Social Science & Medicine 74 (2012) 625–636

Contents lists available at SciVerse ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Health disparities in mid-to-late life: The role of earlier life family and neighborhood socioeconomic conditions☆

Rucker C. Johnson [a,*], Robert F. Schoeni [b], Jeannette A. Rogowski [c]

[a] Goldman School of Public Policy, University of California, 2607 Hearst Ave., Berkeley, CA 94547, USA
[b] Institute for Social Research, University of Michigan, USA
[c] University of Medicine and Dentistry of New Jersey, USA

ARTICLE INFO

Article history:
Available online 14 December 2011

Keywords:
USA
Self-assessed health
Family
Neighborhood effects
Life course
Racial & socioeconomic disparities
Multi-level modeling

ABSTRACT

The relationship between neighborhoods of residence in young adulthood and health in mid-to-late life in the United States are examined using the 1968–2005 waves of the Panel Study of Income Dynamics (PSID). The sample consists of persons who were aged 20–30 in 1968 and are followed for a period of 38 years ($N = 2730$). Four-level hierarchical random effects models of self-assessed general health status as a function of individual, family, and neighborhood factors are estimated. Using the original sampling design of the PSID, we analyze adult health trajectories of married couples and neighbors followed from young adulthood through elderly ages to assess the magnitudes of the possible causal effects of family and neighborhood characteristics in young adulthood on health in mid-to-late life. Estimates suggest disparities in neighborhood conditions in young adulthood account for one-quarter of the variation in mid-to-late life health. Living in poor neighborhoods during young adulthood is strongly associated with negative health outcomes in later-life. This result is robust even in the presence of a reasonably large amount of potential unobservable individual and family factors that may significantly affect both neighborhood of residence and subsequent health status. Racial differences in health status in mid-to-late life are also associated with family and neighborhood socioeconomic conditions earlier in life. Three quarters of the black-white gap in health status at ages over 55 can be accounted for by differences in childhood socioeconomic status and neighborhood and family factors in young adulthood.

Introduction

Increasingly, research on the determinants of adult and old age health is recognizing the need to incorporate earlier life circumstances. The dearth of available longitudinal, nationally-representative US data with extensive information on socioeconomic status and health status has presented formidable challenges to understanding how early, middle, and late life factors influence the life cycle trajectory of health in this country.

We investigate the extent to which later-life health inequality in the United States is related to the different types of contexts into which individuals are born, and within which they grow up, live, and work at earlier stages of the life cycle. Unraveling the sources of racial and socioeconomic differences in adult health outcomes has eluded studies that have attempted to explain them using only individual-level factors and contemporaneous socioeconomic factors. Accordingly, this study uses data from the Panel Study of Income Dynamics (PSID), spanning nearly four decades, to investigate whether and how neighborhoods affect the aging process over the life course.

There is strong theoretical rationale supporting the hypothesis that neighborhood factors influence health, and many empirical studies have found strong associations between health and specific neighborhood variables (see Ellen, Mijanovich, & Dillman, 2001 and Kawachi & Berkman, 2003, for reviews). The critical and unresolved question is whether these health status differences by neighborhood quality can be causally attributed to neighborhood environment, or whether they are simply due to pre-existing differences between individuals living in different neighborhoods. Because differences between individuals and families in poor neighborhoods and those in affluent neighborhoods will bias estimates of

☆ This research was funded by the National Institute on Aging (R01AG024058). Funding for RAND's Center for Population Health and Health Disparities was provided through NIEHS P50 ES12383. Direct correspondence to Rucker C. Johnson, University of California, Berkeley, Goldman School of Public Policy, 2607 Hearst Ave, Berkeley, CA 94547, or email to ruckerj@berkeley.edu. We thank the PSID staff for access to the confidential restricted-use PSID geocode data.

* Corresponding author. Tel.: +1 510 643 0169; fax: +1 510 643 9657.
E-mail address: ruckerj@berkeley.edu (R.C. Johnson).

0277-9536/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.socscimed.2011.10.021

neighborhood effects, it is necessary to control for such differences. This problem has been well-documented and is variously referred to in different disciplines as selection bias, omitted variable bias, confounding, or unobserved heterogeneity (Manski, 1993; Oakes, 2004).

There is little evidence in the literature about how late life health is affected by neighborhoods of residence earlier in life in the United States. Evidence from international studies, however, demonstrates the importance of earlier life neighborhood conditions on health in late life in such countries as England and Canada (Curtis, Setia, & Quesnel-Vallee, 2009; Curtis, Southall, Congdon, & Dodgeon, 2004).

This paper exploits the unique features of the PSID to address these challenges. The analysis is carried out in three stages. We first bound the proportion of inequality in late life health that may be attributed to child socioeconomic status and disparities in neighborhood and family characteristics during early-to-mid adulthood, using an aggregated neighborhoods estimation strategy. Second, we attempt to identify the specific childhood and young adulthood family and neighborhood-level characteristics that influence later-life health. Sensitivity analyses are then used to probe the robustness of the results for causal inference. Specifically, we use a novel empirical approach, recently proposed by Altonji, Elder, and Taber (2005) and Krauth (2006), to test the robustness of the estimated effects of young adult neighborhood poverty to selection bias due to omitted variables. We assess how large the unobservable factors would need to be, relative to the observable factors included, to invalidate the results. Finally, in addition to addressing how neighborhoods of residence over the life course affect elderly health, our empirical analysis makes a further contribution by producing evidence on the role of neighborhood environments in contributing to socioeconomic and racial health disparities later in life.

Conceptual framework

This paper takes a life course perspective on health and hypothesizes that inequalities in health status are a product of the interaction between the developmental opportunities and vulnerabilities at each stage of the life course, on the one hand, with widely varying attributes of socioeconomic neighborhood and family conditions, on the other. In the life course approach, long-term health consequences of earlier life adverse socioeconomic conditions will occur through multiple pathways. While there is no single life course approach that is widely accepted, some of the pathways that have been postulated in the literature include such concepts as “latency”, “cumulative”, and “pathway” processes. For instance, *latency* refers to resultant effects of adverse conditions early in life that cause the body’s physiology and metabolism to be fundamentally altered, with some of the consequences of these changes manifesting in visible health problems much later in life (Barker, 1998). Alternately, the cumulative process explanation posits that persistent exposures to disadvantaged neighborhood and family conditions may have a cumulative toll in the form of “weathering” (Geronimus, 1996). This conceptual framework is consistent with a stress and adaptation perspective on how neighborhood conditions may influence health trajectories. Finally, *pathway* processes refer to linkages between health and socioeconomic attainment where childhood socioeconomic and family factors affect adult socioeconomic outcomes such as education, which in turn influences health later in life (Kuh & Wadsworth, 1993). An additional pathway process may arise because socioeconomic attainment in adulthood affects the quality of residential neighborhood and therefore health. Latent, cumulative, and pathway explanations for the linkages between health and socioeconomic status are not mutually exclusive and may produce synergistic influences.

We extend the SES-health literature and the literature on racial disparities in the United States by increased attention to both family- and neighborhood-level processes. It would be an atomistic fallacy to assume that individual-level processes operate the same at the neighborhood-level. We hypothesize that neighborhood SES conditions may influence health and well-being over and above the independent influence of individual and family-level SES factors. Living in a neighborhood with concentrated poverty may have consequences above and beyond those of growing up in a poor family because of social isolation, crime, weakened social institutions, unrelenting stress, inferior health care accessibility, and other factors. Corroborating evidence has shown that community-wide job losses induce community-wide impacts beyond the individuals that lose jobs; i.e., estimates reflect negative effects on workers and families who maintain employment but are impacted by their friends’ and neighbors’ loss of employment and the resulting changes to their communities (Ananat, Gassman-Pines, & Gibson-Davis, in press). Aggregate changes in family stress and socioeconomic resources may also help explain the effects of community economic conditions on subsequent health. While separately identifying how latent, cumulative, and other pathway effects may combine with current circumstances to impact later-life health is beyond the scope of this paper, the underlying framework informs the modeling approach taken in our empirical work which takes a life course perspective.

