Urban Distress*

Urban Economics: Week 8

Giacomo A. M. Ponzetto

CREI – UPF – Barcelona GSE

27th and 28th February 2012

*I thank Kurt Schmidheiny for sharing his slides on "Segregation"

Sorting and Segregation

- Spatial equilibrium with freely mobile individuals
- Locations have different characteristics
  - Commuting distance from the CBD
  - Real-estate prices
  - Amenities from geography or history
- Heterogeneous agents have different valuations
  - The rich have a higher value of time and a lower value of money
  - Tastes for amenities are idiosyncratic to some extent

Sorting: Each agent locates in the places he prefers

Segregation: Each location hosts the agents who value it the most

Patterns and Sources of Sorting

- We discussed already sorting by income
  - Income elasticities of demand for land and amenities
  - Income elasticity of commuting costs
  - Multiple means of transportation
- Perfect sorting and segregation by concentric rings
- Richer models would yield less than perfect sorting
  - Idiosyncratic tastes for amenities, land, commute times
  - Heterogeneous housing stock in the same location, filtering
- Endogenous amenities are another source of sorting
  - They do not explain where people locate
  - They can explain why there is so much segregation
  - As we discussed in the case of income heterogeneity

Local Public Education

- The quality of public education rises with residents' wealth
  - Neighborhood peer-group effects affect school productivity
  - In the U.S., public schools are financed mainly by local taxes
  - The richer mostly get better schools from the central government too
- Endogenous amenity
  - Everyone prefers to live in a wealthy neighborhood
- Heterogeneous willingness to pay
  - The rich care more about school quality and can afford to
  - The poor are deterred by high housing costs
- Income segregation implies different educational opportunities
  - This certainly seems unfair
  - It can also be inefficient, and even Pareto inefficient

Fernandez and Rogerson (1996)

- Individuals have identical preferences: $u(c_1) + \beta u(c_2)$
- There are I income groups with $y_1 > y_2 > ... > y_I$
  - Group size λ_i with normalized population $\sum_i \lambda_i = 1$
- There are $J < I$ locations with no exogenous characteristics
  - An endogenous fraction ρ_{ij} of income-i agents live in j
  - Endogenous average income $\bar{y}_j = \left(\sum_i \rho_{ij} \lambda_i y_i\right) / \left(\sum_i \rho_{ij} \lambda_i\right)$
  - Endogenous tax rate t_j
  - Endogenous quality of public education $q_j = t_j \bar{y}_j$
- There is no private investment or saving technology

$$V_i^j = u\left(\left(1 - t_j\right) y_i\right) + \beta u\left(f\left(t_j \bar{y}_j\right)\right)$$

  - $f(q)$ is an increasing, concave school production function

Policy Preferences

- Assumptions to generate stratification
  1. Willingness to pay taxes for education rises with income
  2. Each agent wants both c_1 and c_2 to increase when $\bar{y}$ does
     I.e., they do if the tax rate is his preferred $\tilde{t}\left(y_i, \bar{y}_j\right)$ such that

$$y_i u'\left(\left(1-\tilde{t}\right) y_i\right) = \beta \bar{y}_j u'\left(f\left(\tilde{t}\bar{y}_j\right)\right) f'\left(\tilde{t}\bar{y}_j\right)$$

- These assumptions imply that
  1. $\tilde{t}\left(y_i, \bar{y}_j\right)$ is increasing in y_i and decreasing in $\bar{y}_j$
  2. $\tilde{q}\left(y_i, \bar{y}_j\right) \equiv \bar{y}_j \tilde{t}\left(y_i, \bar{y}_j\right)$ is increasing in both y_i and $\bar{y}_j$
  3. Preferences for t_j are single-peaked
- Majority vote results in the tax rate preferred by the resident with the median income within the community
- Spatial equilibrium: every agent is in his favorite community

Equilibrium Stratification

1. Tax rates and school quality must rise together

$$(q_j, t_j) \neq (q_k, t_k) \Rightarrow (q_j, t_j) > (q_k, t_k)$$

   - Nobody tolerates higher taxes and worse schools with free mobility

2. There is perfect stratification by income

$$(q_j, t_j) \neq (q_k, t_k) \Rightarrow \min y_i \in j \geq \max y_i \in k$$

   - By assumption, wealthier people are more willing to pay for education

3. An equilibrium in which $(q_j, t_j) = (q_k, t_k)$ is unstable
   - Unless all residents of j and k belong to a single group i
   - If the wealthiest residents of k move to j
     - ★ Average income $\bar{y}_j$ rises and $\bar{y}_k$ falls
     - ★ Median income cannot fall in j nor rise in k
     - ★ The wealthy movers strictly resist moving back

Inefficient Stratification

- A stable stratified equilibrium with no homogeneous community
- Entirely described by boundaries $\left(\rho_j^b, y_j^b\right)$
  - y_j^b is the minimum income in j and maximum in $j+1$
  - ρ_j^b is the share of y_j^b-earners in j, while $1-\rho_j^b$ are in $j+1$
- Any equilibrium with $\rho_j^b < 1$ is Pareto inefficient
  - Move the marginal agent from j to $j+1$
  - Average incomes $\bar{y}_j$ and $\bar{y}_{j+1}$ both increase
  - Median incomes are unchanged
  - Tax rates t_j and t_{j+1} fall, school qualities q_j and q_{j+1} rise
  - All agents are strictly better off
  - But the welfare increase is greater for y_j^b-earners in j than in $j+1$
- A y_j^b-earner moving from $j+1$ to j imposes negative externalities
  - On the poor in $j+1$ whose education he refuses to finance
  - On the rich in j whom he forces to finance his education

Policy Analysis

- Robust intuitions from two communities and three income groups
  - A share ρ_2 of y_2-earners live in the rich community
- The rich community has a rich median resident
  - Otherwise all y_2-earners would be there to milk the rich
- The poor community could have either median resident
  - A y_2-earners or a poor resident
- Simple intuition for efficient policy
  1. Make the poor community more attractive to y_2-earners
  2. Make the rich community less attractive to y_2-earners

Pareto Efficient Policy Interventions

1. Direct Pigovian taxation
   - Tax y_2-earners who live in the rich community
   - Subsidize y_2-earners who live in the poor community
   - Do not change the income of the median resident of the poor community
2. Redistribute tax revenues from the rich to the poor community
   - The rich respond by raising tax rates to recover some lost revenues
   - y_2-earners move away from tax rates and toward tax revenues
   - Calibrate so that y_2-earners in the rich community are not worse off
3. Stop a poor median voter setting policy in the poor community
   - Binding minimum educational standard to raise q_2
   - Analogous but paradoxical minimum tax rate t_2

Inefficient Policy Interventions

1. Stop the rich median voter setting policy in the rich community
   - Binding maximum tax rate t_1
   - Analogous but perverse maximum school quality q_1
2. Stop a median y_2-earner setting policy in the poor community
   - The same policies that are efficient with a poor median voter

- With both of these, everyone is worse off in equilibrium
- Abolishing local autonomy is merely not Pareto efficient
  - The rich are necessarily worse off as they get milked
  - All the others are better off as they milk the rich

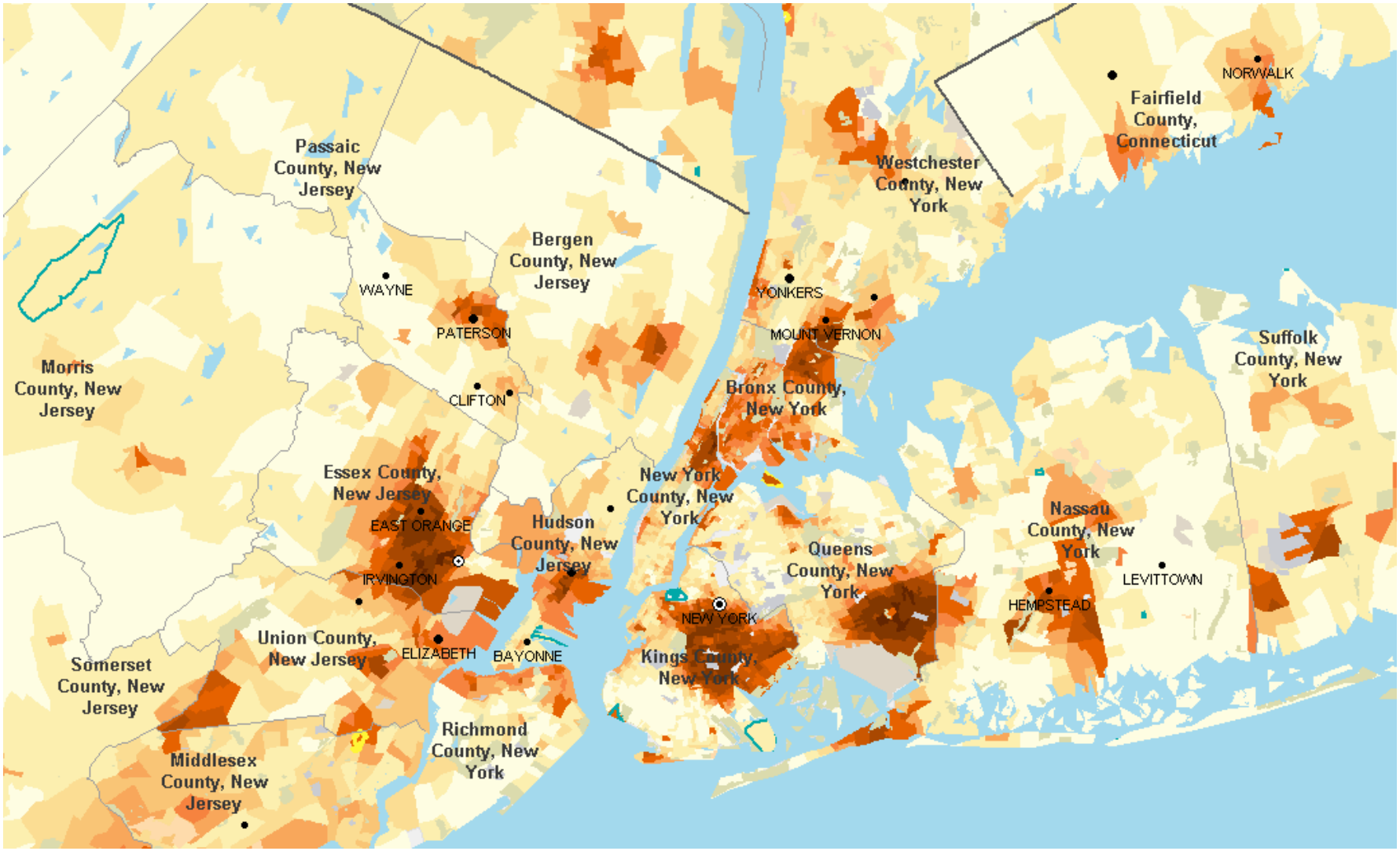
Peer Effects

- Bénabou (1993) has a more complex model without voting
- Education is simply a matter of neighborhood spillovers
  - The more neighbors acquire high skill, the cheaper it is to do so
  - There is lower spillover for acquiring low skill
  - There are no spillovers from acquiring low skill
- Local complementarities in education lead to stratification
  - Rich neighborhoods are homogeneously high skill
  - There is at most one neighborhood with mixed skills
- Segregation can be incredibly costly
  - At worst, no skill acquisition in homogeneous poor neighborhoods
- Global complementarities in production make stratification bad for all
  - The poor are unemployed in ghettos
  - The rich have very few low-skill agents to work with
  - In the absence of a mixed neighborhood, no production is possible

