PH784 Advanced Social Epidemiology: Methods

Week 4: Causal Inference (Continued) & Measurement in Social Epi

Koichiro Shiba

kshiba@bu.edu

Boston University School of Public Health

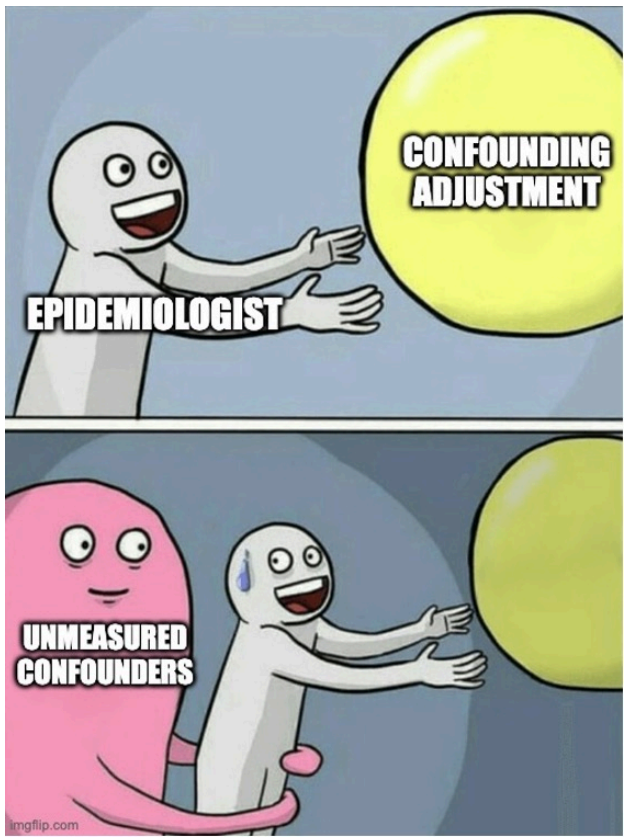
September 29, 2025

Today's Plan

1. **Alternative approaches for exchangeability**
2. Measurement
    - Key considerations
    - Specific measures in social epi
        - → SES
        - → Neighborhood environment

Conditional Exchangeability Assumption in Practice

An epidemiologist's delusion?



- Never achieved perfectly with observational data
  - → Can never be verified empirically
  - → The cliché: "We cannot preclude the possibility of bias due to unmeasured confounders"

What to do instead of giving up

1. **Critical (yet constructive) evaluation**
   - Bias analysis
   - Falsification
2. **Alternative assumptions**

Alternative Assumptions: Approach 1

- Standard approach
  - → Compare outcome values between groups with different exposure levels
  - → **Between-person comparison**
    - People have different backgrounds (i.e., confounding)
      - *Time-invariant*: genes, biological sex, (perhaps personality/education)
      - *Time-varying*: income, depressive symptoms
- **Within-person comparison**
  - → Focus on changes within the same individuals
    - Fixed-effect analysis
  - → Control for ALL time-invariant confounders
    - Including **unobserved** ones
      - Time-invariant: no change during the same time period
    - Still susceptible to confounding by time-varying confounders (e.g., income)





Fixed-Effect Analysis

- Use panel data with multiple waves
  - → So you can observe changes in exposure, outcome, covariates for the same individuals
    - Two common analytical approaches

1. "Change-on-change" model

- $E[\Delta Y_{it}|\Delta A_{it}, \Delta L_{it}] = \beta_0 + \beta_1 \Delta A_{it} + \beta \Delta L_{it}$
  - → $\Delta Y_{it} = Y_{it} - Y_{it-1}$ (a.k.a. first-difference method)
  - → or use "de-meaned data": $\tilde{Y}_{it} = Y_{it} - \bar{Y}_i$

ID (i)	Time (t)	Y_{it}	A_{it}	L_{it}
1	1	100	10	20
1	2	110	12	22
1	3	125	15	25
2	1	80	8	15
2	2	85	9	16
2	3	95	11	18

ID (i)	Time (t)	ΔY_{it}	$\tilde{Y}_{it}$	ΔA_{it}	$\tilde{A}_{it}$
1	1	-	-11.67	-	-2.33
1	2	10	-1.67	2	-0.33
1	3	15	13.33	3	2.67
2	1	-	-6.67	-	-1.33
2	2	5	-1.67	1	-0.33
2	3	10	8.33	2	1.67

Fixed-Effect Analysis

2. Individual dummy variables

ID (i)	Time (t)	Y_{it}	A_{it}	L_{it}
1	1	100	10	20
1	2	110	12	22
1	3	125	15	25
2	1	80	8	15
2	2	85	9	16
2	3	95	11	18

- Regression for person-time data
  - → Adjusting for individual dummies as a covariate
  - → $E[Y_{it}|A_{it}, L_{it}, D_i] = \beta_0 + \beta_1 A_{it} + \beta L_{it} + \gamma D_i$
    - D_i for $ID_i = 1, \ldots, n-1$

- β_1: a confounder-adjusted exposure-outcome association
  - → Conditional on individual IDs
    - Within the same individual
    - Eliminate confounding by confounders that did not change **between** $t-1$ **and** t

Fixed-Effect Analysis

⚠Caveats

1. **Temporal ordering**
    - Within-individual changes in exposure and outcome
        - → Which change occurred first?
            - **Reverse causation**
            - Lagged model (e.g., regress $Y_{i,t+1} - Y_{i,t}$ on $A_{i,t} - A_{i,t-1}$)
2. **Time interval**
    - Short-term changes
        - → Lack of variations for relatively stable exposures
        - → Capture immediate effects
    - Long-term changes
        - → Greater risks of time-varying confounders

