Estimating the effects of time-varying treatments in the presence of time-varying confounding

An application to neighborhood effects on high school graduation

David J. Harding, University of Michigan

(joint work with Geoffrey T. Wodtke, University of Michigan, and Felix Elwert, University of Wisconsin-Madison)

IU Workshop in Methods

March 22, 2013

Goals for Today

1. Understand the challenges of estimating the effects of time-varying treatments in the presence of time-varying confounding
2. Discuss two methodologies developed in Epi/Biostatics for estimating such effects:
   - Main effects: Inverse Probability of Treatment Weighting (IPTW) in a Marginal Structural Model (MSM) – Robins et al. 2000
   - Conditional effects: Regression with residuals (RR) for estimating Structural Nested Mean (SNM) Models – Robins 1994, Almirall et al 2010

Goals for Today

3. Illustrate the application of these models using the case of neighborhood effects on high school graduation
   - Geoffrey Wodtke, David J. Harding, and Felix Elwert. 2011. "Neighborhood Effects in Temporal Perspective." *American Sociological Review* 76(5): 713-736.
   - Geoffrey Wodtke, Felix Elwert, and David J. Harding. 2012. "Poor Families, Poor Neighborhoods: How Family Poverty Intensifies the Impact of Concentrated Disadvantage on High School Graduation." Population Studies Center Research Report 12-776.

Overview

- Motivate our interest in time-varying treatments
- Explain why conventional regression methods will usually produce biased results
- Describe and illustrate IPTW/MSM methodology for estimating main effects
- Describe and illustrate RR/SNM methodology for estimating conditional effects (interactions with time-varying covariates)
- Discuss assumptions and issues in implementation

Motivation: Neighborhood Effects Example

- What is the effect of growing up in a disadvantaged neighborhood on one's probability of graduating from high school?
- Theories
  - Social and cultural isolation: Role models, linguistic isolation
  - Social organization: low social cohesion limits collective supervision of youth behavior; high crime/violence as stressor
  - Institutions and resources: deficient infrastructure, e.g. schools, daycare centers, grocery stores, recreational areas
  - Environmental: health effects of air pollution, housing stock, etc.
- Previous research finds mixed results
  - Little attention to *duration* of exposure to disadvantaged neighborhoods
  - *Over control* of intermediate pathways

Motivation: Neighborhood Effects in Temporal Perspective

- The above theories suggest duration of exposure matters
- Neighborhoods are not a static feature of a child's life; families move and neighborhoods change
- Selection into different neighborhoods across time is based on both time-invariant ("baseline") and time-varying covariates

$$\text{Income}_{T=1} \longrightarrow \text{NH}_{T=2}$$

- Neighborhood context, in turn, may impact many of the same time-varying family characteristics that influence neighborhood selection

$$\text{Income}_{T=1} \longrightarrow \text{NH}_{T=2} \longrightarrow \text{Income}_{T=3}$$

Data for Main Effects Analysis

- 1968-1997 waves of the Panel Study of Income Dynamics (PSID) linked to the Geolytics Neighborhood Change Database (NCDB)
- Analytic sample
  - 4,154 children present at age 1 in PSID family units (FUs) between 1968-1978; subjects observed yearly until age 17 or loss to follow-up
  - 2,380 subjects in final outcomes model, all subjects used for weights
- Weights to adjust for sample attrition – similar to IPT weights (see supplemental slides)
- k=16 usable waves of follow-up
  - Measurements taken once per year, every year from age 1 to 17
  - HS graduation measured at end of follow-up (age 20)

Key Variables

- Time-dependent exposure A_k
  - PCA of tract characteristics used to create neighborhood disadvantage index
  - Ordinal Measure: Residence in a neighborhood in a specific quintile of the index
- Outcome Y: HS graduation by age 20
- Time-invariant (baseline) characteristics L_0
  - Gender, birth weight, mother's age at birth, mother's marital status at birth, "family unit" head's education (measured at baseline), year born
- Time-dependent confounders L_k
  - FU head's marital status, employment, age, and work hours; welfare receipt, homeownership, income, family size, moves, past neighborhood exposure

Time-varying Treatments in the Counterfactual Framework

- 5-category ordinal treatment: $A_k \in \{1,2,\dots,5\}$
- Treatment sequence up to wave k: $\bar{a}_k = (a_1,\dots,a_k)$
- Complete treatment sequence (age 2-17): $\bar{a} = \bar{a}_K$
- $Y_{\bar{a}}$ is potential outcome had child been exposed to the sequence of neighborhood contexts $\bar{a}$
  - $Y_{(4,5,\dots,5)}$ outcome had child been exposed to a 4th quintile neighborhood during the first follow-up wave and neighborhoods in the most disadvantaged quintile thereafter
- average causal effect of neighborhood exposure sequence $\bar{a}$ compared to another exposure sequence $\bar{a}'$:
  - $E(Y_{\bar{a}} - Y_{\bar{a}'}) = E(Y_{\bar{a}}) - E(Y_{\bar{a}'}) = P(Y_{\bar{a}} = 1) - P(Y_{\bar{a}'} = 1)$

Marginal Structural Models (MSMs)

- specify the following parametric model for the counterfactual probabilities:

$$logit\left(P(Y_{\bar{a}} = 1)\right) = \theta_0 + \theta_1 \left(\sum_{k=1}^{16} a_k \Big/ 16 \right)$$

- the probability of high school graduation is a function of duration-weighted exposure to different levels of neighborhood disadvantage
  - i.e., the average of ordinal wave-specific treatments from wave $k = 1$ to 16
  - θ_1 = the effect of growing up in neighborhoods that are, on average, located in quintile q of the disadvantage distribution rather than the less disadvantaged quintile $q - 1$
- “marginal” here refers to population average effects (as opposed to conditional effects)
- “structural” here simply means causal effects

No Unobserved Confounders Assumption

- $Y_{\bar{a}} \perp A_k | \bar{L}_k, \bar{A}_{k-1}$
- In words: the level of neighborhood disadvantage at each wave k is independent of potential outcomes given observed covariate history and past treatments
  - children with the same combination of observed covariate values do not systematically select into different neighborhood contexts based on unobserved factors predictive of the outcome
  - "No unobserved confounding of treatment"
- Not a directly testable assumption

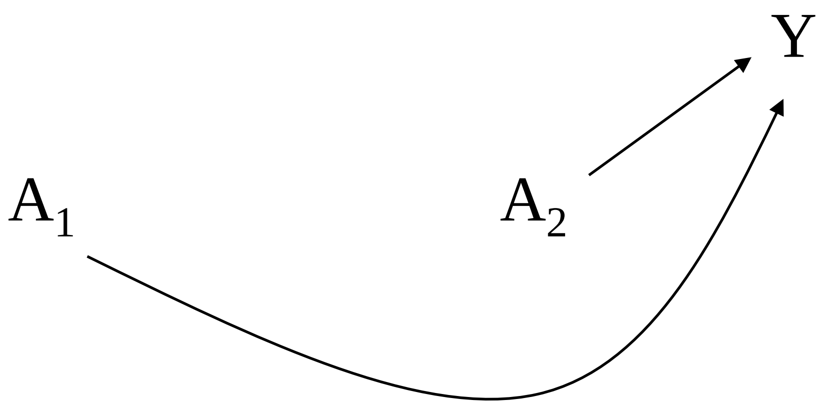
- But how do we control for observed covariates?

Intuition for Methodological Problems

- Consider a world with treatment at two time points
- How would we estimate the effect of a "treatment" like neighborhood disadvantage?
- Standard regression

A_k : NH at time k

Y : HS grad

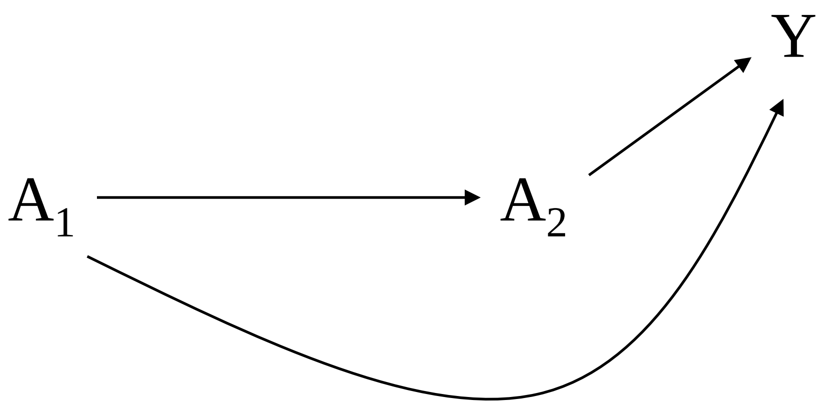


Intuition for Methodological Problems

- Now make things slightly more complicated
- How would we estimate the effect of a "treatment" like neighborhood disadvantage?
- Standard regression still okay

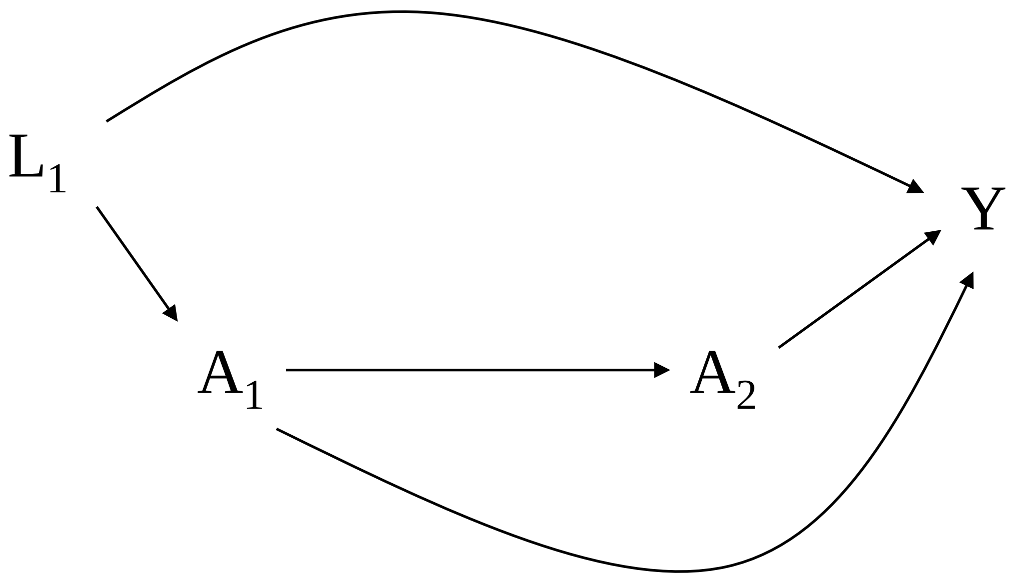
A_k : NH at time k

Y : HS grad



Intuition for Methodological Problems

- Now with static (baseline) selection into treatment
- How would we estimate the effect of a "treatment" like neighborhood disadvantage?
- Standard regression, control for baseline



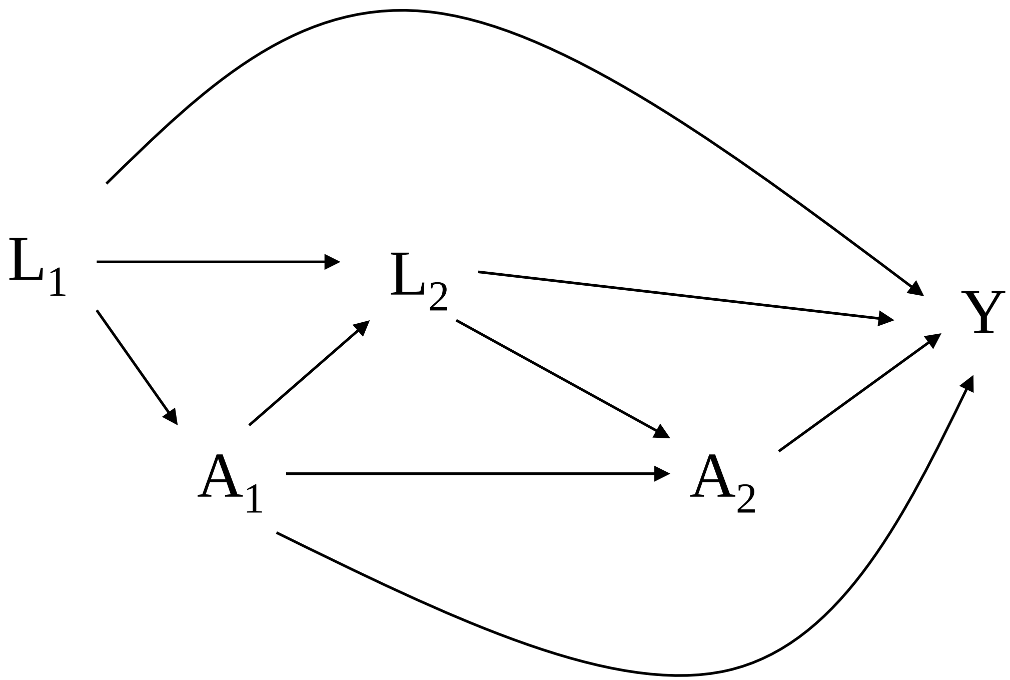
A_k : NH at time k

Y : HS grad

L_k : Observed Confounders

Intuition for Methodological Problems

- Now add dynamic (time-varying) selection into treatment
- Standard regression with control for L_2 "over controls"
  - The part of the effect of A_1 that goes through L_2 is gone



