary school	1.296	2.477	(0.132)	1.296	2.025	(0.110)	−0.368	(0.282)
Bank	3.960	4.694	(0.317)	3.958	4.086	(0.316)	0.223	(0.639)
Observations	494	1,228		490	811		559	

# Results

## 1. Impacts on Cholera Exposure

TABLE 2—CHANGE IN EXPOSURE TO CHOLERA AT BOUNDARY OF BSP CATCHMENT AREA					
	Local linear regression		Polynomial RD		
	Baseline (1)	Controls (2)	Optimal band (3)	Wide band (4)	Segment FE (5)
<i>Panel A. Number of deaths in household</i>					
Inside BSP area	0.548 (0.150)	0.567 (0.149)	0.378 (0.120)	0.639 (0.110)	0.463 (0.111)
Observations	477	485	477	1,284	1,284
Mean outside BSP	0.253	0.260	0.253	0.160	0.160
Bandwidth (meters)	25.00	25.21	25.00	100	100
<i>Panel B. House had at least one death</i>					
Inside BSP area	0.215 (0.068)	0.185 (0.068)	0.138 (0.043)	0.253 (0.034)	0.174 (0.035)
Observations	538	540	538	1,284	1,284
Mean outside BSP	0.172	0.172	0.172	0.0989	0.0989
Bandwidth (meters)	28.49	28.65	28.49	100	100



## 2. Short-run Impacts on property prices

Houses inside the catchment area suffer a roughly 15 percent loss in rental value within a decade of the epidemic.

TABLE 3—BOUNDARY EFFECTS ON RENTAL PRICES (1853, 1864, 1894, 1936)					
	Local linear regression		Polynomial RD		
	Baseline (1)	Controls (2)	Optimal band (3)	Wide band (4)	Segment FE (5)
<i>Panel A. log rental prices, 1853</i>					
Inside BSP area	0.052 (0.124)	0.035 (0.078)	−0.021 (0.074)	−0.041 (0.073)	−0.079 (0.070)
Observations	588	469	588	1,070	1,070
Mean outside BSP	47.01	46.01	47.01	48.63	48.63
Bandwidth (meters)	35.72	27.25	35.72	100	100
<i>Panel B. log rental prices, 1864</i>					
Inside BSP area	−0.186 (0.118)	−0.186 (0.089)	−0.113 (0.068)	−0.118 (0.068)	−0.127 (0.068)
Observations	469	511	469	1,047	1,047
Mean outside BSP	48.48	47.74	48.48	50.24	50.24
Bandwidth (meters)	28.11	31.38	28.11	100	100

Sharp RD vs. Fuzzy RD:

- Treatment is not strictly defined by the BSP boundary (Table 3).  $\Rightarrow$  Fuzzy RD

$$\begin{aligned}
 y_{it} &= \alpha^f + \gamma^f BSP_{fuzzy,i} + f(X_i) + \mathbf{W}'_{it}\beta^f + \epsilon_{it}, & X_i < h, \\
 BSP_{fuzzy,i} &= \delta + \tau BSP_i + g(X_i) + \mathbf{W}'_{it}\mu + \nu_{it}, & X_i < h,
 \end{aligned}$$

where  $BSP_{fuzzy,i} = 1$  if house  $i$  experiences at least one cholera death.

Question here: why aren't estimates using fuzzy RD and IV exactly equivalent!

	Fuzzy RD estimates			IV estimates	
	LLR	Polynomial RD			
	(1)	(2)	(3)	(4)	(5)
At least one death	−0.799 [−0.550] (0.483)	−0.753 [−0.529] (0.392)	−0.767 [−0.536] (0.719)	−0.694 [−0.501] (0.326)	
Number of deaths					−0.293 [−0.254] (0.141)
Observations	512	512	1,024	467	467
Mean outside BSP	48.47	48.47	50.24	48.23	48.23
Bandwidth (meters)	33.02	33.02	100	29.20	29.20
First-stage $F$ -statistic	—	—	—	10.58	11.34
Sargan-Hansen $p$ -value	—	—	—	0.156	0.161

### 3. Short-run Impacts on other attributes

#### Migration

We anticipate that both impoverishment from cholera and also the change in the quality of the neighborhood would have driven residents out.

