HEALTH DYNAMICS AND THE EVOLUTION OF HEALTH INEQUALITY OVER THE LIFE COURSE: THE IMPORTANCE OF NEIGHBORHOOD AND FAMILY BACKGROUND

RUCKER C. JOHNSON*
GOLDMAN SCHOOL OF PUBLIC POLICY
UNIVERSITY OF CALIFORNIA, BERKELEY

May 2009

* Please direct correspondence to Rucker Johnson, University of California, Berkeley, Goldman School of Public Policy, 2607 Hearst Ave, Berkeley, CA 94547, or email to ruckerj@berkeley.edu. This research was supported by funding from the Robert Wood Johnson Foundation and NIA grant R01AG04058-02. I am grateful to Bob Schoeni, Sheldon Danziger, John Bound, Charlie Brown, Gary Solon, Marianne Page, Ken Chay, Greg Duncan, Vicki Freedman, Jeannette Rogowski, Bob Willis, and Bobbi Wolfe for helpful comments on an earlier version of this paper. I wish to thank Jesse Rothstein for providing data on rivers, and the PSID staff for access to the confidential restricted-use PSID geocode data. This work also benefited from the discussion of seminar participants at the NBER Summer Institute, University of Michigan, UC-Berkeley, Cornell, NYU, Washington, the RAND Health Economics conference, and PAA Demography conference.

HEALTH DYNAMICS AND THE EVOLUTION OF HEALTH INEQUALITY OVER THE LIFE COURSE: THE IMPORTANCE OF NEIGHBORHOOD AND FAMILY BACKGROUND

ABSTRACT

This paper investigates the extent and ways in which childhood family and neighborhood quality causally influence later-life health outcomes. The study analyzes the health trajectories of children born between 1950 and 1970 followed through 2005. Data from the Panel Study of Income Dynamics (PSID) spanning four decades are linked with information on neighborhood attributes and school quality resources that prevailed at the time these children were growing up.

There are several key findings. First, estimates of sibling and child neighbor correlations in health are used to bound the proportion of inequality in health status in childhood through mid-life that are attributable to childhood family and neighborhood quality. Estimates based on four-level hierarchical random effects models (neighborhoods, families, individuals, over time) consistently show a significant scope for both childhood family and neighborhood background (including school quality). The results imply substantial persistence in health status across generations that are linked in part to low intergenerational economic mobility. Sibling correlations are large throughout at least the first 50 years of life: roughly three-fifths of adult health disparities may be attributable to family and neighborhood background. Childhood neighbor correlations in adult health are also substantial (net of the similarity arising from similar family characteristics), suggesting that disparities in neighborhood background account for more than one-third of the variation in health status in mid life.

Second, exposure to concentrated neighborhood poverty during childhood has significant deleterious impacts on adult health. I control for the endogeneity of neighborhood location choice using instrumental variables based on political factors, historical migration patterns, and topographical features. The results reveal that even a large amount of selection on unobservable factors does not eliminate the significant effect of child neighborhood poverty on health status later in life. Thus, racial differences in adult health can be accounted for by childhood family, neighborhood, and school quality factors, while contemporaneous economic factors account for relatively little of this gap.

I. INTRODUCTION

Persistent residential segregation of poor and minority populations has spurred a growing literature that investigates the effects of community background on socioeconomic outcomes. However, the effects of the physical and socioeconomic neighborhood on health outcomes have been relatively unexplored. Some studies have demonstrated that health outcomes exhibit a distinctive spatial pattern that mirrors the spatial pattern of physical and socioeconomic disadvantage (e.g., Geronimus et al., 2001; Morenoff and Lynch, 2004; Skinner et al., 2002; Chandra and Skinner, 2003). The similarity of these geographic patterns motivates this paper's investigation into the potential causal effects of neighborhood quality during childhood on adult health status.

This paper examines how and why individual, family, and neighborhood factors produce and reproduce poor health. The principal impact of parents on their children is shaped during childhood. To understand how childhood disadvantage transmits itself into adulthood, we must separate the effects of neighborhood background from parental factors or genetic factors.

Most analyses of health disparities are cross-sectional and do not examine the dynamics of health inequality over the life-course. Do those who are born into disadvantaged neighborhood and family backgrounds persistently have worse health over their lifetime? Or, is the economic mobility process in the U.S. fluid enough to enable those from less advantaged backgrounds to achieve relatively good health and better economic status in adulthood?

The typical analytical approach used in neighborhood studies is to regress individual level outcomes such as education, criminal activity, or health on contemporaneous neighborhood-level factors such as census tract mean income, poverty rates, or rates of single motherhood. Such attempts to estimate causal effects of neighborhood context have faced well-documented challenges of endogeneity (Manski, 1993). The primary difficulty in disentangling the relative importance of childhood family and neighborhood quality factors is isolating variation in neighborhood quality characteristics that are unrelated to family factors. Another obstacle is that available data used in prior studies has rarely measured neighborhood factors at a geographic level (e.g., the block) that is detailed enough to identify

the neighborhood features that affect future health. Few studies have used convincing identification strategies to overcome these challenges, exceptions being experimental evaluations such as Katz, Kling, Liebman (2001) and Leventhal and Brooks-Gunn (2001).