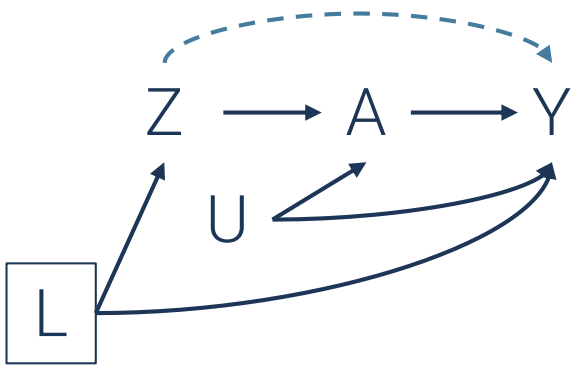
Alternative Assumptions: Approach 2

- **Quasi-experimental approaches**
  - → Use (plausibly) random exposure variations
    - Many of the “econometric” methods
    - Instrumental variable, regression discontinuity, interrupted time series
      - Often viewed as “superior” causal inference

! Not necessarily “superior” as they use:

- Different assumptions
- Different target population (hence, different estimand)

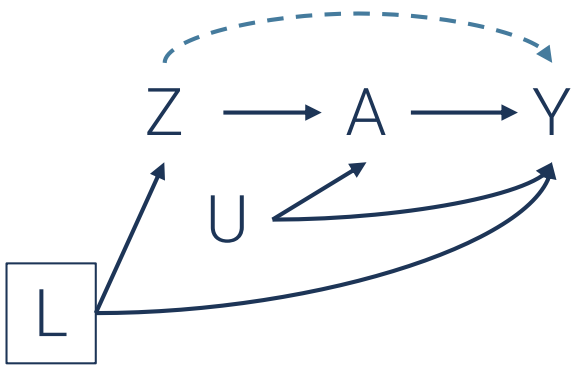
Instrumental Variable



- Effect of participating in community gathering places (A) on functional disability (Y) (Hikichi et al. 2015)
  - → Many confounders (U)
  - → The number of gathering places within a radius of 350m from each respondent's home (Z)
    - which affects participation (A)



Instrumental Variable



- Effect of participating in community gathering places (A) on functional disability (Y) (Hikichi et al. 2015)
  - → Many confounders (U)
  - → The number of gathering places within a radius of 350m from each respondent's home (Z)
    - which affects participation (A)

The effect of A on Y can be identified when:

1. Z is correlated with A
2. Z affects Y only through A
   - No direct effect of Z on Y (i.e., absence of the dotted arrow)
3. Exchangeability for the relationship between Z and Y (conditional on L)
   - **Replacing (conditional) exchangeability assumption**
     - → with extra assumptions
     - → may or may not be more plausible

Instrumental Variable: The Fourth Assumption

- Counterfactual exposure values (A) under an intervention for a proposed instrument ($Z = z$)
  - → A^z
    - Not observable

Types of A^z for binary A and Z

Type	$A^{z=0}$	$A^{z=1}$	Behavior
Never-takers	0	0	Never exposed regardless of Z
Compliers	0	1	Exposed only when $Z = 1$
Always-takers	1	1	Always exposed regardless of Z
~~Defiers~~	~~1~~	~~0~~	~~Would do opposite~~

4. **Monotonicity assumption**
   - There is no "defiers"
     - → If $z < z^*$ then $A_i^z \leq A_i^{z^*}$ for all i
     - → Intuition: There are no individuals who would do the opposite of what the instrument "encourages"

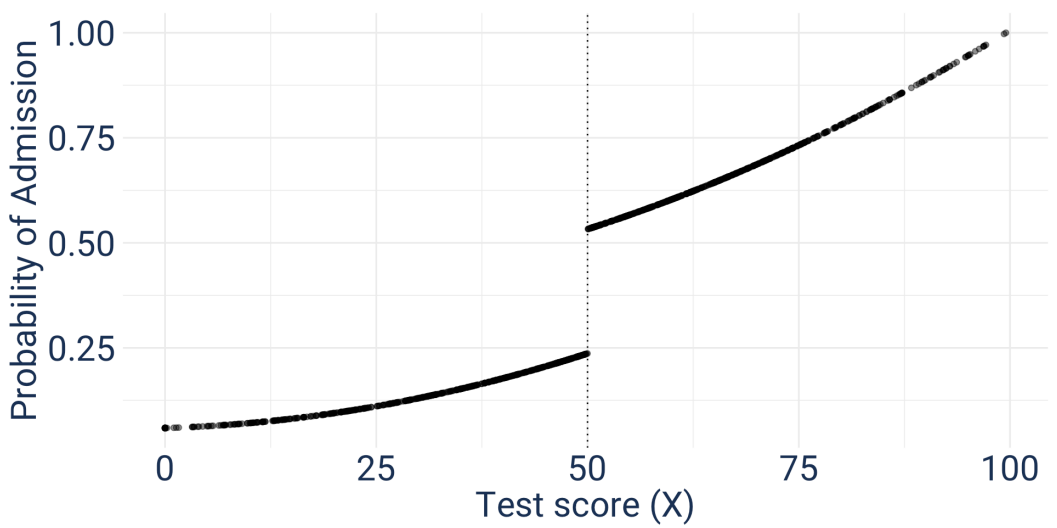
- Under the monotonicity assumption, the instrument Z allows the estimation of effect A on Y among "compliers"
  - → Who are they? External validity?