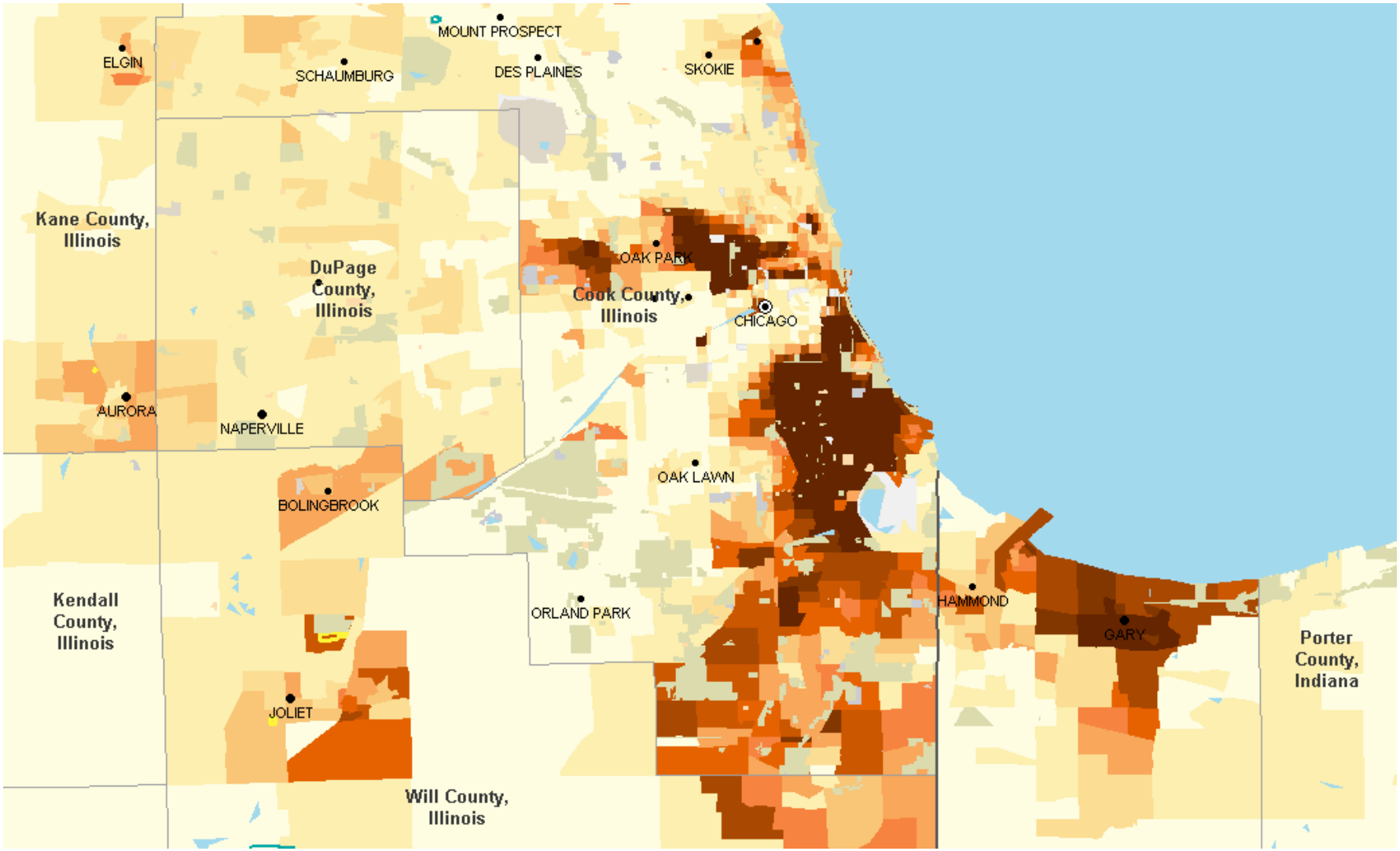
Dynamic Considerations

- Bénabou (1996) extends the analysis to a dynamic setting
- Instead of static inefficiency, a trade off between short and long run
- Stratification is better at processing inequality
  - Local complementarities between household and neighborhood
  - Assortative matching is efficient in the short run
- Integration is better at reducing inequality
  - Global complementarities imply that inequality is costly
  - Integration can be more efficient in the long run
  - Unless local complementarities are much stronger than global ones
- The trade off between local and national school financing is the same

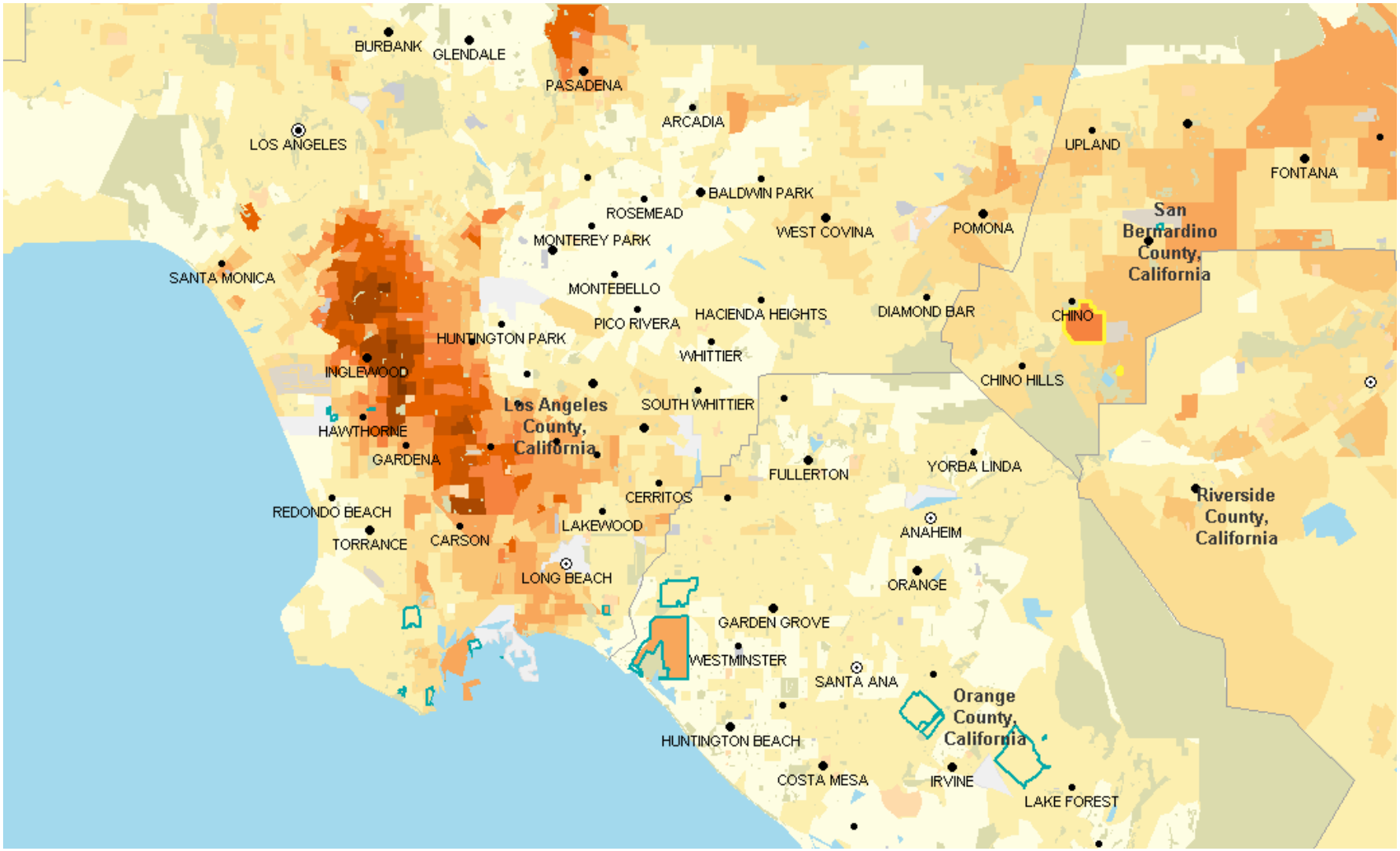
% Black in New York



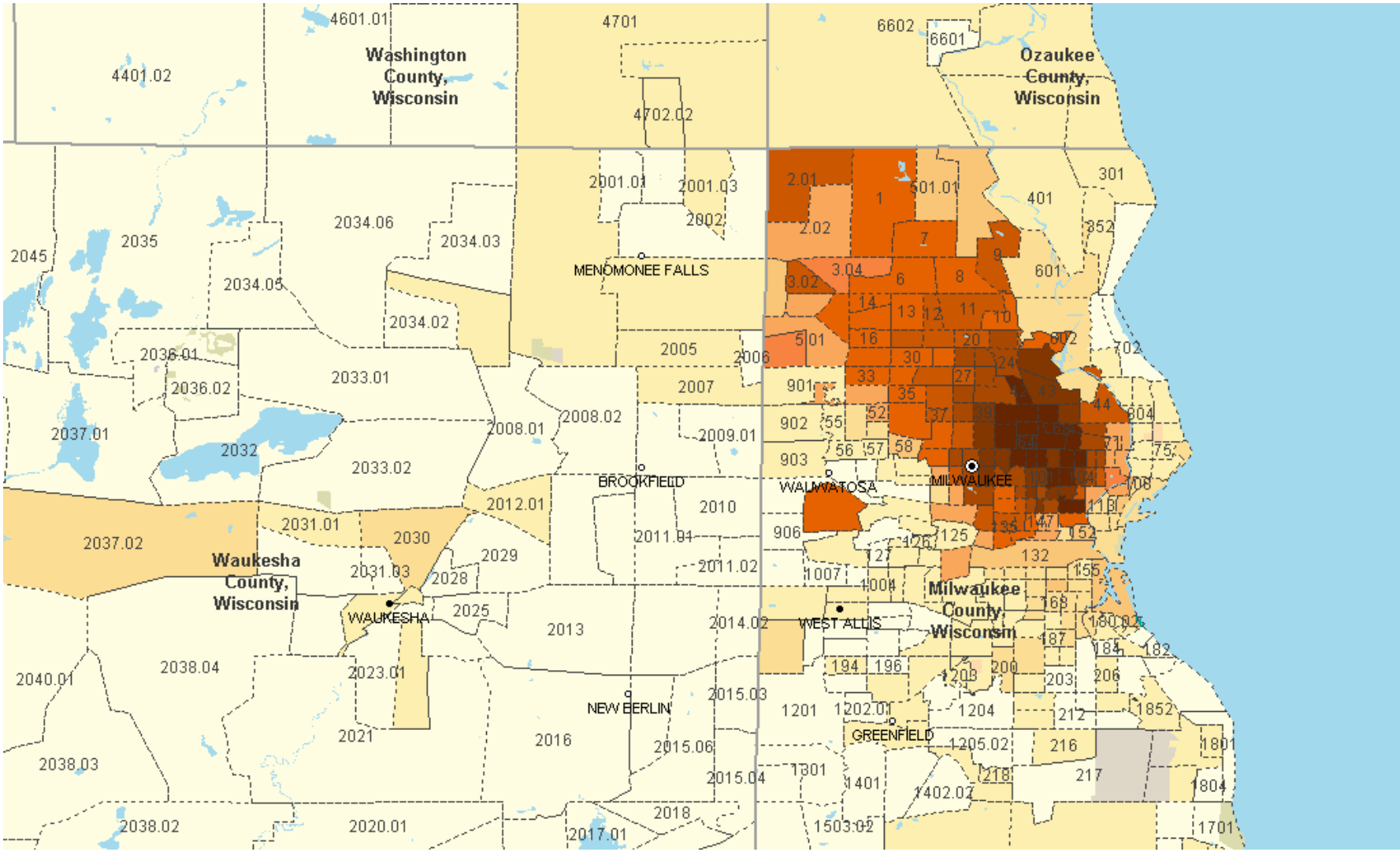
% Black in Chicago



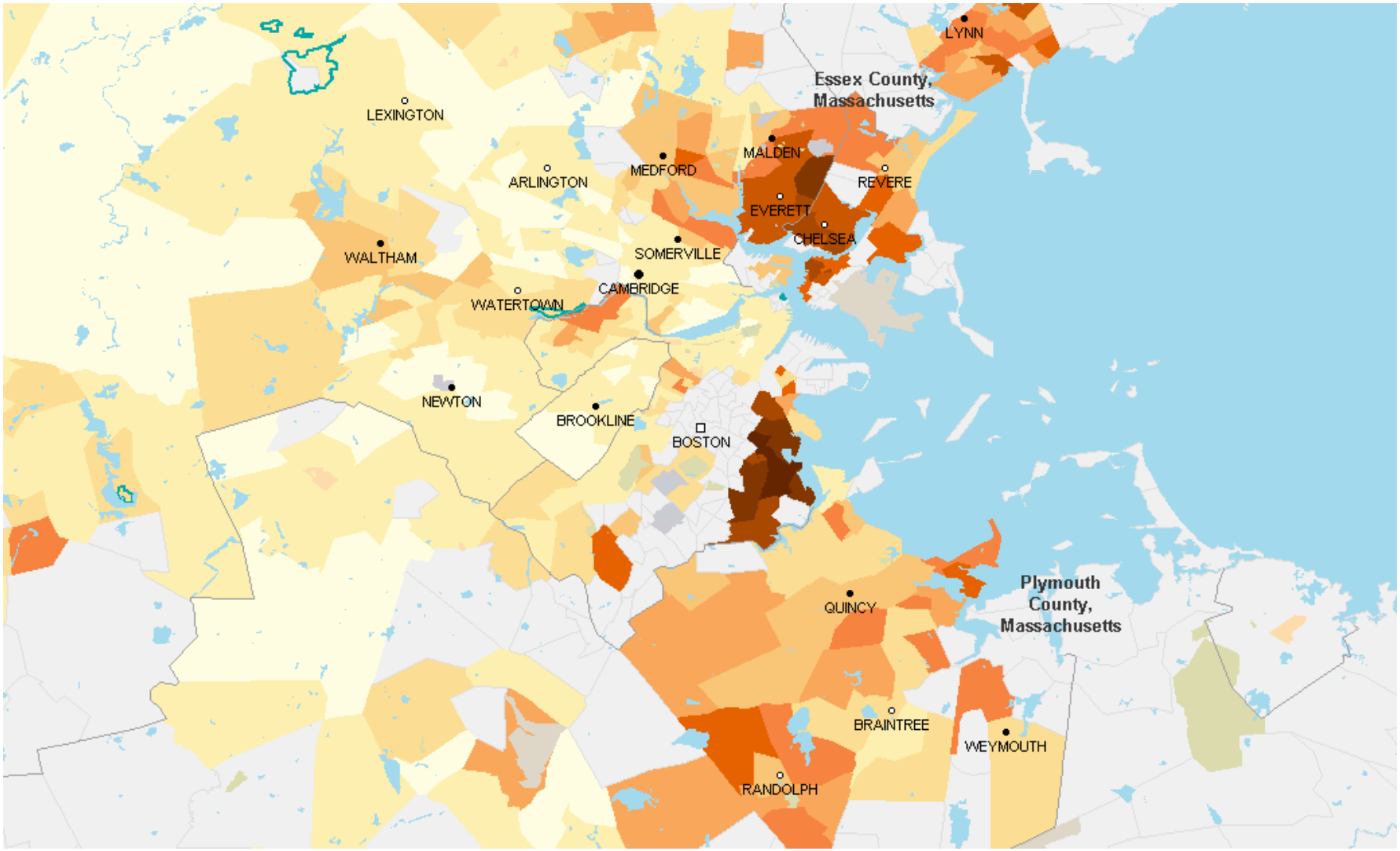
% Black in Los Angeles



The Most Segregated Metropolitan Area in the U.S.