Methodological challenges in estimating neighborhood effects

The primary methodological challenge in estimating the causal effects of neighborhoods on health status is that unobserved factors that affect health may also be correlated with neighborhood factors, leading to biased estimates of neighborhood effects. The problem of endogeneity of residential location (in the form of “selection bias”) is a paramount concern because families choose where to live in part based on the characteristics they value (although factors such as racial discrimination and exclusionary zoning may constrain these decisions). Thus, families who care more about investing in health-promoting activities may be less likely to choose to reside in a community with a poor health care system or high pollution. Many of the multidimensional aspects that influence residential location decisions cannot be easily measured, which makes it particularly difficult to disentangle the causal influences of individual-, family-, and neighborhood-level factors.

Most health outcomes are likely to be a product of cumulative exposures to environments spanning decades or exhibit long latent periods before problems manifest. A problematic characteristic of many neighborhood studies is that they examine the relationship between contemporaneous health and neighborhood conditions without regard to changing circumstances during the life cycle or the persistence of exposure to neighborhood conditions. If exposure to neighborhood conditions at an earlier stage of the life cycle influences current health but is not accounted for, then the estimated relationship between contemporaneous neighborhoods and health may be misleading. Moreover, the typical methods used to address endogeneity in observational studies (e.g., instrumental variables and fixed effect approaches) (Wooldridge, 2002) have significant limitations in this context. Because most methods for overcoming endogenous residential location are based on small, short-run changes in the neighborhood environment, these approaches might be limited to uncovering effects only for rapidly-responding intermediate outcomes such as health behaviors (e.g.,

smoking/drinking, exercise/diet). An additional issue is that neighborhood variables change slowly over time (based on our own analysis and Kunz, Page, & Solon, 2003), implying that much of the measured year-to-year variation may be measurement error.

The most powerful way to address selection is through a randomized trial. But an experimental design where neighborhoods are randomly assigned is rare. A significant exception is the evaluation of the Move to Opportunity (MTO) program, where an experimental design is used to estimate the effects of offering housing assistance that allows individuals to move out of low-income, poor neighborhoods. Several papers demonstrate that MTO had beneficial effects on child and adult health (Katz, Kling, & Liebman, 2001; Leventhal & Brooks-Gunn, 2003).

Data

The PSID began interviewing a national probability sample of families in 1968 and re-interviewed them in each year through 1997, when interviewing became biennial. All individuals are followed until death or attrition, and data through the 2005 wave is utilized in the analysis. The PSID maintains extremely high wave-to-wave core re-interview response rates of 95–98%. Studies have concluded that the PSID remains representative of the U.S. population (Andreski, McGonagle, & Schoeni, 2009; Becketti, Gould, Lillard, & Welch, 1988; Fitzgerald, Gottschalk, & Moffitt, 1998a; 1998b; Lillard & Panis, 1998).

In 1968 the PSID used a cluster sample design which today provides the unusual opportunity to examine health outcomes for adults who were neighbors in 1968. Although the initial sample was clustered in 1968, PSID families were located in 2570 different census tracts in 2005 because families and individuals moved throughout the nation. The major advantage of the PSID for our purposes is that it offers the potential for constructing earlier life neighborhood effects, coupled with rich detail on early life and mid-life socioeconomic data, and data on dimensions of health status over time.

The selected sample consists of PSID original sample members born between 1928 and 1948, which consists of adults who were in their 20s and 30s in the first wave of interviewing in 1968. We then obtain all available information on these individuals for each wave, 1968–2005. Therefore, by 2005 the oldest person in the sample is 77 and the youngest is 57. Due to the complexity of the health status changes for women during the childbearing years, we exclude observations for women in the years they were pregnant.

To increase the sample size as well as the proportion of poor and black families in the sample, we include both the Survey Research Center (SRC) component and the Survey of Economic Opportunity (SEO) component, commonly known as the "poverty sample," of the PSID. We appropriately apply multi-level sample weights at the neighborhood and family levels to produce nationally-representative estimates. Estimates that exclude the SEO sample are similar to those reported in the paper but have much less precision. The sample consists of 36,121 person-year observations from 2730 individuals in 1894 families, 1457 neighborhoods, and 306 US counties. The mean age is 53, with age ranging from 37 to 77, and an average of 13 observations per person. A total of 865 families contained married couples, and a total of 307 neighborhoods contained at least two different unrelated families.

Health outcome

The health outcome examined is self-assessed general health status (GHS) based on the question asked of household heads and spouses (if present) at each wave between 1984 and 2005: "Would you say your health in general is excellent, very good, good, fair, or poor?" This question also was asked of all family members in 1986. GHS is highly predictive of morbidity measured in clinical surveys, and is a powerful predictor of mortality, even when controlling for physician-assessed health status and health-related behaviors (Benyamini, Idler, Leventhal, & Leventhal, 2000).

Differentials in the likelihood of reaching old age are important in understanding the health differences observed in old age (Johnson, 2009). The first wave of PSID interviewing in 1968 included 4354 individuals who were in their 20s and 30s; 321 of these individuals died by 2005; 2871 had at least one valid report of health status in mid-to-late adulthood. These individuals are included in the analyses for the years they are observed alive. Johnson (2009) analyzes the effects of family and neighborhood conditions earlier in life on adult mortality in mid-to-late life. The results suggest that any selective attrition with respect to mortality is likely to lead to an understatement of the impact of adverse conditions in childhood and young adulthood on health status in older adulthood because those who suffered premature death disproportionately grew up in the more disadvantaged family and neighborhood environments.

In order to scale the GHS categories, we use the health utility-based scale that was developed in the construction of the Health and Activity Limitation index (HALex) and used by the National Center for Health Statistics to estimate health-related quality of life measures (Healthy People, 2000). We estimate all of the regression models of health status using the interval regression method. Using a 100-point scale where 100 equals perfect health, the interval health values associated with GHS (based on HALex) that are used in this paper are: [95, 100] for excellent, [85, 95) for very good, [70,85) for good, [30,70) for fair, and [1,30) for poor health. While the HALex approach with interval regressions is superior to alternatives (see Electronic Appendix for further details, we have also estimated identical models to those reported in the tables but using poor/fair health as the dependent variable in a multi-level logit model. The substantive conclusions are unchanged.

Childhood factors

The measures of childhood socioeconomic status come from retrospective self-reports of childhood conditions collected in the early years of the PSID. These measures include childhood poverty (poor and non-poor), mother's and father's education (high school dropout, high school graduate and college educated), father's occupation (white collar and blue collar), region of birth, and the type of community the individual grew up in (farm, large urban area, small town or different places). Additional childhood factors include year of birth, birth order, gender, and race (black and white). The accuracy of retrospective reports of childhood conditions have been examined and found to be reliable (Haas, 2007; Smith, 2009).

Young adult family/individual factors

Socioeconomic and health insurance factors in young adulthood are represented by several variables: average annual family income-to-needs ratio during that period 1967–1971 (ratio$<$1.0, ratio 1.0–3.0, and ratio $>$ 3.0), educational attainment (high school dropout, high school graduate, some college, and college graduate or higher), no private health insurance coverage at any point 1968–1972, cigarette smoking at any point 1968–1972, and annual expenditures on alcohol consumption 1968–1972.

Young adult neighborhood factors

A key aspect of the data is that each individual is geocoded to the census block of residence. We define the neighborhood of residence during young adulthood as the census block where the respondent lived in 1968. In the original wave of the PSID, a represented census block contains four sampled families, on average. The PSID cluster design is discussed in greater detail in Solon, Page, and Duncan (2000). We utilize information on neighborhood characteristics from respondent self-reports in the survey and merged on neighborhood-level variables from the 1970 Decennial Census. The Census-based variables include the neighborhood poverty rate (as defined by the US federal poverty line), and the black-white dissimilarity index as a measure of racial residential segregation in the metropolitan area. We classify neighborhoods with low poverty levels as those in which less than 10% of the households are poor; neighborhoods with high poverty levels are those in which more than 30% of the households are poor; and neighborhoods with medium levels of poverty are those in which 10–30% of households are poor.

