Natural Amenities, Neighborhood Dynamics and Persistence in the Spatial Distribution of Income

Sanghoon Lee   Jeffrey Lin

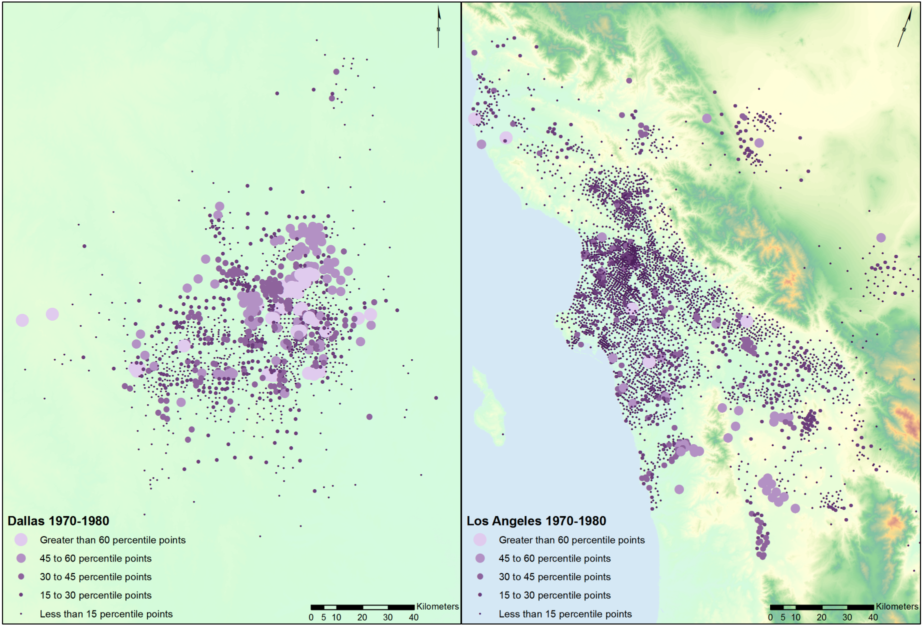
Sauder School of Business, University of British Columbia

Federal Reserve Bank of Philadelphia

November 3, 2015

Neighborhood change

- 2/3 changed income quartile, 1950–2000; 1/5 top $\leftrightarrow$ bot quartile
- Many policies aim to preserve or change neighborhood status
- Patterns of change vary across neighborhoods and even cities



Absolute change in percentile rank by average household income, 1970–1980.
Mean absolute change in Dallas=21pp; Los Angeles=9pp. Neighborhood and metropolitan boundaries held constant.

Why is the spatial distribution of income persistent
for some neighborhoods and cities while others turn over frequently?

- Theory and evidence highlighting the role of persistent natural amenities in neighborhood dynamics and the internal stability of cities
- Persistent natural amenities "anchor" neighborhoods to high incomes as they experience various shocks or interventions over time
- For cities as a whole, the presence of a strong natural amenity slows down neighborhood change, suburbanization and tipping patterns

What we do

A stylized, dynamic theory of neighborhood choice generating key predictions

- Natural and endogenous amenities
- Sorting by income on overall QOL (aggregate amenities)
- Multiple equilibria in each period; selection determined by history
- Amenity shocks that can (might) reverse historical equilibrium

What we do (2)

Empirical results consistent with theory, using consistent-boundary tracts, 1880–2010

1. Neighborhood anchors: Conditioned on current income, neighborhoods with superior natural amenities increase in income more than other neighborhoods
2. Stable cities: Naturally heterogeneous cities (e.g., coastal or hilly cities) have internal spatial distributions of income that are stable over time
3. Anchored downtowns: Cities with coastal downtowns experienced slower suburbanization before 1980 and faster gentrification after 1980

Identification issues

- Unknown or changing natural amenity values
- Role of endogenous factors in reinforcing these patterns
- Growing cities also feature greater churning in the spatial distribution of income

Literature Review

Many, many explanations of neighborhood change

- Filtering of old houses: Brueckner & Rosenthal '09
  Dynamic spillovers: Aaronson '01, Guerreri, Hartley & Hurst '10
  Highways: LeRoy & Sonstelie '83, Baum-Snow '07
  Black migration to northern cities: Boustan '10
  Endogenous amenities: Carlino & Saiz '08
- Evidence for other mechanisms stronger in "flat" cities
- Few papers note that neighborhood change varies across cities and is related to natural heterogeneity (cf. Burchfield et al. '06)
- We analyze a broad set of cities back to 1880

Literature Review (2)

Many (static) theories of sorting (with externalities and multiple equilibria)

- Witherstine '26: "Where the Romans lived in the first century B.C."
  "Certain districts [were] favored more than others; some, because they are accessible; and others, because they are beautiful in themselves or command a fine view. The Aventine, Caelian, Palatine, even the Sacred Way and the Subura, the Carinae, the Esquiline, Quirinal, Viminal, Pincian, the Campus Martius, the Capitoline and the district beyond the Tiber—all these furnish sites for private homes"
- Tiebout '56, Epple & Sieg '99
- Fujita & Ogawa '82, Bond & Coulson '89,
  Brueckner, Thisse & Zenou '99, Redding, Sturm & Wolf '11
- We extend static intuition to a dynamic setting
- We empirically test these implications

Literature Review (3)

Growing evidence of long-run spatial persistence (sometimes)

- Japan: Davis & Weinstein '02, '08
  US: Rappaport & Sachs '03, Bleakley & Lin '12
  Latin America: Maloney & Valencia-Caicedo '15
  Britain vs France: Michaels & Rauch '14
- Los Angeles: Brooks & Lutz '14
  Manhattan: Villareal '14
  San Francisco: Siodla '15
- We focus on within-city distribution of income
- We offer an explanation for differences in persistence across regions

A model of neighborhood choice

- A stylized 2-neighborhood city
- Abstract from many neighborhoods, city growth, housing demand, correlated shocks
- Paper relaxes these assumptions and shows robustness of main implications
- Three testable implications

Geography and households

- A closed city with two neighborhoods $j = b, d$ (beach, desert)
- Each one unit measure of land, owned by absentee landlords

- Two unit measure of workers, heterogenous in income θ
- Consume one unit of land in chosen neighborhood and numeraire c
- Choose neighborhood $j = b, d$ in each t to maximize utility
- No moving costs, no savings: no dynamic problem
- In each t, worker of income θ solves:

$$\max_j A_{j,t} \cdot c_{j,t} \text{ subject to } c_{j,t} + R_{j,t} = \theta$$

Neighborhood aggregate amenity values $A_{j,t}$ determined in each t

Beach	Desert
$A_{b,t} \equiv \alpha_b + \bar{w}_{b,t} + \epsilon_{b,t}$	$A_{d,t} \equiv \alpha_d + \bar{w}_{d,t} + \epsilon_{d,t}$

1. $\alpha_b > \alpha_d$ — Persistent natural amenity advantage of the beach
2. $\bar{w}_{j,t} \equiv E(\theta|j,t) = \bar{w}_H$ or $\bar{w}_L$ — Endogenous amenity value that depends on average income e.g., school quality, safety, shopping (source of multiplicity)
3. $\epsilon_{j,t} \sim G(-\infty, \infty)$ — Idiosyncratic shock to amenity value e.g., natural disasters, changes in local governance, other un-modeled factors (source of churning)

Two possible equilibrium states in each period

- Complementarity between A and c implies perfect sorting (SCP): high-income workers outbid low-income workers for high-amenity nbhd.
- Rents are determined so marginal worker is indifferent between neighborhoods.