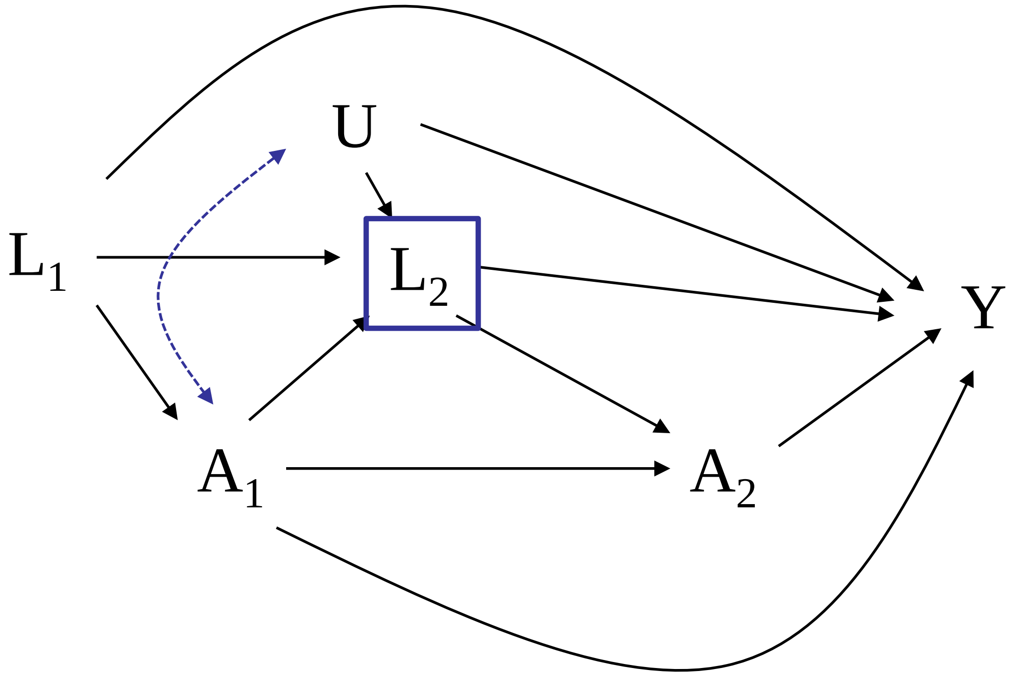
A_k : NH at time k

Y : HS grad

L_k : Observed Confounders

Intuition for Methodological Problems

- Now add unobservables
  - Note: treatment is still unconfounded
- Standard regression induces “endogenous selection” or “collider-stratification” bias
  - Controlling for L_2 also induces association between U and A_1



A_k : NH at time k

Y : HS grad

L_k : Observed Confounders

U : Unobserved Confounders

Inverse probability of treatment weighting

- Resolves the overcontrolling and endogenous selection problems just discussed without making strong assumptions about dynamic selection processes
- Weights are the inverse of the probability of receiving the treatment actually received
- Intuition: weight observations to generate a pseudo-population in which *treatment and observed covariates are no longer correlated*
  - Up-weight observations with low probability of receiving treatment actually received – these observations are important comparisons

Inverse Probability of Treatment Weights

- In words: inverse probability of receiving the treatment actually received, based on prior treatment, baseline confounders, and time-varying confounders

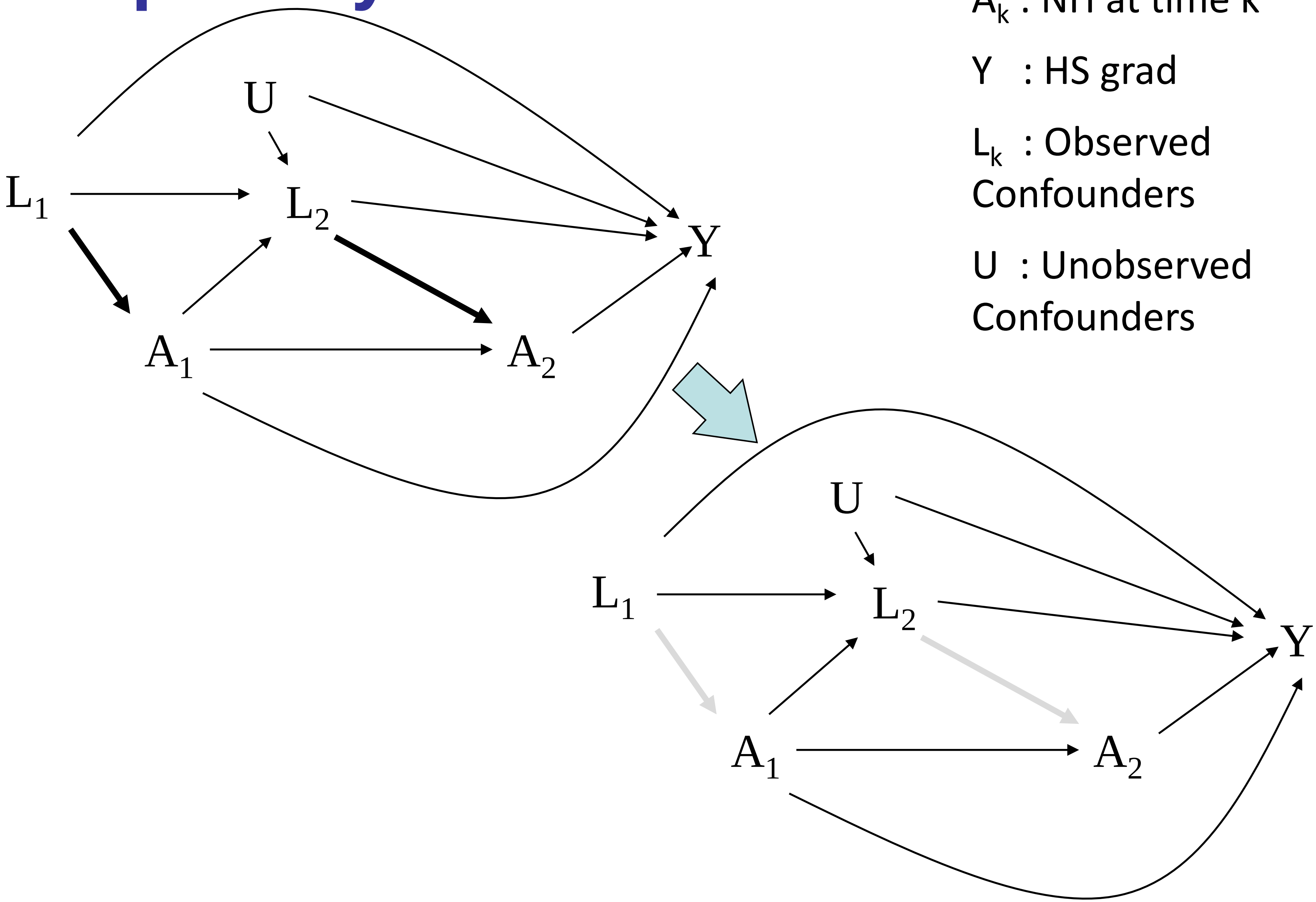
- Weight formula: $$W = \prod_{k=0}^{K} \frac{1}{P(A_k = a_k | \bar{A}_{k-1}, \bar{L}_{k-1}, V)}$$

- Stabilized weight: $$SW = \prod_{k=0}^{K} \frac{P(A_k = a_k | \bar{A}_{k-1}, V)}{P(A_k = a_k | \bar{A}_{k-1}, V, \bar{L}_{k-1})}$$

IPTW: Mechanics

- Estimate a model predicting treatment (here, an ordinal logit) with baseline controls, treatment history, and time-varying controls
  - Use this to estimate predicted probability of treatment -> denominator
- Estimate a model predicting treatment (here, an ordinal logit) with baseline controls and treatment history
  - Use this to estimate predicted probability of treatment -> numerator
- Multiply weights over time to get year-specific cumulative weights
- If necessary: multiply by sampling weight and censoring weight to get final weight
- Weight the regression model, controlling directly for baseline covariates

Graphically



A_k : NH at time k

Y : HS grad

L_k : Observed Confounders

U : Unobserved Confounders

Weights

Table 4. Stabilized Treatment and Attrition Weights

			Percentiles			
Weight	Mean	SD	1st	25th	75th	99th
Blacks (*n* = 834)						
Stabilized treatment weight (SW)	1.03	.58	.27	.73	1.18	4.62
Stabilized attrition weight (CW)	1.00	.12	.76	.92	1.06	1.44
SW x CW	1.04	.61	.26	.71	1.18	4.80
Nonblacks (*n* = 1,259)						
Stabilized treatment weight (SW)	1.00	.31	.32	.84	1.10	2.54
Stabilized attrition weight (CW)	1.00	.14	.71	.93	1.03	1.70
SW x CW	1.00	.37	.32	.81	1.10	3.00

Note: Statistics reported for children not lost to follow-up before age 20 (first imputation dataset).

Results

Table 5. Effects of Duration-Weighted Exposure to Neighborhood Disadvantage on High School Graduation (log odds ratios)

	Blacks (n = 834)			Nonblacks (n = 1,259)		
Model	Coef	SE		Coef	SE	
Unadjusted	−.703	(.170)	***	−.581	(.109)	***
Regression-adjusted	−.416	(.196)	*	−.212	(.125)	
Stabilized IPT-weighted	−.525	(.190)	**	−.274	(.128)	*

Note: Analyses based on children not lost to follow-up before age 20. Coefficients and standard errors are combined estimates from five multiple imputation datasets.
$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$ (two-sided tests of no effect).