The self-reports of housing/neighborhood conditions, which were only collected in 1975, include: whether live in public subsidized housing; poor neighborhood for children, whether there exist plumbing problems, housing structural problems, security problems, cockroach or rat problems, insulation problems, neighborhood cleanliness problems, overcrowding, noise, or traffic problems, burglary, robbery, assault, drug use, or problems related to having too few police. We use these 1975 measures to proxy 1968 neighborhood characteristics, with evidence that 1968 families tended to move to neighborhoods that had observable neighborhood characteristics that were similar to their previous residential location (based on our own examination and evidence from Kunz et al., 2003). We find in our sample that individuals' observable neighborhood characteristics exhibit a great deal of persistence across years during young adulthood. This suggests that the neighborhood an individual inhabits at a particular time in young adulthood is a reasonable proxy for one's long-run environment during that life stage, and that relying on such a proxy produces only modest measurement error biases. Based on this information, we conclude that the neighborhood measures that we use are accurate indicators of the quality of the 1968 neighborhood where respondents lived.

Most prior studies that use survey responses of neighborhood conditions characterize neighborhood features by relying on only a single individual's report of what happens in his/her neighborhood. A more reliable approach to measuring neighborhood socioeconomic conditions is to aggregate the reports of multiple respondents living in the same neighborhood (Kawachi & Berkman, 2003). Thus, the approach taken in this analysis to produce more reliable and valid measures of neighborhood characteristics is to exploit the PSID survey responses about neighborhood conditions from multiple respondents in the same neighborhood and use their collective assessment to build neighborhood indicators, as opposed to a strategy of constructing neighborhood indicators based on a single individual's report of neighborhood characteristics due to same source bias. The mean number of sample families per neighborhood was 4 (ranging from 1 to 32).

Accordingly, a neighborhood with high reported crime levels is defined as one in which the average response among neighboring PSID households is a report of major crime-related problems (e.g., security problems, burglary, robbery, assault, drug use, or problems related to having too few police). Similarly, a "neighborhood housing quality index" is constructed based on the average response of the presence of housing insulation, plumbing, and/or housing structural problems, among all PSID households who live in the same neighborhood. We also make use of a unique set of indices, measured in young adulthood, that capture aspirations/motivation, long-term planning and connectedness to informal sources of help that were collected in the early years of the PSID. The latter measures social supports and social capital. Neighborhood-level measures were obtained by computing the average index scores, respectively, based on responses among neighboring PSID households. This survey information is used along with the 1970 census tract based measures of the neighborhood poverty rate. Appendix Table A0 lists the sources and years of all data elements along with details of the PSID survey questions used to construct these measures. Appendix Table A1 contains descriptive statistics for all childhood and young adult family and neighborhood measures for our sample by race (Appendix).

Ideally, we would also want to incorporate information on detailed measures of the neighborhoods in which individuals grew up, but this information is not available. However, in a subset of analyses, we make use of PSID residential mobility histories over the 1968–2005 period, merged with neighborhood-level variables from the 1970, 1980, 1990 and 2000 Decennial Census, to construct contemporaneous measures of the neighborhood poverty rate and the duration of exposure to concentrated poverty.

Methods

Outline of empirical approach

The first goal of the analysis is focused on an overall assessment of the relative contributions of individual, family and neighborhood influences during childhood and young adulthood on health in mid-to-late life. The overall scope of both young adult family and neighborhood factors on later-life health (implied by the spousal and neighbor correlations in health, reported in Table 2) provide the impetus for further investigation of what aspects of young adult family and neighborhood features influence subsequent health trajectories. We then analyze the relative contribution of a rich array of measured childhood socioeconomic conditions, young adult family and neighborhood, and individual covariates to the total variation from each component, and test hypotheses about the effects of specific characteristics of families and neighborhoods (reported in Table 3). Upon discovering a deleterious relationship between living in concentrated neighborhood poverty during young adulthood and health status in older ages, we probe the robustness of this finding for causal inference. In particular, we employ an innovative empirical approach, recently proposed by Altonji et al. (2005), to gauge how sensitive estimates of the effects of neighborhood poverty are to selection on unobserved variables. The results reveal that even a large amount of selection on unobservable factors does not eliminate the significant effect of neighborhood poverty on health status later in life (reported in Table 4). Finally, we use the estimated models to assess the extent to which earlier life neighborhood and family SES conditions can explain the significant racial health disparities observed at older ages (reported in Table 3). We discuss the methods used for each of these stages of the analysis in turn.

First, we estimated models based on the aggregated neighborhoods approach (Solon et al., 2000) in order to create an upper bound for the size of the neighborhood effect. The basic idea is to estimate the correlation of health status later in life of adults who were neighbors earlier in life. We assess the extent to which health status is correlated among neighbors above and beyond the correlation that arises due to family effects (i.e., net of the resemblance because neighbors have similar family characteristics due to residential sorting). Because all 1968 family members within a given family are followed throughout their lives, the correlation among neighbors can be adjusted to account for family-specific

Table 1
Mid-to-late life health: importance of young adult neighborhood & family background.

(Dependent variable: general health status) Hierarchical random effects interval regression model: 100pt-scale, 100 = perfect health

	All adulthood yrs 35+	Ages 35–55	Over 55
	(1)	(2)	(3)
Constant	81.0112*** (0.1221)	81.1520*** (0.1224)	77.9502*** (0.1631)
Age − 50	−0.3909*** (0.0026)	−0.4340*** (0.0043)	
Age − 60			−0.4451*** (0.0062)
Female	−1.8432*** (0.1295)	−2.1336*** (0.1293)	−1.7299*** (0.1801)
Random effects, unmeasured (std dev)			
Young adult neighborhood component	9.4605*** (0.1262)	9.5284*** (0.1343)	10.5244*** (0.1718)
Family component	5.8912*** (0.1921)	5.6670*** (0.2079)	5.0742*** (0.3605)
Individual component	12.0808*** (0.0727)	11.7314*** (0.0732)	15.0163*** (0.1070)
Transitory error component	7.9767*** (0.0075)	6.7325*** (0.0082)	7.7180*** (0.0122)
Log-likelihood	−2235086	−1338205	−875160.66
Number of counties	306	306	306
Number of neighborhoods	1457	1416	1129
Number of families	1894	1827	1460
Number of individuals	2730	2625	2019
Number of person-year observations	36,121	22,766	13,355

*** $p < 0.01$, ** $p < 0.05$, *$p < 0.10$.
Note: Robust standard errors in parentheses and all standard errors are Huber-corrected, clustered on county. All models control for year of birth and include age squared and age cubed terms (coefficients suppressed to conserve space).

factors. This approach avoids the difficulty of defining neighborhood quality at the outset, and instead asks: relative to the correlation among spouses, do adult neighbors have highly correlated health status? The comparison of spousal correlations with young adult neighbor correlations in mid-to-late life health allows an assessment of the relative magnitudes of the effects of the neighborhood environment in adulthood versus family characteristics in adulthood, placing an upper bound on the neighborhood influence. Small adult neighbor correlations would indicate that adult neighborhood factors can explain only a minor portion of the variation in later-life health outcomes. Large neighbor correlations would leave open the possibility that neighborhoods contribute significantly to inequality in health outcomes, and further analyses of the effects of particular neighborhood characteristics would be warranted.

We begin by estimating the four-level hierarchical random effects model (Kawachi & Berkman, 2003; Rabe-Hesketh & Skrondal, 2008; Raudenbush & Bryk, 2002) given by

$$H^{*}_{\mathrm{tsfn}} = (\beta_{0000} + \beta_{1000}{}^{*}\mathrm{Age_t}) + (\eta_{000\mathrm{n}}) + \left(\varphi_{00\mathrm{fn}}\right) + \left(\delta_{0\mathrm{sfn}}\right) + \varepsilon_{\mathrm{tsfn}} \quad (1)$$

We estimate these models separately at middle age (ages 35–55) and later-life (ages over 55), in order to gain greater insight into the extent to which family and neighborhood factors earlier in life influence the trajectory of health in mid-to-late life. These unconditional baseline models also include controls for year of birth, quadratic terms for age (suppressed in the above notation), and gender. The indices t, s, f, and n denote time, individuals, families, and neighborhoods, respectively, where there are

$t = 1, 2,\ldots, O_{sfn}$ observations over time of individual s in family f in neighborhood n;

$s = 1, 2,\ldots, S_{fn}$ individuals in family f in neighborhood n;

$f = 1, 2,\ldots, F_n$ families in neighborhood n;

$n = 1, 2,\ldots, N$ neighborhoods.