% Vietnamese in Boston



Measuring Segregation

- Many indices of spatial segregation have been proposed over the years
  - The U.S. Census computes 19 and reports 5
- Two indices a particularly common in economics
  1. Dissimilarity index
  2. Isolation index
- Defined for categorical variables (like race, rather than income)
- Focused on one group relative to all others
  - Group $\mathcal{G}$ has G total members and g_i in location i
  - The remainder $\neg\mathcal{G}$ has N total members and n_i in location i
  - Aggregates $G + N = T$ and $g_i + n_i = t_i$
  - Share $G / T = \gamma$
- Very much subject to the modifiable areal unit problem

Dissimilarity Index

$$
\begin{aligned}
D_g &= \frac{1}{2}\sum_i \left|\frac{g_i}{G} - \frac{n_i}{N}\right| \\
&= \frac{1}{2}\frac{1}{1-\gamma}\sum_i \left|\frac{g_i}{G} - \frac{t_i}{T}\right| \\
&= \frac{1}{2}\frac{1}{\gamma\left(1-\gamma\right)}\sum_i \frac{t_i}{T}\left|\frac{g_i}{t_i} - \gamma\right|
\end{aligned}
$$

- Normalized: $D_g \in [0, 1]$
- The share of either group that needs to move to yield a uniform distribution across locations
- Invariant to equiproportional increases in the size of either group in each location

Isolation Index

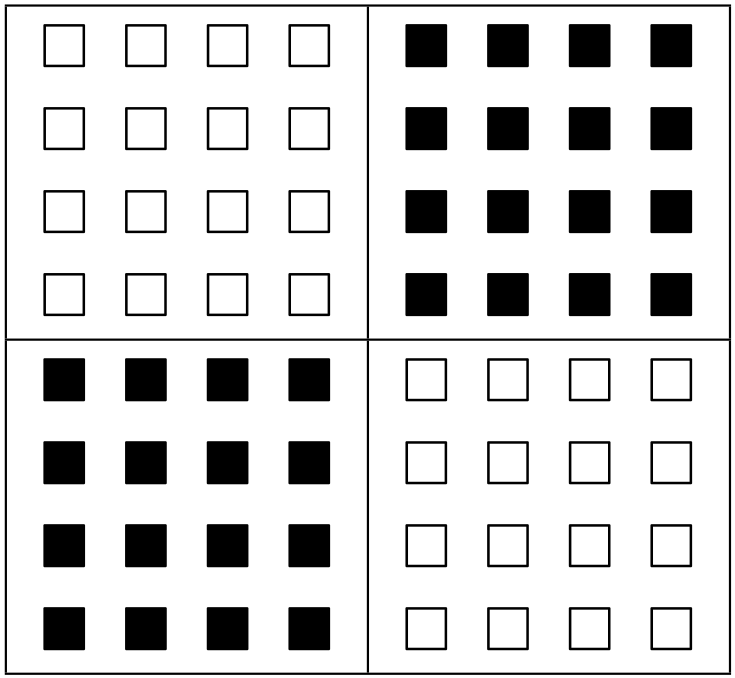
$$I_g = \sum_i \frac{g_i}{G} \times \frac{g_i}{t_i}$$

- The share of group $\mathcal{G}$ in the location where its average member lies
- Not normalized: $I_g \in [\gamma, 1]$
- Normalization yields the correlation ratio

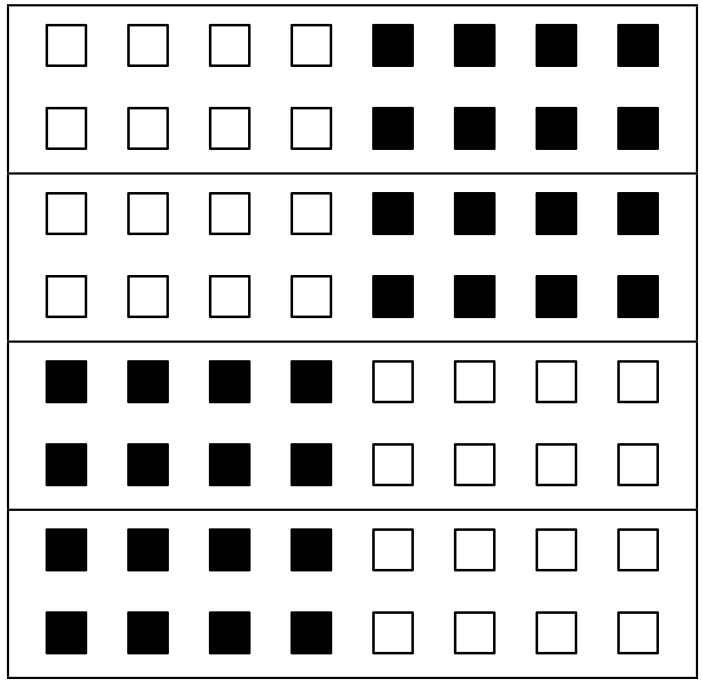
$$\begin{aligned} C_g &= \frac{I_g - \gamma}{1 - \gamma} \\ &= \frac{1}{\gamma (1 - \gamma)} \sum_i \frac{t_i}{T} \left(\frac{g_i}{t_i} - \gamma \right)^2 \end{aligned}$$

- Like the dissimilarity index with the L_2 instead of the L_1 distance
  - More sensitive to extreme observations

Partitioning Matters



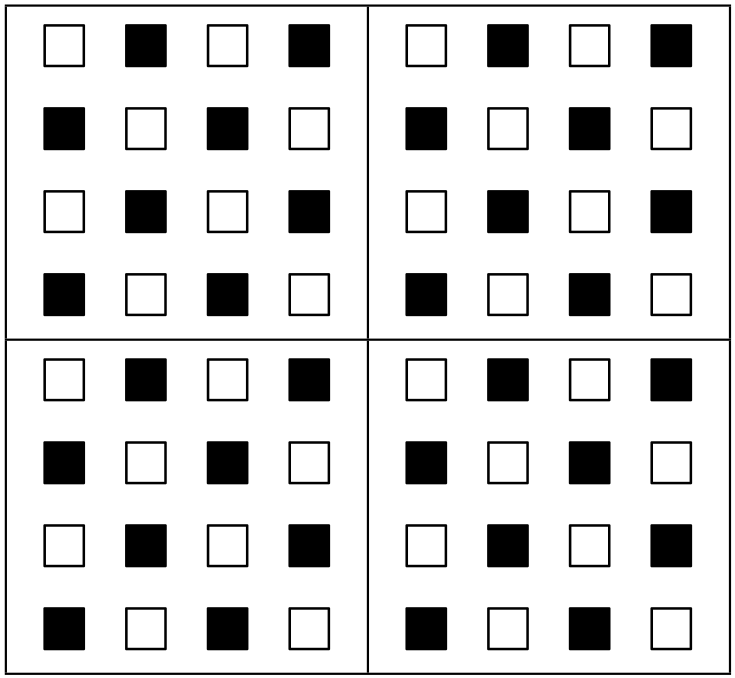
$D_g = I_g = C_g = 1$



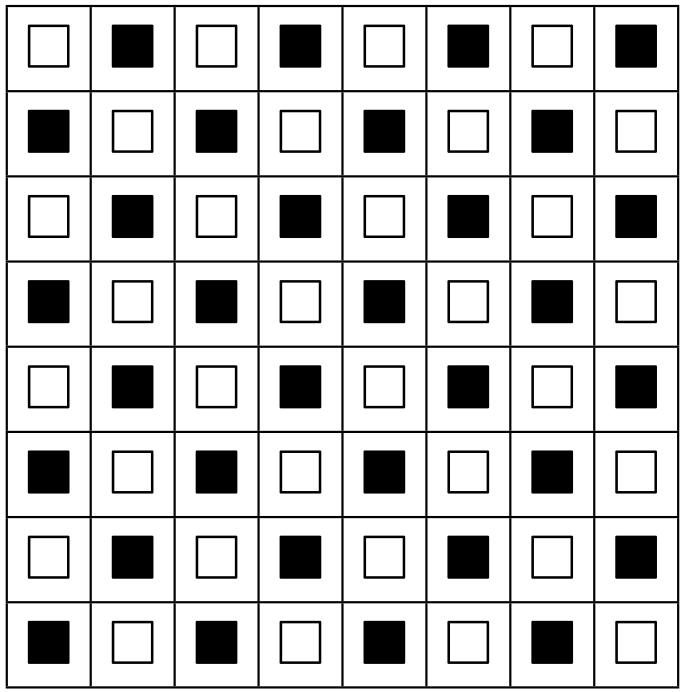
$D_g = C_g = 0$ and $I_g = 1/2$

- You can always gerrymander perfectly segregated areas

Aggregation Matters



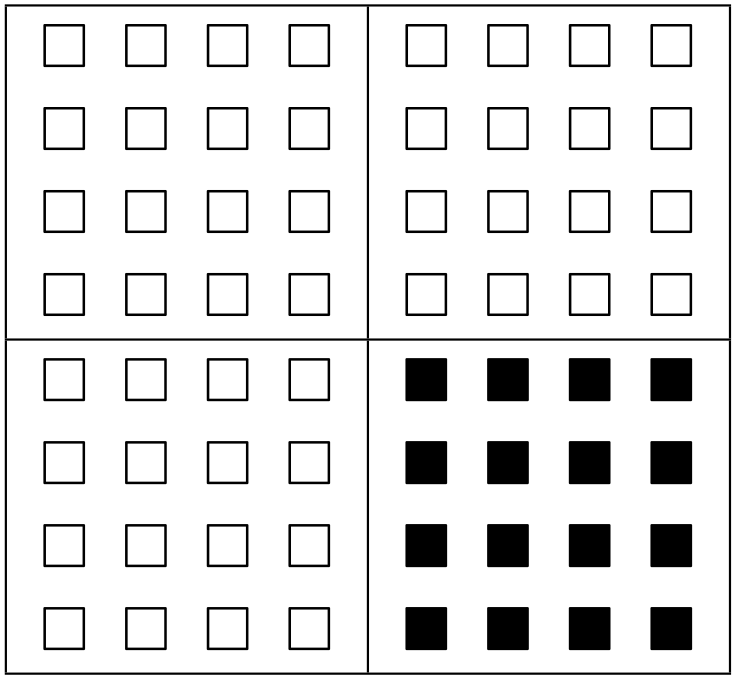
$D_g = C_g = 0$ and $I_g = 1/2$



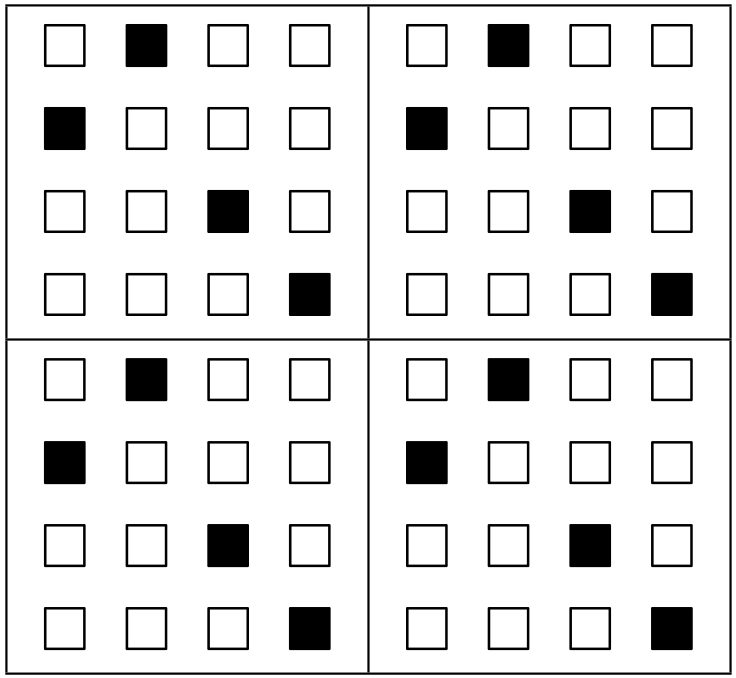
$D_g = I_g = C_g = 1$

- The indices are weakly increasing in the fineness of the partition
- Individuals are perfectly segregated by definition
- The whole is perfectly integrated by definition