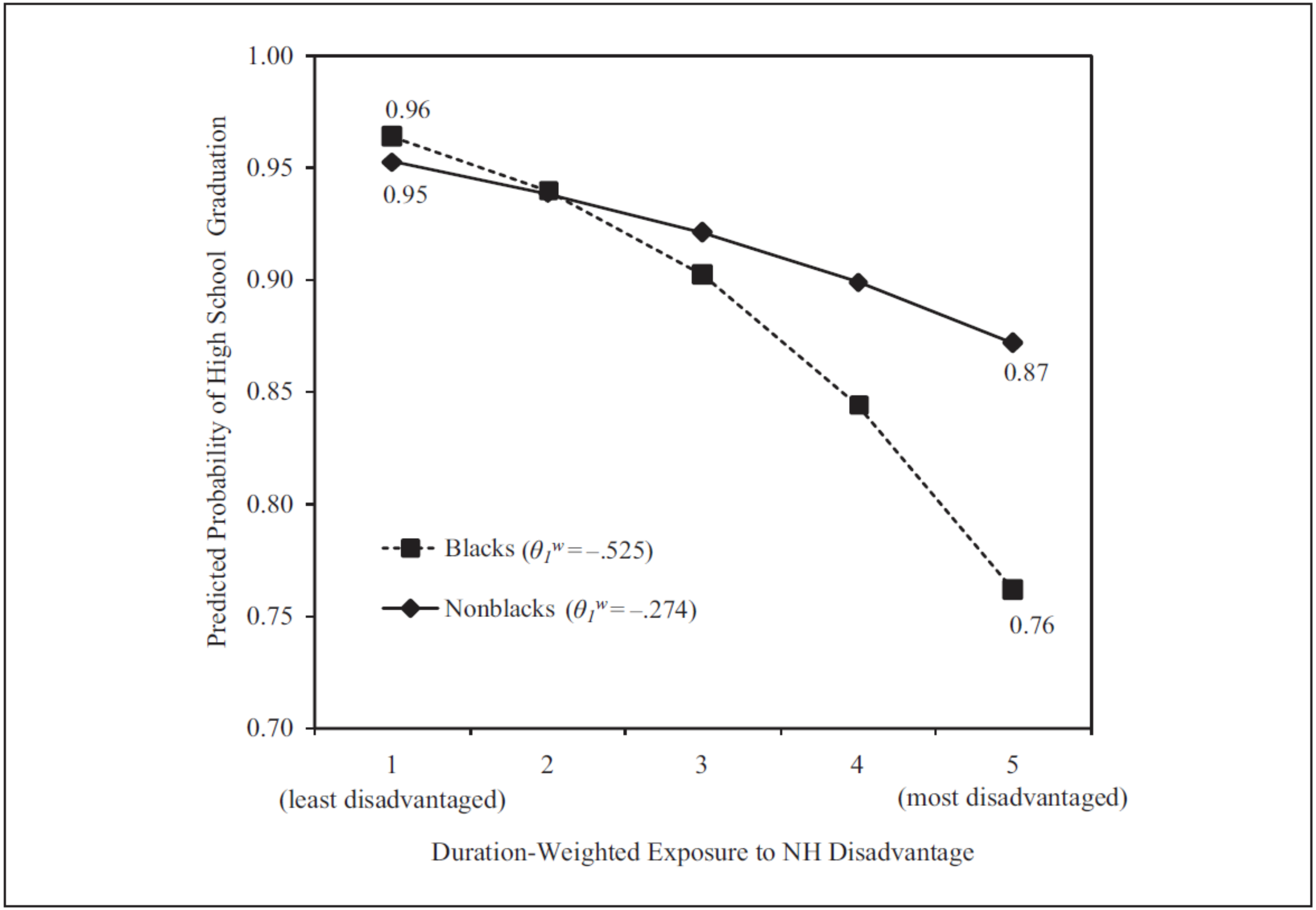


Figure 3. Predicted Probability of High School Graduation by Neighborhood Exposure History
Note: NH = Neighborhood

Key IPTW Assumptions

- No unobserved confounding (a.k.a. "sequential ignorability"), discussed earlier
- No model misspecification
  - Check sensitivity to weight model specification – see supplemental slides
- "positivity" (similar to "common support")
  - nonzero probability of treatment for every level and combination of confounders (Cole and Hernan 2008)
  - Check in data – see supplemental slides
- Do NOT need to assume that observed time-varying confounders are not affected by past treatment

Additional Considerations with IPTW

- In general, weighting increases SEs
- Large weights are indicative of big differences in probability of treatment
  - With many time periods, large weights are often inevitable
- Outlier weights are often removed by "trimming" or "truncation" (top or bottom coding)
  - bias/variance trade-off
- Can combine with multiple imputation
  - Be sure to do the whole procedure for each imputed dataset

Conditional Effects Motivation

- Theory suggests two types of neighborhood effect heterogeneity:
  - Heterogeneity by family poverty
  - Heterogeneity by timing of exposure to different NH contexts
- Limitations of previous studies
  - Focus on marginal, or population average, effects
  - Scant attention to role of timing of neighborhood exposure
  - Improper handling of dynamic neighborhood selection
- Research questions
  - Does impact of neighborhood disadvantage depend on family economic resources? Timing of neighborhood exposure?

Effect Moderation by Family Poverty Status

- Compound disadvantage theory
  - Neighborhood disadvantage has larger impact on children from poor families because of their more limited social networks, greater reliance on neighborhood resources
- Relative deprivation theory
  - Neighborhood disadvantage has larger impact on children from nonpoor families because they can realize benefits of advantaged neighborhoods, unlike poor children

Effect Moderation by Timing of Exposure

- Adolescence:
  - school continuation decisions occur during this period
  - child’s social world begins to incorporate neighborhood
  - peer socialization more important
- Childhood:
  - young children particularly sensitive to environmental inputs
  - later educational outcomes built on foundations laid down early in life

Dynamic neighborhood selection

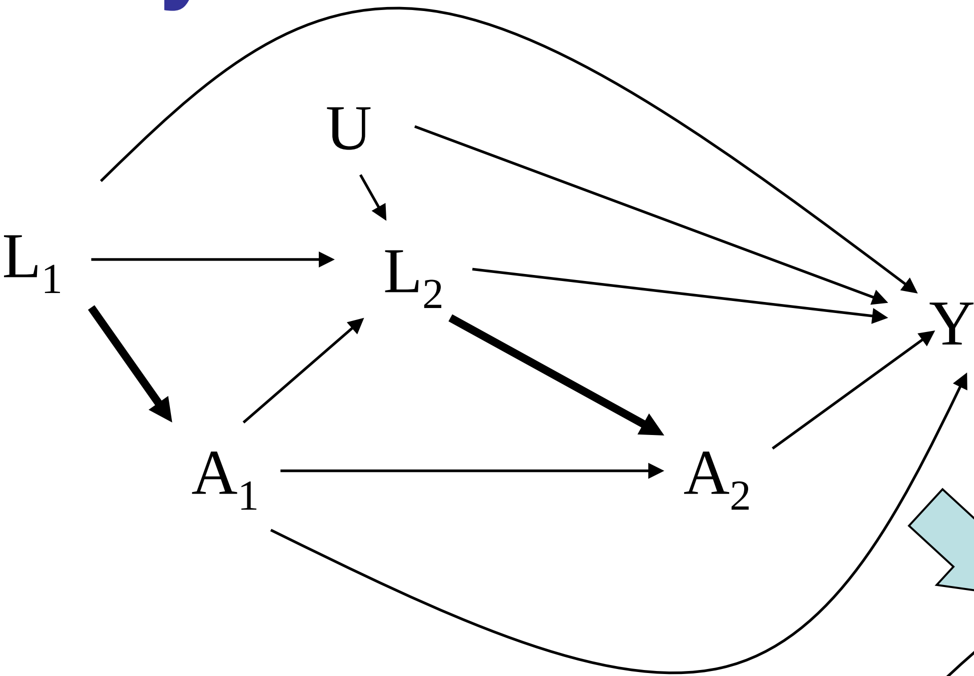
- Impossible to reconcile without temporal framework
- Dynamic neighborhood selection and feedback

family income → NH context → family income → NH context

TIME →

- Family income simultaneously
  - Confounds effects of future NH context
  - Mediates effects of past NH context
  - Moderates effects of NH context????
- Goal
  - Estimate time-dependent NH effects for subgroups of children defined in terms of their family poverty history

Why Doesn’t IPTW Work Here?

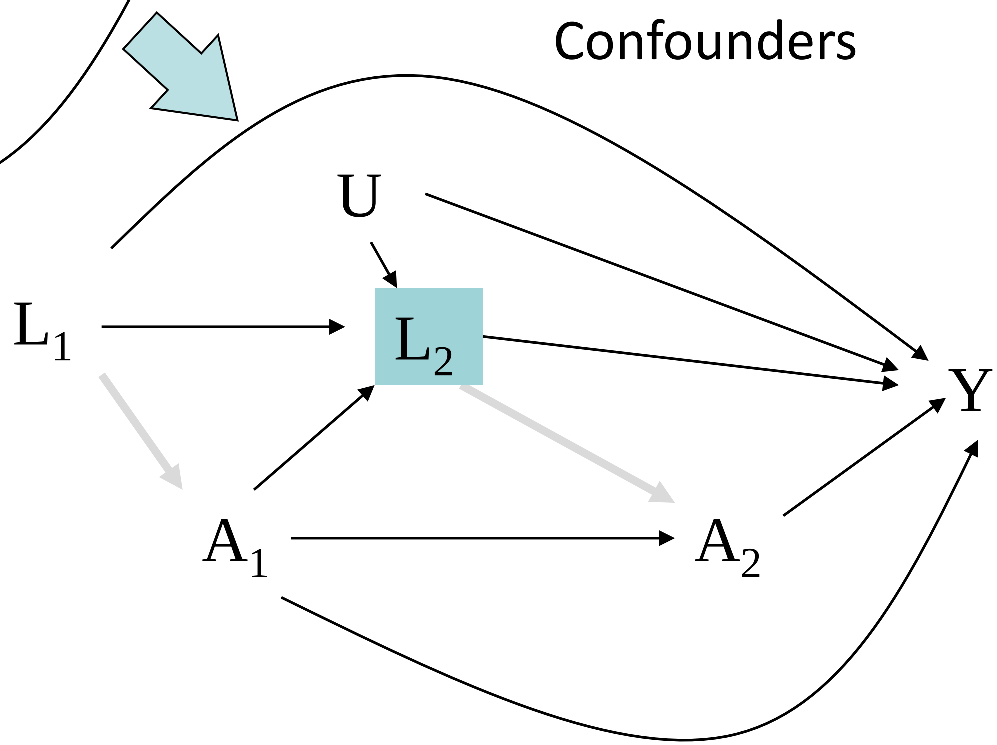


A_k : NH at time k

Y : HS grad

L_k : Observed Confounders

U : Unobserved Confounders



Our effect modifier, L_2, is not in the model!

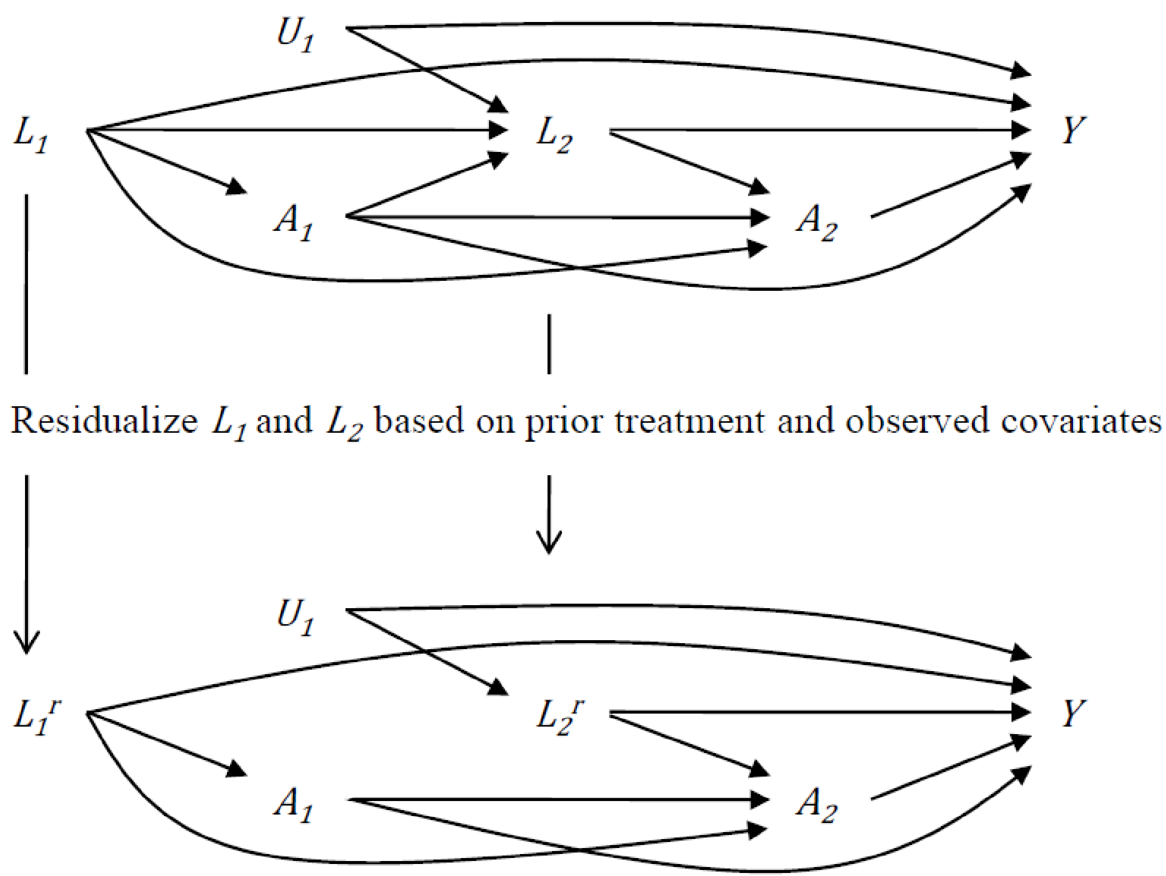
$$logit\big(P(Y_{\bar{a}} = 1)\big)$$
$$= \theta_0 + \theta_1 \left(\sum_{k=1}^{16} a_k \Big/ 16 \right)$$

Intuition of Two-Stage Regression with Residuals (Almirall et al. 2010)

- Problem: we need to
  1. remove confounding by time-varying covariates without introducing bias due to overcontrolling and collider stratification bias; and
  2. preserve our ability to interact our time-varying confounders with treatment
- Intuition for difficulty: we are using a variable that is both a confounder and a mechanism as a moderator
- Solution: Residualize the time-varying confounders to remove their associations with past treatment

Intuition Graphically

Figure 3. Consequences of residualizing time-varying covariates



Notes: A_k = neighborhood disadvantage, L_k = family economic resources and other time-varying covariates, U_k = unobserved factors and Y = high school graduation. L_1 includes time-invariant baseline covariates.