$S1$: High-income households live at the beach only if

$$\alpha_b + \bar{w}_H + \epsilon_{b,t} > \alpha_d + \bar{w}_L + \epsilon_{d,t}$$

$S2$: High-income households live in the desert only if

$$\alpha_b + \bar{w}_L + \epsilon_{b,t} < \alpha_d + \bar{w}_H + \epsilon_{d,t}$$

- Existence: At least one of these states is always an equilibrium
- Multiplicity: Both states may be equilibria (e.g., if $a_b + \varepsilon_{b,t} = a_d + \varepsilon_{d,t}$)

History dependence and dynamics

- If both equilibria are possible, we select historical equilibrium (persistence)
- Since today's eqbm depends on yesterday, state follows a Markov process
- State switches back and forth, depending on realizations of $\epsilon_{j,t+1}$

$S2|S1$ if and only if S1 is no longer a possible equilibrium (i.e., ruled out)

$$\alpha_b + \bar{w}_H + \epsilon_{b,t+1} < \alpha_d + \bar{w}_L + \epsilon_{d,t+1}$$

$S1|S2$ if and only if S2 is no longer a possible equilibrium

$$\alpha_b + \bar{w}_L + \epsilon_{b,t+1} > \alpha_d + \bar{w}_H + \epsilon_{d,t+1}$$

Thus,

$$\Pr(S2|S1) = \Pr(\varepsilon_{d,t+1} - \varepsilon_{b,t+1} > a_b - a_d + \bar{w}_H - \bar{w}_L).$$

$$\Pr(S1|S2) = \Pr(\varepsilon_{b,t+1} - \varepsilon_{d,t+1} > a_d - a_b + \bar{w}_H - \bar{w}_L).$$

Lemma 1: $\Pr(S2|S1) < \Pr(S1|S2)$

The probability of transitioning from S2 to S1 is greater than the probability of transitioning from S1 to S2.

Lemma 2: $\Pr(S1|S2)$ increases with $a_b - a_d$

The probability of transitioning from S2 to S1 increases with natural heterogeneity.

Markov transition matrix:

$$M \equiv \begin{bmatrix} \Pr(S1|S1) & \Pr(S1|S2) \\ \Pr(S2|S1) & \Pr(S2|S2) \end{bmatrix} = \begin{bmatrix} 1 - \Pr(S2|S1) & \Pr(S1|S2) \\ \Pr(S2|S1) & 1 - \Pr(S1|S2) \end{bmatrix}$$

The steady state vector π is defined as: $\pi = M\pi$

Because M is a regular Markov matrix, the probability distribution over states converges to the steady state vector π.

$$\pi \equiv \begin{bmatrix} \Pr(S1)^* \\ \Pr(S2)^* \end{bmatrix} = \frac{1}{\Pr(S2|S1) + \Pr(S1|S2)} \begin{bmatrix} \Pr(S1|S2) \\ \Pr(S2|S1) \end{bmatrix}$$

Three testable implications

Cast in terms of income percentile ranks $r \in \{r_L, r_H\}$

1. Natural amenities are neighborhood anchors
2. Naturally heterogeneous cities are stable
3. Natural amenities are stronger anchors in naturally heterogeneous cities

1. Natural amenities are neighborhood anchors

Conditioned on initial income, neighborhoods with superior natural amenities increase in income *more* than neighborhoods with inferior natural amenities.

Expected change in income for a low-income neighborhood:

$$E(\Delta r | r_L, \text{beach}) = (r_H - r_L) \times \Pr(S1|S2) + 0 \times \Pr(S2|S2)$$
$$E(\Delta r | r_L, \text{desert}) = 0 \times \Pr(S1|S1) + (r_H - r_L) \times \Pr(S2|S1)$$

Expected change in income for a high-income neighborhood:

$$E(\Delta r | r_H, \text{beach}) = 0 \times \Pr(S1|S1) - (r_H - r_L) \times \Pr(S2|S1)$$
$$E(\Delta r | r_H, \text{desert}) = -(r_H - r_L) \times \Pr(S1|S2) + 0 \times \Pr(S2|S2)$$

Thus, $E(\Delta r | r, \text{beach}) > E(\Delta r | r, \text{desert})$

2. Naturally heterogenous cities are stable

The expected over-time variance of neighborhood income, $E[Var(r_{j,t}|j)]$, decreases with cross-sectional heterogeneity in natural amenities, $\alpha_b - \alpha_d$.

$$E[Var(r_{j,t}|j)] = [1 - \Pr(S1)^*] \cdot \Pr(S1)^* \cdot (r_H - r_L)^2$$

- If $\alpha_b - \alpha_d = 0$, then $\Pr(S1)^* = \Pr(S2)^* = 0.5$.
- As $\alpha_b - \alpha_d$ increases, $\Pr(S1|S2)$ increases to 1, and $\Pr(S1)^*$ increases to 1.
- Intuitively, over-time variance in neighborhood income depends on probability of switching between two states

3. Natural amenities are stronger anchors in naturally heterogeneous cities

Conditioned on initial income, the expected difference in income changes between superior and inferior natural amenity neighborhoods increases with cross-sectional heterogeneity in natural amenities, $\alpha_b - \alpha_d$.

- Combines implications 1 and 2

Extensions

Housing demand

- If income elasticity of demand for housing is high, high-income workers may choose to live in inferior aggregate amenity neighborhoods
- Reverses first implication: Conditioned on initial income, neighborhoods with superior natural amenities increase in income *less*

Many neighborhoods

- We extend the equilibrium selection rule—when multiple equilibria are possible, we choose the one that is closest to the historical equilibrium in terms of Euclidean distance between the neighborhood income vectors

Correlated amenity shocks

- Suppose amenity shock follows an AR(1) process, i.e. $\epsilon_{j,t+1} = \rho\epsilon_{jt} + \nu_t$
- Results hold as long as process is stationary ($\rho < 1$)

Data

Neighborhoods and metropolitan areas

- 1880–2000 censuses and 2006–2010 5yr ACS
- Consistent boundary neighborhoods, following 2010 census tracts.