Regression Discontinuity Design (RDD)

Example: Effect of attending an elite university (A) on earnings (Y) (Hoekstra 2009)

- Those who got admitted vs those who did not
  - → Confounding
- Probability of admission "jumps" at the cutoff test score ($X = 50$)
  - → $\Delta Pr[A = 1|X = 50] = \lim_{x \leftarrow 50} Pr[A = 1] - \lim_{x \rightarrow 50} Pr[Y = 1]$
    - Applicants just below vs above the cutoff are similar?
      - Random variation in the treatment A (a special case of the instrumental variable approach)

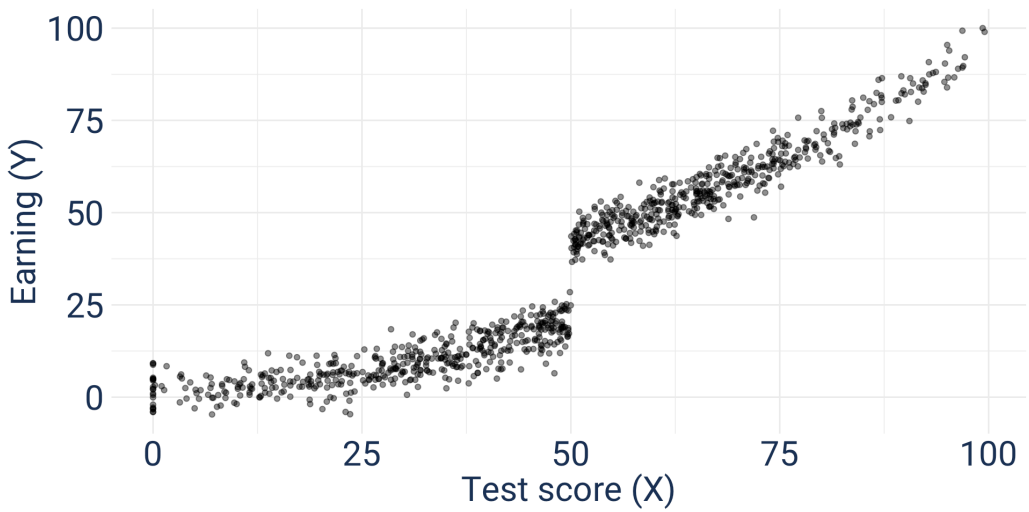
▶ Code

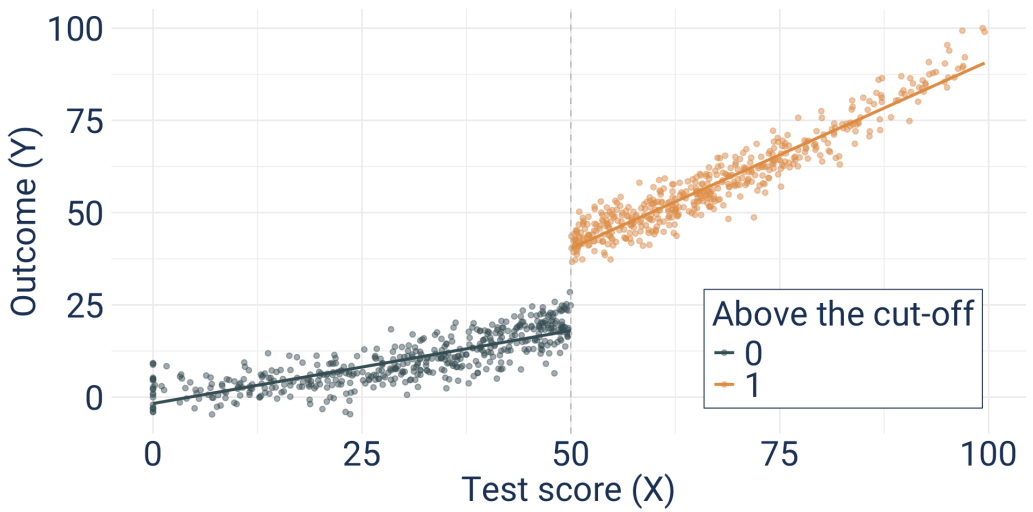


Regression Discontinuity Design (RDD)

- Difference in Y *at the cutoff*
  - → $\Delta E[Y|X=50] = \lim_{x \leftarrow 50} E[Y] - \lim_{x \rightarrow 50} E[Y]$
  - → **Assumption**: Applicants just below vs above the cutoff are similar
    - Attributable to the difference in (the probability of) college admission

▶ Code





- $\Delta E[Y|X=50] / \Delta Pr[A=1|X=50]$
  - → The effect of admission on earnings at the cut-off among compliers
  - → Local average treatment effect (LATE)

