Neighborhood Effects

Econometric Methods for Social Spillovers and Networks

University of St. Gallen, October 3rd to October 10th, 2022

Bryan S. Graham

University of California - Berkeley

Neighborhood effects

- What role do neighborhoods play in shaping long run life outcomes? In ameliorating or perpetuating inequality across generations?
- Large literatures in both economics and sociology.
- Especially salient in the United States where residential segregation by race and racial inequality are pronounced (e.g., Loury, 2002).
- See Newburger et al. (2011), Duncan and Murname (2011), Sharkey and Faber (2014) and Graham (2018) for recent surveys of extant evidence.
- A central issues is that – at least in part – families choose where they live $\Rightarrow$ unobserved family attributes may covary with observed neighborhood attributes (Mayer and Jencks, 1989).
- In part because of difficult identification and specification issues, considerable debate, especially within economics, exists regarding the importance of neighborhood effects (e.g., Angrist, 2014).

Setup & notation

- Population: individuals residing in a given city (e.g., Zurich), composed of $i = 1, \ldots, N$ neighborhoods.
- Let $T \in \{0, 1\}$ denote whether a random draw from this city is minority/disadvantaged $(T = 1)$ or not $(T = 0)$.
- Let Z be an $N \times 1$ vector with a 1 in the i^{th} row and zeros elsewhere when the random draw resides in the i^{th} neighborhood $(i \in \{1, \ldots, N\})$.
- Here N denotes the number of neighborhoods in the city.
- Let $s(z) = \Pr(T = 1 | Z = z)$ be the proportion of neighborhood $Z = z$ that is minority.
- Let $Q = \Pr(T = 1)$ be the city-wide frequency of minority residents.

η^2 segregation index

- (Evidently) introduced by Farley (1977).
- Popularized in economics by Kremer and Maskin (1996).
- First we need to introduce the *isolation* and *exposure* indices.
- Isolation index (I):
$$\mathrm{I} = \mathbb{E}\left[s\left(Z\right) \middle| T = 1 \right].$$
This index equals the fraction of one's neighbors who are minority for the average minority resident in a given city.
- Exposure index (E):
$$\mathrm{E} = \mathbb{E}\left[s\left(Z\right) \middle| T = 0 \right].$$
Exposure coincides with the fraction of neighbors who are minority for the typical non-minority resident.

η^2 segregation index (continued)

- The η^2 index coincides with the coefficient on T in the linear regression of $s(Z)$ onto a constant and T.
- This endows η^2 with a simple predictive interpretation:

$$\begin{aligned}\mathbb{E}^*\left[s(Z)|\,T\right] &= \left[\mathbb{E}\left[s(Z)\right] - \mathbb{E}\left[T\right]\eta^2\right] + \eta^2 T \\ &= \left(1-\eta^2\right)Q + \eta^2 T.\end{aligned}$$

- Our prediction of an individual's neighborhood composition, $s(Z)$, is a weighted average of own race and the city-wide fraction minority.
- In a segregated city own race is very predictive of neighbors' race.

η^2 segregation index (continued)

- First algebraic formulation:

$$
\begin{aligned}
\eta^2 &= \frac{\mathbb{C}\left(T, s\left(Z\right)\right)}{\mathbb{V}\left(T\right)} \\
&= \frac{\mathbb{C}\left(s\left(Z\right) + T - s\left(Z\right), s\left(Z\right)\right)}{\mathbb{V}\left(s\left(Z\right)\right) + \mathbb{E}\left[\mathbb{V}\left(T \mid Z\right)\right]} \\
&= \frac{\mathbb{C}\left(s\left(Z\right), s\left(Z\right)\right)}{\mathbb{V}\left(s\left(Z\right)\right) + \mathbb{E}\left[\mathbb{V}\left(T \mid Z\right)\right]} + \frac{\mathbb{C}\left(T - s\left(Z\right), s\left(Z\right)\right)}{\mathbb{V}\left(s\left(Z\right)\right) + \mathbb{E}\left[\mathbb{V}\left(T \mid Z\right)\right]} \\
&= \frac{\mathbb{V}\left(s\left(Z\right)\right)}{\mathbb{V}\left(s\left(Z\right)\right) + \mathbb{E}\left[\mathbb{V}\left(T \mid Z\right)\right]}
\end{aligned}
$$

- η^2 equals the proportion of variance in T that occurs *between* neighborhoods.

- This formulation does *not* require T to be binary.

η^2 Segregation index (continued)

- Second algebraic formulation:

$$\begin{aligned}
\eta^2 &= \frac{\mathbb{C}\left(T, s\left(Z\right)\right)}{\mathbb{V}\left(T\right)} \\
&= \frac{\mathbb{E}\left[Ts\left(Z\right)\right] - \mathbb{E}\left[T\right]\mathbb{E}\left[s\left(Z\right)\right]}{Q\left(1-Q\right)} \\
&= \frac{Q\mathbb{E}\left[s\left(Z\right)\middle| T=1\right] - Q^2}{Q\left(1-Q\right)} \\
&= \frac{\mathrm{I} - Q}{1-Q}
\end{aligned}$$

- The η^2 index therefore provides a scaled measure of minority isolation.
- It measures the excess isolation of minorities in a city compared to perfect integration.
- This formulation *does* require T to be binary.