Group Size Matters: G = 1/4

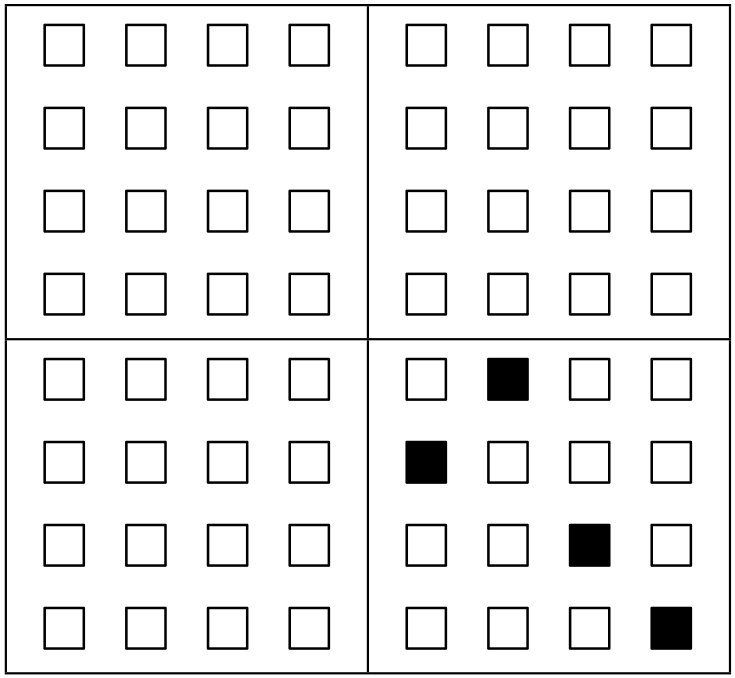


$$D_g = I_g = C_g = 1$$

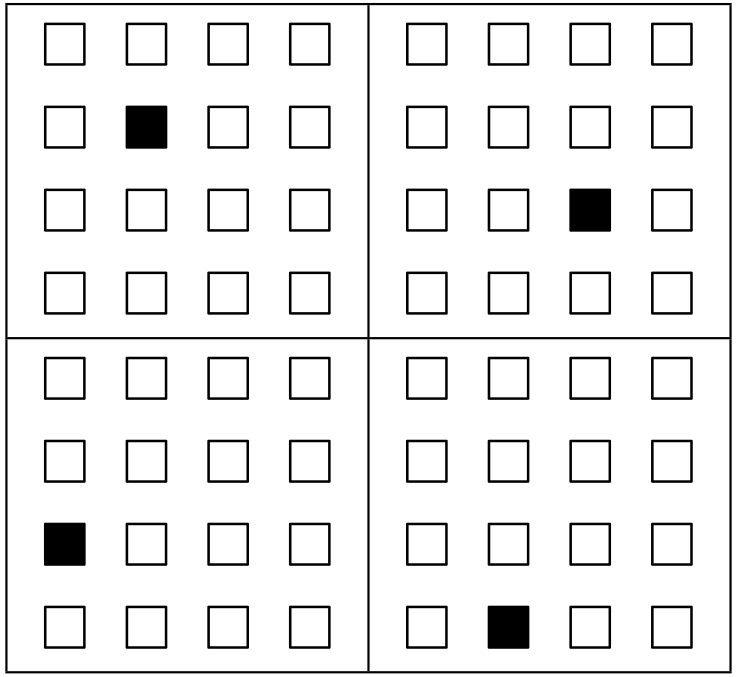


$$D_g = C_g = 0 \text{ and } I_g = 1/4$$

Group Size Matters: G = 1/16



$D_g = 0.8$, $C_g = 0.2$, and $I_g = 0.25$



$D_g = C_g = 0$ and $I_g = 1/16$

- The minority cannot fill an entire neighborhood

More Sophisticated Measurement

Echenique and Fryer's (2007) *Spectral Segregation Index*

- Computed on the basis of individual data
- Not subject to the modifiable areal unit problem
- Grounded axiomatically in the theory of social interactions

Intuition based on social networks

- Map the network of individual social interactions
- Measure if individuals are disproportionately connected to members of their own group, both directly and indirectly

Spectral Segregation Index

- Applications depend on data availability

1. School segregation: precise application
   - Individual survey data
   - Reported friendship links
2. Residential segregation: approximate aplication
   - Census blocks (300 households) as "individuals"
   - Race is assigned from the majority in the block
   - Links are presumed among blocks closer than 1 km

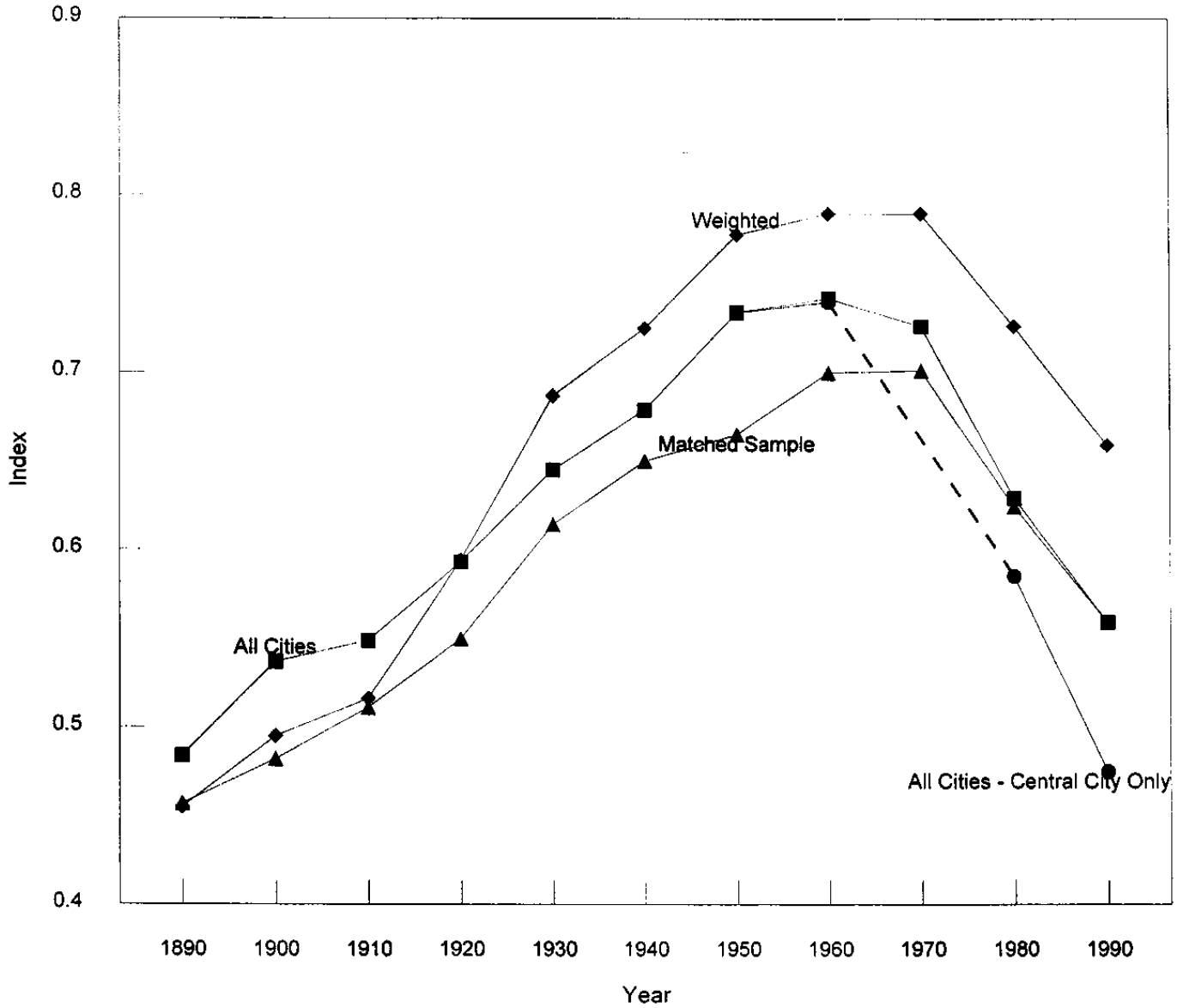
- Possibly the measure of the future, but not of the present
  - 0.93 correlation with the isolation index of residential segregation
  - More complicated intuition and data analysis

Three Periods in the History of U.S. Segregation

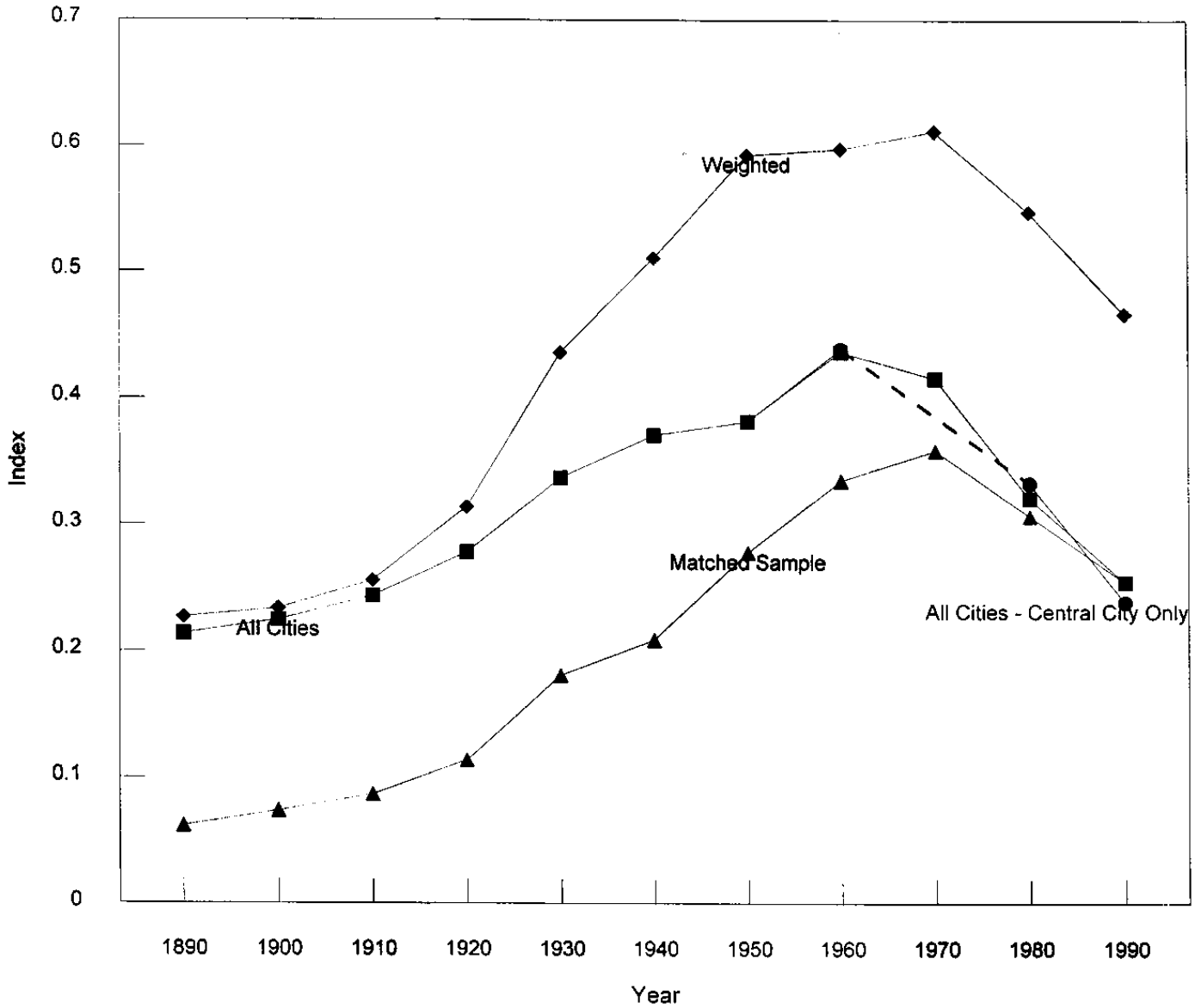
1. The birth of the ghetto, 1890 to 1940
   - First large-scale black migration from rural South to urban North
2. Consolidation and expansion, 1940 to 1970
   - Continued migration, increasing racial tensions
   - Ghettos came to dominate inner cities
3. Decline of segregation, since 1970
   - Particularly strong in the Sun Belt
   - Segregation has been and is declining, but remains high

- The relative segregation of different cities is very stable over time
  - Larger, denser cities always tend to be more segregated

The Rise and Decline of Dissimilarity



The Rise and Decline of Isolation



Tipping into Segregation

- Schelling's (1971, 1978) seminal contribution
- Two groups with mild homophily
  - Everyone wants a minimum share of neighbors from the same group
- Start with a random distribution of individuals over space
- Everyone who's unhappy with his neighborhood moves to an empty lot
- The system converges dynamically to very strong segregation
  - Solve numerically and with cool animations
- Complete segregation in equilibrium although nobody demands it
  - Individuals are happy in fairly mixed neighborhoods
  - Aggregate structure emerges from individual behavior
- The starting point of "agent-based modelling" and complexity
  - Still unorthodox and out of the mainstream, with unbounded ambition

A Simple Model of Sorting by Race

- Two equally sized locations, 1 and 2
- Two unequally sized groups: white majority, black minority
  - A fraction $b < 1/2$ of the population is black
- Idiosyncratic tastes for locations
  - A preference a for location 1
  - Distributed symmetrically around zero
  - Cumulative distribution function $F(a)$ for both groups
- Homophily
  - Let b_i be the share of blacks in location i
  - Blacks derive a benefit $\alpha_B b_i$
  - Whites incur a cost $\alpha_W b_i$
- Discrimination
  - A tax G is levied on blacks in location 1

Indifference of the Marginal Agent

- Since whites are in the majority they must live in both locations
- Equilibrium price premium for location 1

$$P + \alpha_W b_1 - a_W^* = \alpha_W b_2$$

  - a_W^* is the preference of the marginal white resident of location 1
- If blacks also live in both locations

$$P + G - \alpha_B b_1 - a_B^* = -\alpha_B b_2$$

- If blacks only live in location 2

$$P + G + \alpha_B b_2 \geq a_{\max}$$

  - Impossible if a has unbounded support

Taste-Based Sorting

- Within each group, individuals sort by taste
  - For whites

$$\frac{1}{2}(1-b_1) = [1-F(a_W^*)](1-b) \Leftrightarrow a_W^* = F^{-1}\left(1-\frac{1}{2}\frac{1-b_1}{1-b}\right)$$