Neighborhood percentile ranks

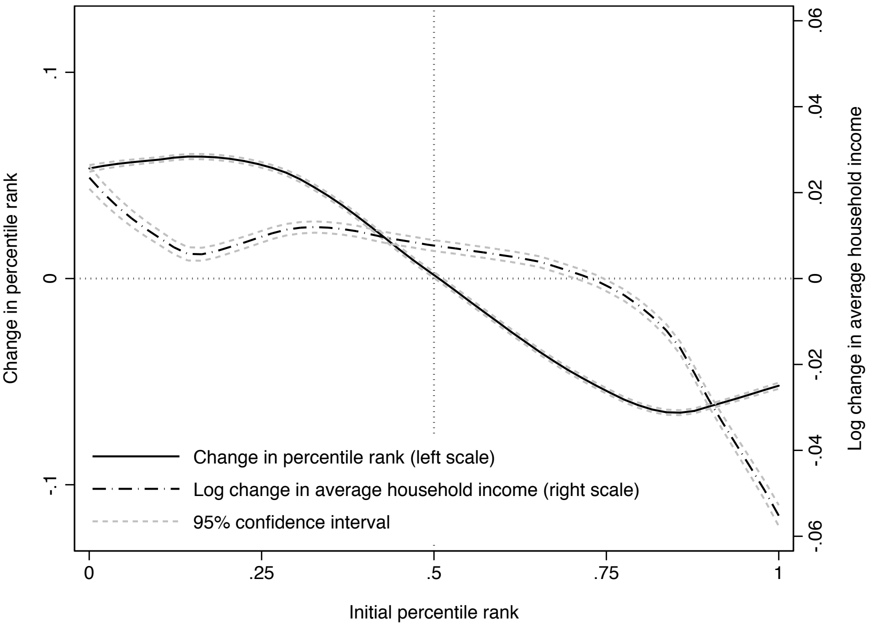
- Average household income (1950–2010)
- Average housing rents (1930–1940)
- Occupational income score or literacy rate (1880–1920)
- Regression of within-metro–year ranks by rent on ranks by income (when both avail.) yields $\hat{\beta} = 0.927$ (crse $= 0.002$), $R^2 = 0.857$

Natural features

- Coastlines, rivers, and lakes from NOAA
- Hills: 90m resolution elevation data from ESRI
- Flood risk from FEMA
- Temperature and precipitation from PRISM/OSU

Year	Census data	Boundaries	Metros	Nbhds
2010	ACS (2006–10)	—	308	60,757
2000	NCDB**	NCDB & LTDB*	308	60,766
1990			308	60,299
1980			259	56,176
1970			229	49,888
1960	Bogue/NHGIS	NHGIS shapefiles†	136	38,669
1950			51	17,681
1940			43	11,527
1930	Beveridge/NHGIS		10	1,962
1920			2	2,505
1910			1	1,748
1880	IPUMS 100% & 10%	UTHGIS shapefiles†	29	3,071

*–weighted by block population, †–land area; **–with censoring corrections in 1980
Data in 1910 for only New York and in 1920 for only New York and Chicago



Kernel-weighted local polynomial smoothing using Stata's `lpoly` function with Epanechnikov kernel, rule-of-thumb bandwidth, and local-mean smoothing. Neighborhood log change in average household income (right scale) is normalized by metropolitan area–year mean.

Conditioned on initial income, neighborhoods with superior natural amenities increase in income *more* than neighborhoods with inferior natural amenities.

$$\Delta r_{i(c),t} = \beta_0 + \beta_1 \mathbf{1}(\mathbf{a}_i) + \beta_2 r_{i,t} + \delta_{c,t} + \epsilon_{i,t}$$

- Neighborhood level, (pooled) cross-sections

$\Delta r_{i,t}$	10-year change (t to $t+1$) in neighborhood percentile rank
$\mathbf{1}(\mathbf{a}_i)$	Indicator for proximity to coast, river, etc. . . .
$r_{i,t}$	Initial percentile rank (within metro) by average income
$\delta_{c,t}$	Metro–year fixed effect

	μ [σ]	(1)
1(Ocean or Gr. L. $<$500m)*,†,‡	0.05	-0.004
	[0.22]	(0.004)
Initial %ile rank by income $(r_{i,t})$	0.50	-0.161^{c}
	[0.29]	(0.007)
Metro–year f.e.		✓
R^2		0.081
Neighborhoods		298,776
Metro–years		1,357

Regressions use pooled observations of 60,872 consistent-boundary neighborhoods over ten census years, 1910–2000.
Dependent variable is 10-year forward change in percentile rank by income $(\Delta r_{i,t})$; mean 0, standard deviation 0.16.
Standard errors, clustered on metropolitan area–year, in parentheses; ^{a}—$p<0.10$, ^{b}—$p<0.05$, ^{c}—$p<0.01$.

Three sources of bias

A. Omitted variables bias
B. Measurement error
C. Dynamic panel bias

A. Omitted variables bias

- Factors omitted from model also affect neighborhood change;
  e.g., old homes, downtown proximity (Brueckner and Rosenthal ’09)
- These factors may be related to natural amenities, too:
  Many cities were founded near natural features, e.g., harbors
- $\hat{\beta}_1$ biased downwards.

	μ [σ]	(1)	(2)	(3)	(4)
1(Ocean or Gr. L. <500m)*,†,‡	0.05 [0.22]	-0.004 (0.004)	0.013^c (0.004)	0.007^a (0.004)	0.014^c (0.003)
Initial %ile rank by income ($r_{i,t}$)	0.50 [0.29]	-0.161^c (0.007)	-0.169^c (0.007)	-0.184^c (0.008)	-0.202^c (0.008)
Log distance to nearest seaport§	5.02 [4.83]		0.028^c (0.004)		-0.004 (0.002)
Log distance to city center	7.51 [1.95]			0.035^c (0.003)	-0.008^c (0.002)
Log population density	9.74 [1.04]				-0.036^c (0.001)
Log average house age	3.00 [0.53]				-0.019^c (0.004)
Metro–year f.e.		✓	✓	✓	✓
R^2		0.081	0.090	0.116	0.202
Neighborhoods		298,776	298,776	297,518	281,321
Metro–years		1,357	1,357	1,313	1,263

Standard errors, clustered on metropolitan area–year, in parentheses; ^{a}—$p<0.10$, ^{b}—$p<0.05$, ^{c}—$p<0.01$.
§—Log distance to nearest seaport times metropolitan indicator for coastal proximity.

B. Measurement error

- Some coastal areas may be disamenable, due to poor climate, drainage, or pollution
- $\hat{\beta}_1$ biased towards zero.