Mechanics of two-stage regression-with-residuals

- First-stage
  - Regress time-varying covariates on past treatment and past covariates; compute residuals

$$L_1^{resid} = L_1 - E(L_1)$$

$$L_2^{resid} = L_2 - E(L_2|L_1,A_1)$$

- Second-stage
  - Enter residuals from first-stage in regression for outcome

$$E(Y|L_1,A_1,L_2,A_2)$$
$$= B_0 + \eta_1 L_1^{resid} + B_1 A_1 + B_2 L_1 A_1 + \eta_2 L_2^{resid} + B_3 A_2 + B_4 L_2 A_2$$

Intuition in terms of hypothetical experiment

- Sequentially randomized experiment with two time points (here, childhood and adolescence)
  - At each time point, randomize treatment (here, neighborhood disadvantage)
- Measure moderator (here, family poverty) before each randomization
- Measure outcome (here, HS graduation) in early adulthood
- To estimate moderated effect of childhood neighborhood disadvantage, compare mean outcomes across randomized childhood treatment categories, separately by childhood family poverty
- To estimate moderated effect of adolescent neighborhood disadvantage, compare mean outcomes across adolescent treatment categories, separately by adolescent family poverty

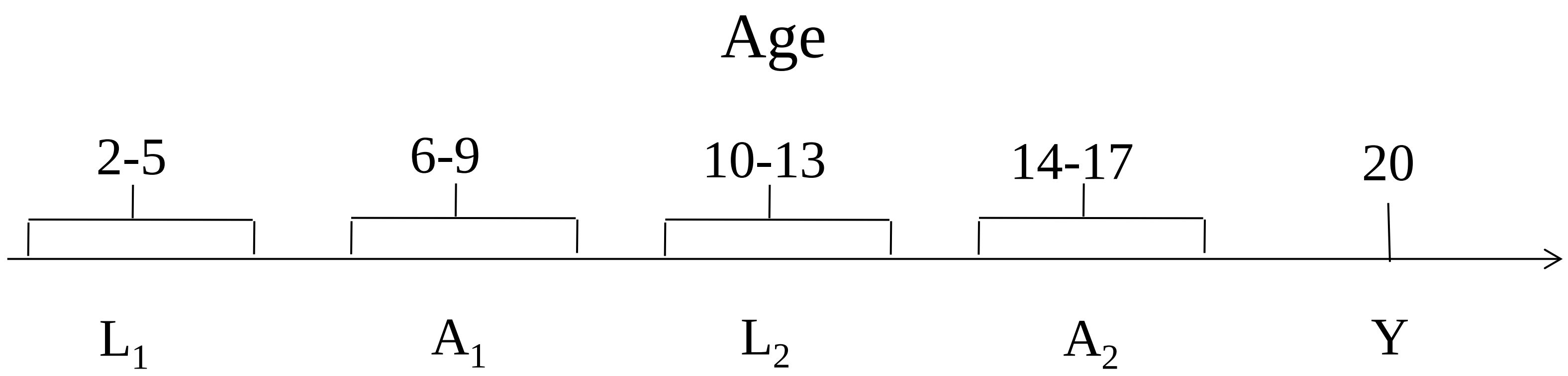
Two-stage regression-with-residuals: assumptions and practicalities

- Unbiased and consistent under assumptions of
  - sequential ignorability (no unobserved time-varying confounding)
  - no model misspecification
  - see supplemental slides for robustness tests
- Requires linear models (decomposition shown below does not work in nonlinear models)
- Model each time-varying confounder at each time point
- Weaker assumptions than conventional regression
- Bootstrap SEs (slightly conservative)
- Multiple imputation for missing data

Same Data and Variables, Except...

- Analytic sample
  - 6,135 children present in PSID at age 2 between 1968 and 1982
  - Children followed from age 2 to 20
- Focus on effect moderation by family income-to-needs ratio (centered at zero) – L_1 and L_2
  - greater than 0 for families with incomes that exceed poverty level
  - less than 0 for families with sub-poverty incomes
  - Categories for ease of presentation:
    - "extremely poor": income-to-needs = –.5
    - "poor": income-to-needs = 0
    - "non-poor": income-to-needs = 2

Measurement strategy



- Treatment, covariate measures based on mutli-wave averages

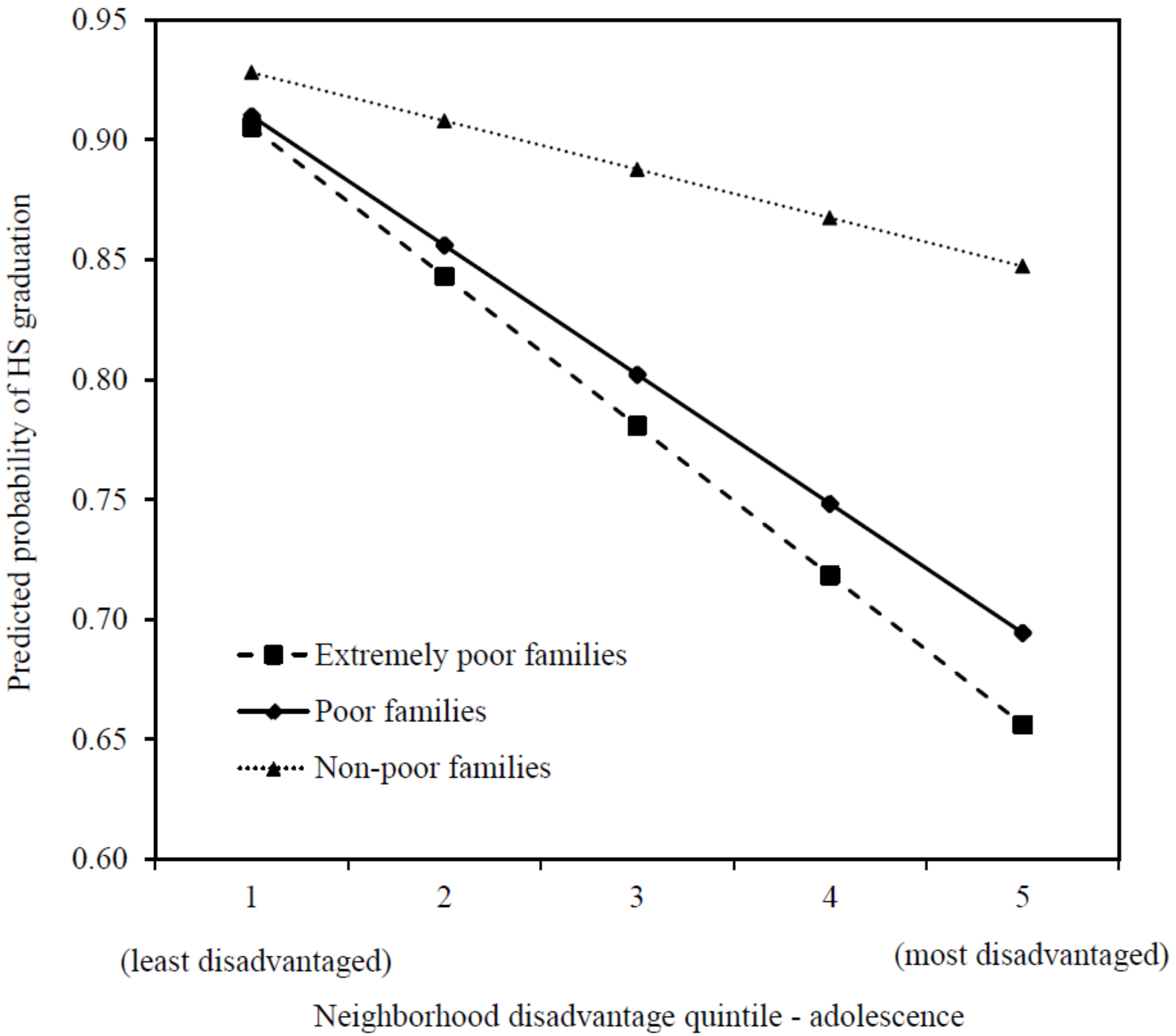
Results

Table 5. Effects of neighborhood disadvantage on high school graduation (two-stage estimates)

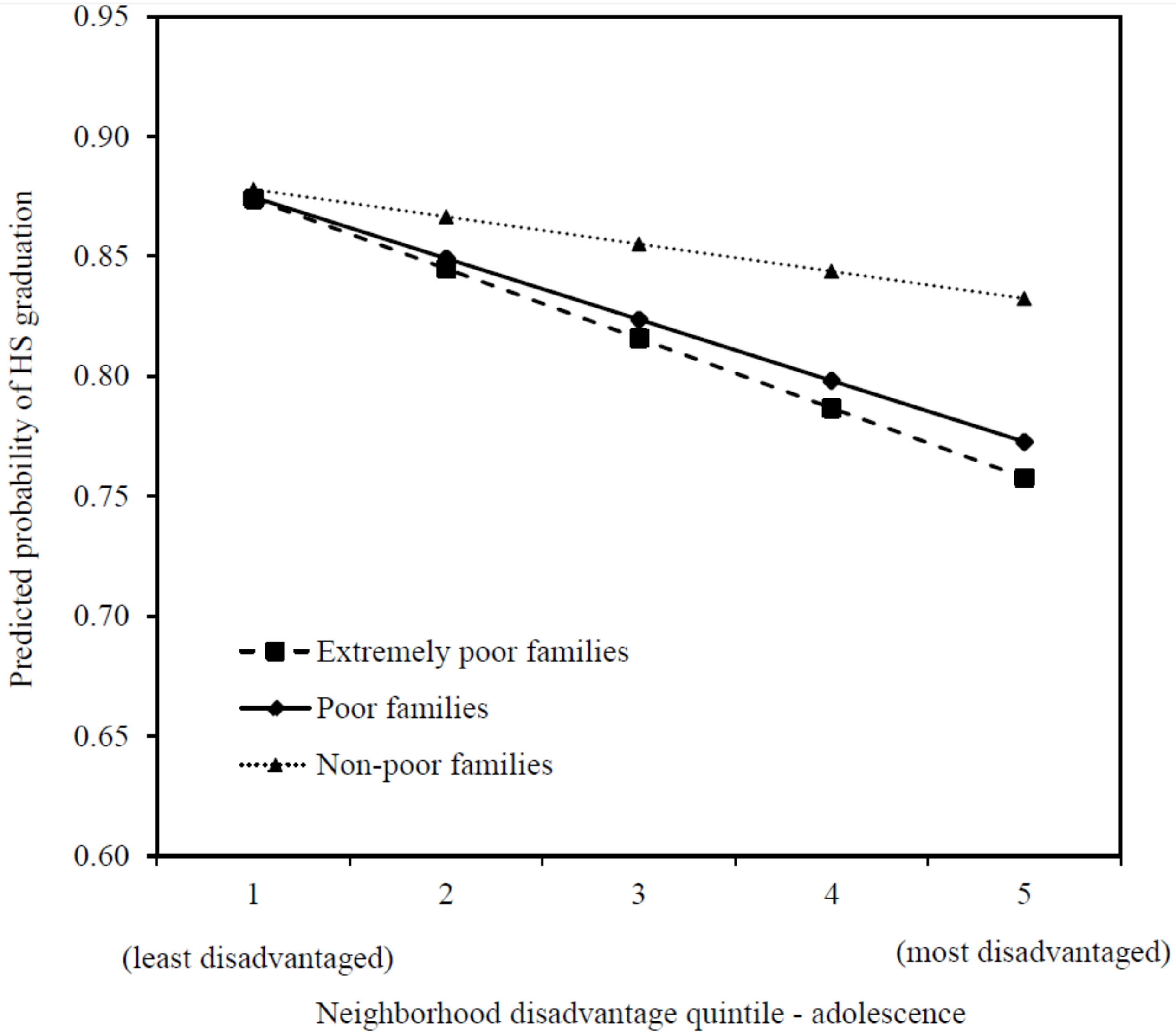
Model	Total			Blacks			Nonblacks		
	coef	se		coef	Se		coef	se	
Intercept	.888	(.021)	***	.916	(.044)	***	.877	(.019)	***
Childhood									
NH dadvg	–.005	(.012)		–.004	(.019)		–.006	(.015)	
NH dadvg x inc-to-needs	.005	(.004)		.005	(.008)		.005	(.005)	
Adolesence									
NH dadvg	–.042	(.010)	***	–.054	(.018)	**	–.026	(.013)	†
NH dadvg x inc-to-needs	.012	(.003)	***	.017	(.006)	**	.007	(.004)	†

Notes: Results are combined estimates from 100 multiple imputation datasets. Standard errors are based on 2000 bootstrap samples.
†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$ for two-sided tests of no effect.