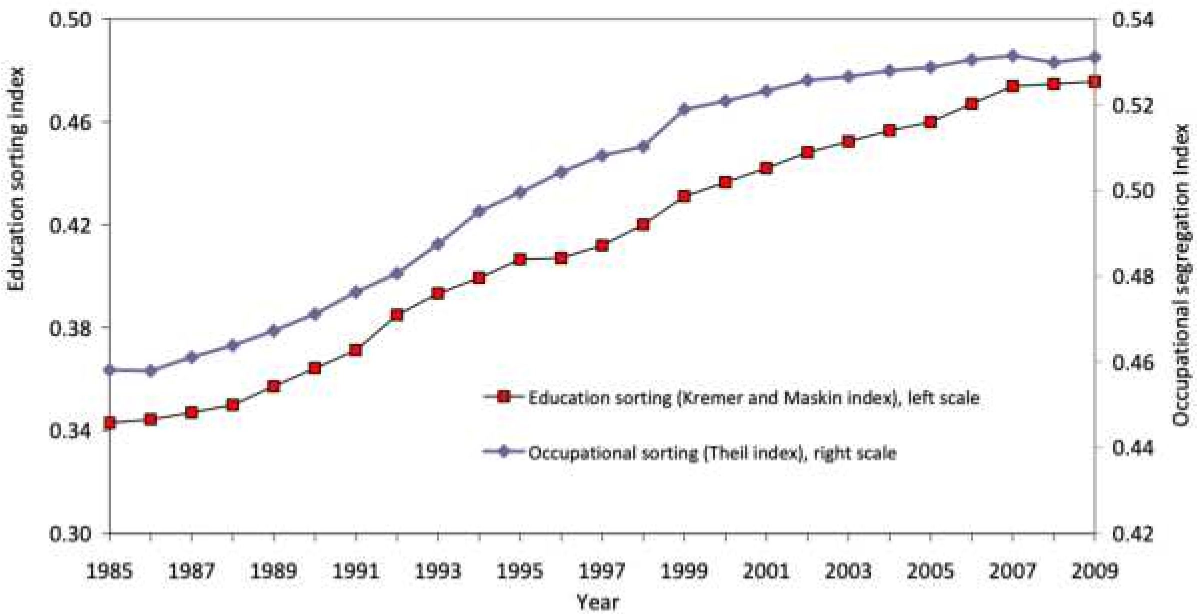
Most segregated metro areas in the US, 2000

	(1) DI	(2) η^2	(3) Frac. URM	(4) I
Detroit, MI	0.790	0.673	0.263	0.759
Newark, NJ	0.724	0.593	0.365	0.741
Milwaukee-Waukesha, WI	0.721	0.537	0.225	0.641
Cleveland-Lorain-Elyria, OH	0.719	0.582	0.224	0.676
Buffalo-Niagara Falls, NY	0.714	0.522	0.151	0.594
New York, NY	0.712	0.591	0.513	0.801
Cincinnati, OH-KY-IN	0.698	0.477	0.145	0.552
St. Louis, MO-IL	0.687	0.531	0.203	0.626
Gary, IN	0.686	0.563	0.306	0.697
All MSAs (N = 319)	0.533	0.337	0.372	0.601
Large MSAs (N = 99)	0.560	0.376	0.381	0.622

Sorting & inequality: other examples

1. Kremer and Maskin (1996)– Workplace segregation and inequality
2. Jargowsky (1996) – Segregation by income
3. Card et al. (2013) – Evidence using West German data (educational sorting across firms)
4. Comment:
   (a) some types of sorting seem to be declining (e.g., racial segregation),
   (b) others increasing (e.g., socio-economic segregation, sorting in labor markets, martial sorting).

Figure 5: Sorting Across Establishments of Workers in Different Education and Occupation Groups



Notes: figure shows two measures of sorting of full time male workers across establishments. See text for definitions of indices.

Neighborhood effects: regression analysis

- We assume that individuals are heterogeneous according to minority status, T, and other background attributes, A (*unobserved*).
- Joint distribution of $(A', T)'$ is given and *invariant* across policies.
- In practice this means that the elements of $(A', T)'$ are non-manipulable (over the time frame in which the outcome is being measured).
  - We conceptualize $(A', T')'$ as a bundle of fixed characteristics that a household brings with them as they move from neighborhood to neighborhood.
  - Behaviors, for example parenting style, which may change with neighborhood of residence, are *not* elements of A.

Neighborhood effects: regression analysis (continued)

- Neighborhoods vary in their composition. Specifically the distribution $(A', T)'$ may vary across neighborhoods due to sorting.
- Neighborhoods may be heterogeneous in other ways (e.g., close or far to employment centers; near a superfund clean-up site).
- Let U be additional *unobserved* neighborhood-level characteristics (that are not a function of residential make-up of the neighborhood).

Neighborhood effects: regression analysis (continued)

- An ideal *long regression* function takes the form

$$\mathbb{E}\left[Y|\,T,s\left(Z\right),m_A\left(Z\right),U,A\right]=\alpha_0+\beta_0T+\gamma_0s\left(Z\right)+m_A\left(Z\right)'\delta_0+U'\kappa_0+A'\lambda_0.$$

- We will endow this regression with structural/causal significance.