	μ [σ]	(1)	(2)	(3)	(4)	$r_{i,t} > 0.9$	Names
1(Ocean or Gr. L. <500m)[*,†,‡]	0.05 [0.22]	-0.004 (0.004)	0.013[c] (0.004)	0.007[a] (0.004)	0.014[c] (0.003)	0.045[c] (0.005)	0.031[c] (0.005)
Initial %ile rank by income ($r_{i,t}$)	0.50 [0.29]	-0.161[c] (0.007)	-0.169[c] (0.007)	-0.184[c] (0.008)	-0.202[c] (0.008)	-0.204[c] (0.008)	-0.203[c] (0.008)
Log distance to nearest seaport[§]	5.02 [4.83]		0.028[c] (0.004)		-0.004 (0.002)	-0.004[a] (0.002)	-0.004[a] (0.002)
Log distance to city center	7.51 [1.95]			0.035[c] (0.003)	-0.008[c] (0.002)	-0.008[c] (0.002)	-0.008[c] (0.002)
Log population density	9.74 [1.04]				-0.036[c] (0.001)	-0.036[c] (0.001)	-0.036[c] (0.001)
Log average house age	3.00 [0.53]				-0.019[c] (0.004)	-0.019[c] (0.004)	-0.019[c] (0.004)
Metro–year f.e.		✓	✓	✓	✓	✓	✓
R^2		0.081	0.090	0.116	0.202	0.202	0.202
Neighborhoods		298,776	298,776	297,518	281,321	281,321	281,321
Metro–years		1,357	1,357	1,313	1,263	1,263	1,263

†—Explanatory variable in column (5) is neighborhood centroid is within 500 meters of ocean, Gulf of Mexico, or Great Lake *and* $r_{i,t} > 0.9$.

‡—Explanatory variable in column (6) is neighborhood centroid is within 500 meters of ocean, Gulf of Mexico, or Great Lake *and* neighborhood name includes “beach,” “coast,” “bay,” “cove,” “lagoon,” “ocean,” or “shore.”

C. Dynamic panel bias

$$\Delta r_{i(c),t} = \beta_0 + \beta_1 \mathbf{1}(\mathbf{a}_i) + \beta_2 r_{i,t} + \delta_{c,t} + v_i + \epsilon_{i,t}$$

- $Corr(v_i, r_{i,t}) \neq 0$ by construction; likely > 0
- $Corr(v_i, \mathbf{1}(\mathbf{a}_i)) \neq 0$; likely < 0
- Intuition: Given two neighborhoods with varying (measured) natural amenities but the same initial income $r_{i,t}$, the neighborhood with inferior measured natural amenities is likely to have other unobserved characteristics v_i that are amenable.
- $\hat{\beta}_1$ biased downwards; $\hat{\beta}_2$ biased upwards.

- Estimators that use first-differencing confound identification of β_1.
  Arellano-Bond estimate of $\hat{\beta}_2 \approx -0.3$ vs. $\hat{\beta}_2^{OLS} \approx -0.2$
- IV? Natural amenity shifters are also likely to affect (initial) income.

Robustness

Other measures of natural amenity

- Lakes, rivers, hills: Results do not depend on inclusion of historical controls
- Hedonic rent index to measure aggregate natural value
- Varying threshholds to define indicator variables

Separate estimates by year and time horizon

Non-parametric estimation

Endogenous factors: housing, zoning

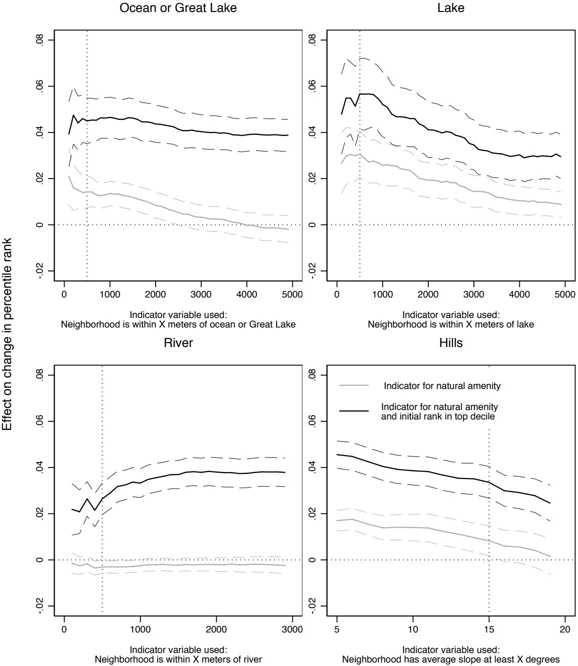
- High-quality housing or more restrictive zoning may reinforce our results
- Little evidence that this plays a big role

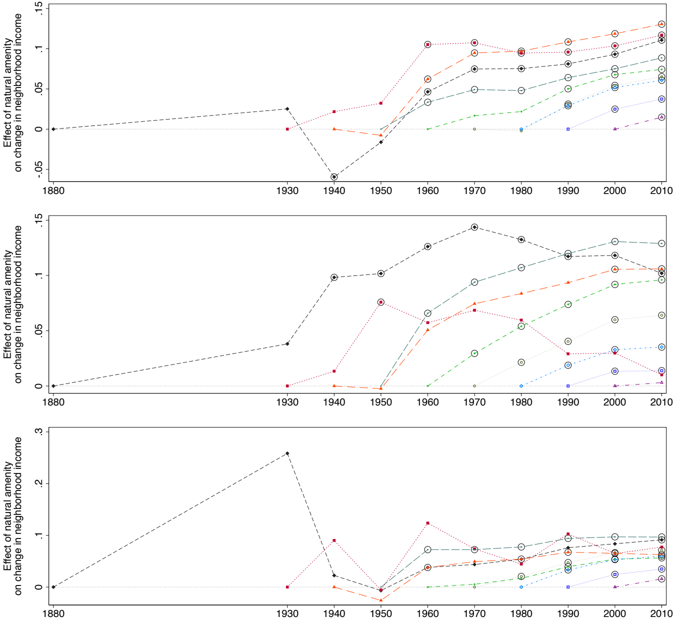
Increasing valuation of natural amenities over time