The neighborhood-, family-, and individual-level random effects capture unobserved characteristics of the neighborhood, family, and individual. The neighborhood random intercept coefficient is represented by $\eta_{000\mathrm{n}}$; the family random intercept coefficient is represented by $\varphi_{00\mathrm{fn}}$; the individual random intercept coefficient is represented by $\delta_{0\mathrm{sfn}}$; and $\varepsilon_{\mathrm{tsfn}}$ represents the individual transitory component of self-reported health (which includes measurement error). Each of these random effects are assumed to be normally distributed with a mean of 0, and $\mathrm{var}(\eta_{000\mathrm{n}}) = \sigma^2_{0n}$, $\mathrm{var}(\varphi_{00fn}) = \sigma^2_{0\mathrm{fn}}$, $\mathrm{var}(\delta_{0\mathrm{sfn}}) = \sigma^2_{0\mathrm{sfn}}$, and $\mathrm{var}(\varepsilon_{tsfn}) = \sigma^2_{tsfn}$. Age_t is the individual's actual age at time t centered around the mean age in the sample. All standard errors are Huber-corrected, clustered on county (Rabe-Hesketh & Skrondal, 2008; Wooldridge, 2002).

Since neighborhoods are nested within counties, we also estimated five-level hierarchical models, where the hierarchical levels represented counties, neighborhoods, families, and individuals over time. Those models were estimated as a robustness check to ensure that the young adult neighborhood random effects components were not primarily driven by effects operating at higher geographic levels of aggregation. However, those models did not significantly improve the fit and the between-county random

Table 2
Spousal and young adult neighbor correlations in later-life health status.

	All adulthood yrs (35+)		Age 35–55		Over 55	
	Spousal correlation	Young adult neighbor correlation	Spousal correlation	Young adult neighbor correlation	Spousal correlation	Young adult neighbor correlation
Unconditional	0.4598 (0.0067)	0.3313 (0.0077)	0.4718 (0.0068)	0.3485 (0.0085)	0.3771 (0.0089)	0.3060 (0.0089)
Adjusted[a] (net of residential sorting of HHs w/similar family bckgrd)	–	0.2677				
Conditional, control for childhood SES	0.3605 (0.0075)	0.1988 (0.0083)	0.3739 (0.0076)	0.2116 (0.0098)	0.2593 (0.0098)	0.1592 (0.0095)
Conditional, control for childhood + young adult family/neighborhood factors	0.2760 (0.0080)	0.1287 (0.0079)	0.3003 (0.0081)	0.1409 (0.0094)	0.1658 (0.0104)	0.0880 (0.0093)

[a] To compute the adjusted neighbor correlations, we first estimated within-neighborhood estimates of the effects of family income, education, race, family structure, health insurance coverage, health behaviors, connectedness to informal sources of help and housing quality on health in adulthood. Then, we used the within-neighborhood estimates of the later-life health effects of this array of family characteristics to assess how much of the raw neighbor correlation is due to young adult neighbors having similar (observable) family characteristics as opposed to neighborhood effects per se. The estimation procedures are described in detail in the methods section of the paper. (The full model results used to compute the adjusted neighbor correlations are not shown to conserve space, but are available from the authors upon request).

Table 3
General health status in mid-to-late life.

	Age 35–55			Over 55		
	Controls for childhood SES + young adult family	Controls for childhood SES + young adult nhood + fam	Controls for childhood SES + young adult backgrd + contemporaneous nhood	Controls for childhood SES + young adult family	Controls for childhood SES + young adult nhood + fam	Controls for childhood SES + young adult backgrd + contemporaneous nhood
Childhood factors						
Black	−6.4238***	−5.2458***	−5.3266***	−5.2408***	−3.6323***	−3.8721***
Non-Hispanic white (reference category)	(0.3035)	(0.3210)	(0.3218)	(0.3851)	(0.4049)	(0.4058)
Childhood poverty	−0.9583***	−0.8027***	−0.8010***	−3.0631***	−2.9577***	−2.9434***
Non-poor (reference category)	(0.1670)	(0.1662)	(0.1663)	(0.2197)	(0.2193)	(0.2196)
Mother's education:						
High school dropout	−1.8096***	−1.8141***	−1.8138***	−2.3982***	−2.4046***	−2.4144***
High school graduate (reference category)	(0.1660)	(0.1659)	(0.1660)	(0.2227)	(0.2241)	(0.2244)
College educated	1.6477***	1.9316***	1.8945***	0.9967***	1.2704***	1.2114***
	(0.2504)	(0.2487)	(0.2489)	(0.3407)	(0.3399)	(0.3404)
Father's education:						
High school dropout	−2.0027***	−1.9255***	−1.8857***	−1.6665***	−1.7798***	−1.7860***
High school graduate (reference category)	(0.1785)	(0.1773)	(0.1774)	(0.2414)	(0.2405)	(0.2409)
College educated	−0.5665**	−0.3117	−0.2771	0.3004	0.1198	0.1359
	(0.2626)	(0.2608)	(0.2611)	(0.3593)	(0.3574)	(0.3578)
Father's occupation:						
White collar	0.2164	0.3893+	0.3786+	0.4631+	0.7858**	0.7271**
Blue collar (reference category)	(0.2441)	(0.2427)	(0.2429)	(0.3334)	(0.3324)	(0.3329)
Grew up on farm	−0.9223***	−1.0149***	−1.1212***	0.8425***	0.7363***	0.6976**
Grew up in large, urban MSA (ref. category)	(0.2153)	(0.2184)	(0.2187)	(0.2743)	(0.2764)	(0.2768)
Grew up in small town	−1.6966***	−1.6357***	−1.7062***	−1.6318***	−1.7885***	−1.7948***
	(0.1879)	(0.1885)	(0.1886)	(0.2407)	(0.2419)	(0.2422)
Grew up in different places	1.3571*	2.1853***	2.1411***	−0.8392	0.1602	0.1722
	(0.7486)	(0.7418)	(0.7421)	(0.9055)	(0.8995)	(0.9004)
Young adulthood factors						
Family income-to-needs ratio (avg during 1967–1972), spline:						
Income-to-needs ratio*ratio is <1	12.6425***	12.0062***	12.0201***	14.4338***	15.1875***	15.4341***
	(1.1252)	(1.1222)	(1.1229)	(1.5567)	(1.5577)	(1.5593)
Income-to-needs ratio*ratio is 1–3	2.2339***	1.6325***	1.6691***	2.6004***	2.0381***	2.0775***
	(0.1510)	(0.1515)	(0.1516)	(0.1937)	(0.1938)	(0.1940)
Income-to-needs ratio* ratio is >3	0.1179*	0.0799	0.0822	0.2545***	0.2012**	0.2175***
	(0.0673)	(0.0666)	(0.0667)	(0.0844)	(0.0837)	(0.0838)
Educational attainment:						
High school dropout	−6.4392***	−6.1404***	−6.1589***	−8.0484***	−7.2543***	−7.2735***
High school graduate (reference category)	(0.2142)	(0.2136)	(0.2138)	(0.2871)	(0.2881)	(0.2885)
Some college	0.4847**	0.6371***	0.6620***	0.9893***	1.1771***	1.1769***
	(0.1964)	(0.1957)	(0.1958)	(0.2626)	(0.2628)	(0.2632)
College graduate or higher	3.7414***	4.0982***	4.1290***	5.0618***	5.4568***	5.4627***
	(0.2039)	(0.2026)	(0.2028)	(0.2717)	(0.2715)	(0.2718)
No Private HI coverage, 1968–1972	−2.3005***	−1.7638***	−1.7383***	−3.2845***	−2.5922***	−2.6260***
	(0.1696)	(0.1691)	(0.1692)	(0.2183)	(0.2192)	(0.2194)
Smoked cigarettes at some point, 1968–1972	−2.4748***	−2.2742***	−2.2738***	−2.7553***	−2.6894***	−2.7168***
	(0.1747)	(0.1737)	(0.1738)	(0.2195)	(0.2180)	(0.2182)
Annual alcohol expenditures (in $100's),	0.0185+	0.0052	0.0050	0.0341**	0.0198	0.0197
5-year average 1968–1972	(0.0122)	(0.0122)	(0.0122)	(0.0155)	(0.0155)	(0.0155)
Young adult neighborhood factors						
Neighborhood poverty rate (1970), spline:						
Low poverty neighborhood (reference category)						
Medium poverty neighborhood		−2.2588***	−2.3401***		−2.7542***	−3.0075***
		(0.2752)	(0.2759)		(0.3558)	(0.3568)
(Neighborhood poverty rate − 20)* rate is 10 to 30%		−1.1892***	−1.2353***		−1.4913***	−1.7541***
		(0.3765)	(0.3772)		(0.4985)	(0.4996)
High poverty neighborhood		−0.5325	−0.5253		−6.3390***	−6.4341***
		(0.4500)	(0.4507)		(0.5921)	(0.5929)