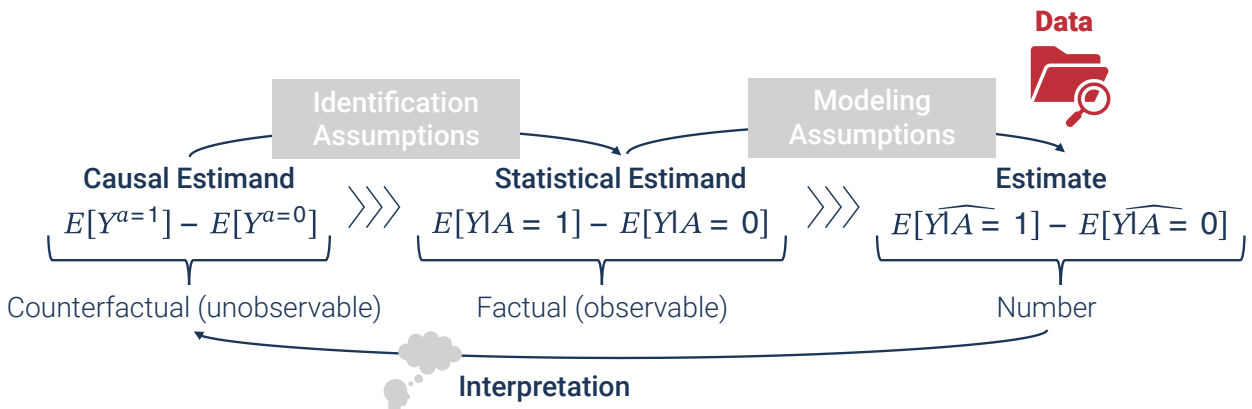
Summary 1

- **Between-person comparison**
  - → Unmeasured confounders but flexible
- **Within-person comparison**
  - → Time-varying confounders
  - → Data limitation, lack of variation, short-term effect
- **Quasi-experimental**
  - → Alternative assumptions
    - Still not verifiable
    - Not always feasible
  - → Different and vague target population (e.g., compliers)
    - External validity issue
    - When results are different from the covariate-adjsutment approaches
      - More rigorous causal inference?
      - Comparing 🍎 and 🍊 ?

Today's Plan

1. ~~Alternative approaches for exchangeability~~
2. Measurement
    - **Key considerations**
    - Specific measures in social epi
        - → SES
        - → Neighborhood environment

Measurement: Why Does It Matter?



- **Wrong estimates**
  - → Imprecise (reliability)
  - → Something different (validity)
    - Wrong interpretation
- **Unsuccessful Identification**
  - → Incomplete adjustment if poorly measured
    - Residual confouding

Measuring social variables

- Harder as they tend to be intangible, subjective, and complex
- But critical in ANY epi research
  - → as exposures and as covariates

Reliability

- Intuition: “Measured values are consistent”
  - → You get similar value every time you measure
    - i.e., Precision
  - → Can be an issue due to random measurement errors
- Implications
  - → Residual confounding
  - → Large noise & uncertainty
  - → Can bias results (Yland et al. 2022)



What do we do?

1. (Measure well)
2. Check reliability
   - Test-retest, Cronbach’s alpha, etc
     - → Context dependent
3. (Quantitative) bias analysis

Validity

- Intuition: “Measuring what you are trying to measure”
  - → Measured values captures truth on average

Validity can be limited due to:

- Information bias
  - → Systematic measurement errors/misclassification
    - e.g., Consistent underestimation due to social desirability bias
  - → Measurement errors in two variables
    - Differential vs Non-differential
    - Dependent vs Independent
- Conceptual violation

Lack of Conceptual Validity

Discordance between what you say you measured vs what you measured



- BMI as a measure of obesity
- Step count as a measure of physical activity
- Disease “diagnosis”
  - → Disease itself?
  - → Use of/access to healthcare services?
  - → Clinical practice?
    - Patterned by social factors

Always keep in mind:

- What exactly are you interested in conceptually?
  - → Match measures and concepts

Lack of Conceptual Validity

Discordance between what you say you measured vs what you measured

- Particularly challenging in social epi
  - → "Socioeconomic status" vs "total household income"
    - household size? living cost?
  - → Smart device-based measures of loneliness (e.g., Jafarlou et al. 2024)
    - "Objective" does not mean more valid

What do we do?

- Theoretical assessment
- Validation
  - → ~~Gold standard?~~
  - → Correlation with related variables
- Quantitative bias analysis (only for information bias)

Validity and Reliability

Valid and reliable



Valid but not reliable



Reliable but not valid



Systematic measurement errors



Conceptual errors



Crude Categories

- Some social measures use crude categories
  - → e.g., Income groups
  - → e.g., “Other” race categories
    - Asian vs NHPI
      - East Asia vs South Asia
        - Japanese American vs Korean American
  - → Not really measurement “error”



Strengths

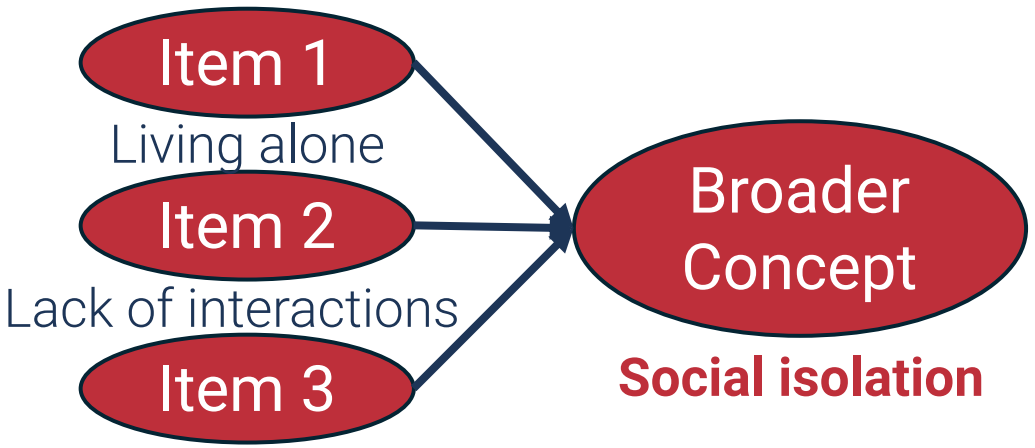
- Practical with a finite sample size
- Interpretability and comparability
- Non-linear

Limitations

- Lacks granular variations
- Assumed variation irrelevance within categories
- Residual confounding