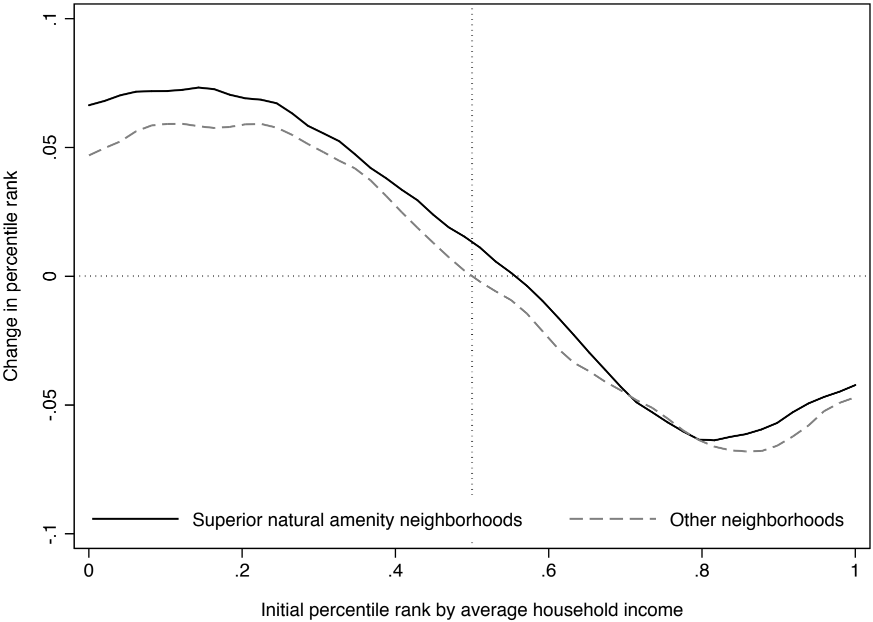
- No evidence that high-income HHs have increased valuation of coastal proximity over time

	Coast	Lake	River	Hills	Temp. & dry	Not flood	Nat'l val. > p(95)
A. Indicator for natural feature							
	-0.004	0.044^c	0.004^c	0.050^c	-0.033^c	-0.025^c	0.013^b
	(0.004)	(0.006)	(0.002)	(0.004)	(0.013)	(0.003)	(0.005)
B. With controls for historical factors							
	0.014^c	0.031^c	-0.003^b	0.008^b	0.015	-0.001	0.028^c
	(0.003)	(0.005)	(0.001)	(0.003)	(0.010)	(0.002)	(0.004)
C. Indicator for natural feature and $r_{i,t} > 0.9$							
	0.045^c	0.057^c	0.026^c	0.034^c	0.031^c	0.042^c	0.040^c
	(0.005)	(0.014)	(0.005)	(0.004)	(0.009)	(0.003)	(0.005)
D. Indicator for natural feature and place names							
	0.031^c	0.030^c	-0.004	-0.005	0.025^c	0.005^b	0.019^c
	(0.005)	(0.007)	(0.003)	(0.003)	(0.007)	(0.002)	(0.005)
E. Sample means of natural amenity indicator							
	0.054	0.006	0.094	0.064	0.072	0.641	0.050

Standard errors, clustered on metropolitan area, in parentheses; ^{b}—$p<0.05$, ^{c}—$p<0.01$.
Explanatory variable is indicator for proximity within 500m in columns (1)–(3), average slope greater than 15 degrees in column (4), mean January minimum temperature between 0 and 18 degrees Celsius and mean July maximum temperature between 10 and 30 degrees Celsius and mean annual precipitation less than 800mm in column (5), mean annual flood probability less than 1% in column (6), and top 5% in natural value estimated using hedonic weights as described in the text in column (7).

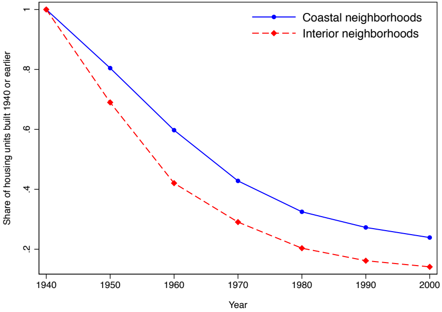

Ocean or Great Lake
Lake
River
Hills
Effect on change in percentile rank
Indicator for natural amenity
Indicator for natural amenity
and initial rank in top decile
.08
.06
.04
.02
0
-.02
0
1000
2000
3000
4000
5000
5
10
15
20
Indicator variable used:
Neighborhood is within X meters of ocean or Great Lake
Indicator variable used:
Neighborhood is within X meters of lake
Indicator variable used:
Neighborhood is within X meters of river
Indicator variable used:
Neighborhood has average slope at least X degrees


Effect of natural amenity on change in neighborhood income
.15
.1
.05
0
-.05
1880
1930
1940
1950
1960
1970
1980
1990
2000
2010
Effect of natural amenity on change in neighborhood income
.15
.1
.05
0
1880
1930
1940
1950
1960
1970
1980
1990
2000
2010
Effect of natural amenity on change in neighborhood income
.3
.2
.1
0
1880
1930
1940
1950
1960
1970
1980
1990
2000
2010

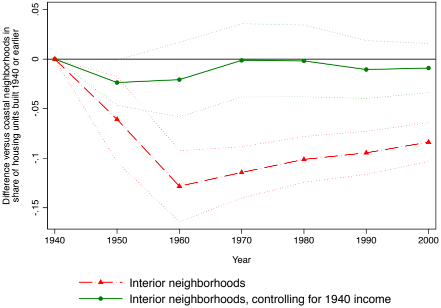

.1
.05
0
-.05
-.1
Change in percentile rank
0
.2
.4
.6
.8
1
Initial percentile rank by average household income
Superior natural amenity neighborhoods
Other neighborhoods

Survival of pre-1940 homes, coastal vs. interior neighborhoods

A. Share of housing units built before 1940



B. Difference in pre-1940 housing share vs. coastal neighborhoods



	μ [σ]	(1)	(2)	(3)	(4)	(5)	(6)
1(Coast)	0.05 [0.21]	0.007^b (0.004)	0.010^c (0.004)	0.011^c (0.004)	0.010^b (0.004)	0.014^c (0.005)	0.015^c (0.005)
Initial %ile rank by income $(r_{i,t})$	0.50 [0.29]	-0.097^c (0.004)	-0.124^c (0.004)	-0.129^c (0.005)	-0.125^c (0.005)	-0.119^c (0.004)	-0.120^c (0.004)
Log distance to seaport	4.73 [4.88]		-0.006^c (0.002)	-0.006^c (0.002)	-0.004^b (0.002)	-0.001 (0.003)	-0.002 (0.003)
Log distance to city center	9.87 [1.04]		-0.010^c (0.002)	-0.011^c (0.002)	-0.009^c (0.002)	-0.011^c (0.002)	-0.010^c (0.002)
Log population density	7.53 [1.79]		-0.022^c (0.001)	-0.022^c (0.001)	-0.022^c (0.001)	-0.024^c (0.002)	-0.024^c (0.002)
Log average house age	3.37 [0.48]			-0.008^c (0.003)	-0.020^c (0.003)	-0.017^c (0.005)	-0.016^c (0.005)
Share houses built before 1940	0.15 [0.19]				0.070^c (0.010)	0.092^c (0.010)	0.092^c (0.010)
Wharton residential land-use reg. index	0.13 [0.98]						0.004^b (0.002)
R^2		0.049	0.095	0.095	0.100	0.113	0.113
Neighborhoods		60,073	60,073	60,073	60,073	22,591	22,591
Metro areas		293	293	293	293	247	247

Regressions use cross-section of consistent-boundary neighborhoods, 2000–2010.
Standard errors, clustered on metropolitan area, in parentheses; ^{a}—$p<0.10$ ^{b}—$p<0.05$, ^{c}—$p<0.01$.