- Direct estimation of the parameters of this regression function is not feasible because individual background, A, neighbors' background, $m_A\left(Z\right)$, and neighborhood attributes, U, are unobserved.

- It is important that all three of these latent variables enter the long regression; they each have different implications for developing intuitions about identification.

- Instead we estimate the *short regression* (e.g., Brooks-Gunn et al., 1993)

$$\mathbb{E}^*\left[Y|\,T,s\left(Z\right)\right]=a_0+b_0T+c_0s\left(Z\right). \tag{1}$$

- When does $c_0=\gamma_0$ or, rather, how do these two coefficient differ from one another in general?

Sorting into neighborhoods

- Consider the following mean regression representation of A

$$A = \pi_0 + \phi_0 T + B, \ \mathbb{E}[B|T] = 0. \tag{2}$$

- Here B is the component of "background" that does not vary, on average, with race.
- It is convenient to partition unobserved agent-level heterogeneity into a component that varies with race and one which is (mean) independent of it. This is nothing more than a decomposition.
- Note that the "background gap" between minorities and non-minorities is given by $\phi_0 = \mathbb{E}[A|T=1] - \mathbb{E}[A|T=0]$.

Sorting (continued)

- There is *no sorting on unobservables* if an individual's neighborhood of residence, Z, is not predictive of B conditional on her observed type, T:

$$\mathbb{E}\left[B \middle| T, Z\right] = \mathbb{E}\left[B \middle| T\right] = 0. \tag{3}$$

- Condition (3) implies that the distribution of "background" (specifically its mean) among, say, minorities is similar across neighborhoods.
- If this is not the case, then we say there is sorting on unobservables.
- For example, it may be that minority families living in predominately non-minority neighborhoods differ systematically, in terms of A, from their counterparts in predominately minority neighborhoods (e.g., their adult members may have graduated from more elite colleges).
- This captures the intuition that observed characteristics of a neighborhood may be correlated with the unobserved attributes of its residents (e.g., Sampson et al., 2002).

Matching

- Matching is distinct from sorting.
- There is matching when the unobserved *exogenous* attributes of one's neighborhood (e.g., it proximity to the city-center, environmental conditions) can be predicted by own type.
- There is *no matching on U* if

$$\mathbb{E}\left[U \mid T\right] = \mathbb{E}\left[U\right]. \tag{4}$$

- If, for example, minorities are less likely to live in neighborhoods adjacent to employment districts, then we say there is matching on U.
- Matching is about whether the demographic structure of a neighborhood is predictive of other "fixed" neighborhood attributes (e.g., geography, environment etc.).

Neighborhood effects: regression analysis (continued)

- After tedious manipulation it is possible to show that the two slope coefficients in (1) equal:

$$\begin{aligned} b_0 &= \beta_0 + \phi_0' \lambda_0 \\ &\quad - \frac{1}{Q(1-Q)} \mathbb{C}\left(s(Z), \mathbb{E}[B|Z]\right)' \frac{\lambda_0}{1-\eta^2}. \end{aligned} \tag{5}$$

$$\begin{aligned} c_0 &= \gamma_0 + \phi_0' \delta_0 + \frac{1}{\eta^2} \left\{ \frac{1}{Q(1-Q)} \mathbb{C}\left(s(Z), \mathbb{E}[B|Z]\right)' \left(\delta_0 + \frac{\lambda_0}{1-\eta^2} \right) \right. \\ &\quad \left. + \left(\mathbb{E}[U|T=1] - \mathbb{E}[U|T=0] \right)' \kappa_0 \right\}. \end{aligned} \tag{6}$$

- Here $Q = \mathbb{E}[T]$ is the population fraction minority, η^2 the eta-squared index of segregation defined earlier and $\mathbb{C}(X, Y)$ denotes the covariance of X with Y.

Neighborhood effects: regression analysis (continued)

- Coefficient on $s(Z)$ is biased by *both* sorting and matching.
- Consider a city where low B households, irrespectively of race, sort into predominately minority neighborhoods and high B households into predominately non-minority neighborhoods.
  - Under this type of sorting pattern $\mathbb{C}(s(Z), \mathbb{E}[B|Z]) < 0$
    - (note: $\mathbb{C}(s(Z), \mathbb{E}[B|Z]) = Q(1-Q)[\mathbb{E}[\mathbb{E}[B|Z]|T=1] - \mathbb{E}[\mathbb{E}[B|Z]|T=0]]$; small typo on p. 474 of Graham (2018)).
  - This biases c_0 downward relative to $\gamma_0 + \phi_0'\delta_0$, making exposure to minority neighbors appear more detrimental for Y than it would be in the absence of sorting.
- Similarly if minority households move into "worse" neighborhoods (e.g., far from employment centers, or with bad air quality), then c_0 may be further biased downwards because $\mathbb{C}(s(Z), U) < 0$.
- These biases imply that we may overstate the magnitude of any neighborhood effects.