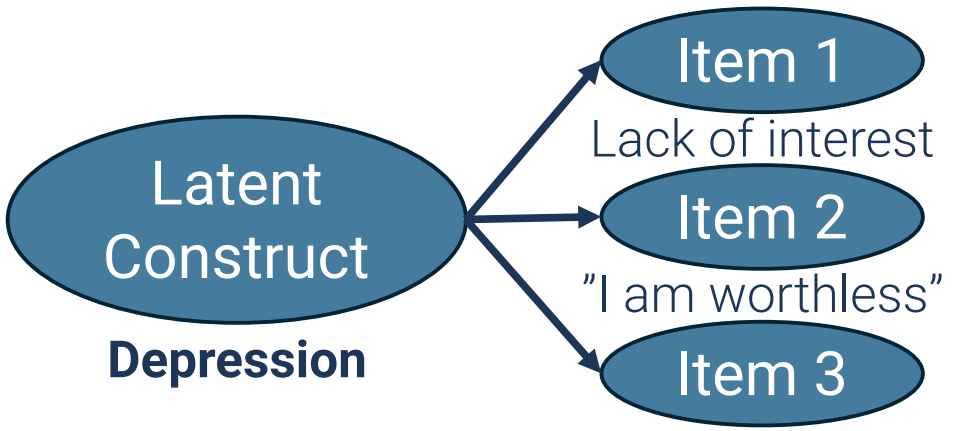
Multi-item Measures of Social Factors

Index



- Composite measures
  - → Totality of the exposures
  - → e.g., social isolation

Scale



- Measures a latent construct
  - → which manifests as each item
  - → e.g., loneliness, depression

Summary 2

- Measurement
  - → Reliability
  - → Validity
    - **Conceptual concordance between measures and target constructs**

Poor measures can give you:

- Biased results
  - → Residual confounding
  - → Information bias
- An answer to a question that you are not even asking

Today's Plan

1. ~~Alternative approaches for exchangeability~~
2. Measurement
    - ~~Key considerations~~
    - **Specific measures in social epi**
        - → SES
        - → Neighborhood environment

Measuring Socioeconomic Status (SES)

- Critical in population research
  - → SES virtually predicts EVERYTHING

Measures of SES

- Composite measures
  - → Can be less interpretable and policy relevant
- Univariate measures
  - → Specific aspects of SES
    - e.g., Education, income
  - → No single best
    - Measurement properties
    - Conceptual properties

Measures of SES 1: Education



- School year, credentials
  - → Degrees tend to matter more
  - → Quality of education?

Strengths

- Short
- Relatively reliable
- Captures social resources

Limitations

- Lacks granularity
- Hard to assess “changes”
- Varies by time & populations
- But poor proxy of economic resources
  - → Disproprotionate return by race & gender

Measures of SES 2: Income



- Often, annual household income
  - → Exact values or groups
    - Living cost
    - Household size
      - Equivalised income: e.g., $Income/\sqrt{HHS}$

Strengths

- Captures material resources
- Money is money
- More granularity

Limitations

- High missing, errors
- Money can't buy friends and wisdom
- Unstable
- Age dependent (e.g., retirees)

Measures of SES 3: Wealth



- Total wealth that combines:
  - → Real estate, cars, etc
  - → Liquid assets (cash, stocks, bonds)
  - → Liabilities (debt, loans, mortgage)

Strengths

- Captures intergenerational effects
- Less age dependent (e.g., retirees)

Limitations

- Can be hard to measure and compare
- Skewed distribution
  - → as if binary?

Other Measures of SES

Occupation

- Managers vs Professionals vs …

- Prestige, power & authority
  - → & social & material resources

- Challenging classification
- Not useful for unemployed, home-makers, retirees

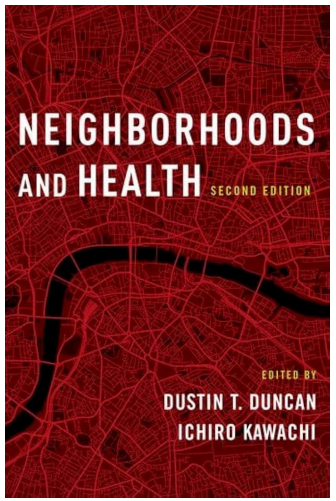
Consumption

- “Do you have a refrigerator, TV, etc?”

- Useful in settings where:
  - → Unemployment is common
  - → Eduction is a poor predictor
  - → Unstable income (e.g., seasonal work)
    - e.g., Global health

- Lacks variation in other contexts

Measuring Neighborhood Environment



- Neighborhood environment exposure affects health
  - → Food environment (e.g., food desert)
  - → Build environment (e.g., green space)
  - → Social environment (e.g., social capital)
    - Comprehensive reviews available in Duncan and Kawachi (2018)

! Measuring "neighborhood environment"

1. Defining neighborhoods
2. Defining (environmental) exposures

Operationalizing Neighborhoods

- Many ways to define “neighborhoods”
  - → No single best
- **Residential trap** (Chaix 2009)
  - → Home vs school, workplace
- **Uncertain geographic context problem (UGCoP)** (Kwan 2012)
  - → Relevant spatial areas, time, and duration for each environmental factor of interest to exert contextual influences on outcomes?

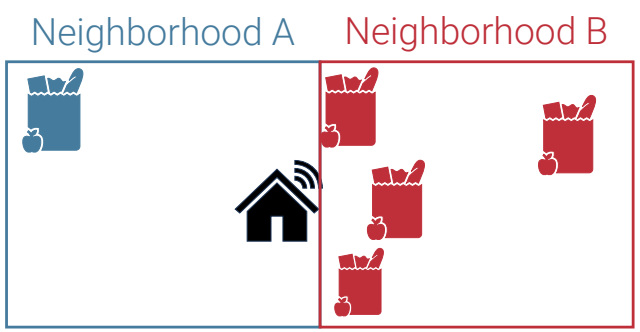
Which Approaches to Use?