Table 3 (*continued*)

	Age 35–55			Over 55		
	Controls for childhood SES + young adult family	Controls for childhood SES + young adult nhood + fam	Controls for childhood SES + young adult backgrd + contemporaneous nhood	Controls for childhood SES + young adult family	Controls for childhood SES + young adult nhood + fam	Controls for childhood SES + young adult backgrd + contemporaneous nhood
Neighborhood crime problem		−0.6165**	−0.6449**		−1.2592***	−1.2929***
		(0.2918)	(0.2920)		(0.3713)	(0.3717)
Residential segregation dissimilarity index, 1970 (MSA)		−2.0344+	−2.4281*		−4.9635***	−5.2839***
		(1.4509)	(1.4528)		(1.8422)	(1.8447)
Neighborhood housing quality index		−1.1363***	−1.1406***		−0.7885***	−0.7852***
		(0.0538)	(0.0538)		(0.0724)	(0.0725)
Neighborhood connectedness to informal sources of help		0.5998***	0.6015***		0.7270***	0.7260***
		(0.0589)	(0.0589)		(0.0750)	(0.0751)
Average aspirations index in neighborhood		0.4689***	0.4689***		0.5967***	0.5995***
		(0.0530)	(0.0531)		(0.0677)	(0.0678)
Contemporaneous neighborhood						
Neighborhood poverty rate			0.0714**			0.5512***
			(0.0322)			(0.0459)
Random effects, unmeasured (std dev)						
Young adult neighborhood component	5.3427***	5.1491***	5.1865***	5.2475***	4.8412***	4.8621***
	(0.1794)	(0.1766)	(0.1760)	(0.2416)	(0.2576)	(0.2575)
Family component	5.6989***	5.4752***	5.4415***	4.8759***	4.5505***	4.5092***
	(0.1948)	(0.1954)	(0.1972)	(0.3473)	(0.3707)	(0.3749)
Individual component	11.5072***	11.4734***	11.4871***	14.8695***	14.9012***	14.9303***
	(0.0711)	(0.0705)	(0.0706)	(0.1040)	(0.1043)	(0.1045)
Transitory error component	6.7333***	6.7333***	6.7304***	7.7175***	7.7177***	7.7130***
	(0.0082)	(0.0082)	(0.0082)	(0.0122)	(0.01218)	(0.0122)
Log-likelihood	−1333818.6	−1333338.8	−1333230.3	−871477.14	−871153.91	−871078.29
Number of counties	306	306	306	306	306	306
Number of neighborhoods	1416	1416	1416	1129	1129	1129
Number of families	1827	1827	1827	1460	1460	1460
Number of individuals	2625	2625	2625	2019	2019	2019
Number of person-year observations	22,766	22,766	22,766	13,355	13,355	13,355

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Note: All models include a constant and controls for age, age squared, age cubed, year of birth, gender, birth order, region of birth, and indices intended to capture long-term planning horizon (coefficients suppressed to conserve space).

effects component was not statistically significant, which supports the use of the four-level hierarchical model.

Of primary interest is the decomposition of the variance of the level of health in mid-to-late life into their within-family, between-family within-neighborhood, and between-neighborhood components. In this model, individuals from the same neighborhood but not in the same family (i.e., neighbors) are correlated because they share the neighborhood random effect η_{000n}, and married couples are correlated because they share the neighborhood and family random effects η_{000n} and φ_{00fn}. We evaluate the health correlation between spouses at the same age, and evaluate the health correlation between neighbors at the same age. In this model, the spousal correlation and neighbor correlation in the level of health can be computed, respectively, as:

$$\rho\,\text{spousal, healthlevel(age)} = \frac{(\sigma^2_{0n}) + \left(\sigma^2_{0fn}\right)}{(\sigma^2_{0n}) + \left(\sigma^2_{0fn}\right) + \left(\sigma^2_{0sfn}\right)}$$

$$\rho\,\text{neighbor, healhlevel(age)} = \frac{(\sigma^2_{0n})}{(\sigma^2_{0n}) + \left(\sigma^2_{0fn}\right) + \left(\sigma^2_{0sfn}\right)}$$

The spousal correlation is between H^*_{sfn} and $H^*_{s'E}$, evaluated at the same age; the neighbor correlation is between H^*_{sfn} and $H^*_{s'f'n}$, evaluated at the same age. Our interest is in the permanent (rather than the transitory) component of health, so we do not include the temporal variation of health in the denominator.

We then use the estimated spousal and young adult neighbor correlations in health at mid-life and later-life to construct an age-profile of spousal and neighbor health correlations. This enables us to better assess the later-life health consequences of neighborhood and family influences experienced earlier in the life cycle.

"Adjusted neighbor correlations"

We then estimate "adjusted neighbor correlations", which are net of the similarity arising from young adult neighbors having similar observed family characteristics. To extract the impact of similar family characteristics out of the neighbor correlation, we first estimate the following regression, where for ease of exposition we omit the random effects terms that are included in the estimated model:

$$H^*_{tsfn} = \alpha_0 age_{tsfn} + \alpha_1 \text{gender}_{sfn} + \alpha'_2 X_{\bullet\bullet fn} + \alpha'_3 \left(\overline{X_{\bullet\bullet\bullet n}}\right) + \varepsilon_{tsfn}, \quad (2)$$

where $X_{\bullet\bullet fn}$ is the vector of young adult family/individual characteristics. $\overline{X_{\bullet\bullet\bullet n}}$ is a vector of the 1968 neighborhood-level means of the same above variables.

Inclusion of family-level and neighborhood-level variables measuring the same concepts enables the vector α_2 of coefficients to capture within-neighborhood effects of family characteristics.

Table 4
Estimated effect of living in high poverty neighborhood during young adulthood on later-life health for a proportional correlation model with varying values of the relative correlation.

Relative correlation	Estimated effect of high poverty neighborhood during young adulthood (reference cat: low poverty neighborhood)
0 (Exogeneity)	−4.7344*** (0.1198)
0.2	−5.1847*** (0.2804)
0.4	−2.4631*** (0.2194)
0.8	−13.1834*** (0.2136)
1	−18.6636*** (0.7760)

Using the within-neighborhood estimates of the family effects of income, education, race, family structure, health insurance coverage, health behaviors, connectedness to informal sources of help and housing quality on health in adulthood, will ensure the coefficients (α_2) will not be biased by omitted neighborhood variables. This follows from the fact that the neighborhood-level unmeasured factors can only be correlated with the neighborhood-level mean of the covariates. In combination, the resulting estimates can be taken as a conservative estimate of the independent effects of family characteristics.

We then estimate the between-neighborhood variance in $\hat{\alpha}'X_{\text{fn}}$ by estimating a hierarchical random effects model of $\hat{\alpha}'X_{\text{fn}}$ on neighborhood-level, family-level, and individual-level random effects. We then subtract the estimate of the between-neighborhood variance in $\hat{\alpha}'X_{\text{fn}}$ from the estimate of the overall between-neighborhood variance in H^*_{sfn}. Dividing the resulting quantity by $\hat{\text{Var}}(H^*_{\text{sfn}})$ yields a tighter upper bound on the proportion of $\text{Var}(H^*_{\text{sfn}})$ that can be attributed to neighborhood effects.

The estimates of "adjusted neighbor correlation" enable us to ascertain how much of the raw neighbor correlation is due to young adult neighbors having similar (observable) family characteristics. We then investigate to what extent observable childhood socioeconomic status and young adult neighborhood- and family-level characteristics explain the observed spousal and neighbor correlations at middle age and late life.