  - For blacks

$$\frac{1}{2}b_1 = [1-F(a_B^*)]\,b \Leftrightarrow a_B^* = F^{-1}\left(1-\frac{1}{2}\frac{b_1}{b}\right)$$

- Adding-up constraint

$$\frac{1}{2}(b_1+b_2) = b \Leftrightarrow b_2 = 2b-b_1$$

Equilibrium Segregation

- Price P is consistent with equilibrium if and only if

$$2\alpha_W (b - b_1) + F^{-1}\left(1 - \frac{1}{2}\frac{1 - b_1}{1 - b}\right) = P$$

$$= -2\alpha_B (b - b_1) - G + F^{-1}\left(1 - \frac{1}{2}\frac{b_1}{b}\right)$$

  - The left-hand side is the marginal white's willingness to pay
  - The right-hand side is the marginal black's willingness to pay

- Heuristic equilibrium stability
  - If one more black is moved into region 1 he wants to leave
  - The left-hand side rises with b_1 faster than the right-hand side

$$-2\alpha_W + \frac{1}{2(1 - b) f\left(a_W^*\right)} > 2\alpha_B - \frac{1}{2bf\left(a_B^*\right)}$$

  - Exogenous location preferences matter more than homophily
  - Eventually $\lim_{a_B^* \to \infty} f\left(a_B^*\right) = 0$ assures stability

Basic Causes of Racial Segregation

- Locally, for a stable equilibrium

$$\frac{\partial b_1}{\partial \alpha_W} < 0, \frac{\partial b_1}{\partial \alpha_B} < 0, \frac{\partial b_1}{\partial G} < 0$$

1. The share of blacks in the "white location" falls with homophily
   - White racism α_W, black clannishness α_B? Not necessarily
   - Consumption patterns vary systematically by ethnic group
   - Segregation can enable better input-sharing for consumers
   - Waldfogel's evidence for various demographic groups
     - ★ Tastes correlate with income, education, age, family size, ethnicity ...
2. The share of blacks in the "white location" falls with discrimination
   - Institutionalized racism

Price Diagnostics

- Suppose that demand for a location by each group slopes down
  - Stronger than, but similar to, the stability condition
- When the right-hand side falls with b_1, $\frac{\partial P}{\partial \alpha_W} > 0$
- When the left-hand side rises with b_1, $\frac{\partial P}{\partial \alpha_B} < 0$ and $\frac{\partial P}{\partial G} < 0$

1. Decentralized homophily raises the price of the "white location"
   - It need not be racism and an unfair, oppressive phenomenon
2. Institutionalized racism raises the price of the "black location"
   - It is unambiguously racist, unfair, and oppressive

- Price gradients only distinguish between modes of white racism
  - Black homophily is a confound for institutionalized white racism
  - A second-order phenomenon in the light of U.S. history

Evidence of Institutionalized Racism

- Kain and Quigley (1972) studied the housing market in St. Louis
  - A sample of 401 black and 784 households in 1967

1. Blacks faced higher quality-adjusted housing prices
   - The discrimination markup was around 7%
2. Blacks were 9% less likely to own, given household characteristics
   - True of those who moved too
   - Being forced to rent is another large discrimination markup

- Consistent with barriers to moving out of the ghetto
  - Uncontroversial today that such barriers existed into the 1960s
  - Fair Housing in Title VII of the Civil Rights Act of 1968
- Others found instead that blacks paid less (decentralized racism)
  - Kain would and did argue they were not controlling properly for quality
  - The literature has abandoned hope of properly measuring quality

Cross-City Evidence

- Cutler, Glaeser and Vigdor (1999) avoid measuring quality
- Difference in differences across metropolitan areas
  - How does the racial price gradient vary with segregation?

1. Institutionalized racism in the mid-twentieth century
   - Blacks paid relatively more in more segregated cities
   - Collective action by whites to keep blacks out of their neighborhoods
2. Decentralized racism in the present day
   - Whites pay relatively more in more segregated cities
   - Differential willingness to pay remains perfectly legal

- It was not and is not a matter of black preferences
  - The differential was not higher for newly immigrated blacks
  - 67% of blacks prefer neighborhoods that are not majority black

Segregation and House Rents

Dependent Variable: ln(Annual Rental Payment)

Independent Variable	1940 City (1)	1940 City (2)	1970 MSA (3)	1970 MSA (4)	1970 Tract (5)	1990 MSA (6)
Black head of household	−1.301**	−1.443**	−.359**	−.416**	−.063**	.155*
	(.258)	(.369)	(.159)	(.132)	(.005)	(.075)
Dissimilarity × black head of household	1.262**	1.328**	.261	.377**	···	−.354*
	(.320)	(.516)	(.207)	(.163)		(.110)
Black × population growth rate past 30 years	−4.290	−3.787	1.034	.652	···	2.072
	(2.812)	(2.431)	(1.854)	(1.871)		(2.358)
Born in different state	···	.089	···	···	···	···
		(.097)				
Born in different state × dissimilarity	···	−.134	···	···	···	···
		(.113)				
Born in different state × black	···	.310	···	···	···	···
		(.465)				
Born in different state × black × dissimilarity	···	−.246	···	···	···	···
		(.636)				
Percentage of renter households living in structure built in past 30 years × black	···	···	−.052**	−.050**	···	−.042*
			(.022)	(.020)		(.021)
Share of MSA population in suburbs × black	···	···	−.081	−.021	···	−.174*
			(.062)	(.062)		(.064)
Log(median family income in tract)	···	···	···	···	.720**	···
					(.010)	
Percentage black in tract	···	···	···	···	.123**	···
					(.009)	
Public housing units per capita × black	···	···	···	···	···	−6.70**
						(3.15)
Section 8 rent subsidy payments per capita × black	···	···	···	···	···	−.000
						(.000
City/MSA fixed effects	yes	yes	yes	yes	no	yes
Structural characteristics	no	no	no	yes	yes	no
R^2	.199	.200	.126	.290	.490	.194
Observations	61,180	61,180	145,236	145,236	156,369	193,61
Number of cities/MSAs	40	40	111	111	···	237

Taste-Based Discrimination

- Becker (1957) for employment, but with broader applicability

1. Employers' tastes: sacrifice profits to avoid hiring minority workers
   - Equilibrium discrimination requires some monopoly power
   - Incentives for segregation to avoid the cost of mismatch
   - Minority workers tend to be paid less for equal productivity
2. Workers' tastes: demand higher wages to tolerate minority colleagues
   - Very strong force for segregation across employers
   - Not driven out by competitive markets
   - Also: majority workers are less productive with minority colleagues
3. Consumers' tastes: accept higher prices to avoid minority suppliers
   - Segregation by occupation, not by employer
   - Not driven out by competitive markets
   - High-status customer-service professions: doctors, lawyers, ...

Classical Statistical Discrimination

- A signal-extraction model (Arrow 1972, 1973; Phelps 1972)

1. Different means of the distribution of ability
   - On average, all workers are paid according to productivity
   - Each worker is paid based on both own and group productivity
   - An individual from the worse group is paid less for identical productivity
2. Different precision of the signal of ability (Aigner and Cain 1977)
   - Rational stereotyping of workers from the minority groups
   - Lower wages for the same productivity if employers are risk averse
3. Self-fulfilling expectations (Akerlof 1976; Coate and Loury 1993)
   - A reasonable reading of Myrdal (1944) though not his whole story
   - Employers expect one group to have lower skill
   - Individual investment in skill has lower returns for workers in that group
   - The group endogenously acquires lower skills

Collective Discrimination

- Most attention in economics has gone to competitive discrimination
- Collective action enables the worst forms of discrimination
  - Institutionalized racism and the cost G in our simple model

1. Bottom-up conspiracy by the majority
   - Extract rents by removing competition from the minority
   - A plausible component of discrimination in housing markets
   - In labor markets, white males were (like) a monopsonist trade union
   - Something of this in Akerlof's (1976) model of social stigma
2. Top-down conspiracy by the elite
   - Glaeser's (2005) model of incorrect statistical discrimination
   - An old Marxian and Marxist idea
   - The elite fosters ethnic conflict to fool the masses out of class conflict
     - ★ Reactionary European monarchists vs. Jews
     - ★ Southern U.S. plutocrats vs. blacks
     - ★ Some Democrats' view of the "culture wars" today

Are Ethnic Neighborhoods Bad for Their Residents?

Potential *costs* of segregation of ethnic minorities in ghettos

- Peer group effects, social interactions, neighborhood effects
- Income segregation of poor minorities
- Negative role models and norms
  - Organized crime: Italian mafia, inner-city gangs
- Ghetto residents don't acquire mainstream skills and norms
  - The community cannot help and may intentionally hinder
  - Fryer et al. on the social stigma against "acting white"
- Spatial mismatch: ghettos may be far from job opportunities

Are Ethnic Neighborhoods Good for Their Residents?

Potential *benefits* of segregation of ethnic minorities in ghettos

- Peer group effects, social interactions, neighborhood effects
- Ethnic segregation can counteract income stratification
- Positive role models and norms
  - Jewish ghettos allowed community leaders to punish misbehavior
- Ghetto residents learn more easily from members of their own group
  - Immigrant communities sharing language and broader culture
  - Historically a pathway to assimilation in the mainstream
- Also learning about job opportunities

Assessing the Impact of Ethnic Segregation

Ultimately the question is empirical

- Not necessarily the same answer across space and over time

Basic empirical strategy

1. Collect outcomes of individuals from an ethnic minority
2. Compare outcomes of those living in and out of the ghetto

Basic identification challenge

- Individual residents choose where to live, leading to sorting
- More successful people are more likely to leave the ghetto
  - Income stratification is the main driving force
  - Assimilation into the mainstream also plays a role

$\Rightarrow$ Negative bias in naive estimates of the effect of the ghetto

Spatial Mismatch

Before spillovers and peer-group effects there was Kain (1968)

1. Good jobs are physically too distant to commute to from the ghetto
2. Ghetto residents have less information about distant good jobs
3. Residential integration may break employment discrimination

- Spatial mismatch strictly speaking refers to the first hypothesis
- It has not fared particularly well over the decades
  1. Blacks were not and are not more distant from jobs than whites
  2. Physical distance from "good jobs" has not proved very important
- But the other hypotheses point towards the modern literature
  - In 1968 economists would have resisted sociological explanations

Cross-City Analysis

- Cutler and Glaeser (1997) compare across but not within cities
  - Avoids the problem of intra-city stratification
- Identification problem: poor citywide outcomes may cause segregation
- Instrumental variable approach
  1. Structure of local government finance
     - ★ Number of local municipal governments
     - ★ Share of local revenue from intergovernmental sources
  2. Topography of the metropolitan area
     - ★ Number of rivers (Hoxby 2000)
- Focus on young people born in the U.S.
  - Theories about learning and peer effect apply more strongly
  - Chances to move across cities mechanically increase with age
  - Robust to using segregation in movers' city of origin

Raw Differences in Differences

	Age 20–24				
	Education		Income		Social
	High school graduate	College graduate	Idle	ln(earn)	Single mother
Black					
Low segregation	79.5%	4.4%	15.4%	8.77	36.7%
High segregation	74.0	4.9	21.6	8.61	39.9
Difference	−5.5	0.5	6.2	−0.16	3.2
Nonblack					
Low segregation	86.7%	10.6%	7.0%	9.03	10.8%
High segregation	87.3	14.7	6.6	9.05	9.4
Difference	0.6	4.1	−0.4	0.02	−1.4
Difference-in-difference (B−W)	−6.1% (0.7%)	−3.7% (0.7%)	6.6% (0.6%)	−0.17 (0.03)	4.6% (0.9%)

OLS Estimates

	Age 20–24				
	Education		Income		Social
Independent variable	High school graduate	College graduate	Idle	ln(earn)	Single mother
Segregation					
Segregation	.016	0.67	−.006	−.060	.008
	(.033)	(.040)	(.019)	(.069)	(.030)
Segregation ∗ black	−.323	−.081	.324	−.740	.355
	(.044)	(.035)	(.044)	(.150)	(.063)
Demographics					
Black	−.599	.018	.388	−1.682	.650
	(.283)	(.327)	(.313)	(.772)	(.356)
Asian	.042	.064	−.013	−.034	−.007
	(.012)	(.027)	(.008)	(.051)	(.019)
Other nonwhite	−.134	−.088	.092	−.260	.193
	(.018)	(.010)	(.018)	(.045)	(.026)
Hispanic	−.161	−.090	.086	−.152	.129
	(.013)	(.012)	(.010)	(.022)	(.015)
Female	.029	.026	.054	−.278	—
	(.002)	(.003)	(.003)	(.015)	

OLS Estimates (Continued)