#### Interesting things:

- The exit is not just driven by mortality, but also neighborhood deaths. (Table 4)
- Two working effects: (一句话：没死人的家庭走了，死人的家庭留下来了)
  - Externalities from poor neighbors drive people away;
  - Rental price responses to neighborhood impoverishment increased the affordability of housing in affected areas.

	Resident is different in 1864 than in 1853					log rental price, 1864	
	Polynomial RD						
	LLR (1)	Optimal band (2)	Wide band (3)	Optimal band (4)	Optimal band (5)	Optimal band (6)	Optimal band (7)
Inside BSP area	0.075 (0.083)	0.101 (0.046)	0.103 (0.040)	0.094 (0.049)	0.010 (0.066)	−0.160 (0.070)	−0.044 (0.072)
Household death				0.072 (0.044)	0.053 (0.043)	−0.075 (0.049)	−0.047 (0.049)
Neighborhood deaths					0.003 (0.002)		−0.005 (0.003)
Houses in neighborhood					−0.002 (0.001)		0.004 (0.002)
Observations	700	915	1,312	894	894	455	455
Mean outside BSP	0.594	0.517	0.515	0.517	0.517	0.594	0.594
Bandwidth (meters)	37.98	52.82	100	52.82	52.82	28.11	28.11



TABLE 5—MIGRATION PATTERNS BY CHOLERA EXPOSURE: BSP CATCHMENT AREA		
	Resident is different in 1864 than in 1853	
	(1)	(2)
Household death	0.271 (0.083)	0.330 (0.093)
Neighborhood deaths	0.004 (0.001)	
Household death $\times$ neighborhood deaths	−0.005 (0.002)	
Neighborhood houses with death		0.010 (0.003)
Household death $\times$ neighborhood houses with death		−0.015 (0.004)
Total number of houses in neighborhood	−0.001 (0.001)	−0.001 (0.002)
Number of deaths in household	0.014 (0.020)	0.014 (0.020)
log total sums assessed, 1853	−0.129 (0.077)	−0.130 (0.076)
Observations	489	489
Mean outside BSP	0.504	0.504

## Residential Crowding

TABLE 6—BOUNDARY EFFECTS ON HOUSE OCCUPANCY CHARACTERISTICS						
	Number of occupants at address		Number of immigrant families at address		Proportion of immigrant families at address	
	LLR (1)	Optimal band (2)	LLR (3)	Optimal band (4)	LLR (5)	Optimal band (6)
<i>Panel A. Census data (1851)</i>						
Inside BSP area	2.591 (2.334)	3.131 (2.997)	0.089 (0.141)	0.085 (0.098)	0.060 (0.039)	0.020 (0.026)
Observations	490	490	556	556	515	515
Mean outside BSP	13.502	13.502	0.324	0.324	0.095	0.095
Bandwidth (meters)	34.81	34.81	34.07	34.07	30.94	30.94
<i>Panel B. Census data (1861)</i>						
Inside BSP area	4.725 (2.425)	4.900 (3.891)	0.239 (0.189)	0.211 (0.104)	0.093 (0.052)	0.085 (0.041)
Observations	524	524	482	482	501	501
Mean outside BSP	13.62	13.62	0.391	0.391	0.126	0.126
Bandwidth (meters)	37.38	37.38	32.01	32.01	33.86	33.86

## 4. Long-run Impacts on property prices & socioeconomic status

TABLE 7—BOUNDARY EFFECTS ON HOUSE SOCIOECONOMIC STATUS, 1899								
	Very poor		Poor		Working poor		Middle class	
	LLR (1)	Optimal band (2)	LLR (3)	Optimal band (4)	LLR (5)	Optimal band (6)	LLR (7)	Optimal band (8)
Inside BSP area	0.092 (0.050)	0.142 (0.092)	0.043 (0.047)	−0.012 (0.051)	0.067 (0.044)	0.035 (0.103)	−0.236 (0.062)	−0.222 (0.100)
Observations	523	597	581	649	571	503	483	493
Mean outside BSP	0.220	0.233	0.250	0.232	0.378	0.347	0.164	0.168
Bandwidth (meters)	25.37	30.17	29.19	34.34	28.41	24.26	22.37	23.24