Sensitivity analysis

As a preview of our results, we find concentrated neighborhood poverty is associated with significantly elevated risks of problematic health, even after controlling for an extensive set of individual and family-level factors. Rather than using traditional regression models with statistical controls, we provide a new approach to the study of neighborhood effects that allows for the assessment of the sensitivity of results to selection bias. We use sensitivity analysis to assess the impact of selection bias from unobserved covariates on estimates of neighborhood effects—in this case, the later-life health impacts of living in a poor neighborhood. Intuitively, these methods can be thought of as estimating neighborhood effects by comparing individuals living in poor and non-poor neighborhoods, who are otherwise identical on observable characteristics, and then testing the robustness of the results to the presence of unobserved covariates. The goal of the sensitivity analysis here is to assess how an unobserved covariate that affects both neighborhood of residence and adult health would alter our conclusions about the neighborhood effect. Sensitivity analysis asks, How do our inferences change given various hypothetical unobserved covariates?

Our aim is to assess how the point estimate and confidence interval of the effect of neighborhood poverty change under the presence of selection bias of varying strengths. We use an approach proposed by Altonji et al. (2005) and Krauth (2006). This sensitivity analysis allows one to determine the threshold of selection on unobservable factors, if any, at which neighborhood poverty during young adulthood no longer has a significant effect on adult health. The approach uses the statistical relationship between observed explanatory variables as a guide to generate plausible estimates about the relationship between observed and unobserved variables. The sensitivity parameter, θ, can be defined as

$$\text{corr}(X_k, u) = \theta \text{corr}(X_k, X\beta - X_k\beta_k)$$

where θ indexes the magnitude of the correlation between observables and unobservables relative to the analogous correlation among observables themselves. In other words, the correlation between the neighborhood poverty rate and the (effect-weighted) unobservables is proportional to the correlation between the neighborhood poverty rate and the effect-weighted observables. The standard exogeneity assumption is the special case of $\theta = 0$. This approach provides a way to construct bounds on the effect of neighborhood poverty during young adulthood on mid-to-late life health based on the bounds one places on the sensitivity parameter θ (i.e., the relative correlation).

Altonji et al. (2005) argue that if the observable determinants of an outcome are truly just a random subset of the complete determinants, selection on observable characteristics must be equal to selection on unobservables. Because the PSID was conducted to study family factors that affect well-being, we would expect selection on observables to be greater than selection on unobservable factors. In other words, the extensive measures of family and neighborhoods captured in the PSID are likely to be the most important determinants of adult health. Thus, estimates obtained under the assumption of equal selection, i.e., θ=1, will be biased downwards.

Results

The existence of a relationship between neighborhood poverty status in young adulthood and health status in mid-to-late life are evident in simple bi-variate comparisons. In the PSID sample, 45% of adulthood years are spent in fair or poor health among those who lived in poor neighborhoods in young adulthood, while that proportion is 15% among those who lived in non-poor neighborhoods. Further, the gaps in health status between groups living in more and less socioeconomically advantaged areas are large and increase across the life course.

We explore these relationships further by examining the unadjusted spousal and young adult neighbor correlations of health in mid-to-late life. We then examine how much of the adult neighbor correlations can be explained by the fact that families in a neighborhood tend to be similar as opposed to emanating from neighborhood effects *per se*. We then present estimates of the magnitude of the effects and attempt to explain the life cycle pattern of spousal and neighbor correlations and to explore potential mechanisms that underlie the relative roles of neighborhood and family factors on the health trajectory over the course of adulthood. Finally, we explore the role of factors over the life course in explaining racial disparities in health in mid-to-late life.

Unadjusted spousal and young adult neighbor correlations in later-life health

Table 1 presents the estimates from the baseline four-level hierarchical random effects models, separately at ages 35–55 and at ages over 55—these models control only for age, gender, and year of birth. The random effects estimates are all significant at each of the young adulthood neighborhood, family and individual levels.

The baseline models in Table 1 measure the overall magnitude of variation in mid-to-late life health-related to young adult

neighborhood, family, and individual-level factors. The spousal and neighbor correlation estimates are based on the decomposition of variance over time into the fraction that lies between neighborhoods, families, and individuals. The unadjusted spousal and neighbor correlations calculated from the baseline models are summarized in the first row of Table 2. The spousal correlation in general health status is 0.46, on average, across ages 35 and older, and declines somewhat during ages after 55.

Spousal correlations by themselves cannot disentangle how much of the resemblance among their health outcomes is due to the effects of shared household environment and family background and how much is due to the effects of neighborhoods. Augmenting the spousal correlation estimates with corresponding neighbor correlation estimates reveals young adult neighborhood and family factors may both be important determinants of general health status in mid-to-late life. While the young adult neighbor correlations are smaller than the spousal correlations, they are significant through middle age and old age. In particular, the young adult neighbor correlation in health at ages 35 and older averages 0.33 (Table 2). These magnitudes are particularly noteworthy given the fact that the vast majority of these individuals no longer live in the same neighborhoods they once shared as young adults (in this paper young adulthood is defined as individuals in their 20s and 30s). In the final section we will return to the issue of the extent of residential mobility and persistence in neighborhood quality across the life course, and its implications for these results.

Adjusted spousal and young adult neighbor correlations in later-life health

From the adjusted neighbor correlation estimates, we find that observable family sorting (controlling for a broad array of young adult family characteristics including income, education, race, family structure, health insurance coverage, health behaviors, connectedness to informal sources of help and housing quality) explain some but not all of the resemblance in adulthood health status among individuals who lived in the same neighborhood as young adults. (The full model results used to compute the adjusted neighbor correlations are not shown to conserve space, but are available from the authors upon request). Specifically, the adjusted neighbor correlation is roughly 20% lower than the unadjusted neighbor correlation (Table 2, row 2), suggesting that differences in neighborhood quality during young adulthood may account for up to one-quarter of health disparities in mid-to-late life.

Magnitude of effects and potential pathways

What do these correlation estimates mean in terms of the absolute size of the effects of family and neighborhood factors, and what aspects of family socioeconomic status and young adult neighborhood matter to explain the lasting importance of earlier life conditions on late life health? Estimates of the neighborhood random components (σ_n) indicate that neighborhood quality during young adulthood has significant and enduring association with general health status over the course of adulthood. It is also possible that some of this underlying relationship may be due to differences in childhood neighborhood quality and the correlation between childhood neighborhood and young adult neighborhood factors. (We cannot directly investigate this issue for the older birth cohorts analyzed in this paper because we do not have information on detailed neighborhood-level factors experienced during childhood).

With these questions in mind, we next estimate a series of models, reported in Table 3, building toward a full model specification that includes a rich array of observable child socioeconomic status measures and young adult family-level and neighborhood-level characteristics to attempt to identify earlier life determinants of health in mid-to-late life. Family income and neighborhood poverty are dimensions of family and neighborhood background that are a key focus of the analysis. We include neighborhood measures of crime, connectedness to informal sources of help, aspirations/motivation, long-term planning, and racial residential segregation in young adulthood. These factors may themselves be the product of living in a very poor neighborhood and may represent pathways through which exposure to depressed neighborhood environments earlier in life affect health trajectories later in life. However, controlling for this myriad of ways in which those who reside in very poor neighborhoods may differ from individuals who live in affluent neighborhood environments allows one to generate a more conservative estimate of the effect of neighborhood poverty itself, as well as shed light on the factors that affect adult health status.

Table 3 reports the regression results for adult ages 35–55, and ages 55 and older, respectively. For each age group, the series of results reported include models that control for childhood SES and young adult family characteristics (first column) and then adds young adult neighborhood characteristics (second column), and a full model that adds contemporaneous socioeconomic and neighborhood characteristics (third column). Separately identifying the causal pathways through which earlier life socioeconomic factors influence late life health is an arduous task, and one must use caution in drawing causal inferences from these coefficient estimates. The estimates summarize the relationships between the health trajectory in adulthood with various dimensions of neighborhood and family background. The robustness of the results for causal inference is examined in detail in the final section of the paper.