Relative likelihood that high-income neighborhood is coastal

This figure shows, for each census year, the relative likelihood that a high-income neighborhood (versus a randomly-selected neighborhood) is within 500m of an ocean, the Gulf of Mexico, or a Great Lake.

Recap (1)

- Natural amenities *anchor* neighborhoods to high incomes over time.
- Conditioned on initial income, a superior natural amenity neighborhood increases 1–6+ percentile points more than other neighborhoods.
- The effect of coastal proximity on neighborhood change is confounded by historical factors, but the effects of other natural amenities are not.
- Estimated effects of natural amenities are biased downwards.
- Robust to other measures of natural amenities, start years, time horizon.
- Effects not due to observed endogenous factors like housing, zoning.
- Effects not due to increased valuation of natural amenities over time.
- There is mean reversion in neighborhood rank.

The expected over-time variance of neighborhood income, $E[Var(r_{j,t}|j)]$, decreases with cross-sectional heterogeneity in natural amenities, $\alpha_b - \alpha_d$.

$$\text{Var}(r_{i(c)}|i) = \delta_m + \varepsilon_{i(c)}$$
$$\hat{\delta}_c = \gamma_0 + \gamma_1 \mathbf{\Gamma}_c + \mathbf{Z}'_c \gamma_3 + \mu_c$$

- Metro level, cross sectional regression
  e.g., base year 1960, variance over 1960–2010

Var(r_i)	Over-time variance in neighborhood rank from 1960 to 2010
$\mathbf{\Gamma}_c$	Initial city variation in natural amenities
	A. Metro indicator for coastal/interior, hilly/flat, etc.
	B. Metro variation in log distance to coast, elevation, etc.
	C. Metro indicator × variation in log distance to coast, etc.
Z	Metro variation in income, log growth in population and land area

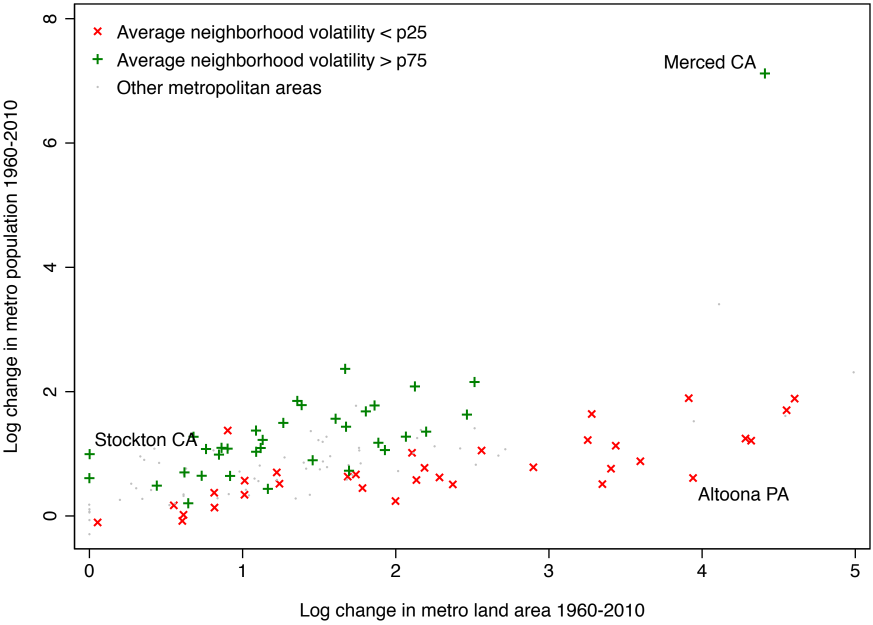
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Measure of variation in coastal proximity:		A. Metro coastal indicator			B. Metro s.d. nbhd. log dist. to coast		C. Metro indicator × s.d. dist. coast	
	μ [σ]	0.289 [0.455]			0.420 [0.573]		0.344 [0.570]	
Variation in coastal proximity		-0.363[a] (0.184)	-0.341[b] (0.148)	-0.303[a] (0.157)	-0.365[c] (0.122)	-0.334[b] (0.138)	-0.254[b] (0.117)	-0.259[b] (0.118)
Metro log change in population, 1960–2010	0.94 [0.78]		1.131[c] (0.220)	1.142[c] (0.218)	1.109[c] (0.217)	1.119[c] (0.213)	1.135[c] (0.214)	1.136[c] (0.213)
Metro log change in land area, 1960–2010	1.61 [1.15]		-0.492[c] (0.104)	-0.491[c] (0.103)	-0.547[c] (0.115)	-0.542[c] (0.117)	-0.496[c] (0.104)	-0.493[c] (0.104)
Within-metro s.d. in nbhd. income (thous.)	1.91 [0.43]			-0.220 (0.165)		-0.181 (0.184)	-0.203 (0.168)	-0.206 (0.168)
Within-metro s.d. in nbhd. avg. house age	3.69 [0.78]			0.066 (0.104)		0.052 (0.102)	0.057 (0.101)	0.050 (0.102)
1st-level controls								
Initial rank decile								✓
R^2		0.057	0.451	0.463	0.476	0.484	0.471	0.473
Metro areas		135	135	135	135	135	135	135

First-level OLS regressions (unreported) use neighborhood observations in census year 1960 to estimate 135 metropolitan area means and cluster-robust standard errors.