Results: effects of neighborhood disadvantage during adolescence, blacks



Results: effects of neighborhood disadvantage during adolescence, nonblacks



Substantive Conclusions

- Negative effect of neighborhood disadvantage is moderated by family poverty
  - Impact much more severe for children in families at or below poverty level
- Adolescent exposure to neighborhood disadvantage is much more consequential than childhood exposure
- Studies of neighborhood effects must investigate temporal dependency and subgroup heterogeneity
- Growth in income inequality and income segregation mutually reinforcing

More formally: counterfactual model

- As before, causal effects are defined as differences in potential outcomes
- $Y(a_1,a_2)$ is subject's HS graduation outcome had she been exposed to sequence of NHs (a_1,a_2) – note only two time periods
  - 25 potential education outcomes
- $L_2(a_1)$ is a subject's family income-to-needs ratio in adolescence had she been exposed to NH context (a_1) during childhood
  - 5 potential income-to-needs outcomes
  - Reflects dynamic NH selection process (L_2 is a function of a_1)

A Structural Nested Mean Model (Robins 1994, 1999)

- Decompose the conditional expectation of Y into five components (following Almirall et al. 2010):

$$E(Y(a_1, a_2)|L_1, L_2(a_1)) \\ = \beta_0 + \varepsilon_1(L_1) + u_1(L_1, a_1) + \varepsilon_2(L_1, a_1, L_2(a_1)) + u_2(L_2(a_1), a_2)$$

$\beta_0 = E\big(Y(1,1)\big)$: intercept

$u_1(L_1, a_1)$ and $u_2(L_2(a_1), a_2)$: causal functions of interest (capture association between treatment and outcome)

$\varepsilon_1(L_1)$ and $\varepsilon_2(L_1, a_1, L_2(a_1))$: “nuisance” functions (capture association between moderators and outcome)

- Note on Terminology:
  - “structural” refers to causal
  - “nested mean” refers to decomposition of overall mean into component parts

Causal Functions

- Moderated effect of neighborhood disadvantage in childhood (main effect and interaction) holding adolescent treatment constant

$$u_1(L_1, a_1) = E(Y(a_1, 1) - Y(1,1)|L_1) = (a_1 - 1)(\beta_1 + \beta_2 L_1)$$

Average causal effect of exposure to treatment sequence (a_1, 1) versus (1,1) within levels of L_1

- Moderated effect of neighborhood disadvantage in adolescence (main effect and interaction) holding childhood treatment constant

$$u_2(L_2(a_1), a_2) = E(Y(a_1, a_2) - Y(a_1, 1)|L_2(a_1)) = (a_2 - 1)(\beta_3 + \beta_4 L_2(a_1))$$

Average causal effect of exposure to treatment sequence (a_1, a_2) versus (a_1, 1) within levels of $L_2(a_1)$

- Note these are linear parametric functions

Nuisance Functions

- Capture the association between moderators and outcome

$$\varepsilon_1(L_1) = E(Y(1,1)|L_1) - E\big(Y(1,1)\big)$$
$$= \eta_1\ (L_1 - E(L_1))$$

$$\varepsilon_2(L_1, a_1, L_2(a_1)) = E\big(Y(a_1, 1)|L_1, L_2(a_1)\big) - E(Y(a_1, 1)|L_1)$$
$$= \eta_2(L_2(a_1) - E(L_2(a_1)|L_1))$$

- Notice that these are residuals of L at each time point
- Must have mean zero if we want the decomposition to work

Putting it all back together...

- Original decomposition equation:

$$E(Y(a_1, a_2)|L_1, L_2(a_1)) \\ = \beta_0 + \varepsilon_1(L_1) + u_1(L_1, a_1) + \varepsilon_2(L_1, a_1, L_2(a_1)) \\ + u_2(L_2(a_1), a_2)$$

- Residualized time-varying covariates:

$$L_1^r = L_1 - E(L_1)$$
$$L_2^r = L_2 - E(L_2|L_1, A_1)$$

- Estimated model:

$$Y = \beta_0 + \eta_1 L_1^r + (A_1 - 1)(\beta_1 + \beta_2 L_1) + \eta_2 L_2^r \\ + (A_2 - 1)(\beta_3 + \beta_4 L_2) + e$$

Identification

- Assume sequential ignorability of treatment (aka no unobserved confounding)

$$Y(a_1, a_2) \perp A_1 | L_1 \text{ and } Y(a_1, a_2) \perp A_2 | L_1, A_1, L_2$$

- If sequential ignorability holds, $u_1(L_1, a_1)$ and $u_2(L_2(a_1), a_2)$ can be identified from observed data

- Goal is to estimate $u_1(L_1, a_1)$ and $u_2(L_2(a_1), a_2)$

A few concluding thoughts

- Once you start thinking in terms of time-varying treatments and time-varying confounding, many longitudinal analysis problems can be understood in this way
- Mechanics of both methods are relatively easy to implement
- Assumptions are important, but fewer than conventional methods, and testable to some degree
- Be careful of poorly defined estimands
- Be careful of estimands that can't be identified

Be careful of poorly defined estimands

- Example: What is the effect of continuously living in the most disadvantaged quintile of neighborhoods, rather than the least disadvantaged quintile, among subjects whose families stay poor throughout the study?
  - $E(Y(5,5) - Y(1,1) | L_1 = 0, L_2 = 0))$
  - Requires comparison of those whose families would have stayed poor had they experienced the most disadvantaged neighborhoods with those who would have stayed poor had they experienced the least disadvantaged neighborhoods
  - Not a proper counterfactual: Compares two different groups of people, not counterfactuals of same people

Be careful of estimands that can't be identified without further assumptions

- Example: What is the effect of continuously living in the most disadvantaged quintile of neighborhoods, rather than the least disadvantaged quintile, *among subjects whose families would stay poor regardless of treatment received*?
  - $E(Y(5,5) - Y(1,1) | L_1 = 0, L_2(5) = L_2(1) = 0))$
  - Cannot be identified: we cannot tell who would have stayed poor regardless of neighborhood disadvantage
  - Not substantively interesting: involves an unobserved subpopulation for whom one the hypothesized mechanisms does not operate by definition

Some Key References: IPTW/MSMs

Brumback, B. B., E. D. Bouldin, H. W. Zheng, M. B. Cannell, and E. M. Andresen. 2010. “Testing and Estimating Model-Adjusted Effect-Measure Modification Using Marginal Structural Models and Complex Survey Data.” *American Journal of Epidemiology* 172(9): 1085–1091

Cole, S. R. and M. A. Hernan. 2008. "Constructing Inverse Probability of Treatment Weights for Marginal Structural Models." *American Journal of Epidemiology* 168:656-664.

Greenland, S. 2003. "Quantifying biases in causal models: Classical confounding vs collider-stratification bias." *Epidemiology* 14:300-306.

Hernan, M. A., B. A. Brumback, and J. M. Robins. 2002. "Estimating the Causal Effect of Zudovudine on CD4 Count with a Marginal Structural Model for Repeated Measures." *Statistics in Medicine* 21:1689-1709.

Robins, J. M. 1999. "Association, Causation, and Marginal Structural Models." *Synthese* 121:151-179.

Robins, J. M., M. A. Hernan, and B. Brumback. 2000. "Marginal Structural Models and Causal Inference in Epidemiology." *Epidemiology* 11:550-560.

Robins, J. M., A. Rotnitzky, and D. Scharfstein. 1999. "Sensitivity Analysis for Selection Bias and Unmeasured Confounding in Missing Data and Causal Inference Models." Pp. 1-94 in *Statistical Models in Epidemiology*, edited by E. Halloran. New York: Springer-Verlag.

Sociological Examples of IPTW/MSMs

Barber, J. S., S. A. Murphy, and N. Verbitsky. 2004. "Adjusting for Time-Varying Confounding in Survival Analysis." *Sociological Methodology* 34:163-192.

Hong, Guanglei and Stephen W. Raudenbush. 2008. “Causal Inference for Time-Varying Instructional Treatments.” *Journal of Educational and Behavioral Statistics* 81:333-362.

Sampson, R. J., J. H. Laub, and C. Wimer. 2006. "Does Marriage Reduce Crime? A Counterfactual Approach to Within-Individual Causal Effects." *Criminology* 44:465-508.

Sampson, R. J., P. Sharkey, and S. W. Raudenbush. 2008. "Durable Effects of Concentrated Disadvantage on Verbal Ability among African-American Children." *Proceedings of the National Academy of Sciences* 105:845-852.

Sharkey, Patrick and Felix Elwert. 2011. "The Legacy of Disadvantage: Multigenerational Neighborhood Effects on Cognitive Ability." *American Journal of Sociology* 116:1934-81.

Wodtke, Geoffrey T. Forthcoming, October 2013. “Duration and Timing of Exposure to Neighborhood Poverty and the Risk of Adolescent Parenthood.” *Demography*

Some Key References: RR/SNM

Almirall, Daniel, Cynthia J. Coffman, William S. Yancy, and Susan A. Murphy. 2010. "Structurnal Nested Models." Pp. 231-61 in *Analysis of Observational Health Care Data Using SAS*, edited by D. Faries, A. C. Leon, J. M. Haro, and R. L. Obenchain. Cary, NC: SAS Institute.

Almirall, Daniel, Daniel F. McCaffrey, Rajeev Ramchand, and Susan A. Murphy. 2011. "Subgroups Analysis when Treatment and Moderators are Time-varying." *Prevention Science* (electronic publication ahead of print): http://dx.doi.org/10.1007/s11121-011-0208-7.

Almirall, Daniel, Thomas Ten Have, and Susan A. Murphy. 2010. "Structural Nested Mean Models for Assessing Time-Varying Effect Moderation." *Biometrics* 66:131-9.

Robins, James M. 1987. "A New Approach to Causal Inference in Mortality Studies with a Sustained Exposure Period--Application to Control of the Healthy Worker Survivor Effect." *Mathematical Modeling* 7:1393-512.

Robins, James M. 1994. "Correcting for Noncompliance in Randomized Trials Using Structural Nested Mean Models." *Communications in Statistics-Theory and Methods* 23:2379-412.

Robins, James M. 1999b. "Marginal Structural Models versus Structural Nested Models as Tools for Causal Inference." Pp. 95-134 in *Statistical Models in Epidemiology*, edited by E. Halloran. New York: Springer-Verlag.