The results demonstrate that the most salient young adult neighborhood factor is concentrated neighborhood poverty. To put the magnitudes in perspective, it is useful to consider that the estimated later-life health differences between those who lived in a neighborhood with medium poverty levels versus low poverty levels during young adulthood are on par with the estimated impacts of smoking as a young adult. As shown in column 5 of Table 3, the results for health at ages over 55 show that a 10 percentage-point increase in the neighborhood poverty rate (in young adulthood) from 10 to 20 percent is related to a 1.49 reduction in GHS, and living in a neighborhood with a high poverty level corresponds with a 6.34 lower GHS score in later-life, relative to living in a neighborhood with a low poverty level as a young adult. This latter effect is equivalent to reaching a level of health deterioration roughly 14 years sooner for an individual who lived in a neighborhood with a high poverty level, relative to the simulated health trajectory experienced by one who lived in a neighborhood with a low poverty level earlier in adulthood. For purposes of causal inference, the robustness of this result to alternative thresholds of selection on unobservable factors is analyzed in the following section.

Several other related dimensions of neighborhood disadvantage experienced in young adulthood had substantial, independent associations with the health trajectory over the course of adulthood, including high crime, racial residential segregation, neighborhood housing problems, neighborhood connectedness to informal sources of support (which may serve as a proxy for social cohesion), and neighborhood-level average aspirations for socioeconomic attainment. These factors appear to have stronger relationships with health over the life course, with stronger links to health at ages over 55 relative to middle age. The age-profile of these estimated effects suggest that the linkages may be the result of cumulative exposure to disadvantaged environments taking a toll on health later in life that may be reinforced with how these

factors earlier in life influence the socioeconomic mobility process. For example, as shown in column 2 of Table 3, for health status at ages over 55, living in a neighborhood with high reported crime as a young adult reduces GHS by 1.3 points; a 25 percentage-point increase in childhood residential segregation (dissimilarity index increase from 50 to 75) is related to a 1.2 point reduction in GHS (similar effects for both blacks and whites (results not shown)). Taken together, the cumulative set of childhood SES measures, and young adult neighborhood and family factors account for 54 percent of the neighborhood-level variance at ages over 55 (implied quasi-R^2 at the neighborhood-level). After controlling for observable childhood and young adult SES, the similarity of both spouses' and young adult neighbors' later-life health outcomes are less marked. Namely, conditional on childhood SES, and young adult neighborhood and family factors, the spousal correlation is 0.28 and the young adult neighbor correlation in mid-to-late life health is reduced to 0.13, which is roughly one-half the size of the adjusted neighbor correlations (see the third row of Table 2). Thus, we see that our observable measures of childhood SES and young adult neighborhood factors account for a sizable share of the patterns of health inequality in older adulthood and explain a significant portion of the resemblance of young adult neighbors' later-life health outcomes.

Most of the neighborhood effects literature that has examined health outcomes has investigated whether contemporaneous neighborhood factors are significantly associated with health at a point in time (see, for example, literature reviews by Ellen et al. (2001) and Kawachi & Berkman, 2003). In columns 3 and 6 of Table 3, we use our models to re-examine this issue. The final model includes both the full set of childhood SES and young adult neighborhood and family factors along with the contemporaneous adult neighborhood poverty rate. (Differences in the functional form specified for contemporaneous neighborhood poverty yielded nearly identical results). Contemporaneous adult neighborhood poverty was only weakly related to health status in mid-to-late life. The coefficient estimates on the childhood SES and young adult neighborhood and family factors remain significant and are not significantly reduced with the inclusion of the contemporaneous adult neighborhood poverty rate.

Results from sensitivity analysis

The estimates of the significant effects of neighborhood poverty during young adulthood on later-life health reported in Table 3 are based on models in which exogeneity is assumed. We conduct a sensitivity analysis to test the robustness of the estimated effects of young adult neighborhood poverty to selection bias due to an omitted variable. Table 4 presents the range of estimated coefficients and standard errors on young adult neighborhood poverty as a function of the ratio of selection on unobservables to selection on observables, i.e., θ. Under the assumption of exogeneity, i.e. , $\theta = 0$, the estimated effect is -4.7344 and highly significant. If instead, $\theta = 0.4$, the estimated effect is smaller but still substantial at -2.4631 and highly significant. In sum, we find that the effect of young adult neighborhood poverty on health status later in life remains large and significant even with a large amount of selection on unobservable factors.

While there is no single perfect solution to address the endogeneity of residential location, there are additional tests we employed to determine whether selection bias is driving the results. In additional model results not shown, we find that the strength of the estimated effects of neighborhood poverty during young adulthood vary with the number of known neighbors and duration of exposure to concentrated neighborhood poverty. The fact that the effect of concentrated neighborhood poverty is weaker when the duration of exposure is brief and individuals know few of their neighbors suggests that selection bias is not driving these results, and that the observed patterns of health inequality are not solely attributable to pre-existing differences between individuals living in different neighborhoods. If effects simply represented unmeasured family factors, then the number of years in the neighborhood and the number of neighbors known by name should not be associated with the strength of these effects. But that is not the case here, and the evidence is not consistent with alternative omitted variables counter-explanations of the results, nor with the hypothesis that families predisposed to poor outcomes select into poor neighborhoods. Instead, the pattern of results, taken together, point towards true 'causal' effects of concentrated neighborhood poverty, rather than results purely driven by selection effects.

Health disparities

Residential segregation patterns that lead to racial differences in neighborhood quality have been cited as a factor that perpetuates racial health disparities (Williams & Collins, 1995). We also used the PSID data to examine racial disparities in health over the life course. Gaps in health between blacks and whites are large and exist at all stages in life. The general health status (GHS) index in adulthood, on average, based on the PSID data is 13.9 points lower for blacks and this gap increases in levels and in proportionate terms in adulthood. A useful way to interpret the estimate is in relationship to the size of the effect of age on health, with the race gap in ages over 55 equivalent to blacks (on average) reaching a level of health deterioration about 30 years prior to their white counterparts. That is, GHS is 13.9 points lower for black adults, which is equal to roughly 30 years evaluated at an effect of age of -0.45. Thus, substantial disparities in health exist between blacks and whites over the life course.

In this section, we explore the role of childhood and young adult family/neighborhood measures on the black-white difference in health status later in life (at ages 55 and older). We find that these factors have a substantial influence. The black-white difference in health status at ages over 55 in models that include childhood SES and young adult family measures is reduced to one-third of its original raw age-adjusted gap. Most prominent among these young adult family factors is family income, with substantially larger impacts of income in the lower tail of the distribution. For example, a one-unit increase in the family income-to-needs ratio (during young adulthood) from half of the poverty line to 1.5 times the poverty line translates into an 8.5 point increase in adult GHS at ages over 55 ($0.5 \times 14.4338 + 0.5 \times 2.6004$), which is equivalent to 18.9 years younger. One's own educational attainment, health insurance coverage, cigarette smoking and alcohol use in young adulthood are each also shown to be strongly associated with later-life health.

Further evidence of the influence of early life SES on late life health disparities is found in the estimates in columns 2 and 5 of Table 3. Once we control for child SES and both young adult neighborhood and family factors, the black-white difference in health status at ages over 55 is reduced by 76 percent. Further, the coefficients on the childhood SES measures are reduced significantly with the inclusion of the set of young adult family and neighborhood factors, but the childhood factors remain large and significant.

Finally, as shown in these columns 3 and 6 of Table 3, the results reveal that while three quarters of racial differences in health at ages over 55 can be accounted for by childhood SES and young adult family and neighborhood factors, contemporaneous neighborhood factors independently account for little of this gap. Taken together, these results underscore the importance of earlier life family and neighborhood SES conditions as keys to understand sources of health disparities at older ages.

Limitations of the study

Although the forty-year panel of data from the PSID provides a rich source of data from which to study the effects of neighborhoods on health, the analyses here have some limitations. While it is important to theorize potential pathways through which neighborhood influences earlier in life may have long-run impacts on later-life health, it is also important to bear in mind that the empirical work does not aim to separately identify all the causal pathways through which, for example, living in concentrated poverty neighborhoods matter. Instead, we produce estimates of the total (direct and indirect) impacts of these neighborhood conditions, and show that these neighborhood effects persist even in the presence of a reasonably large amount of potential unobservable individual and family factors that may significantly affect both neighborhood of residence and subsequent health status. Understanding and uncovering all of the potential causal mechanisms through which neighborhood SES matters remains fertile ground for future research. While we have a more extensive set of measures than most previous studies have employed across the life course regarding earlier life family and neighborhood influences, we are nonetheless restricted to available measures in the PSID and Census data. Therefore, our measures may proxy for key underlying constructs of neighborhood SES and social capital such as community resources, social cohesion, and stressful environments. We have limited measures of childhood neighborhood conditions for this older birth cohort. Research using new measures of family- and neighborhood-level stress and potential interventions to moderate it are an important and promising area for future study. Finally, although our research design permits us to produce upper bounds of the potential scope of neighborhood influences, and sensitivity analyses allowed us to assess the robustness of results to varying degrees of selection, it remains difficult to fully address endogeneity issues in this work.