- Theory-driven selection
  - → Is it capturing people’s lived experience and relevant mechanisms?
- Sensitivity analysis with various operationalization approaches

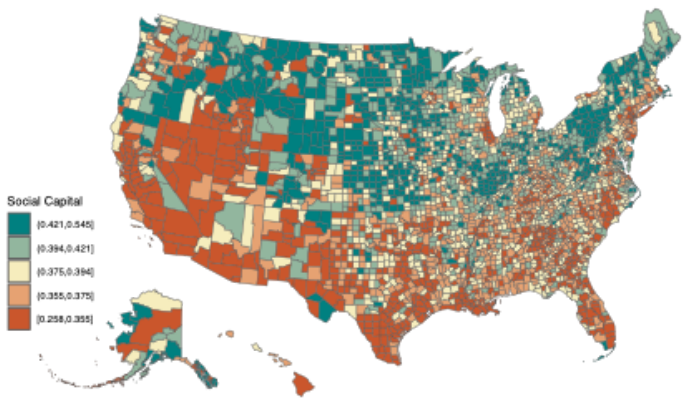
Operationalizing Neighborhoods

Approaches

- **Perceived boundaries**
  - → "In your neighborhood, …."
    - Strengths: may capture more relevant individual exposures (less UGCoP)
    - Limitations: hard to alter via policies, measurement errors, residential trap
- **Administrative boundaries**
  - → e.g., County, census tracts, ZIP
    - Strengths: policy relevant, data availability
    - Limitations: may not reflect true exposures, UGCoP & residential trap, arbitrary and changing boundaries[1], spatial misclassification