Neighborhood effects: regression analysis (continued)

- Under the no sorting we have $\mathbb{C}\left(s\left(Z\right),\mathbb{E}\left[\left.B\right|Z\right]\right)=0$ since

$$\begin{aligned}\left(s\left(Z\right)-Q\right)\mathbb{E}\left[\left.B\right|Z\right] &= \left(s\left(Z\right)-Q\right)\mathbb{E}\left[\left.\mathbb{E}\left[\left.B\right|T,Z\right]\right|Z\right] \\ &= \left(s\left(Z\right)-Q\right)\mathbb{E}\left[\left.\mathbb{E}\left[\left.B\right|T\right]\right|Z\right] \\ &= \left(s\left(Z\right)-Q\right)\mathbb{E}\left[\left.0\right|Z\right]=0\end{aligned}$$

- Under no matching we have that $\mathbb{E}\left[\left.U\right|T=1\right]-\mathbb{E}\left[\left.U\right|T=0\right]=0$.

- With no sorting and no matching the short regression coefficients on T and $s\left(Z\right)$ equal

$$\begin{aligned}b_0 &= \beta_0+\phi_0'\lambda_0 \\ c_0 &= \gamma_0+\phi_0'\delta_0.\end{aligned}$$

- Although $b_0\neq\beta_0$and $c_0\neq\gamma_0$, the above expressions are still useful for policy analysis; particularly for understanding the effects of "associational redistributions" (cf., Durlauf, 1996).

Doubly Randomized Assignments

- Social planner specifies a *feasible* distribution of $s(Z)$ (i.e., chooses the fraction minority for each neighborhood).
- She then assigns the appropriate number of minority and non-minority households to each neighborhood by taking independent random draws from these two household sub-populations (ensures no sorting on unobservables condition).
  - If a neighborhood consists of 100 minority and 100 non-minority families, then each of these families correspond to random draws from the underlying minority and non-minority populations.
- She also assigns neighborhood compositions to locations at random (ensures no matching on unobservables condition).
- In this world, if we move a minority household from a neighborhood with $s(Z) = s$ to one with $s(Z) = s'$, the expected change in the outcome will be: $c_0(s' - s)$.

Double Randomization (continued)

- It doesn't really matter if the underlying spillover operates via fraction minority ($\gamma_0 \neq 0$) or via background characteristics ($\delta_0 \neq 0$), the distribution of which varies across minority and non-minority populations ($\phi_0 \neq 0$).
- In the absence of sorting and matching the canonical neighborhood effects regression can be used to understand how the distribution of outcomes would be altered by changes in neighborhood structure.
- Unfortunately the no matching and sorting conditions are rather strong. Although they can be weakened by conditioning on covariates (cf., Graham, 2018), they are difficult to satisfy in many (most?) empirical settings of interest.
- A setting where *they are* satisfied is provided by the Project STAR experiment (e.g., Graham, 2008).

Moving to Opportunity (MTO) experiment

- Housing mobility experiment conducted by HUD in the five cities in the United States.
- See Chetty et al. (2016) for a recent analysis of MTO data and additional references.
- Residents in public housing were randomly assigned rental vouchers which allowed them to move to "better" neighborhoods.
- Long term follow-up of households; linkage of survey data with administrative records, etc.

Moving to Opportunity Experiment (continued)

- Let $W \in \{0, 1\}$ now be an indicator for whether a household has access to a rent subsidy.
- To keep the exposition more policy-relevant, and in line with the MTO example, let T now indicate whether a household falls below (some multiple of) the poverty line or not.
- Assume that that W is independent of B.
- This allows for random assignment of vouchers to households...
- ...but with differing assignment probabilities across poor and non-poor households
  - (including an assignment probability of zero for non-poor households).

Moving to Opportunity Experiment (continued)

- Let $\Xi_S = (\mathbb{E}[s(Z)|W=1] - \mathbb{E}[s(Z)|W=0])$ be the average difference in neighborhood poverty across voucher and non-voucher households.

- Similarly define

$$\begin{aligned} \Xi_{\overline{B}} &= (\mathbb{E}[\mathbb{E}[B|Z]|W=1] - \mathbb{E}[\mathbb{E}[B|Z]|W=0]) \\ \Xi_U &= \mathbb{E}[U|W=1] - \mathbb{E}[U|W=0] \end{aligned}$$

  as the average difference in (i) neighbors' background and (ii) neighborhood amenities respectively across voucher and non-voucher households.

- All these contrasts are post-assignment, and hence post-move (if applicable).

- Let $p_W = \mathbb{E}[W]$ equal the marginal frequency of voucher receipt.