Supplemental Slides
Main Effects Analysis

Sample Attrition and Censoring Weights

- Let C_k be a binary variable equal to 1 if a child drops out of the study at wave k and 0 otherwise
- Estimate logit models predicting C_k
- Generate predicted probabilities
- Stabilized weight that adjusts for nonrandom attrition based on observed covariates:

$$cw_i = \prod_{k=1}^{K} \frac{P\left(C_k = 0 \middle| \bar{C}_{k-1} = 0,\ \bar{A}_{k-1} = \bar{a}_{(k-1)i}, L_0 = l_0\ \right)}{P\left(C_k = 0 \middle| \bar{C}_{k-1} = 0, \bar{A}_{k-1} = \bar{a}_{(k-1)i},\ \bar{L}_k = \bar{l}_{ki}\right)}$$

Sample Characteristics

Table 1. Time-Invariant Sample Characteristics

Variable	Blacks (n = 834)	Nonblacks (n = 1,259)
High school graduation, percent		
Did not graduate high school	23.38	11.60
Graduated high school	76.62	88.40
Gender, percent		
Male	52.04	50.99
Female	47.96	49.01
Birth weight, percent		
5.5 lbs or more	90.77	94.52
Less than 5.5 lbs	9.23	5.48
Mother's marital status at birth, percent		
Unmarried	41.97	5.56
Married	58.03	94.44
FU head's education, percent		
Less than high school	55.28	22.72
High school graduate	26.26	26.29
At least some college	18.46	50.99
Mother's age at birth, mean	24.34	26.12

Note: FU = family unit. Statistics reported for children not lost to follow-up before age 20 (first imputation dataset).

Table 2. Time-Dependent Sample Characteristics

Variable	Blacks (*n* = 834)			Nonblacks (*n* = 1,259)		
	Age 1	Age 10	Age 17	Age 1	Age 10	Age 17
NH disadvantage index, percent						
1st quintile	3.48	3.60	3.48	13.34	19.14	20.65
2nd quintile	3.24	3.72	6.00	19.46	18.67	21.84
3rd quintile	5.28	5.88	7.79	26.13	23.27	22.48
4th quintile	14.87	18.11	18.47	26.13	23.99	21.13
5th quintile	73.14	68.71	64.27	14.93	14.93	13.90
FU head's marital status, percent						
Unmarried	33.93	44.84	52.04	5.88	11.36	15.09
Married	66.07	55.16	47.96	94.12	88.64	84.91
FU head's employment status, percent						
Unemployed	27.22	32.61	33.09	8.10	8.02	9.69
Employed	72.78	67.39	66.91	91.90	91.98	90.31
Public assistance receipt, percent						
Did not receive AFDC	81.06	75.66	82.37	96.27	96.19	97.93
Received AFDC	18.94	24.34	17.63	3.73	3.81	2.07
Homeownership, percent						
Do not own home	69.66	53.48	50.12	40.19	22.32	20.73
Own home	30.34	46.52	49.88	59.81	77.68	79.27
FU income in $1,000s, mean	19.68	25.04	27.45	32.59	46.65	57.50
FU head's work hours, mean	30.08	26.82	27.51	42.65	40.84	40.68
FU size, mean	5.75	5.32	4.81	4.22	4.69	4.33
Cum. residential moves, mean	.32	2.48	3.64	.32	2.16	3.02

Note: NH = neighborhood; FU = family unit. Statistics reported for children not lost to follow-up before age 20 (first imputation dataset).

Neighborhood Mobility

Table 3. Exposure to Neighborhood Disadvantage from Age 2 to 17 Years

Variable	Blacks (n = 834)	Nonblacks (n = 1,259)
Duration-weighted exposure to NH disadvantage, percent		
1.0 to 1.4 (least disadvantaged NHs)	.84	12.31
1.5 to 2.4	2.64	20.57
2.5 to 3.4	6.24	30.26
3.5 to 4.4	24.82	28.12
4.5 to 5.0 (most disadvantaged NHs)	65.47	8.74
Number of moves between exposure levels, percent		
0	37.53	16.52
1	12.83	22.40
2	19.78	16.84
3+	29.86	44.24

Note: NH = neighborhood. Statistics reported for children not lost to follow-up before age 20 (first imputation dataset). NH disadvantage quintiles are based on distribution of the NH disadvantage index across all U.S. census tracts between 1970 and 2000.

Neighborhood Disadvantage Index

Table S1. Component Weights and Correlations from Principal Component Analysis (PCA) of Neighborhood Characteristics, U.S. Census Data 1970 to 2000

	1st PC	
Variable	Weight	Corr
Percent poverty	.408	.861
Percent unemployed	.371	.783
Percent receiving welfare	.412	.868
Percent female-headed households	.337	.711
Percent without high school diploma	.378	.798
Percent college graduates	–.348	–.735
Percent mgr/prof workers	–.385	–.812
Component variance	4.449	
Proportion total variance explained	.636	

Note: PCA is based on the correlation matrix. Analysis includes all tract-year observations between 1970 and 2000.

Table S2. Neighborhood (NH) Characteristics by Disadvantage Index Quintiles, U.S. Census Data 1970 to 2000

	NH Disadvantage Index									
	1st Quintile		2nd Quintile		3rd Quintile		4th Quintile		5th Quintile	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Percent poverty	4.30	3.32	6.44	3.96	9.14	4.49	13.68	5.29	28.31	11.65
Percent unemployed	3.16	1.33	4.27	1.70	5.30	2.12	6.91	2.72	12.12	6.23
Percent receiving welfare	2.31	1.40	3.71	1.87	5.21	2.41	8.01	3.26	18.98	9.73
Percent female-headed households	11.70	6.64	14.22	7.56	16.05	8.57	20.20	10.15	37.89	17.99
Percent without high school diploma	10.55	5.46	20.54	7.48	29.02	9.93	37.99	12.25	50.53	14.17
Percent college graduates	39.74	13.00	20.83	7.96	13.69	6.69	10.05	5.86	6.93	4.86
Percent mgr/prof workers	30.67	6.83	19.47	4.34	14.16	3.88	10.98	3.55	7.60	3.38

Models of Treatment (Neighborhood), Analysis 1

Table S3. Models of Neighborhood (NH) Exposure Status

	Blacks (person-years = 20,953)				Nonblacks (person-years = 28,084)			
	Model 1		Model 2		Model 1		Model 2	
Covariate	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Time-invariant characteristics								
Gender								
Male	ref	ref	ref	ref	ref	ref	ref	ref
Female	–.030	(.050)	–.039	(.051)	.044	(.029)	.042	(.029)
Birth weight								
≥5.5 lbs	ref	ref	ref	ref	ref	ref	ref	ref
<5.5 lbs	–.067	(.090)	–.068	(.091)	.060	(.072)	.084	(.072)
Mother's marital status at birth								
Unmarried	ref	ref	ref	ref	ref	ref	ref	ref
Married	–.239	(.070)	–.202	(.069)	–.129	(.065)	–.083	(.066)
Mother's age at birth (years)	0.019	(.006)	.017	(.006)	–.010	(.004)	–.012	(.004)
Year born								
1968 to 1969	ref	ref	ref	ref	ref	ref	ref	ref
1970 to 1972	–.134	(.076)	–.144	(.077)	.012	(.048)	.023	(.048)
1973 to 1975	–.080	(.083)	–.117	(.085)	.026	(.047)	.020	(.046)
1976 to 1978	–.153	(.076)	–.194	(.077)	.026	(.046)	.015	(.047)
Time								
Wave 2 to 6	ref	ref	ref	ref	ref	ref	ref	ref
Wave 7 to 11	–.128	(.042)	–.060	(.046)	–.003	(.023)	.042	(.027)
Wave 12 to 17	–.205	(.042)	–.067	(.054)	–.064	(.023)	.019	(.030)

(continued on next page)

Table S3. *continued*

Covariate	Blacks (person-years = 20,953)				Nonblacks (person-years = 28,084)			
	Model 1		Model 2		Model 1		Model 2	
	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Time-dependent characteristics measured at baseline ($k = 0$)								
Ordinal NH disadvantage	.129	(.030)	.141	(.028)	.149	(.022)	.146	(.022)
FU head's education								
Less than high school	.184	(.068)	.145	(.070)	.248	(.041)	.185	(.041)
High school graduate	.122	(.066)	.120	(.066)	.154	(.035)	.120	(.036)
At least some college	ref	ref	ref	ref	ref	ref	ref	ref
FU head's marital status								
Unmarried	ref	ref	ref	ref	ref	ref	ref	ref
Married	.139	(.078)	.214	(.081)	.246	(.079)	.248	(.081)
FU head's employment status								
Unemployed	ref	ref	ref	ref	ref	ref	ref	ref
Employed	–.047	(.094)	–.039	(.098)	.003	(.065)	.032	(.065)
Home ownership								
Do not own home	ref	ref	ref	ref	ref	ref	ref	ref
Own home	–.106	(.054)	–.066	(.060)	–.077	(.033)	–.080	(.035)
Public assistance receipt in past year								
Did not receive AFDC	ref	ref	ref	ref	ref	ref	ref	ref
Received AFDC	.403	(.094)	.330	(.096)	–.046	(.089)	–.073	(.090)
FU income in past year (log $)	–.005	(.044)	.051	(.046)	–.163	(.056)	–.093	(.047)
FU head's work hours in past year (hrs)	–.005	(.002)	–.003	(.002)	–.003	(.001)	–.003	(.001)
FU size	.001	(.010)	–.013	(.012)	.074	(.012)	.066	(.015)

(continued on next page)

Table S3. *continued*

	Blacks (person-years = 20,953)				Nonblacks (person-years = 28,084)			
	Model 1		Model 2		Model 1		Model 2	
Variable	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Time-dependent characteristics measured at wave $k-1$								
Ordinal NH disadvantage	2.193	(.084)	2.150	(.086)	2.113	(.050)	2.096	(.050)
FU head's marital status								
Unmarried			ref	ref			ref	ref
Married			–.147	(.127)			.067	(.116)
FU head's employment status								
Unemployed			ref	ref			ref	ref
Employed			.074	(.100)			–.113	(.103)
Public assistance receipt in past year								
Did not receive AFDC			ref	ref			ref	ref
Received AFDC			.106	(.080)			–.073	(.105)
Home ownership								
Do not own home			ref	ref			ref	ref
Own home			.103	(.087)			.044	(.086)
FU income in past year (log $)			.013	(.041)			–.032	(.024)
FU head's work hours in past year (hrs)			–.001	(.002)			.002	(.001)
FU size			–.029	(.028)			–.012	(.034)

(continued on next page)

Table S3. *continued*

	Blacks (person-years = 20,953)				Nonblacks (person-years = 28,084)			
	Model 1		Model 2		Model 1		Model 2	
Variable	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Time-dependent characteristics measured at wave *k*								
FU head's marital status								
Unmarried			ref	ref			ref	ref
Married			–.324	(.133)			.134	(.136)
FU head's employment status								
Unemployed			ref	ref			ref	ref
Employed			–.046	(.109)			–.367	(.097)
Public assistance receipt in past year								
Did not receive AFDC			ref	ref			ref	ref
Received AFDC			.141	(.083)			.066	(.105)
Home ownership								
Do not own home			ref	ref			ref	ref
Own home			–.254	(.096)			–.116	(.086)
FU income in past year (log $)			–.149	(.047)			–.165	(.031)
FU head's work hours in past year (hrs)			–.002	(.002)			–.002	(.001)
FU size			.026	(.029)			.018	(.033)
Cumulative residential moves			–.056	(.012)			–.023	(.008)
Interactions								
FU head's marital status (wave k x k – 1)			.375	(.182)			–.125	(.142)
FU head's employment status (wave k x k – 1)			–.104	(.121)			.270	(.108)

Note: Models 1 and 2 are ordinal logistic regressions for the numerator and denominator of the stabilized IPT weight. Analyses are based on all person-years contributed by children present at age 1 in a PSID core family unit between 1968 and 1978. Coefficients and standard errors are combined estimates from five multiple imputation datasets.