	Education		Income		Social
Independent variable	High school graduate	College graduate	Idle	ln(earn)	Single mother
MSA characteristics					
ln(population)	.005	.016	−.003	−.002	−.004
	(.003)	(.005)	(.003)	(.008)	(.003)
ln(population) ∗ black	.007	−.010	−.006	.045	−.032
	(.005)	(.004)	(.005)	(.016)	(.006)
Percent black	−.062	.078	.007	.432	−.098
	(.043)	(.061)	(.021)	(.104)	(.028)
Percent black ∗ black	.008	−.106	−.004	−.387	−.012
	(.071)	(.049)	(.055)	(.177)	(.068)
ln(median household income)	.028	−.001	−.060	.597	−.008
	(.020)	(.042)	(.009)	(.051)	(.013)
ln(median household income) ∗ black	.054	.008	−.036	.129	−.009
	(.024)	(.033)	(.028)	(.064)	(.030)
Manufacturing share	−.149	−.152	.049	.251	.087
	(.067)	(.082)	(.035)	(.166)	(.047)
Manufacturing share ∗ black	.108	.128	.035	−.795	−.085
	(.103)	(.071)	(.093)	(.308)	(.143)
Summary statistics					
N	97,976	97,976	97,976	56,627	49,038
σ^2_ε	.121	.096	.076	.868	.111
R^2	.034	.093	.050	.090	.108

IV Estimates

	Age 20–24					Age 25–30				
	Education		Income		Social	Education		Income		Social
Independent variable	High school graduate	College graduate	Idle	ln(earn)	Single mother	High school graduate	College graduate	Idle	ln(earn)	Single mother
				A. Fiscal variables as instruments						
Segregation	.129	.211	−.046	−.042	−.051	.076	.095	.005	−.005	−.108
	(.044)	(.053)	(.025)	(.095)	(.038)	(.032)	(.077)	(.028)	(.090)	(.035)
Segregation ∗ black	−.405	−.201	.317	−.921	.326	−.231	−.121	.295	−.532	.583
	(.085)	(.056)	(.087)	(.236)	(.101)	(.076)	(.069)	(.062)	(.196)	(.116)
N	97,976	97,976	97,976	56,627	49,038	139,715	139,715	139,715	105,997	71,531
σ^2_ε	.121	.096	.076	.868	.111	.107	.181	.092	.835	.123
				B. Topographical data as instruments						
Segregation	.040	.122	.018	−.208	.105	.003	.034	−.068	−.126	−.020
	(.078)	(.099)	(.051)	(.217)	(.066)	(.063)	(.150)	(.052)	(.238)	(.065)
Segregation ∗ black	−.579	−.168	.329	−1.100	.261	−.291	−.149	.558	−.719	1.030
	(.199)	(.109)	(.173)	(.602)	(.217)	(.135)	(.140)	(.184)	(.299)	(.242)
N	90,684	90,684	90,684	52,281	45,442	129,324	129,324	129,324	97,973	66,276
σ^2_ε	.122	.096	.076	.873	.112	.107	.180	.093	.838	.124
				C. Fiscal variables from city of residence five years previously						
Segregation	.189	.238	−.060	.082	−.077	.112	.232	−.018	.143	−.137
	(.044)	(.051)	(.022)	(.108)	(.040)	(.036)	(.079)	(.024)	(.094)	(.038)
Segregation ∗ black	−.265	−.229	.197	−.791	.109	−.231	−.255	.311	−.566	.480
	(.083)	(.053)	(.086)	(.266)	(.091)	(.076)	(.073)	(.062)	(.188)	(.107)
N	95,955	95,955	95,955	54,084	47,950	137,496	137,496	137,496	104,078	70,596
σ^2_ε	.124	.092	.078	.895	.112	.110	.177	.095	.869	.126

How Does Segregation Affect Outcomes?

- Blacks are significantly *worse off* in segregated metropolitan areas
  - $\sigma = 13\%$ fall in segregation $\Rightarrow$ eliminate 1/3 of the racial gap
- Negligible effect for whites (though positive with IV)
- Theoretical channels to explain the effect (OLS only)
  1. No evidence that the effect works through income stratification
  2. Some evidence of lack of education spillovers for all outcomes
     - ★ Partly interpreted as worse parents' education
  3. Some evidence of spatial mismatch for employment outcomes
- All these channels account for at most a third of the impact

$\Rightarrow$ Segregation is extremely harmful for blacks
No exact understanding of why this is true

Experimental Evidence

- The U.S. government ran an actual randomized field experiment
  - *Moving to Opportunity*, 1994–97
  - Baltimore, Boston, Chicago, Los Angeles, New York
- Lottery for resident of public housing in poor neighborhoods

S. Section 8: rent-subsidy voucher to move to any other neighborhood

E. Experimental: voucher to move to a non-poor neighborhood
  - Required to move within 1 year; received mobility counseling

C. Control group: no voucher, remain in public housing

- Voucher recipients still decide whether to move
  - 60% compliance in S, 47% compliance in E

Move to Better Neighborhoods

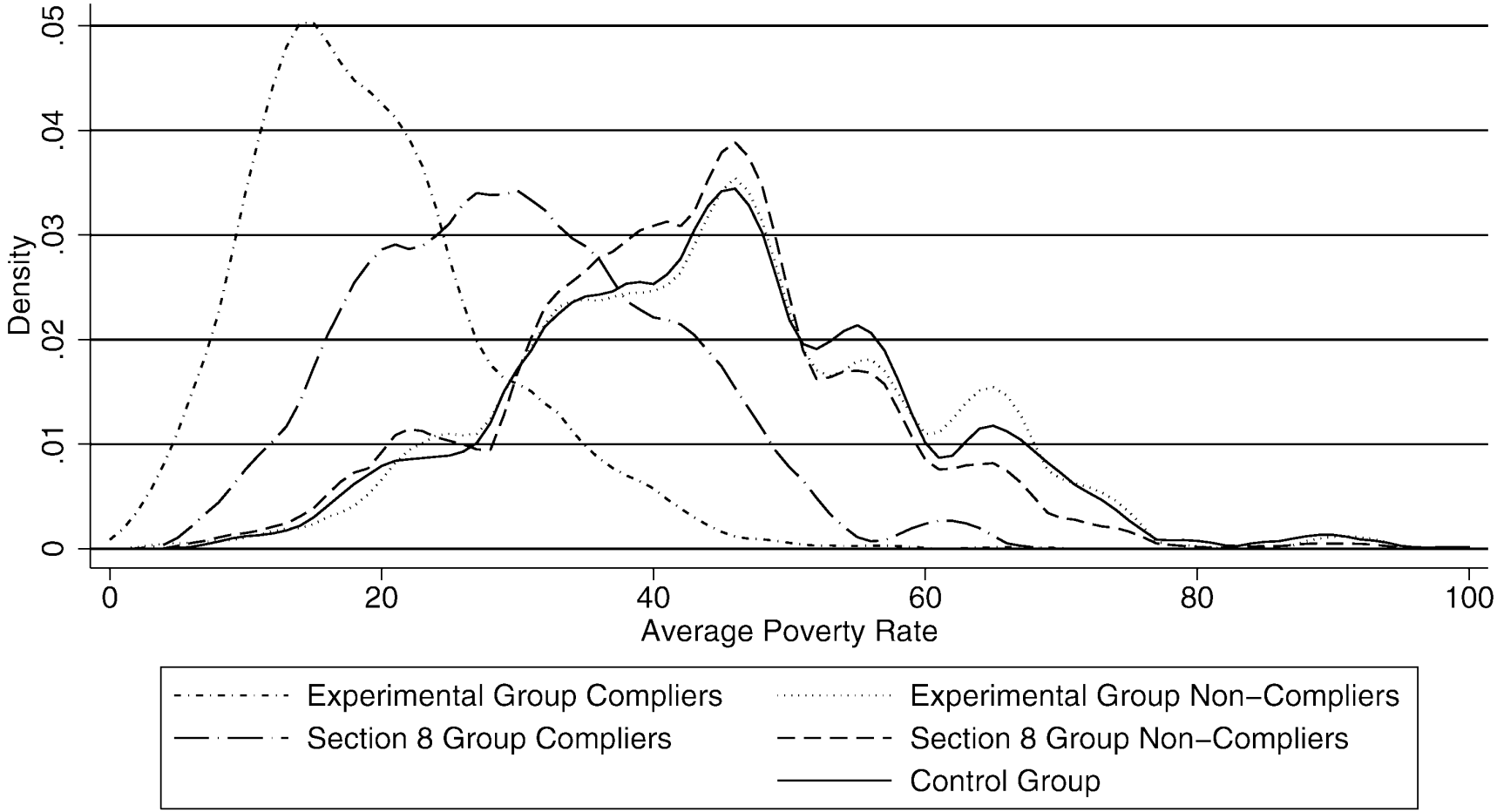


FIGURE 1.—Densities of average poverty rate, by group. Average poverty rate is a duration-weighted average of tract locations from random assignment through 12/31/2001. Poverty rate is based on linear interpolation of 1990 and 2000 Censuses. Density estimates used an Epanechnikov kernel with a half-width of 2.

Estimation

1. Intent-to-treat effect (ITT): reduced form OLS

$$Outcome_i = \pi_1 Voucher_i + \beta_1 Controls_i + \varepsilon_i$$

2. Treatment-on-treated effect (TOT): 2SLS

$$Outcome_i = \gamma_2 Moved_i + \beta_2 Controls_i + \varepsilon_i$$

   - Instrument $Moved_i$ with $Voucher_i$
   - Construct the control complier mean (CCM): mean outcome for those in C who would have moved if they had received a voucher

3. Effect of neighborhood quality: 2SLS

$$Outcome_i = \gamma_3 Poverty_i + \beta_3 Controls_i + \varepsilon_i$$

   - Instrument $Poverty_i$ with $Voucher_i$ interacted with neighborhoods

Treament Outcomes

SPECIFIC OUTCOMES WITH EFFECTS SIGNIFICANT AT 5 PERCENT LEVEL[a]

	E/S (i)	CM (ii)	ITT (iii)	TOT (iv)	CCM (v)
A. Adult outcomes					
Obese, BMI $\geq$ 30	E – C	0.468	−0.048	−0.103	0.502
			(0.022)	(0.047)	
Calm and peaceful	E – C	0.466	0.061	0.131	0.443
			(0.022)	(0.047)	
Psychological distress, K6 z-score	E – C	0.050	−0.092	−0.196	0.150
			(0.046)	(0.099)	
B. Youth (female and male) outcomes					
Ever had generalized anxiety symptoms	E – C	0.089	−0.044	−0.099	0.164
			(0.019)	(0.042)	
	S – C	0.089	−0.063	−0.114	0.147
			(0.019)	(0.035)	
Ever had depression symptoms	S – C	0.121	−0.039	−0.069	0.134
			(0.019)	(0.035)	

[a] E/S: indicates whether the row is experimental − control (E – C) or Section 8 − control (S – C). CM, control mean; ITT, intent-to-treat, from Equation (1); TOT, treatment-on-treated, from Equation (2); CCM, control complier mean. Robust standard errors adjusted for household clustering are in parentheses. The estimated equations all include site indicators and the baseline covariates listed in Appendix A with those in Table A1 included for adults and those in Tables A1 and A2 for youth. Rows shown in the table to illustrate magnitudes were selected based on ITT p-values < 0.05 and are 17 of 120 from the set of specific contrasts (E − C, S − C), based on the outcomes (15 for adults and 15 for youth) and subgroups—adults, youth (female and male), female youth, and male youth—described in the notes to Table II.

Treatment Outcomes by Gender

C. Female youth outcomes					
Psychological distress, K6 scale z-score	E – C	0.268	−0.289	−0.586	0.634
			(0.094)	(0.197)	
Ever had generalized anxiety symptoms	E – C	0.121	−0.069	−0.138	0.207
			(0.027)	(0.055)	
	S – C	0.121	−0.075	−0.131	0.168
			(0.029)	(0.051)	
Used marijuana in the past 30 days	E – C	0.131	−0.065	−0.130	0.202
			(0.029)	(0.059)	
	S – C	0.131	−0.072	−0.124	0.209
			(0.032)	(0.056)	
Used alcohol in past 30 days	S – C	0.206	−0.091	−0.155	0.306
			(0.038)	(0.056)	
D. Male youth outcomes					
Serious nonsports accident or injury in past year	E – C	0.062	0.087	0.215	0
			(0.026)	(0.064)	
	S – C	0.062	0.080	0.157	0
			(0.028)	(0.058)	
Ever had generalized anxiety symptoms	S – C	0.055	−0.049	−0.098	0.126
			(0.024)	(0.047)	
Smoked in past 30 days	E – C	0.125	0.103	0.257	0
			(0.032)	(0.084)	
	S – C	0.125	0.151	0.293	0.014
			(0.037)	(0.073)	

Move to Better Neighborhoods

EFFECTS OF NEIGHBORHOOD POVERTY RATES ON SELECTED OUTCOMES[a]

		Models			
		OLS	2SLS	2SLS	
Variables	Group	Poverty (i)	Poverty (ii)	Poverty (iii)	Compliance (iv)
Mental health	Adult	0.13	−0.62*	−1.35*	−0.17
		(0.17)	(0.24)	(0.60)	(0.13)
	Youth (female and male)	0.57	−0.97*	−0.18	0.20
		(0.34)	(0.41)	(0.87)	(0.21)
	Female youth	0.99	−1.84*	−1.88	−0.01
		(0.61)	(0.50)	(1.09)	(0.25)
Risky behavior	Female youth	−0.61	−0.94*	−1.03	−0.02
		(0.42)	(0.39)	(0.85)	(0.19)
Overall	Female youth	−0.03	−0.90*	−1.03	−0.03
		(0.28)	(0.26)	(0.56)	(0.12)
Physical health	Male youth	−0.84*	1.07*	1.77	0.18
		(0.35)	(0.49)	(1.09)	(0.26)
Risky behavior	Male youth	−0.06	1.46*	0.94	−0.13
		(0.42)	(0.54)	(1.29)	(0.31)
Overall	Male youth	−0.13	0.80*	1.47*	0.17
		(0.23)	(0.28)	(0.68)	(0.16)

Kling, Liebman, and Katz's (2007) Findings

- No significant effects on economic self-sufficiency
  - Unsuccessful as a direct antipoverty program
- No pattern of improvement in adult physical health
  - The t-statistic on obesity should be taken with a pinch of salt
- Substantial mental health benefits for adults and girls
  - Removes the stress of living in a dangerous violent neighborhood
  - Probably sufficient to make the program welfare-improving
- Teenage girls derive benefit across the board
  - Physical and mental health, risky behavior, education
- Teenage boys suffer adverse effects across the board
  - The gender difference runs counter to initial expectations
  - Ex post we can find sociological rationalizations

A European Natural Experiment

- Edin, Fredriksson, and Åslund (2003)
- Random initial placement of refugees in Sweden, 1985–91
  - Based on availability of housing
  - Unrelated to individual characteristics
- Some immigrants are assigned to ethnic enclaves, some aren't
- Later, everyone can freely choose where to locate
- Initial location is a valid instrument
  - Exogenous because of quasi-random assignment
  - Relevant because of persistence: jobs, social networks, ...