TABLE 3—BOUNDARY EFFECTS ON RENTAL PRICES (1853, 1864, 1894, 1936)					
	Local linear regression		Polynomial RD		
	Baseline (1)	Controls (2)	Optimal band (3)	Wide band (4)	Segment FE (5)
<i>Panel C. log rental prices, 1894</i>					
Inside BSP area	−0.255 (0.234)	−0.278 (0.147)	−0.242 (0.119)	−0.218 (0.076)	−0.217 (0.115)
Observations	370	365	370	799	794
Mean outside BSP	119.41	119.19	119.41	121.11	121.11
Bandwidth (meters)	29.19	28.54	29.19	100	100
<i>Panel D. log rental prices, 1936</i>					
Inside BSP area	−0.299 (0.311)	−0.375 (0.280)	−0.325 (0.147)	−0.458 (0.144)	−0.271 (0.150)
Observations	221	221	221	354	354
Mean outside BSP	454.49	454.49	454.49	451.43	451.43
Bandwidth (meters)	37.26	37.30	37.26	100	100

TABLE 8—BOUNDARY EFFECTS ON HOUSE PRICES, ZOOPLA HOUSE VALUE ESTIMATES, AND RENTAL PRICES, 1995–2013, 2015						
	House prices and Zoopla estimates		Zoopla estimates only		House rental prices	
	LLR (1)	Optimal band (2)	LLR (3)	Optimal band (4)	LLR (5)	Optimal band (6)
Inside BSP area	−0.290 (0.097)	−0.383 (0.22)	−0.324 (0.15)	−0.293 (0.15)	−0.248 (0.21)	−0.150 (0.069)
Observations	173	163	115	104	150	150
Mean outside BSP	1,005	1,005	1,049	1,049	746.9	746.9
Bandwidth (meters)	26.60	24.15	31.11	30.75	49.10	41.15

# 5. Robustness Checks

## BSP Boundary Definition

- Using the alternative definition of the BSP catchment area boundary proposed by Snow (Snow 1855).

## Falsification Tests

- It is possible that **cross-block price differences of the magnitude of our estimates are sufficiently common** (confound the boundary effect) in this setting to raise concerns about the validity of our interpretation.

⇒ **We examine the distribution of differences in average house prices** across neighboring cells of a grid covering the Greater London area.

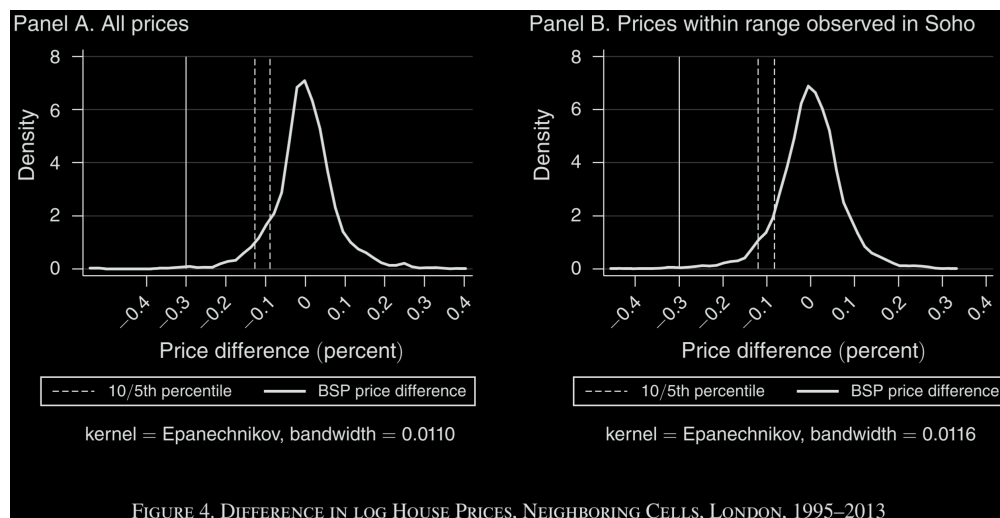
- If sharp differences in prices across neighboring blocks are common, the differences across neighboring blocks within the bandwidth of our estimates should fall near the center of such distributions.