Moving to Opportunity Experiment (continued)

- Consider the linear instrumental variables fit of Y onto a constant, T and $s(Z)$, using W as an excluded instrument for $s(Z)$.
- Since W is randomly assigned conditional on T, this would seem to be a promising and valid application of the method of instrumental variables.
- The coefficient on $s(Z)$ in this fit has a probability limit of (Graham, 2018)

$$
\begin{aligned}
c_{IV} = (\gamma_0 + \phi_0' \delta_0) & \\
& + \frac{p_W(1-p_W)\,\Xi_{\overline{B}} - (\mathbb{E}[\mathbb{E}[B|Z]|T=1] - \mathbb{E}[\mathbb{E}[B|Z]|T=0])\,Q\,(\mathbb{E}[W|T=1]-p_W)'}{p_W(1-p_W)\,\Xi_S - \eta^2 Q\,(\mathbb{E}[W|T=1]-p_W)}\delta_0 \\
& \qquad + \frac{p_W(1-p_W)\,\Xi_U - (\mathbb{E}[U|T=1] - \mathbb{E}[U|T=0])\,Q(\mathbb{E}[W|T=1]-p_W)'}{p_W(1-p_W)\,\Xi_S - \eta^2 Q\,(\mathbb{E}[W|T=1]-p_W)}\kappa_0
\end{aligned}
$$

which is *not* free of sorting and matching bias.

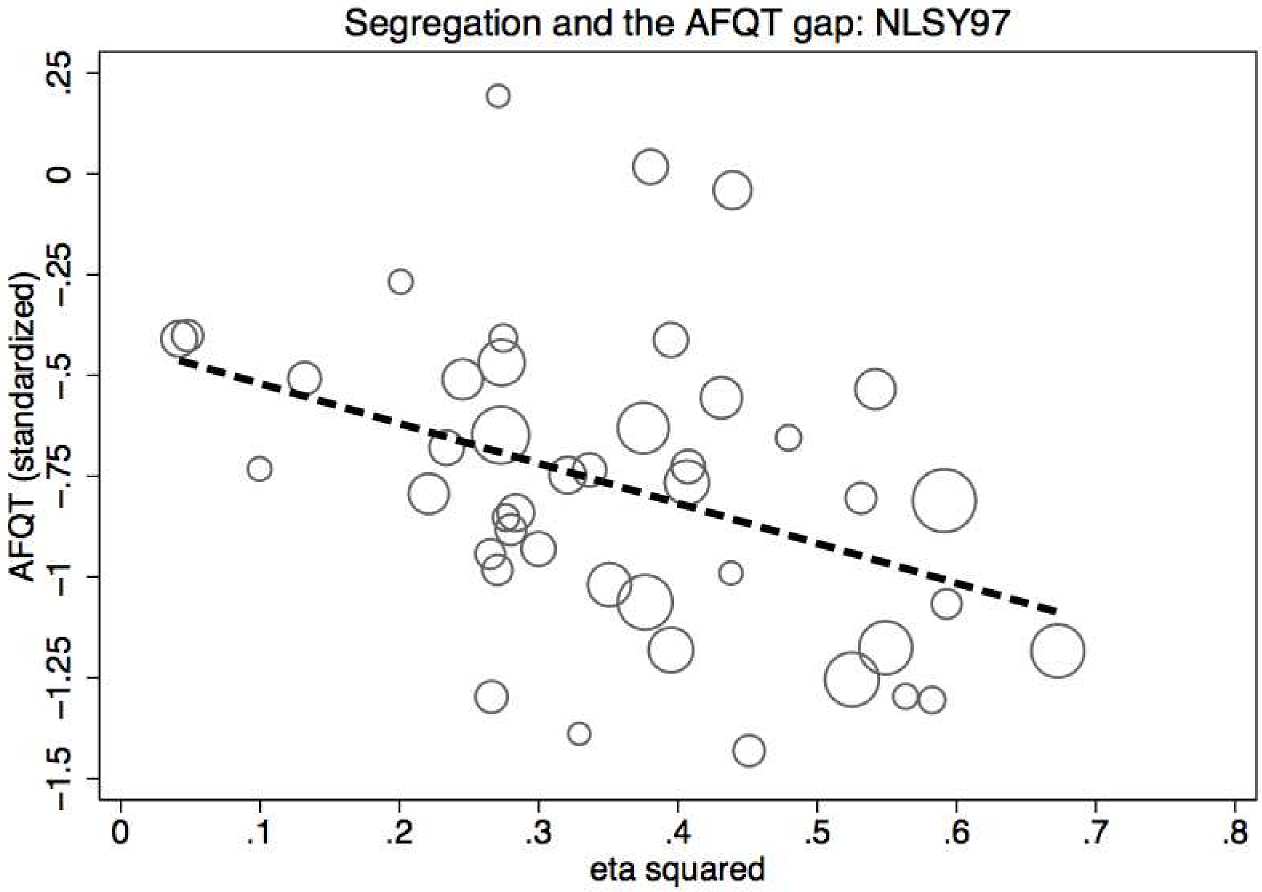
Moving to Opportunity Experiment (continued)

- Although W is conditionally randomly assigned given T *and* the underlying outcome model is linear, the instrumental variables fit in this setting does not recover an interpretable causal parameter.
- The reason is that T is not exogenous.
- Under sorting and matching a household's type will covary with unobserved neighbors' attributes $m_A(Z)$ and neighborhood amenities U.
- In the Moving To Opportunity (MTO) demonstration experiment, vouchers were only available with positive probability to $T = 1$ households. In case we have

$$c_{MTO} = (\gamma_0 + \phi_0'\delta_0) + \frac{\Xi_{\overline{B}}'\delta_0}{\Xi_S} + + \frac{\Xi_U'\kappa_0}{\Xi_S}. \tag{7}$$

- Equation (7) *is* a measure of the "effect of place" *free* of sorting and matching bias.