Conclusion

Previous research demonstrates that health status in the US varies strongly across local, state, and regional settings (Murray, Michaud, McKenna, & Marks, 1998). This paper investigates multi-level causation of these differences over the course of adulthood and provides insights into how family and neighborhood environments earlier in life influence the health trajectories of individuals in ways that cannot be reduced to the characteristics of the individuals themselves.

In this paper, we used correlations based on a nationally-representative longitudinal sample of married couples and neighbors followed from young adulthood through elderly ages to estimate bounds on the possible causal effects of young adult family and neighborhood characteristics on general health status in mid-to-late life. The results show a significant scope for both young adult family and neighborhood factors. The estimates suggest that disparities in neighborhood conditions experienced in young adulthood account for up to one-quarter of the variation in health status in mid-to-late life. While the neighbor correlations must be strictly interpreted as upper bounds, the estimates suggest that neighborhood factors experienced earlier in the life course influence both contemporaneous and future health outcomes.

Research on how neighborhood and family background influence later-life health is one with potential endogeneity issues that are not amenable to the usual microeconometric corrections through use of instrumental variables or fixed effect approaches, and for which the extant experimental evidence is likely too short a time horizon to detect effects on overall health status. Instead of attempting to remove or avoid selection bias caused by unobserved factors, the methods employed in this paper assess how the presence of varying levels of selection bias would alter conclusions about the effect of living in a neighborhood with a high poverty level during young adulthood on mid-to-late life health. The results reveal that even a large amount of selection on unobservable factors does not eliminate the significant effect of neighborhood poverty on health status later in life. This evidence indicates further research on the effects of particular neighborhood characteristics is strongly warranted to identify the causal mechanisms through which concentrated neighborhood poverty effects operate.

The findings in this paper further our understanding of the underlying processes that produce health disparities between different racial and socioeconomic groups. The study finds that racial differences in later-life health can be largely accounted for by childhood socioeconomic status and young adult family and neighborhood factors, while contemporaneous neighborhood factors account for relatively little of this gap. The paper investigates differential effects of socioeconomic conditions on health over the life cycle and provides evidence that dimensions of neighborhood disadvantage that occur earlier in life are more predictive of morbidity during mid-to-late life than are contemporaneous neighborhood measures. The cumulative and dynamic nature of the relationship between socioeconomic conditions and health over the life course applies to dimensions of both family and neighborhood characteristics. The findings of this paper challenge future research to further our understanding of the pathways through which neighborhoods influence health and how the effects of neighborhood conditions differ over the life cycle. This knowledge is critical to help policy makers develop interventions (e.g., early childhood interventions or targeted policies for the geographic deconcentration of the poor) that build a bridge between childhood and early adulthood for impoverished families, so fewer individuals arrive at the doorstep of retirement with accumulated exposures that are irreversible.

Appendix. Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.socscimed.2011.10.021.

References

Altonji, J., Elder, T., & Taber, C. (2005). Selection on observed and unobserved variables: assessing the effectiveness of catholic schools. *Journal of Political Economy, 113*(1), 151–184.

Ananat, E., Gassman-Pines, A..Gibson-Davis, C. M. The effect of local employment losses on children's educational attainment. In G. J. Duncan & R. Murnane (Eds.), *Social inequality and educational disadvantage*. New York: Russell Sage Foundation Press, in press.

Andreski, P., McGonagle, K., & Schoeni, R. (2009). *An analysis of the quality of the health data in the panel study of income dynamics*. PSID Technical Series Paper #09–02.

Barker, D. (1998). *Mothers, babies and health in later life*. Edinburgh: Churchill Livingstone.

Benyamini, Y., Idler, E., Leventhal, H., & Leventhal, E. (2000). Positive health and function as influences on self-assessments of health: expanding our view beyond illness and disability. *Journal of Gerontology: Psychological Sciences, 55B*, 107–116.

Becketti, S., Gould, W., Lillard, L., & Welch, F. (1988). The PSID after fourteen years: an evaluation. *Journal of Labor Economics, 6*(4), 472–492.

Curtis, S., Setia, M., & Quesnel-Vallee, A. (2009). Socio-geographic mobility and health status: a longitudinal analysis using the National Population Health Survey of Canada. *Social Science and Medicine, 69*(12), 1845–1853.

Curtis, S., Southall, H., Congdon, P., & Dodgeon, B. (2004). Area effects on health variation over the life-course: analysis of the longitudinal study sample in England using new data on area of residence in childhood. *Social Science & Medicine, 58*, 57–74.

Ellen, I., Mijanovich, T., & Dillman, K. (2001). Neighborhood effects on health: exploring the links and assessing the evidence. *Journal of Urban Affairs, 23*(3), 391–408.

Fitzgerald, J., Gottschalk, P., & Moffitt, R. (1998a). An analysis of sample attrition in panel data. *The Journal of Human Resources, 33*(2), 251–299.

Fitzgerald, J., Gottschalk, P., & Moffitt, R. (1998b). An analysis of the impact of sample attrition on the second generation of respondents in the Michigan panel study of income dynamics. *The Journal of Human Resources, 33*(2), 300–344.

Geronimus, A. (1996). Black/White differences in the relationship of maternal age to birthweight: a population-based test of the weathering hypothesis. *Social Science & Medicine, 42*(4), 589–597.

Haas, S. (2007). The long-term effects of poor childhood health: an assessment & application of retrospective reports. *Demography, 44*(1), 113–135.

Healthy People. (2000). United States Department of Health and Human Services. www.healthypeople.gov.

Johnson. (2009). *Inequality in men's mortality: The influence of childhood SES and young adult neighborhood and family factors*. Working paper. University of California, Berkeley.

Katz, L., Kling, J., & Liebman, J. (2001). Moving to opportunity in Boston: early results of a randomized mobility experiment. *Quarterly Journal of Economics, 116*(2), 607–654.

Kawachi, I., & Berkman, L. (2003). *Neighborhoods and health*. New York, New York: Oxford University Press.

Krauth, B. (2006). Simulation-based estimation of peer effects. *Journal of Econometrics, 133*(1), 243–271.

Kuh, D., & Wadsworth, M. (1993). Physical health status at 36 years in the British national birth cohort. *Social Science & Medicine, 37*(7), 905–916.

Kunz, J., Page, M., & Solon, G. (2003). Are point-in-time measures of neighborhood characteristics useful proxies for children's long-run neighborhood environment? *Economics Letters, 79*(2), 231–237.

Lillard, L., & Panis, C. (1998). Panel attrition from the panel study of income dynamics. *The Journal of Human Resources, 33*(2), 437–457.

Leventhal, T., & Brooks-Gunn, J. (2003). New York city site findings: the early impacts of moving to opportunity on children and youth. In J. Goering, & J. D. Feins (Eds.), *Choosing a better life: Evaluating the moving to opportunity social experiment* (pp. 213–243). Washington, DC: Urban Institute Press.

Oakes, J. (2004). The (Mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Social Science & Medicine, 58*(10), 1929–1952.

Manski, C. (1993). Identification of endogenous social effects: the reflection problem. *The Review of Economic Studies, 60*(3), 531–542.

Murray, C., Michaud, C., McKenna, M., & Marks, J. (1998). *U.S. patterns of mortality by county and race: 1965–1994*. Cambridge, MA: Harvard Center for Population and Development Studies (Monograph Series).

Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modeling using stata* (2nd ed.). College Station, TX: Stata Press.

Raudenbush, S., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks: Sage Publications.

Smith, J. (2009). Reconstructing childhood health histories. *Demography, 46*(2), 387–403.

Solon, G., Page, M., & Duncan, G. (2000). Correlations between neighboring children in their subsequent educational attainment. *The Review of Economics and Statistics, 82*(3), 383–392.

Williams, D., & Collins, C. (1995). US socioeconomic and racial differences in health: patterns and explanations. *Annual Review of Sociology, 21*, 349–386.

Wooldridge, J. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.