### Three Steps:

**Step 1:** we overlay geocoded addresses on a grid covering the Greater London area with each graticule cell having a perimeter equal to Soho's perimeter;

**Step 2:** within each of these cells, we randomly draw 40 pairs of subcells with length equal to the optimal bandwidth used to calculate boundary effects on contemporary real estate prices (43 meters);

**Step 3:** for each pair, we randomly assign one of the two subcells into treatment and calculate the average difference in log prices between treated and nontreated subcells for each Soho-sized cell.



## 6. Theoretical Framework

The empirical findings suggest a “**tipping**” story, according to which

- if a shock to a block is severe enough, it pushes it over a threshold, and the block tips into an equilibrium with relatively poorer tenants;
- if the shock is moderate, the block converges back to the original equilibrium.

### Some features to exclude confounding channels

- The shock only directly affected people, not infrastructure;
- The small size of the affected area and that affected blocks are in close proximity and relatively intertwined with unaffected blocks;
- Apartments in the affected blocks were in the hands of landlords who by screening tenants via prices could potentially control the composition of blocks.

### Three key elements of "Tipping"

tipping的理解：稍微倾斜会恢复，倾斜过度就会倒。

- A block-level negative externality from living near poor neighbors;
- A higher willingness to pay to avoid this externality for the rich than the poor;
- Some friction that imposes a cost on landlords when they change tenants in a block.

When the **frictions associated with changing tenants are not too small and the landlords' discount factor is not too high**, although the landlords could make pricing decisions that would take back the block to its original tenant composition, if the initial shock is too severe, it would be too costly for them, and instead it is optimal for them to let the block transition to having poorer tenants.

- The main direct effect of the cholera epidemic was increasing the share of poor tenants in affected blocks through the death of wage-earners.
- There are contracting frictions (landlords have to wait until existing tenants' contracts expire before kicking them out) and that in the meantime they have to offer discounts for rich tenants to be willing to stay in a block with poor neighbors.

## Reading Notes

一般认为，局部地理范围内的收入冲击只会影响一时，人口流动最终会抹平这种影响。然而在现实中，却经常可以观察到，相邻地区的土地价值长期存在差异。Attila Ambrus, Erica Field和Robert Gonzalez三位作者2020年在AER上发表的论文 *Loss in the Time of Cholera: Long-run Impact of a Disease Epidemic on the Urban Landscape* 使用伦敦小范围瘟疫爆发的事件，检验了局部范围内的收入冲击对当地房价的长期影响及其形成机制。

1854年，伦敦爆发了一场小范围霍乱，造成当地大量居民死亡，霍乱的病原体来自水源，靠近受污染水泵的街区病情远比其他地方严重，于是作者利用空间模糊断点回归，考察了霍乱对房价的影响。首先，在短期内，霍乱显著降低了房价，居民迁出本地的概率下降，居民区变得更拥挤。利用1899年至2015年间不同时间段的房价数据，作者检验了霍乱对房价的长期影响，发现当年靠近该水泵的街区仍然有着更低的房价，当地居住的穷人更多、富人更少。

随后，作者探索了这种长期影响的形成机制，一些家庭在霍乱中失去了劳动力，因此选择将房屋租给更多人居住，由此带来了拥挤等负外部性，让富人避之不及。作者用模型刻画了这一过程，相较于降低租金吸引富人入住，将房子持续租给穷人是房东的占优决策，于是从长期来看，房价错配、贫困集聚的情况再难逆转。