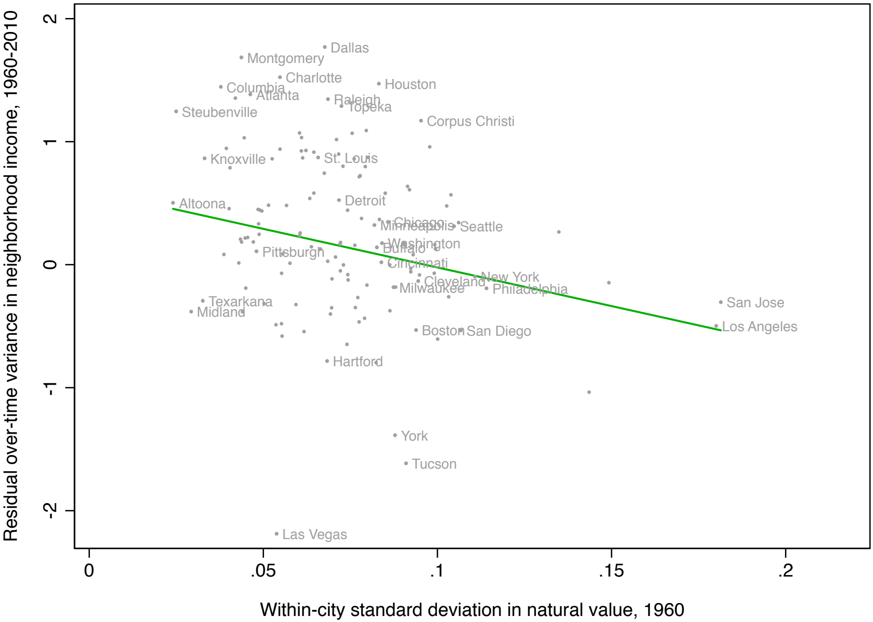
Dependent variable is over-time variance in percentile rank $\times 100$, 1960–2010; mean 2.29, standard deviation 2.63 in balanced panel of 38,293 neighborhoods over six census years.

Second-level WLS regressions use 135 metropolitan areas. Dependent variable is estimated metropolitan area means from first level and weights are inverse estimated variance from first level; mean 2.3, standard deviation 0.9.

Robust standard errors in parentheses; ^{a}—$p<0.10$, ^{b}—$p<0.05$, ^{c}—$p<0.01$.



Average neighborhood volatility < p25
Average neighborhood volatility > p75
Other metropolitan areas
Merced CA
Stockton CA
Altoona PA
Log change in metro population 1960-2010
Log change in metro land area 1960-2010
0
2
4
6
8
0
1
2
3
4
5


Residual over-time variance in neighborhood income, 1960-2010
Within-city standard deviation in natural value, 1960
2
1
0
-1
-2
0
.05
.1
.15
.2
Dallas
Montgomery
Charlotte
Houston
Columbia
Atlanta
Raleigh
Topeka
Steubenville
Corpus Christi
Knoxville
St. Louis
Altoona
Detroit
Chicago
Minneapolis
Seattle
Washington
Buffalo
Pittsburgh
Cincinnati
New York
Cleveland
Milwaukee
Philadelphia
Texarkana
Midland
San Jose
Los Angeles
Boston
San Diego
Hartford
York
Tucson
Las Vegas

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Base year:		1880	1930	1940	1950	1960	1970	1980	1990
	μ	5.13	3.48	2.95	2.41	2.29	1.75	1.21	0.98
	[σ]	[3.00]	[3.31]	[2.97]	[2.72]	[2.63]	[2.32]	[1.83]	[1.71]
Within-metro s.d. in	0.42	0.855^b	1.367^a	-0.073	-0.078	-0.365^c	-0.189^c	-0.049	-0.016
log nbhd. distance to coast	[0.62]	(0.359)	(0.687)	(0.227)	(0.168)	(0.122)	(0.059)	(0.038)	(0.032)
Within-metro s.d. in	0.40	2.149	2.999	-0.021	-1.168^a	-2.011^c	-1.157^c	-0.464^c	-0.315^c
nbhd. average slope	[0.22]	(1.471)	(4.115)	(0.554)	(0.619)	(0.369)	(0.265)	(0.129)	(0.112)
Within-metro s.d. in	0.08	21.078^c	29.105^a	-0.949	-7.543^b	-9.387^c	-6.049^c	-2.294^c	-1.250^b
nbhd. natural value	[0.03]	(4.744)	(12.909)	(2.825)	(3.296)	(1.608)	(1.011)	(0.594)	(0.616)
Census years*		9	9	8	7	6	5	4	3
Metropolitan areas		25	10	38	51	135	227	277	308
Neighborhoods (1st stage)		2,938	1,935	11,167	17420	38,293	49,660	55,911	60,063

Dependent variable is over-time variance in percentile rank $\times 100$, between base year and 2010; metropolitan-level means and standard deviations in first row.

Second-level WLS regressions use metropolitan areas. Dependent variable is estimated metropolitan area means from first level and weights are inverse estimated variance from first level.

Robust standard errors in parentheses; ^{a}—$p<0.10$, ^{b}—$p<0.05$, ^{c}—$p<0.01$. *—For each base year, we balance our neighborhood panel to calculate over-time variances. Thus, historical base years have fewer than expected time periods and metropolitan areas, since we drop metro–years with missing data.

Recap (2)

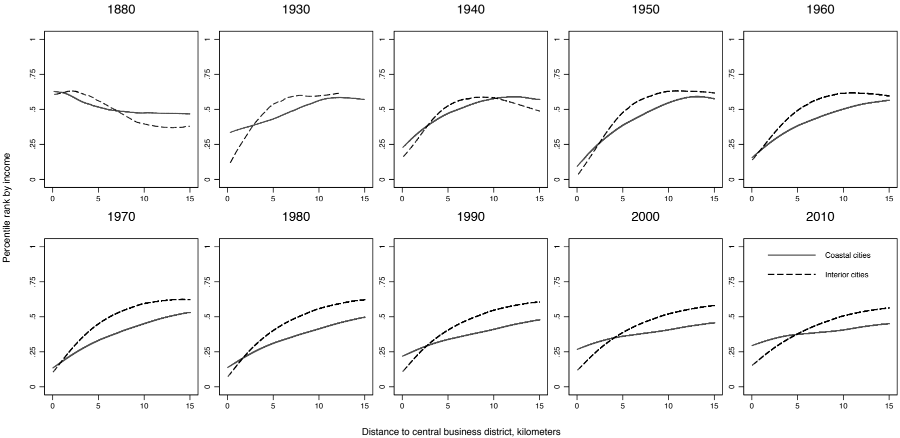
- Neighborhood incomes are less volatile—i.e., the spatial distribution of income is more persistent—in naturally heterogenous (coastal, hilly) cities.
- Neighborhood incomes are less volatile in cities with low population growth or high land area growth.
- These results are robust to other start years 1940 and later.

3. Natural amenities are stronger anchors in naturally heterogeneous cities

$$\Delta r_{i(m),t} = \zeta_0 + \zeta_1 \mathbf{1}(CBD_i) + \zeta_2 \mathbf{1}(CBD_i) \times \mathbf{1}(Coast_m) + \zeta_2 \ln r_{i,t} + \mathbf{X}'_{i,t}\zeta_3 + \delta_{m,t} + \epsilon_{i,t}$$

- Application: downtown neighborhoods (1982 Census of Retail Trade)
- In U.S. cities, downtowns typically near natural amenities, e.g. coasts or rivers
- Given greater amenity value of coasts, coastal metros more naturally heterogeneous

Natural amenities anchor downtowns to high incomes



Lowess regressions. Coastal cities are solid line; Interior cities are dashed line.