Model Specification Tests (Analysis 1)

Table S4. Stabilized IPT-Weighted Estimates by Different Specifications of the Treatment Model

		Blacks (*n* = 834)				Nonblacks (*n* = 1,259)			
		Weights		Effect Estimates		Weights		Effect Estimates	
Model	Description	Mean	SD	Coef	SE	Mean	SD	Coef	SE
A	reported treatment model	1.04	.61	–.525	(.190)	1.00	.37	–.274	(.128)
B	(A) – birth weight, female	1.04	.61	–.537	(.192)	1.00	.37	–.268	(.128)
C	(B) – marital and employment interactions	1.04	.61	–.522	(.192)	1.00	.36	–.284	(.129)
D	(C) – covariates at wave *k* – 1	1.04	.61	–.513	(.190)	1.00	.35	–.280	(.128)
E	(A) + income x family size interaction	1.03	.61	–.507	(.193)	1.00	.38	–.288	(.128)
F	(E) + income x homeowner interaction	1.03	.61	–.512	(.195)	1.00	.39	–.282	(.129)
G	(A) + income x time interaction	1.04	.64	–.536	(.191)	1.00	.39	–.284	(.128)
H	(G) + marital status x time interaction	1.04	.64	–.535	(.192)	1.00	.39	–.279	(.127)
I	(A) + quad. terms for income, family size	1.04	.65	–.540	(.203)	1.00	.42	–.294	(.130)
J	(I) + quad. terms for work hours, cum moves	1.04	.65	–.566	(.202)	1.00	.44	–.282	(.131)
K	(J) + income interaction (wave *k* x *k* – 1)	1.04	.65	–.562	(.201)	1.00	.43	–.279	(.130)
L	(K) + family size interaction (wave *k* x *k* – 1)	1.04	.67	–.579	(.201)	1.00	.43	–.280	(.131)
M	(L) + AFDC receipt interaction (wave *k* x *k* – 1)	1.04	.66	–.588	(.202)	1.01	.44	–.280	(.131)
N	(M) + homeowner interaction (wave k x *k* – 1)	1.05	.70	–.523	(.209)	1.01	.45	–.280	(.130)

Note: Analyses based on children who were not lost to follow-up before age 20. Weights truncated at 1st and 99th percentiles. Coefficients and standard errors are combined estimates from five multiple imputation datasets.

Positivity Checks (Analysis 1)

Table S5. Treatment Distribution by Selected Covariates

			NH Disadvantage Quintiles				
Family Income	Homeownership	Marital Status	1	2	3	4	5
0 to 15K	Not homeowner	Unmarried	.03	.03	.05	.13	.77
		Married	.02	.05	.07	.17	.69
	Homeowner	Unmarried	.02	.05	.07	.18	.68
		Married	.05	.07	.15	.26	.48
15K to 30K	Not homeowner	Unmarried	.07	.08	.11	.15	.59
		Married	.03	.09	.13	.24	.51
	Homeowner	Unmarried	.04	.08	.11	.19	.57
		Married	.06	.09	.19	.30	.36
30K to 45K	Not homeowner	Unmarried	.13	.10	.16	.17	.44
		Married	.07	.13	.17	.26	.38
	Homeowner	Unmarried	.07	.15	.21	.20	.37
		Married	.09	.15	.25	.26	.24
45K to 60K	Not homeowner	Unmarried	.19	.15	.16	.18	.32
		Married	.12	.20	.18	.20	.29
	Homeowner	Unmarried	.18	.26	.23	.14	.18
		Married	.16	.22	.23	.20	.19
60K+	Not homeowner	Unmarried	.09	.23	.28	.12	.28
		Married	.20	.24	.20	.16	.19
	Homeowner	Unmarried	.40	.21	.10	.07	.22
		Married	.38	.22	.16	.14	.10

Note: Analyses are based on all person-year observations contributed by children present at age 1 in a PSID core family unit between 1968 and 1978. Cells contain the row proportion of person-years exposed to different quintiles of the neighborhood disadvantage index.

Table S6. Treatment Distribution by Selected Covariates

Cumulative Moves	Education	Employment Status	NH Disadvantage Quintiles 1	2	3	4	5
0	Less than HS	Unemployed	.01	.02	.05	.14	.77
		Employed	.04	.05	.14	.23	.55
	HS graduate	Unemployed	.01	.04	.09	.10	.76
		Employed	.05	.12	.21	.30	.32
	Some college	Unemployed	.11	.14	.16	.25	.34
		Employed	.22	.22	.20	.19	.17
1	Less than HS	Unemployed	.02	.05	.06	.11	.75
		Employed	.04	.07	.13	.26	.51
	HS graduate	Unemployed	.02	.07	.16	.17	.58
		Employed	.09	.15	.21	.24	.32
	Some college	Unemployed	.09	.11	.13	.21	.47
		Employed	.25	.20	.18	.19	.18
2	Less than HS	Unemployed	.02	.03	.04	.14	.78
		Employed	.03	.05	.14	.24	.53
	HS graduate	Unemployed	.03	.07	.08	.15	.67
		Employed	.08	.14	.20	.24	.34
	Some college	Unemployed	.08	.08	.09	.19	.56
		Employed	.24	.22	.23	.16	.15
3+	Less than HS	Unemployed	.02	.03	.06	.19	.70
		Employed	.04	.10	.14	.24	.47
	HS graduate	Unemployed	.02	.06	.09	.19	.64
		Employed	.13	.15	.18	.22	.32
	Some college	Unemployed	.06	.09	.09	.18	.59
		Employed	.20	.19	.24	.19	.19

Note: Analyses are based on all person-year observations contributed by children present at age 1 in a PSID core family unit between 1968 and 1978. Cells contain the row proportion of person-years exposed to different quintiles of the neighborhood disadvantage index.

Supplemental Slides
Conditional Effects Analysis

Table 1. Time-invariant sample characteristics

Variable	Total			Blacks		Nonblacks	
	% miss	mean	sd	mean	sd	mean	sd
R - high school graduate	43.0	.80	(.40)	.75	(.44)	.85	(.36)
R - female	0.0	.48	(.50)	.49	(.50)	.48	(.50)
M - age at childbirth	23.4	24.79	(5.56)	23.78	(5.62)	25.70	(5.35)
M - married at childbirth	25.8	.71	(.45)	.50	(.50)	.90	(.30)
H - high school graduate	2.9	.24	(.43)	.25	(.43)	.24	(.43)
H - some college	2.9	.35	(.48)	.22	(.41)	.48	(.50)

Notes: Results are combined estimates from 100 multiple imputation datasets. R, M and H indicate respondent, mother of respondent and household head, respectively.

Table 2. Time-varying sample characteristics

Variable	Total			Blacks		Nonblacks	
	% miss	mean	sd	mean	sd	mean	sd
Childhood							
H - married	0.0	.73	(.40)	.58	(.45)	.87	(.29)
H - employed	0.0	.79	(.35)	.67	(.40)	.90	(.24)
FU - owns home	0.0	.46	(.45)	.30	(.41)	.61	(.44)
FU - size	0.0	4.85	(1.78)	5.23	(2.06)	4.51	(1.38)
FU - number of moves	13.1	1.15	(1.13)	1.20	(1.12)	1.11	(1.14)
FU - inc-to-needs ratio	0.0	.89	(1.22)	.35	(.92)	1.37	(1.26)
Adolescence							
H - married	23.8	.67	(.44)	.49	(.47)	.82	(.34)
H - employed	23.8	.78	(.37)	.65	(.42)	.89	(.25)
FU - owns home	23.8	.57	(.46)	.40	(.46)	.72	(.41)
FU - size	23.8	4.86	(1.57)	5.09	(1.83)	4.65	(1.25)
FU - number of moves	29.8	.76	(1.01)	.83	(1.03)	.69	(.98)
FU - inc-to-needs ratio	23.8	1.28	(1.66)	.55	(1.14)	1.95	(1.76)

Notes: Results are combined estimates from 100 multiple imputation datasets. R, M and H indicate respondent, mother of respondent and household head, respectively.

Table 3. Joint treatment distribution

n / row / cell	NH disadvantage quintile - childhood	Blacks: NH disadvantage quintile - adolescence 1	2	3	4	5	Nonblacks: NH disadvantage quintile - adolescence 1	2	3	4	5
n	1	38	11	6	8	5	358	49	23	15	3
row	1	.56	.16	.09	.12	.07	.80	.11	.05	.03	.01
cell	1	.01	.00	.00	.00	.00	.11	.02	.01	.00	.00
n	2	19	26	28	12	6	169	279	87	31	6
row	2	.21	.29	.31	.13	.07	.30	.49	.15	.05	.01
cell	2	.01	.01	.01	.00	.00	.05	.09	.03	.01	.00
n	3	20	37	62	39	38	48	245	356	107	34
row	3	.10	.19	.32	.20	.19	.06	.31	.45	.14	.04
cell	3	.01	.01	.02	.01	.01	.01	.08	.11	.03	.01
n	4	15	24	75	180	152	34	61	229	425	130
row	4	.03	.05	.17	.40	.34	.04	.07	.26	.48	.15
cell	4	.01	.01	.03	.06	.05	.01	.02	.07	.13	.04
n	5	14	33	76	239	1738	8	13	49	144	331
row	5	.01	.02	.04	.11	.83	.01	.02	.09	.26	.61
cell	5	.00	.01	.03	.08	.60	.00	.00	.02	.04	.10

Notes: Results based on first imputation dataset.

Table 4. Effects of neighborhood disadvantage on high school graduation (two-stage estimates)

Model	Total			Blacks			Nonblacks		
	coef	se		coef	se		coef	se	
Intercept	.888	(.021)	***	.916	(.044)	***	.877	(.019)	***
Childhood									
NH dadvg	–.005	(.012)		–.004	(.019)		–.006	(.015)	
NH dadvg x inc-to-needs	.005	(.004)		.005	(.008)		.005	(.005)	
Adolesence									
NH dadvg	–.042	(.010)	***	–.054	(.018)	**	–.026	(.013)	†
NH dadvg x inc-to-needs	.012	(.003)	***	.017	(.006)	**	.007	(.004)	†

Notes: Results are combined estimates from 100 multiple imputation datasets. Standard errors are based on 2000 bootstrap samples.
†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$ for two-sided tests of no effect.

Table A.1 Principal component weights and correlations

Variable	1st PC	
	Weight	Corr
Percent poverty	.408	.861
Percent unemployed	.371	.783
Percent receiving welfare	.412	.868
Percent female-headed households	.337	.711
Percent without high school diploma	.378	.798
Percent college graduates	–.348	–.735
Percent mgr/prof workers	–.385	–.812
Component variance	4.449	
Proportion total variance explained	.636	

Notes: Principal component analysis based on correlation matrix. Analysis includes all tract-year observations from the 1970 to 2000 U.S. censuses.

Figure A.1. Neighborhood socioeconomic characteristics by disadvantage index quintile