Random Initial Location

INDIVIDUAL CHARACTERISTICS BY INITIAL PLACEMENT

	Initial placement	
	Enclave	No enclave
Female	.44	.45
Age	37.3	37.6
	(7.7)	(7.4)
Years of schooling	11.3	11.7
	(3.0)	(2.9)
Married	.63	.62
Kid $\leq$15 years of age	.55	.57
No. of individuals	3094	3324

Standard deviations are in parentheses. An enclave is defined as described in the main text. Years of schooling are imputed from highest degree attained. Individuals with missing information on education were given the same number of years of schooling as those with less than nine years of schooling. All characteristics are measured eight years after immigration. The sample is restricted to those with positive earnings at that point in time.

Little Subsequent Sorting on Observables

INDIVIDUAL AND LOCAL CHARACTERISTICS BY MOBILITY STATUS

	Mobility status	
	Stayer	Mover
Individual characteristics		
Female	.48	.41
Age	38.1	36.6
	(7.9)	(7.0)
Years of schooling	11.2	11.8
	(2.9)	(3.0)
Married	.66	.59
Kid ≤15 years of age	.58	.53
Eastern Europe	.24	.15
Africa	.10	.13
Middle East	.33	.49
Asia	.10	.08
South America	.23	.14
Local characteristics		
Ethnic concentration (percent)	.32	.34
Immigrant density (percent)	7.85	8.38
Population size (1000)	217.0	234.8
Unemployment rate (percent)	5.87	5.76
No. of individuals	3492	2926

Standard deviations are in parentheses. Years of schooling are imputed from highest degree attained. Individuals with missing information on education were given the same number of years of schooling as those with less than nine years of schooling. All characteristics are measured eight years after immigration. The sample is restricted to those with positive earnings at that point in time.

Ethnic Concentration and Earnings

BASELINE ESTIMATES—DEPENDENT VARIABLE: ln(EARNINGS)

	Full sample		Low education (10 years or less)		High education (more than 10 years)	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
ln(ethnic group)	−.056	.012	−.053	.174	−.050	−.057
	(.022)	(.050)	(.024)	(.088)	(.030)	(.080)
...						
Immigration year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Country of origin dummies	Yes	Yes	Yes	Yes	Yes	Yes
Municipality dummies	Yes	Yes	Yes	Yes	Yes	Yes
No. of individuals	6393	6393	2205	2205	4188	4188
Standard error of regression	1.44	.145	1.44	1.48	1.42	1.42

Benefits of Ethnic Enclaves

- Evidence on sorting by unobservables across locations
- Naive OLS estimates suggest losses from greater segregation
- The instrument uncovers the opposite effect
- Earnings of the low-skilled actually rise with ethnic concentration
  - σ increase in concentration $\Rightarrow$ 13% increase in earnings
- The quality of the enclave matters
  - Large benefits for immigrants from ethnic groups with higher earnings
  - Immigrants from poorer ethnic groups may actually lose
- Suggestive of learning spillovers from human capital in the enclave
  - The effect also seems to be cumulative over time

Peer Effects

- Manski introduced the terminology for social interaction effects
- The outcome for individual i in neighborhood j is

$$y_i = \alpha + \beta \bar{y}_j + \gamma \bar{x}_j + \delta x_i + \varepsilon_i$$

1. *Endogenous* effect β
   - What you do depends on what your neighbors are doing
2. *Contextual* effect γ
   - What you do depends on who your neighbors are
3. *Exogenous* or *correlated* effect δ
   - What you do depends on who you are
   - This is the basic effect, but a confound in this context

The Reflection Problem

- A problem of linearity, not endogeneity

$$\bar{y}_j = \frac{1}{N_j}\sum_{i\in j} y_i = \alpha + \beta\bar{y}_j + (\gamma+\delta)\,\bar{x}_j + \frac{1}{N_j}\sum_{i\in j}\varepsilon$$

- In a large population, the variance of the error terms goes to zero

$$(1-\beta)\,\bar{y}_j = \alpha + (\gamma+\delta)\,\bar{x}_j$$

  - Manski assumed that $\bar{y}_j$ is a rational expectation, not a sample mean

$\Rightarrow$ Perfect multicollinearity between $\bar{y}_j$ and $\bar{x}_j$

- Impossible to separate endogenous and contextual effects
  - This is disturbing because they conceptually different phenomena
- Confounded neighborhood effect

$$y_i = \alpha + \zeta\bar{x}_j + \delta x_i + \varepsilon_i$$

The Social Multiplier

- Stronger relationship between x and y higher levels of aggregation

1. Endogenous effect: $y_i = \alpha + \beta \bar{y}_j + \delta x_i + \varepsilon_i$
   - Within-group effect: $\partial y_i / \partial x_i = \delta$
   - Between-group effect: $\partial \bar{y}_i / \partial \bar{x}_i = \delta / (1 - \beta)$

   $\Rightarrow$ Social multiplier $1 / (1 - \beta)$
2. Contextual effect: $y_i = \alpha + \gamma \bar{x}_j + \delta x_i + \varepsilon_i$
   - Within-group effect: $\partial y_i / \partial x_i = \delta$
   - Between-group effect: $\partial \bar{y}_i / \partial \bar{x}_i = \gamma + \delta$

   $\Rightarrow$ Social addend γ?

- The two models are confounded by the reflection problem
- The intuition for a multiplier is local complementarity of choices y

Variance Magnification

- Individuals simultaneously choose y_i to solve

$$\max_y \left\{ (\alpha + \beta \bar{y} + \delta x_i + \varepsilon_i) y - \frac{1}{2} y^2 \right\}$$

  - Stategic complementarities with reverse-engineered linearity

- Individual outcome with endogenous effect

$$y_i = \alpha + \beta \bar{y} + \delta x_i + \varepsilon_i$$

- Average outcome

$$\bar{y} = \frac{1}{1-\beta} (\alpha + \delta \bar{x} + \bar{\varepsilon})$$

$\Rightarrow$ Individual outcome with confounded neighborhood effect

$$y_i = \frac{\alpha}{1-\beta} + \frac{\beta}{(1-\beta) N} \sum_{j \neq i} (\delta x_j + \varepsilon_j) + \left[1 + \frac{\beta}{(1-\beta) N} \right] (\delta x_i + \varepsilon_i)$$

Uncorrelated Unobservables

- Idiosyncratic shocks: $\mathbb{E}\left(\varepsilon_i^2\right) = \sigma_\varepsilon^2$ and $\mathbb{E}\left(\varepsilon_i \varepsilon_j\right) = 0$
- Individual-level variance with endogenous effect

$$Var\left(y_i|x_i\right) = \left[1 + \frac{\beta\left(2-\beta\right)}{\left(1-\beta\right)^2 N}\right]\sigma_\varepsilon^2$$

- Group-level variance with social multiplier

$$Var\left(\bar{y}|\bar{x}\right) = \frac{1}{\left(1-\beta\right)^2 N}\sigma_\varepsilon^2$$

- Variance multiplier

$$\frac{Var\left(\bar{y}|\bar{x}\right)}{Var\left(y_i|x_i\right)} = \frac{1}{1+\left(1-\beta\right)^2\left(N-1\right)} \Rightarrow \frac{\partial}{\partial\beta}\frac{Var\left(\bar{y}\right)}{Var\left(y_i\right)} > 0$$

Correlated Unobservables

- Correlated shocks $\varepsilon_i = \mu + \nu_i$
  - Common component $\mathbb{E}\left(\mu^2\right) = \lambda\sigma_\varepsilon^2$
  - Idiosyncratic component $\mathbb{E}\left(\nu_i^2\right) = (1-\lambda)\,\sigma_\varepsilon^2$
- Individual-level variance with endogenous effect

$$Var\left(y_i|x_i\right) = \left\{\frac{\beta^2}{(1-\beta)^2}\lambda + \left[1 + \frac{\beta\,(2-\beta)}{(1-\beta)^2\,N}\right](1-\lambda)\right\}\sigma_\varepsilon^2$$

- Group-level variance with social multiplier

$$Var\left(\bar{y}|\bar{x}\right) = \frac{1}{(1-\beta)^2}\left(\lambda + \frac{1-\lambda}{N}\right)\sigma_\varepsilon^2$$

- Variance multiplier

$$\frac{Var\left(\bar{y}|\bar{x}\right)}{Var\left(y_i|x_i\right)} = \frac{1 + \lambda\,(N-1)}{1 - \lambda + \beta^2\lambda N + (1-\beta)^2\,(1-\lambda)\,(N-1)}$$

  - Now increasing in both β and λ: a confound

The Selection Problem

- The usual endogeneity problem
  - Context is a choice variable, not an exogenous variable
  - Selection on the basis of unobservables generates upward bias

1. Random or quasi-random assignment
   - If you can find the natural experiment or run a field experiment
2. Structural estimation of location choice
   - IO models in the Berry-Levinsohn-Pakes tradition
   - Some arbitrariness in choosing moment restrictions
   - Tendency toward ad hoc error terms wherever useful
   - Not properly identified with with endogenous location attributes
3. Instrument with predetermined individual characteristics
   - Plagued by correlations between observables and unobservables

Crime and Social Interactions

- Social multiplier for crime (Glaeser, Sacerdote, and Scheinkman 1996)
- Estimate from the variance of crime rates across cities
- Weak identification strategy
  - Control for observable city characteristics
  - Make structural assumptions about unobservables
  - Time variation with city fixed effects
  - Placebo test with mortality from disease and suicide
- Strength of social interactions declines with seriousness of crime
  1. Pettier property crimes: larceny, auto theft
  2. More violent crimes: assault, burglary, robbery
  3. Most serious crimes: arson, murder, rape

The Rational Crime Model

- Becker (1968): individuals optimally choose to be criminals iff

$$\theta_i < B - pC$$

  - Crime pays B
  - Individual cost θ_i: opportunity cost, criminal productivity, morals
  - Probability of being caught p
  - Expected punishment when caught C

- Two (in)famous predictions

1. 100% recidivism
   - Prison is more likely to raise B and lower θ_i than viceversa
   - Admittedly C may be higher for repeat offenders
2. "Boil 'em in oil!" p is expensive but C is cheap
   - Not so cheap if there are costs of punishing the innocent
   - Tremendous incentives for corruption
   - Marginal deterrence

Supply Elasticity of Crime

- Decompose $\theta_i = \bar{\theta} + \varepsilon_i$
  - City-specific mean $\bar{\theta}$
  - Idiosyncratic mean-zero shock ε_i with CDF $G(\varepsilon_i)$
- The amount of crime is

$$Q = G\left(B - pC - \bar{\theta}\right)$$

- Its elasticity with respect to city charactieristic X is

$$\frac{d \log Q}{d \log X} = \frac{g}{G}\left[B\frac{\partial \log B}{\partial \log X} - pC\left(\frac{\partial \log p}{\partial \log X} + \frac{\partial \log C}{\partial \log X}\right) - \bar{\theta}\frac{\partial \log \bar{\theta}}{\partial \log X}\right]$$