Cities

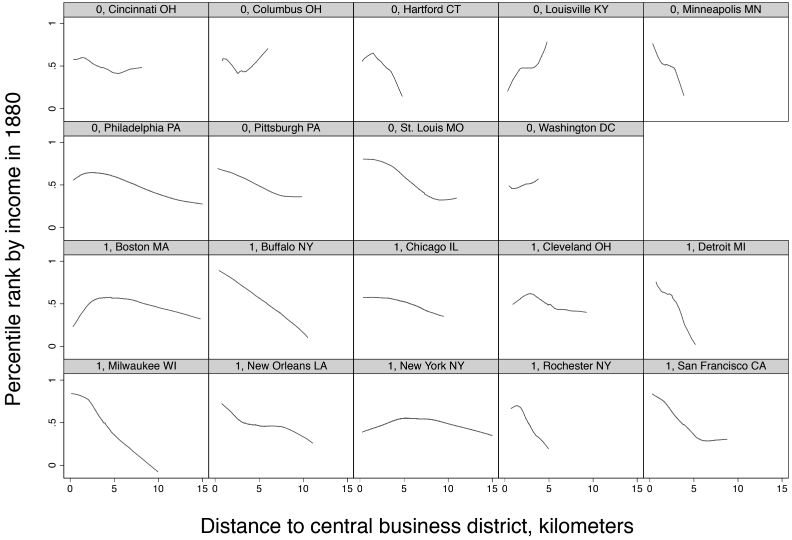
	(1)	(2)	(3)	(4) 1950–1980	(5) 1980–2010
Sample years:		1910–2010¶			
1(CBD)*	-0.040^c	-0.085^c	-0.015^c	-0.023^c	0.005^b
	(0.004)	(0.004)	(0.003)	(0.005)	(0.003)
1(CBD) × 1(Coastal metro)†	0.032^c	0.034^c	0.020^b	0.030^c	0.029^c
	(0.009)	(0.008)	(0.008)	(0.009)	(0.009)
Initial %ile rank by income ($r_{i,t}$)		-0.181^c	-0.205^c	-0.313^c	-0.124^c
		(0.009)	(0.006)	(0.010)	(0.003)
Log distance to seaport			-0.003	-0.003	-0.008^c
			(0.003)	(0.005)	(0.002)
Log population density			-0.035^c	-0.048^c	-0.021^c
			(0.001)	(0.001)	(0.001)
Log average house age‡				-0.075^c	0.005^a
				(0.010)	(0.002)
Metro–year f.e.	✓	✓	✓	✓	✓
R^2	0.004	0.098	0.207	0.338	0.107
Neighborhoods	297,522	297,522	297,520	105,529	175,794
Metros	293	293	293	224	293

Dependent variable is 10-year change in percentile rank by income ($\Delta r_{i,t}$); mean 0, standard deviation 0.16.

Standard errors, clustered on metropolitan area–year, in parentheses; ^{a}—$p<0.10$, ^{b}—$p<0.05$, ^{c}—$p<0.01$.

*—Neighborhood is within 5 km of central-city CBD. †—Metropolitan area CBD is within 1 km of ocean or Great Lake. ‡—Available 1950 and later.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				1950–1980	1980–2010	1990–2010 (Kneebone)	
Sample years:		1910–2010[¶]					
1(CBD)*	-0.040^c	-0.085^c	-0.015^c	-0.023^c	0.005^b	0.018^c	0.018^b
	(0.004)	(0.004)	(0.003)	(0.005)	(0.003)	(0.004)	(0.008)
1(CBD) $\times$ 1(Coastal metro)[†]	0.032^c	0.034^c	0.020^b	0.030^c	0.029^c	0.025^b	0.025^b
	(0.009)	(0.008)	(0.008)	(0.009)	(0.009)	(0.011)	(0.011)
Initial %ile rank by income $(r_{i,t})$		-0.181^c	-0.205^c	-0.313^c	-0.124^c	-0.108^c	-0.108^c
		(0.009)	(0.006)	(0.010)	(0.003)	(0.003)	(0.003)
Log distance to seaport			-0.003	-0.003	-0.008^c	-0.007^c	-0.007^c
			(0.003)	(0.005)	(0.002)	(0.002)	(0.002)
Log population density			-0.035^c	-0.048^c	-0.021^c	-0.020^c	-0.020^c
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log average house age[‡]				-0.075^c	0.005^a	0.008^c	0.008^c
				(0.010)	(0.002)	(0.003)	(0.003)
1(CBD) $\times$ ΔCBD job share[§]							0.034
							(0.242)
1(CBD) $\times$ Δ3–10 mi job share[||]							-0.378^b
							(0.148)
Metro–year f.e.	✓	✓	✓	✓	✓	✓	✓
R^2	0.004	0.098	0.207	0.338	0.107	0.099	0.099
Neighborhoods	297,522	297,522	297,520	105,529	175,794	98,006	98,006
Metros	293	293	293	224	293	86	86



This figure shows the pattern of neighborhood average household income on the vertical axis versus the neighborhood distance to the city center (up to 15km) on the horizontal axis, for the 19 largest cities in our 1880 sample. Ten metropolitan areas with 20 or fewer neighborhoods in 1880 are not shown. Cities are organized by coastal status (0=interior, 1=coastal) then alphabetically. The plotted lines are results from lowess smoothing with bandwidth 0.9.

Recap (3)

- Downtown neighborhoods in both coastal and interior cities featured higher incomes until the late 19th century.
- Some cities saw a reversal in the relative status of core and peripheral neighborhoods by 1880. All cities by 1930.
- Downtowns in coastal cities experienced smaller declines before 1980 and larger gains after 1980 vs. downtowns in interior cities.
- Varying patterns of job decentralization do not drive these results.

Conclusions

- Persistent natural amenities affect both neighborhood dynamics and the dynamic stability of the spatial distribution of income within cities
- Downtown neighborhoods in cities with superior natural amenities resisted the suburbanization of income and gentrified more quickly
- Evidence that heterogeneity in fundamentals affects uniqueness of equilibrium and thus persistence in spatial distribution of activity

Coastal	Interior
Boston, MA	Albany, NY
Buffalo, NY	Atlanta, GA
Charleston, SC	Cincinnati, OH
Chicago, IL	Columbus, OH
Cleveland, OH	Hartford, CT
Detroit, MI	Indianapolis, IN
Milwaukee, WI	Kansas City, MO
Mobile, AL	Louisville, KY
New Orleans, LA	Memphis, TN
New York, NY	Minneapolis, MN
Rochester, NY	Nashville, TN
San Francisco, CA	Omaha, NE
	Philadelphia, PA
	Richmond, VA
	Pittsburgh, PA
	St. Louis, MO
	Washington, DC