Cross-city research designs



- FACT: Minority/non-minority outcome gaps are larger in more segregated metro areas.

Cross-city research designs (continued)

- Several researchers have proposed estimation methods involving aggregation to the city-level as a remedy for sorting bias
  - Cutler and Glaeser (1997)
  - Card and Rothstein (1997)
- This approach requires observations from a cross-section of cities, typically operationalized by a metropolitan statistical area (MSA).
- Also required are restrictions on how the joint distribution of $(A', T)'$ various across cities.

Cross-city research designs (continued)

- The researcher has access to a nationally representative sample of (T, Y) pairs.
- Assume further that this sample is geocoded, such that each observation can be assigned to a specific MSA.
- As a concrete example, the NLSY79 and NLSY97 restricted-use geocode files may be used to assign each respondent to an MSA of residence (at the time of adolescence).
- Index MSAs by $c = 1, \ldots, N$ and sampled respondents within a city by $i = 1, \ldots, M_c$.
- Measures of residential segregation by race, corresponding to the period coinciding with the respondent's adolescence, are available for each city (e.g., from the Neighborhood Change Database (NCDB)).

Cross-city research designs (continued)

- We begin by modifying our long regression to incorporate a city-specific intercept

$$\begin{aligned}\mathbb{E}\left[Y_{ci}\middle| T_{ci}, s\left(Z_{ci}\right), m_A\left(Z_{ci}\right), U_{ci}, A_{ci}\right] &= \alpha_c + \beta_0 T_{ci} + \gamma_0 s\left(Z_{ci}\right) \\ &\quad + m_A\left(Z_{ci}\right)' \delta_0 + U'_{ci}\kappa_0 + A'_{ci}\lambda_0.\end{aligned} \tag{8}$$

- The presence of α_c allows the mean outcome to vary across cities for reasons unrelated to segregation.
- Let $\mathbb{E}^*\left[Y|\, X; c\right]$ denote the best linear predictor of Y given X conditional on residence in city c (cf., Wooldridge, 1999).
- Let $\mathbb{V}\left(Y|\, c\right)$ and $C\left(X, Y|\, c\right)$ denote city-specific variances and covariances.

Cross-city research designs (continued)

- Recall that

$$\mathbb{E}^*\left[s\left(Z_{ci}\right) \middle| T_{ci}; c\right] = \left(1-\eta_c^2\right) p_c + \eta_c^2 T_{ci} \tag{9}$$

where η_c^2 is the eta squared segregation measure for city c and $p_c = \Pr\left(T=1 \middle| c\right)$ is the city-wide fraction Minority.

- In highly segregated cities ($\eta_c^2 \to 1$) own race is very predictive of neighbors' race.

- In integrated cities ($\eta_c^2 \to 0$) the city-wide fraction Minority, p_c, has more predictive value.

Cross-city research designs (continued)

- Define

$$\begin{aligned}
\phi_c &= \mathbb{E}\left[A_{ci}\middle| T_{ci}=1,c\right]-\mathbb{E}\left[A_{ci}\middle| T_{ci}=0,c\right]\\
\upsilon_c &= \mathbb{E}\left[\mathbb{E}\left[B_{ci}\middle| Z_{ci}\right]\middle| T_{ci}=1,c\right]-\mathbb{E}\left[\mathbb{E}\left[B_{ci}\middle| Z_{ci}\right]\middle| T_{ci}=0,c\right]\\
\tau_c &= \mathbb{E}\left[U_{ci}\middle| T_{ci}=1,c\right]-\mathbb{E}\left[U_{ci}\middle| T_{ci}=0,c\right]
\end{aligned}$$

  to be the minority/non-minority gaps within city c in

  - "background" (ϕ_c),
  - neighbors' background (υ_c),
  - neighborhood amenities (τ_c) respectively.

- The first of these terms is a vector of average differences between minorities and non-minorities.

Cross-city research designs (continued)

- The latter two terms are vectors of average differences in features of their *neighborhoods*.
- Specifically in the unobserved attributes of their neighbors...
- ...and the unobserved non-composition-based characteristics of their neighborhoods.
- Both of these are measures of average differences in neighborhood quality between minorities and non-minorities.
- To the extent that these components of neighborhood quality directly influence the outcome of interest (i.e., $\delta_0 \neq 0$ and/or $\kappa_0 \neq 0$), they both are drivers of neighborhood effects, broadly defined.

Cross-city research designs (continued)

- Using (8), (9) and the notation defined above we get a city-specific linear regression of Y_{ci} onto a constant and T_{ci}, deviated from city-specific means, of

$$\begin{aligned} \mathbb{E}\left[\left.Y_{ci}\right| T_{ci}; c\right] - \mathbb{E}\left[\left.Y_{ci}\right| c\right] &= \left\{\beta_0 + \left(\gamma_0 + \phi_c'\delta_0\right)\eta_c^2 \right. \\ &\quad \left. + \upsilon_c'\delta_0 + \tau_c'\kappa_0 + \phi_c'\lambda_0\right\} \\ &\quad \times \left(T_{ci} - p_c\right). \end{aligned} \tag{10}$$

- Let ϕ_0 now equal the mean of ϕ_c across cities (i.e., $\phi_0 = \mathbb{E}\left[\phi_c\right]$) with ν_0 and τ_0 analogously defined.