- Glaeser and Sacerdote (1999) consider city size

Crime and City Size

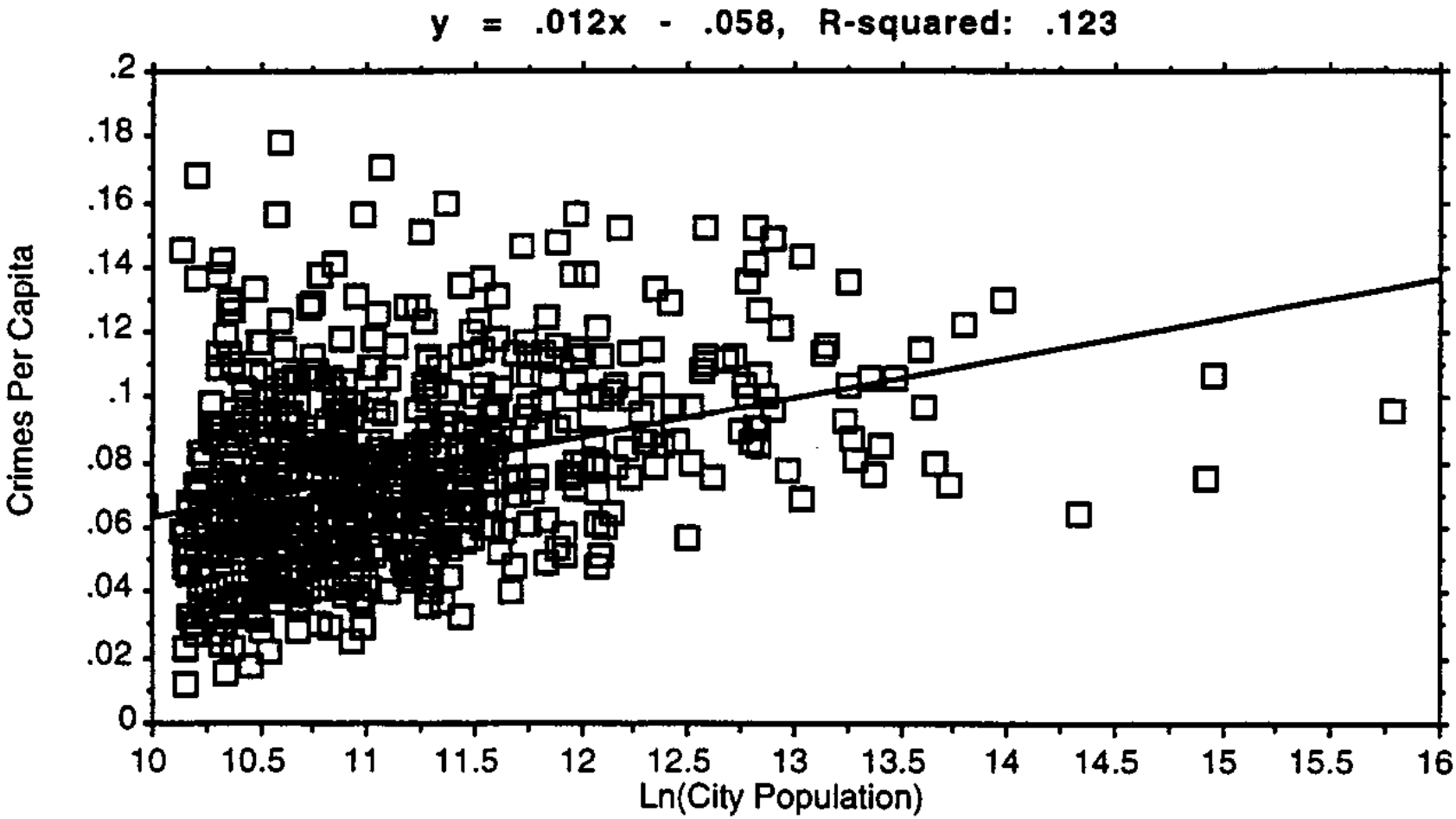


FIG. 1.—Crime and city population: relationship between crime and city population taken from the 1982 Uniform Crime Reports and the 1980 census. *t*-statistic is 9.44.

Why Is There More Crime in Large Cities?

B. Agglomeration economies increase productivity for crime too
   - Greater density of victims and wealthier victims to prey upon

p. A larger population makes it harder to catch criminals
   - More anonymity, greater number of suspects to monitor
   - Possibly being undone by economies of scale in new policing technology

(C.) Sentencing of criminals may be more lenient in large cities
   - A little bit of suggestive evidence that it is so

$\bar{\theta}$. Large cities attract or create crime-prone individuals
   - Strong evidence on the prevalence of female-headed households
   - Instrument with lagged welfare benefits to alleviate reverse causation

- About half of the correlation remains unexplained
   - Social interactions could be the remaining cause

Decomposition of the City Size–Crime Connection

IMPLEMENTING THE DECOMPOSITION

	Effect	Percentage of City Size–Crime Connection Explained by Effect
Initial city size–crime connection	.24	
Effect of deterrence:		
$\epsilon_P^Q = -.2$	.02	8.33
$\epsilon_P^Q = -.5$	.05	20.8
Effect of pecuniary returns:		
$\epsilon_P^Q = -.2$	.032	13.33
$\epsilon_P^Q = -.5$	.080	33.33
Effect of city composition	.07	29.2
Unexplained city size–crime connection:		
$\epsilon_P^Q = -.2$		49.14
$\epsilon_P^Q = -.5$		16.67

Crime with Strong Social Interactions

- Each rioter is less likely to be punished when there are more rioters
- Equilibrium size of the riot

$$R = G\left(B - p\left(R\right)C - \bar{\theta}\right)$$

- Well-behaved function $p\left(R\right)$ and $G\left(\varepsilon_i\right)$

1. Usually there are no riots: $B - p\left(0\right)C - \bar{\theta} < \min \varepsilon_i$
2. There are no universal riots: $G\left(B - p\left(1\right)C - \bar{\theta}\right) < 1$

- Multiple equilibria

$$p\left(R\right)C = B - G^{-1}\left(R\right) - \bar{\theta}$$

Costs and Benefits of Rioting

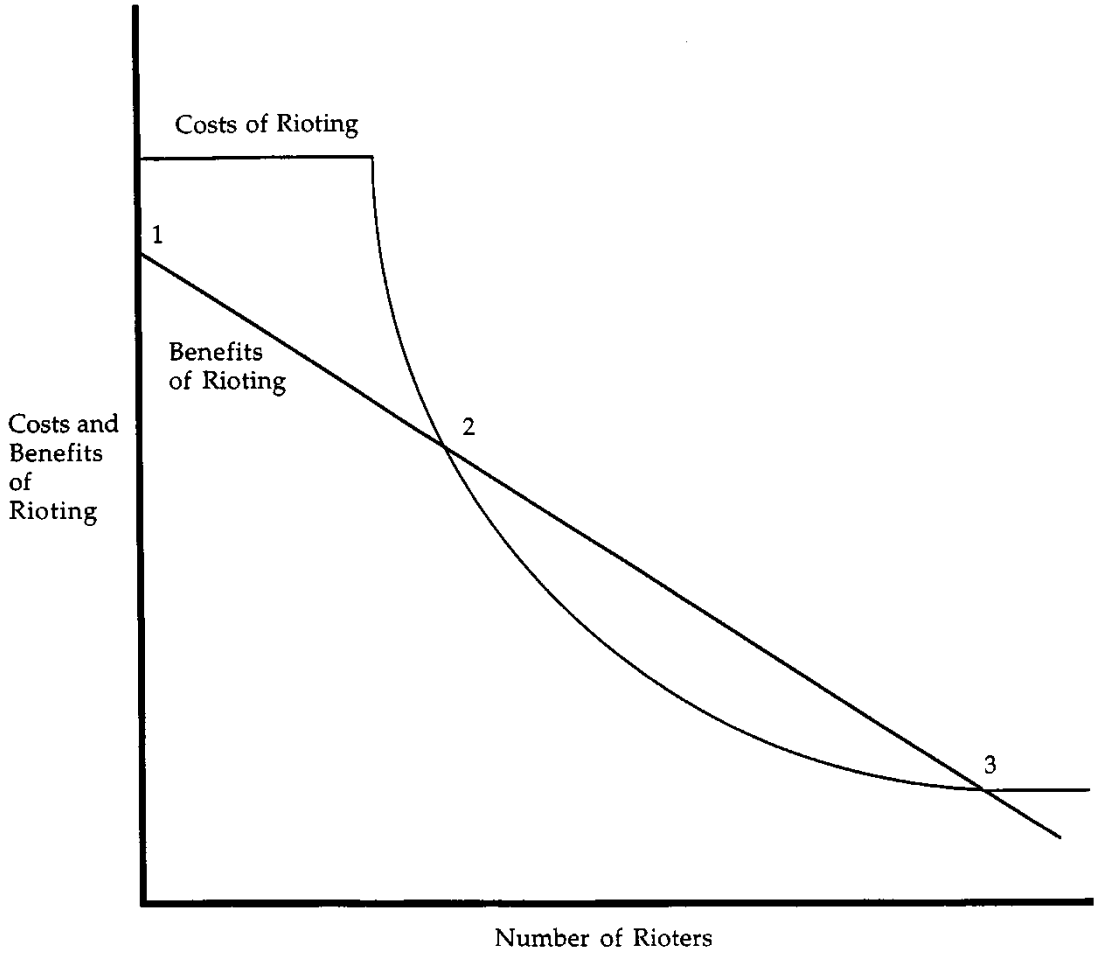


FIG. 1. The costs and benefits of rioting. The benefits of rioting curve is downward sloping because as the number of rioters increases, the marginal rioter receives lower benefits from rioting. The costs curve is downward sloping because more rioters decrease the probability that the marginal rioter will be arrested. Point 1 is the no riot equilibrium, point 2 is the unstable mid-level riot equilibrium, and point 3 is the high riot equilibrium.

Riot Tipping Point

- The lowest value

$$R^* = \min R > 0 : p(R) C = B - G^{-1}(R) - \bar{\theta}$$

is an unstable equilibrium

$$p'(R^*) C < -\frac{1}{g\left(B - p(R^*) C - \bar{\theta}\right)}$$

- If the mass of rioters reaches R^*, self-sustaining riot growth ensues

1. Direct organization by political actors
2. A large gathering causing congestion in policing
3. A focal event creating self-fulfilling expectations

Comparative Statics for Riots

- The second positive equilibrium R^{**} is stable

$$p'\left(R^{**}\right)C > -\frac{1}{g\left(B - p\left(R^{**}\right)C - \bar{\theta}\right)}$$

- Parameter changes that shift R^{*} down shift R^{**} up
  - They shift up every stable equilibrium

$\Rightarrow$ Riot probability and intensity have the same determinants

1. Weaker punishment
   - Lower probability $p\left(R\right)$ for all R
   - Lower penalties C
2. Lower opportunity cost $\bar{\theta}$

- You could add information that allows people to join the riot

Incidence of Riots Across Countries

	Regression 1	Regression 2
Dependent variable	Log of riots	Log of riots
Ethnic heterogeneity (log)	0.057*	−0.075
	(0.030)	(0.053)
Real per capita GDP 1970	−0.049**	−0.052**
	(0.025)	(0.024)
Urbanized population/Total population 1970	0.005**	0.010**
	(0.002)	(0.003)
Population 1970 (log)	0.265**	0.273**
	(0.026)	(0.026)
Dictatorship dummy 1970	−0.249**	−0.258**
	(0.095)	(0.092)
Latin American country dummy	0.266**	0.231**
	(0.085)	(0.083)
Ethnicity (log)* urbanization		0.003**
		(0.001)
Constant	−0.110	−0.265*
	(0.127)	(0.135)
Pseudo *R*-squared	0.566	0.618
Observations	102	102

Note. Estimated as tobit regressions. Standard errors appear in parentheses.

Occurrence of Riots in the U.S. in the 1960s

Dependent variable: Occurrence	Regression 1	Regression 2
Southern city dummy	−0.777**	−0.899**
	(0.349)	(0.375)
Segregation index 1960	0.015	0.022
	(0.018)	(0.018)
Log of total population 1960	0.029	0.025
	(0.222)	(0.227)
Log of non-white population 1960	0.569**	0.554**
	(0.190)	(0.203)
Age of non-white community	−0.518	−0.413
	(0.543)	0.550
Non-white unemployment rate 1960	5.308*	5.601*
	(3.214)	(3.238)
Relative poverty rate 1960 (non-white/total)	0.095	0.178
	(0.263)	(0.266)
Relative homeownership rate (non-white/total)	−1.346**	−1.212*
	(0.637)	(0.643)
Police expenditures per capita 1960		−0.013
		(0.033)
Non-police government expenditures per capita 1960		0.005**
		(0.003)
Constant	−6.514**	−7.637**
	(2.262)	(2.422)
Pseudo R-squared	0.219	0.237
Observations	192	192

Note. Estimated as probit regressions. Standard errors appear in parentheses.

Intensity of Riots in the U.S. in the 1960s

	Regression 1	Regression 2	Regression 3
Dependent variables	Log of arrests	Log of arsons	Log of injuries
Southern city dummy	−0.275	−0.598	−0.504
	(0.579)	(0.749)	(0.707)
Segregation index 1960	−0.058**	−0.028	−0.040
	(0.024)	(0.031)	(0.029)
Log of population 1960	−0.488	−0.368	0.473
	(0.332)	(0.429)	(0.407)
Log of non-white population 1960	1.434**	1.185**	0.748**
	(0.302)	(0.392)	(0.369)
Age of the non-white community	−1.010	−0.587	−0.971
	(0.696)	(0.902)	(0.852)
Non-white unemployment rate 1960	3.761	9.840*	−2.140
	(3.963)	(5.128)	(4.861)
Relative poverty rate 1960 (non-white/total)	0.821*	−0.562	0.463
	(0.432)	(0.559)	(0.528)
Relative homeownership rate (non-white/total)	−0.579	−0.359	−0.135
	(0.892)	(1.159)	(1.098)
Police expenditures per capita 1960	−0.059**	−0.100**	−0.038
	(0.026)	(0.035)	(0.032)
Non-police government expenditures per capita 1960	0.004	0.002	0.004
	(0.003)	(0.004)	(0.004)
Constant	0.425	0.096	−7.262**
	(2.625)	(3.414)	(3.239)
Pseudo R-squared	0.194	0.118	0.164
Observations	83	83	83

Note. Estimated as tobit regressions. Standard errors appear in parentheses.

Salient Facts About Riots

1. Neoclassical incentives matter
2. Repression works
   - Fewer riots are started in dictatorships
   - Riots are ended in democracies by police and military force
3. Poverty is not a major determinant of riots
4. Ethnic diversity combined with density is the key driver

- DiPasquale and Glaeser (1998) have no identification strategy