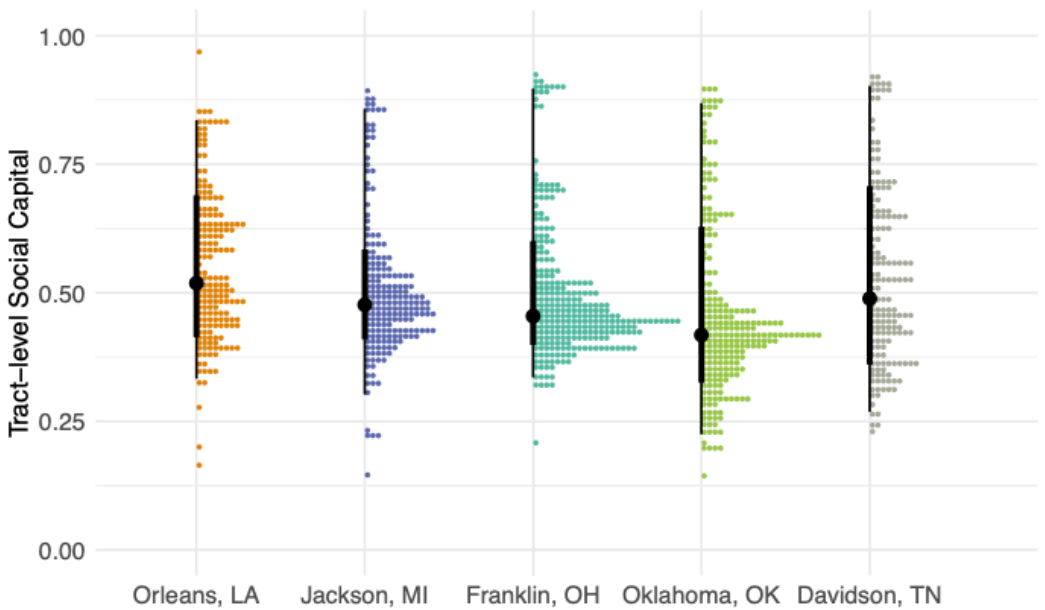


Administrative Boundaries



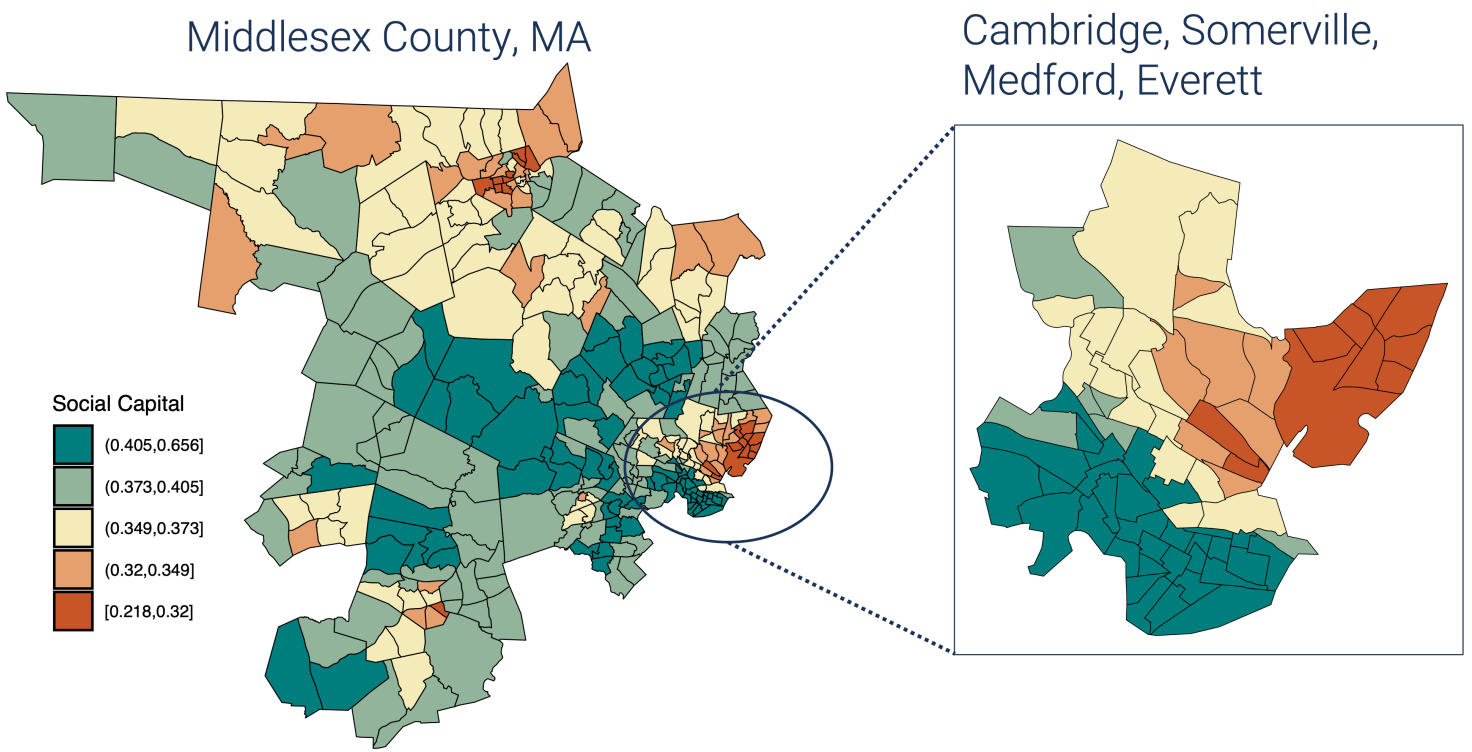
- Social capital
  - → Levels of mutual trust, reciprocity, social cohesion
  - → Often measured at the state/county levels
  - → New tract-level measures (Kyne and Aldrich 2020)
    - Based on publicly available data (e.g., # of volunteer organizations)

- Within-county variations in social capital across census tracts



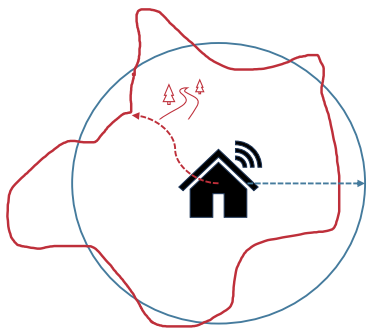
Administrative Boundaries

Within-County Variation in Social Capital



- **Does this mean the county-level measure is "wrong"?**
  - → Depends on what you are trying to capture conceptually
    - e.g., Daily exchange of social support vs collective efficacy

Operationalizing Neighborhoods



- **GIS-based buffer**
  - → e.g., Euclidian vs network distance from home address
    - Network distance vs travel time
  - → Strengths: Individualized exposures, minimized spatial misclassification
  - → Limitations: Still subject to the UGCoP & residential trap, rural vs local, requires exact address info



- **Activity space**
  - → Mobility patterns (e.g., GPS data through smart phones)
    - e.g., Experienced economic diversity of urban encounters (Yabe et al. 2023)
  - → Strengths: Total environmental (vs residential trap), less susceptible to UGCoP & spatial misclassification
  - → Limitations: Expensive, measurement errors, unstable

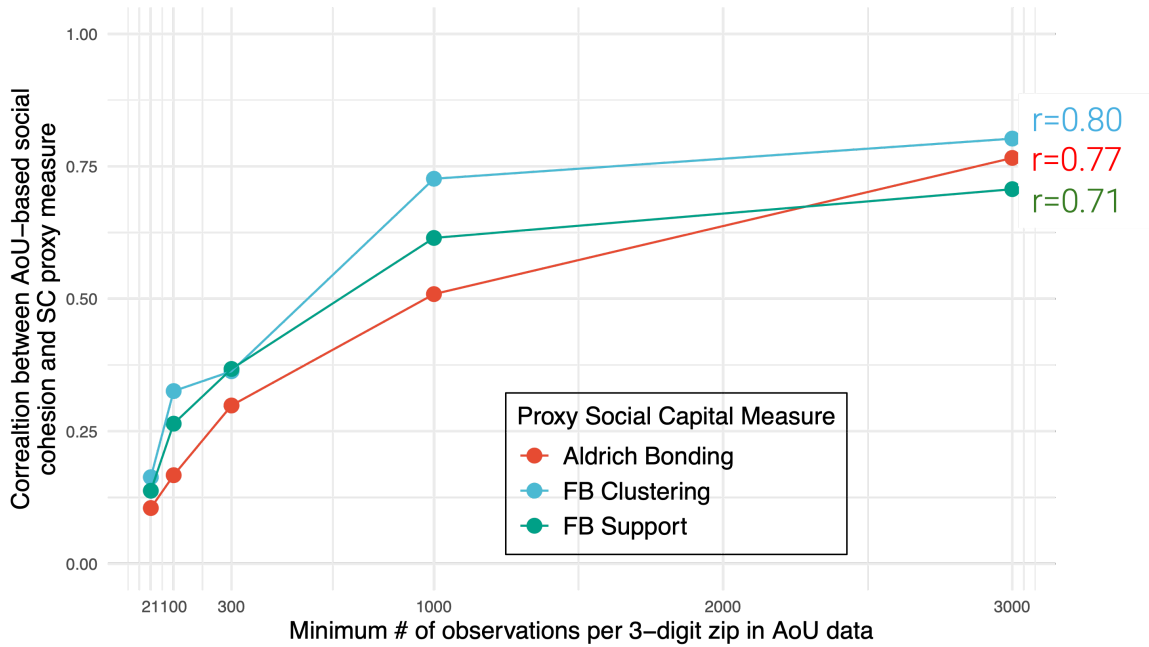
Once you define "neighborhoods", measure neighborhood-level "exposures"