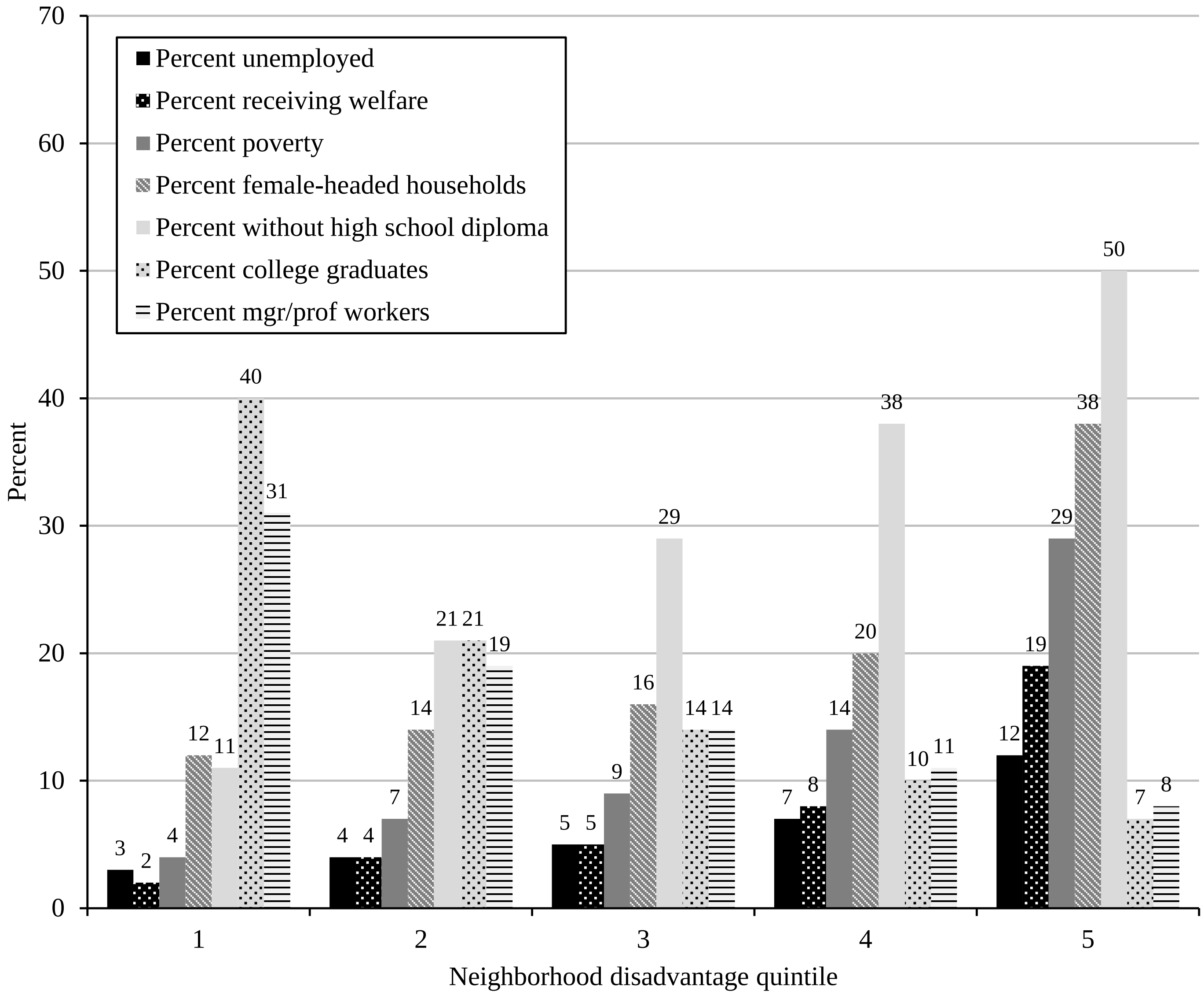


Table B.1. Two-stage estimates with different specifications of SNMM causal functions

Model	A (base)			B			C			D		
	coef	se		coef	se		coef	se		coef	se	
Intercept	.888	(.021)	***	.890	(.025)	***	.880	(.024)	***	.882	(.024)	***
Childhood												
NH dadvg	–.005	(.012)		–.006	(.014)		.022	(.028)		.027	(.034)	
NH dadvg x inc-to-needs	.005	(.004)		.005	(.004)		.007	(.005)		.008	(.005)	
NH dadvg x H-less than HS							–.010	(.021)		–.007	(.021)	
NH dadvg x H-some college							–.016	(.019)		–.012	(.020)	
NH dadvg x H-married							.005	(.018)		.003	(.019)	
NH dadvg x H-employed							–.034	(.027)		–.034	(.026)	
NH dadvg x H-homeowner							.010	(.013)		.003	(.014)	
NH dadvg x family size										.004	(.004)	
NH dadvg x num. moves										–.003	(.005)	
Adolesence												
NH dadvg	–.042	(.010)	***	–.043	(.016)	**	–.044	(.023)	†	–.047	(.030)	
NH dadvg x inc-to-needs	.012	(.003)	***	.012	(.003)	***	.011	(.007)	**	.010	(.004)	**
NH dadvg x H-less than HS							.003	(.018)		.001	(.018)	
NH dadvg x H-some college							.005	(.017)		.004	(.017)	
NH dadvg x H-married							.002	(.013)		.009	(.014)	
NH dadvg x H-employed							–.008	(.020)		–.009	(.020)	
NH dadvg x H-homeowner							.009	(.012)		.010	(.012)	
NH dadvg x family size										–.005	(.004)	
NH dadvg x num. moves										.000	(.005)	
Chld x Adl NH dadvg				.000	(.004)							

Notes: Results are combined estimates from 100 multiple imputation datasets. Standard errors are based on 2000 bootstrap samples.
†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$ for two-sided tests of no effect.

Table B.2. Two-stage estimates with different specifications of SNMM nuisance functions

Model	A (base)			B			C			D		
	coef	se		coef	se		coef	se		coef	se	
Intercept	.888	(.021)		.886	(.021)		.879	(.021)		.876	(.021)	
Childhood												
NH dadvg	–.005	(.012)		–.005	(.012)		.000	(.012)		–.005	(.012)	
NH dadvg x inc-to-needs	.005	(.004)		.006	(.004)		.002	(.004)		.005	(.004)	
Adolesence												
NH dadvg	–.042	(.010)	***	–.042	(.010)	***	–.041	(.010)	***	–.033	(.011)	**
NH dadvg x inc-to-needs	.012	(.003)	***	.013	(.003)	***	.012	(.003)	***	.007	(.003)	*
Description												
Num. of 2nd stage parameters	25			39			69			99		
Nuissance functions	main effects for V, L_1 and L_2			A + all two-way interactions btw elements of V			B + all two-way interactions btw V and L_1			C + all two-way interactions btw V and L_2		

Notes: Results are combined estimates from 100 multiple imputation datasets. Standard errors are based on 2000 bootstrap samples.
†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$ for two-sided tests of no effect.

Table B.3. Two-stage estimates with different specifications of SNMM nuisance functions continued

Model	E coef	E se		F coef	F se		G coef	G se	
Intercept	.882	(.021)		.883	(.021)		.882	(.021)	
Childhood									
NH dadvg	–.001	(.012)		–.006	(.012)		–.006	(.012)	
NH dadvg x inc-to-needs	.002	(.004)		.005	(.005)		.005	(.005)	
Adolesence									
NH dadvg	–.041	(.010)	***	–.037	(.011)	***	–.035	(.011)	**
NH dadvg x inc-to-needs	.012	(.003)	***	.009	(.003)	**	.008	(.004)	*
Description									
Num. of 2nd stage parameters	40			55			91		
Nuissance functions	A + all two-way interactions btw elements of L_1			E + all two-way interactions btw elements of L_2			F + all two-way interactions btw L_1 and L_2		

Notes: Results are combined estimates from 100 multiple imputation datasets. Standard errors are based on 2000 bootstrap samples.
†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$ for two-sided tests of no effect.

Table D.1. Two-stage estimates under different methods of adjusting for missing data/sample attrition

Model	MI (base)			MID			SI			CC		
	coef	se		coef	se		coef	se		coef	se	
Intercept	.888	(.021)	***	.906	(.018)	***	.896	(.014)	***	.915	(.019)	***
Childhood												
NH dadvg	–.005	(.012)		–.008	(.012)		.006	(.008)		–.004	(.013)	
NH dadvg x inc-to-needs	.005	(.004)		.007	(.004)		.001	(.003)		.007	(.005)	
Adolesence												
NH dadvg	–.042	(.010)	***	–.040	(.010)	***	–.055	(.007)	***	–.051	(.011)	***
NH dadvg x inc-to-needs	.012	(.003)	***	.011	(.003)	***	.016	(.002)	***	.014	(.004)	***
Description												
Num. of observations	6135			3500			6135			2626		
Num. of replications	100			100			1			0		

Notes: MI = multiple imputation, MID = multiple imputation then deletion, SI = single imputation, and CC = complete case analysis. Standard errors are based on 2000 bootstrap samples.
†*p* < 0.10, **p* < 0.05, ***p* < 0.01, and ****p* < 0.001 for two-sided tests of no effect.

Sensitivity to Unobserved Confounding

- Compute bias-adjusted effect estimates under various assumptions about unobserved confounding, separately by treatment period
- Generate a bias-adjusted outcome (Y^C) and re-run the outcome model (SNMM)
- Bias-adjusted outcome derived from hypothetical counterfactual outcomes
- Sensitivity parameter α calibrated to observed confounding (α < 0 implies upward bias in neighborhood effect)

Table C.1. Potential outcomes from hypothetical neighborhood experiment

Observed	Mean Potential Outcome, $E(Y(a)\|A)$				
Treatment	$E(Y(1)\|A)$	$E(Y(2)\|A)$	$E(Y(3)\|A)$	$E(Y(4)\|A)$	$E(Y(5)\|A)$
$A = 1$	E	(J)	(O)	(T)	(EE)
$A = 2$	(F)	K	(P)	(U)	(FF)
$A = 3$	(G)	(L)	Q	(V)	(GG)
$A = 4$	(H)	(M)	(R)	W	(HH)
$A = 5$	(I)	(N)	(S)	(X)	II

Figure C.1. Sensitivity of direct effect estimates for childhood neighborhood disadvantage to hypothetical unobserved confounding

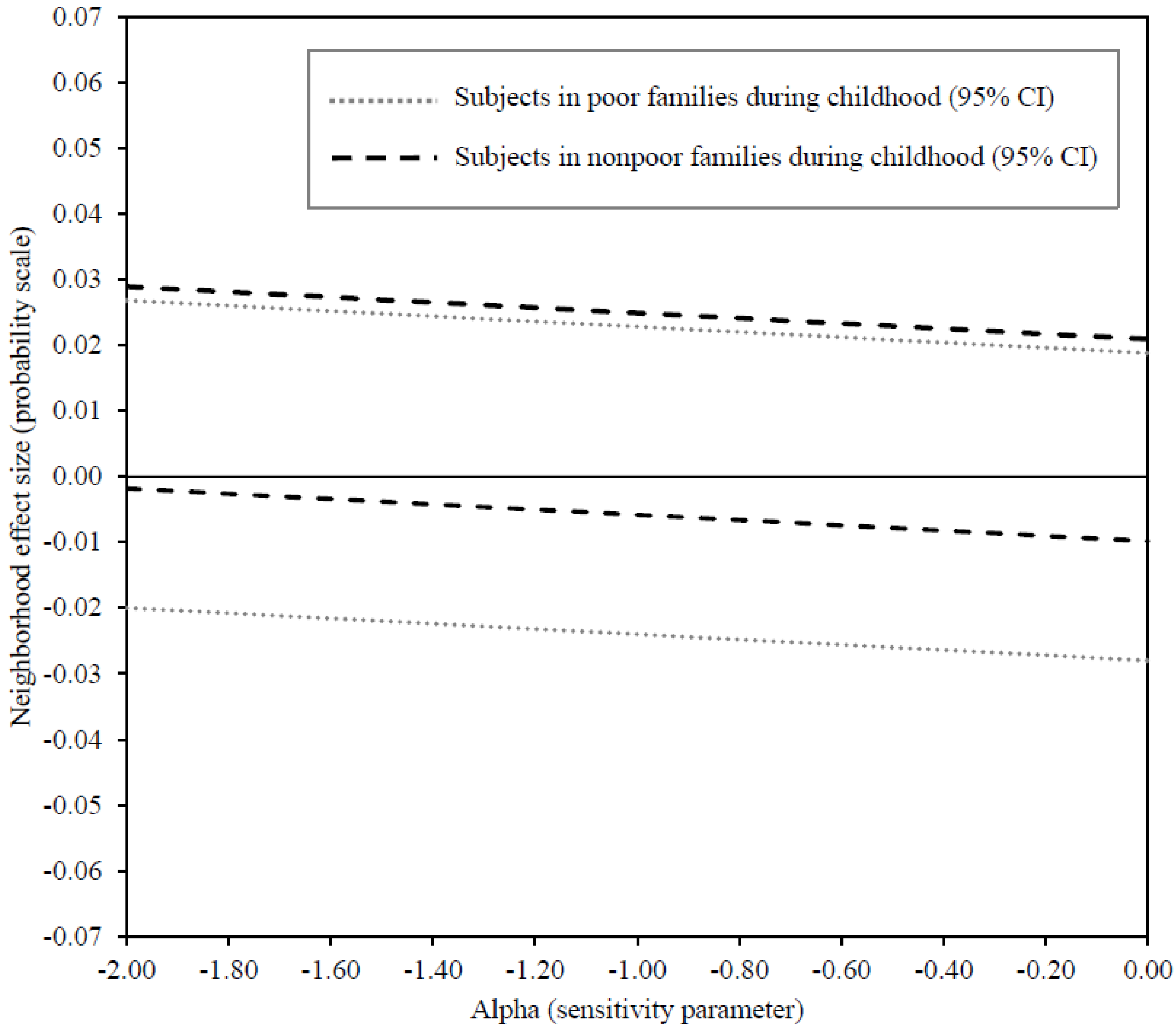


Figure C.2. Sensitivity of effect estimates for adolescent neighborhood disadvantage to hypothetical unobserved confounding