Cross-city research designs (continued)

- Equation (10) indicates that the Minority-White *outcome* gap in city c – $\mathrm{GAP}_c = \mathbb{E}\left[Y_{ci}|\, T_{ci}=1;c\right] - \mathbb{E}\left[Y_{ci}|\, T_{ci}=0;c\right]$ – varies with its degree of segregation, as measured by the eta-squared, η_c^2, index:

$$\mathrm{GAP}_c = a_0 + \left(\gamma_0 + \phi_0'\delta_0\right)\eta_c^2 + V_c \tag{11}$$

with

$$\begin{aligned} a_0 &= \beta_0 + \upsilon_0'\delta_0 + \tau_0'\kappa_0 + \phi_0'\lambda_0 \\ V_c &= \left(\upsilon_c - \upsilon_0\right)'\delta_0 + \left(\tau_c - \tau_0\right)'\kappa_0 + \left(\phi_c - \phi_0\right)'\lambda_0 + \left(\phi_c - \phi_0\right)'\delta_0\eta_c^2. \end{aligned}$$

Cross-city research designs (continued)

- V_c varies with a city's minority/non-minority gap in
  1. neighbors' unobserved "background",
  2. exogenous neighborhood attributes,
  3. own "background", and
  4. the interaction of own-background with measured segregation.

- Under the orthogonality condition

$$\mathbb{E}\left[V_c \eta_c^2\right] = 0 \tag{12}$$

  then, by equations (10) and (11), a least squares fit of the city-specific measure of the minority/non-minority outcome gap, GAP_c, onto a constant and η_c^2 will provide a consistent estimate of $\gamma_0 + \phi_0' \delta_0$.

- Observe that this coincides with the coefficient on fraction minority in the prototypical neighborhood effects regression under no sorting and matching.

Cross-city research designs (continued)

- Restriction (12) is often referred to as a no "differential sorting" across cities assumption (e.g., Card and Rothstein, 1997).

- To understand this language consider the common case where δ_0 and κ_0 are *a priori* presumed equal to zero (virtually all empirical work makes this assumption implicitly!). such that (12) requires that, across cities, variation in the "background gap" is uncorrelated with variation in racial segregation (i.e, $\mathbb{C}\left(\phi_c, \eta_c^2\right) = 0$).

  - Sorting across cities *is* allowed in the sense that differences in average "background" across cities do not threaten identification.
  - However the background-gap across cities, if it varies, must do so independently of measured segregation.

Cross-city research designs (continued)

- Note that, again *in the special case* where $\delta_0 = \kappa_0 = 0$, aggregation *does* eliminate biases due to sorting on unobservables. In that case the coefficient on $s(Z)$ in the *within-city* neighborhood effects regression is (see (6) above)

$$c_0 = \gamma_0 + \frac{1}{\eta^2(1-\eta^2)}\left(\mathbb{E}\left[\mathbb{E}\left[B|Z\right]|T=1\right] - \mathbb{E}\left[\mathbb{E}\left[B|Z\right]|T=0\right]\right)'\lambda_0$$

with the second term due to sorting and where I have also made use of the equality

$$\frac{\mathbb{C}\left(s(Z), \mathbb{E}\left[B|Z\right]\right)}{Q(1-Q)} = \mathbb{E}\left[\mathbb{E}\left[B|Z\right]|T=1\right] - \mathbb{E}\left[\mathbb{E}\left[B|Z\right]|T=0\right].$$

- In contrast, the coefficient on η_c^2 in the *cross-city* regression of the minority/non-minority outcome gap onto a constant and η_c^2 is equal to γ_0 alone.

- Under these conditions aggregation does eliminate biases due to sorting, as is often asserted in empirical work.

Cross-city research designs (continued)

- When δ_0 and κ_0 possibly differ from zero, condition (12) requires maintaining additional (strong) assumptions.
- First, for $\delta_0 \neq 0$, we require zero covariance between measured racial segregation and $(\upsilon_c - \upsilon_0)' \delta_0$, the first element of V_c.
  1. This implies that the minority/non-minority gap in neighbors' background is uncorrelated with the degree of metropolitan-area segregation.
  2. Does not rule out within-city sorting on unobservables, but constrains such sorting to be similar across cities.
  3. When $\delta_0 \neq 0$ we also require strengthening the zero covariance condition $\mathbb{C}\left(\phi_c, \eta_c^2\right) = 0$. For example if ϕ_c is mean-independent of η_c^2, then $\left\{(\phi_c - \phi_0)' \delta_0 \eta_c^2\right\} \eta_c^2$, the last term in $V_c \eta_c^2$, will be mean zero (along with the third term).

Cross-city research designs (continued)

- Second, for $\kappa_0 \neq 0$, we require zero covariance between η_c^2 and $(\tau_c - \tau_0)' \kappa_0$, the second element of V_c.
  1. This implies that the degree to which race predicts exogenous neighborhood characteristics is uncorrelated with the level of racial segregation.
  2. In practice, it seems likely that own race will be a better predictor of unobserved neighborhood attributes in highly segregated cities, for no other reason than in such cities minorities/non-minorities live apart.
- These considerations suggest that recovering a consistent estimate of γ_0 from the cross-city correlation of GAP_c and η_c^2 is difficult.