Measuring Neighborhood Characteristics

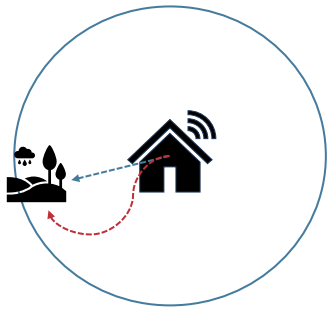
- **Self-report data**
  - → Individual perception: e.g., “Most people in this area can be trusted”
    - May not reflect the true neighborhood level
  - → Often aggregated: e.g., mean response across all respondents from the “neighborhood”
    - Non-representativeness
    - Can be highly unstable due to small sample size
- **Ecological data**
  - → Data availability
    - Administrative data (e.g., American Community Survey)
  - → Composite measures
    - e.g., “Social vulnerability index” (Flanagan et al. 2011)
    - Often proxy at best but relatively precise even in small neighborhoods

Aggregated Perception vs Ecological Data: An Example

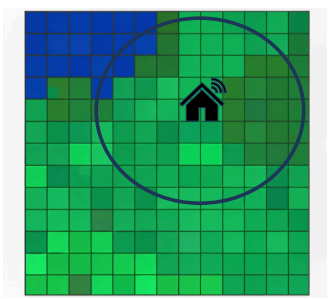
- **Self-report**: Social cohesion scale
  - → All of Us Research Program ($n \approx 260,000$)
  - → Aggregated at the 3-digit ZIP level with sample weights
- Correlation with social capital **ecological data** (Kyne and Aldrich 2020) [1]
  - → Strong only when #obs per neighborhood is large
    - "Does the aggregated perception of 30 respondents represent the neighborhood?"



Measuring Neighborhood Characteristics



- **Point of interest (POI) data** (e.g., locations of food outlets)
  - → Density (availability)
  - → Proximity (access)
    - Euclidian vs network distance



- **Raster data** (e.g., green space)
  - → Remote sensing, Google Street View, etc
  - → Composite (e.g., Google's PDFM)
    - e.g., median, centroid value, weighted average



- **Human mobility data** (Yabe et al. 2024)
  - → Activity pattern
  - → Can be combined with other data
    - e.g., Active usage, diversity of visitors

Summary 3

- **SES**
  - → Education, income, wealth, occupation, consumption
    - Unique strengths and limitations
    - Capture conceptually different aspects of SES
- **Neighborhood**
  - → Operationalizing neighborhoods
  - → Assessing neighborhood exposures
    - Unclear which neighborhood exposures are most relevant
    - Choose based on a theoretical model & perform sensitivity analyses

Assignment

- Problem set #4
- Reading quiz based on:
    1. Westreich, D. From patients to policy: population intervention effects in epidemiology. Epidemiology. 2017 July; 28(4):525-528.
    2. Shiba K, Torres JM, Daoud A, Inoue K, Kanamori S, Tsuji T, Kamada M, Kondo K, Kawachi I. Estimating the impact of sustained social participation on depressive symptoms in older adults. Epidemiology. 2021 Nov 1;32(6):886-95.

References

Chaix B. Geographic life environments and coronary heart disease: A literature review, theoretical contributions, methodological updates, and a research agenda. Annual review of public health. 2009;30(1):81–105.

Chetty R, Jackson MO, Kuchler T, Stroebel J, Hendren N, Fluegge RB, et al. Social capital i: Measurement and associations with economic mobility. Nature. 2022;608(7921):108–21.

Duncan DT, Kawachi I. Neighborhoods and health. Oxford University Press; 2018.

Flanagan BE, Gregory EW, Hallisey EJ, Heitgerd JL, Lewis B. A social vulnerability index for disaster management. Journal of homeland security and emergency management. 2011;8(1).

Hikichi H, Kondo N, Kondo K, Aida J, Takeda T, Kawachi I. Effect of a community intervention programme promoting social interactions on functional disability prevention for older adults: Propensity score matching and instrumental variable analyses, JAGES taketoyo study. J Epidemiol Community Health. 2015;69(9):905–10.

Hoekstra M. The effect of attending the flagship state university on earnings: A discontinuity-based approach. The review of economics and statistics. 2009;91(4):717–24.

Jafarlou S, Azimi I, Lai J, Wang Y, Labbaf S, Nguyen B, et al. Objective monitoring of loneliness levels using smart devices: A multi-device approach for mental health applications. Plos one. 2024;19(6):e0298949.

Kwan MP. The uncertain geographic context problem. Annals of the Association of American Geographers. 2012;102(5):958–68.

Kyne D, Aldrich DP. Capturing bonding, bridging, and linking social capital through publicly available data. Risk, Hazards & Crisis in Public Policy. 2020;11(1):61–86.

Yabe T, Bueno BGB, Dong X, Pentland A, Moro E. Behavioral changes during the COVID-19 pandemic decreased income diversity of urban encounters. Nature communications. 2023;14(1):2310.

Yabe T, Luca M, Tsubouchi K, Lepri B, Gonzalez MC, Moro E. Enhancing human mobility research with open and standardized datasets. Nature Computational Science. 2024;4(7):469–72.

Yland JJ, Wesselink AK, Lash TL, Fox MP. Misconceptions about the direction of bias from nondifferential misclassification. American journal of epidemiology. 2022;191(8):1485–95.