Cross-city research designs (continued)

- One may be able to learn about the "effects of place" broadly defined from cross-city analyses.
- For example, if $\mathbb{E}\left[\phi_c \middle| \eta_c^2\right] = \phi_0$, then

$$\begin{aligned}\mathbb{E}^*\left[\mathrm{GAP}_c \middle| \eta_c^2\right] =& a_0 + (\gamma_0 + \phi_0'\delta_0)\,\eta_c^2 \\ &+ \mathbb{E}^*\left[\upsilon_c - \upsilon_0 \middle| \eta_c^2\right]'\delta_0 + \mathbb{E}^*\left[\tau_c - \tau_0 \middle| \eta_c^2\right]'\kappa_0\end{aligned}$$

  so that the coefficient on η_c^2 is only non-zero if "place matters".
- The intuition is simple: if inequality is greater in segregated cities this suggests "place matters".

Twilight Zone

- Some interesting empirical papers which don't neatly fall into the framework sketched above:

  1. Bayer et al. (2008)
  2. Wodtke et al. (2011)
  3. Altonji and Mansfied (2018)
  4. Chetty and Hendren (2018)

- None of these papers is perfect, but they each bring interesting ideas and/or data which merit further consideration/development.

References

Altonji, J. G. and Mansfied, R. K. (2018). Estimating group effects using averages of observables to control for sorting on unobservables: school and neighborhood effects. *American Economic Review*, 108(10):2902 – 2946.

Angrist, J. (2014). The perils of peer effects. *Labour Economics*, 30:98–108.

Bayer, P., Ross, S. L., and Topa, G. (2008). Place of work and place of residence: informal hiring networks and labor market outcomes. *Journal of Political Economy*, 116(6):1150 – 1196.

Brooks-Gunn, J., Duncan, G. J., Klebanov, P. K., and Sealand, N. (1993). Do neighborhoods influence child and adolescent development? *American Journal of Sociology*, 99(2):353 – 395.

Card, D., Heining, J., and Kline, P. (2013). Workplace heterogeneity and the rise of west german wage inequality. *Quarterly Journal of Economics*, 183(3):967 – 1015.

Card, D. and Rothstein, J. (1997). Racial segregation and the black–white test score gap. *Journal of Public Economics*, 91(11-12):2158 – 2184.

Chetty, R. and Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility i: childhood exposure effects. *Quarterly Journal of Economics*, 113(3):1107 – 1162.

Chetty, R., Hendren, N., and Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: new evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855 – 902.

Cutler, D. M. and Glaeser, E. L. (1997). Are ghettos good or bad? *Quarterly Journal of Economics*, 112(3):827 – 872.

Duncan, G. J. and Murname, R. J., editors (2011). *Whither Opportunity?: Rising Inequality, Schools, and Children's Life Chances*. Russell Sage Foundation.

Durlauf, S. N. (1996). Associational redistribution: A defense. *Politics & Society*, 24(4):391 – 410.

Farley, R. (1977). Residential segregation in urbanized areas of the united states in 1970: an analysis of social class and racial differences. *Demography*, 14(4):497 – 518.

Graham, B. (2008). Identifying social interactions through conditional variance restrictions. *Econometrica*, 76(3):643–660.

Graham, B. S. (2018). Identifying and estimating neighborhood effects. *Journal of Economic Literature*, 56(2):450 – 500.

Jargowsky, P. A. (1996). Take the money and run: economic segregation in u.s. metropolitan areas. *American Sociological Review*, 61(6):984 – 998.

Kremer, M. and Maskin, E. (1996). Wage inequality and segregation by skill. NBER Working Paper 5718, National Bureau of Economic Research.

Loury, G. C. (2002). *The Anatomy of Racial Inequality*. Harvard University Press, Cambridge, MA.

Mayer, S. E. and Jencks, C. (1989). Growing up in poor neighborhoods: how much does it matter? *Science*, 243(4897):1441 – 1445.

Newburger, H. B., Birch, E. L., and Wachter, S. M., editors (2011). *Neighborhood and Life Chances: How Place Matters in Modern America*. The City in the Twenty-First Century. University of Pennsylvania Press, Philadelphia.

Sampson, R. J., Morenoff, J. D., and Gannon-Rowley, T. (2002). Assessing "neighborhood effects": social processes and new directions in research. *Annual Review of Sociology*, 28:443 – 478.

Sharkey, P. and Faber, J. W. (2014). Where, when, why, and for whom do residential contexts matter? moving away from the dichotomous understanding of neighborhood effects. *Annual Review of Sociology*, 40:559 – 579.

Wodtke, G. T., Harding, D. J., and Elwert, F. (2011). Neighborhood effects in temporal perspective. *American Sociological Review*, 76(5):713 – 736.

Wooldridge, J. M. (1999). Distribution-free estimation of some nonlinear panel data models. *Journal of Econometrics*, 90(1):77 – 97.