This paper exploits unique features of the Panel Study of Income Dynamics (PSID) linked with multiple data sources to investigate the long-run consequences of dimensions of childhood neighborhood quality on adult health status. I control for the potential endogenous selection of families into neighborhoods using instrumental variables based on political factors, historical patterns of migration, and topographical features and assess the extent and ways in which childhood family and neighborhood quality factors causally influence later-life health outcomes. The sample includes children born between 1950 and 1970 who are followed through 2005.

The analysis proceeds in three stages. I first bound the proportion of inequality in adult health that may be attributed to disparities in family and neighborhood quality characteristics (observed and unobserved) during childhood. The research strategy exploits the fact that the initial PSID sample in 1968 was highly clustered with most PSID families having several other sample families living on the same block. This survey design allows a comparison of the similarity in adulthood health between siblings who grew up together, versus unrelated individuals who grew up in the same narrowly defined neighborhood. I use correlations between neighboring children's subsequent health in adulthood to bound the proportion of inequality in health outcomes that can be attributed to disparities in neighborhood background. The comparison of sibling and child neighbor correlations in adult health status allows an assessment of the relative magnitudes of the effects of the childhood neighborhood and family environments. The findings are based on the estimation of four-level hierarchical random effects models of health status over the life course.

Second, after documenting substantial child neighbor correlations in adult health outcomes, the paper analyzes the relative contribution of a rich array of measured individual, family, neighborhood, and school characteristics to the total variation from each component, and tests hypotheses about the effects of specific characteristics of families, neighborhoods and schools. I find that growing up in a neighborhood

with concentrated poverty substantially increases the likelihood of having problematic health at mid-life, in ways that cannot be reduced to the characteristics of the individuals and families themselves.

Third, to assess the robustness of the results for causal inference, I assess how large the unobservables would need to be, relative to the observable factors included, to invalidate the results. As an alternative approach, I control for the endogeneity of neighborhood location choice using instrumental variables based on political factors, historical migration patterns, and topographical features. The analysis attempts to disentangle the effects of neighborhood and school quality. The effects of childhood school quality factors are analyzed but presented in detail in a companion paper by Johnson (2009). Finally, the paper assesses the extent to which race differences in childhood families, neighborhoods, and schools account for racial health disparities in adulthood.

The remainder of the paper is organized as follows. I next discuss how family and neighborhood factors during childhood may affect an individual's health in adulthood. I describe an economic model of health that provides the theoretical framework, highlights the relevant theoretical issues, and motivates the empirical analyses to follow. Section III lays out the methodological challenges in estimating neighborhood effects. I outline my empirical approach in section IV. The data and descriptive results are presented in section V. Section VI discusses the econometric model and estimation methods. The regression results are presented in section VII, with concluding statements in section VIII.

II. WHY MIGHT NEIGHBORHOOD AND FAMILY BACKGROUND MATTER?

Family background can have direct effects on health status over the life course through several mechanisms. Transmission of genetic traits from parents to children plays an important role. Parental socio-economic and demographic factors most likely influence children's health status (Case, Lubotsky, and Paxson, 2002), which in turn carries through to health in adulthood. The transmission of health lifestyles – eating habits, exercise and smoking behaviors, for example – across generations may also affect adult health.

Similarly, it has been hypothesized that childhood neighborhood factors such as water and air quality, sanitation, pollution and environmental toxins, crime, health care and social services, and public schools most likely influence childhood health. Health lifestyles may also have a neighborhood component, with peer groups and role models within neighborhoods influencing children's opportunities and preferences (Johnson, 2008).

Neighborhood and family background may also have indirect effects on health over the life course through their effects on socioeconomic mobility. The degree of mobility has direct implications on the resemblance of an individual's childhood and adulthood family characteristics, such as income and education, which may in turn affect health. Because economic status is a major determinant of residential choice, persistence in economic status is likely to lead to persistence in neighborhood quality.

Theoretical Framework/Considerations. I present a two-period overlapping generations model of the transmission of health and economic status from parents to children to motivate the empirical analyses. The model adopts a simplified version of the basic framework of Becker and Tomes (1986). Some children have an advantage because they are born into families with favorable genetic attributes, which I refer to as the endowment component. Assume endowments are only partially inherited and parents cannot control endowment transmission, but can influence the adult human capital of their children through investment expenditures on their health, learning, and motivation. For example, while the child is in utero, the mother can invest in prenatal care or refrain from smoking. In the model, the central role the parent plays in determining the well-being of their children is to guide the level and allocation of investment in the child until the child is sufficiently mature to make independent decisions. Assume parents are altruistic toward their children in that their children's lifetime utility is a branch of the parents' utility function.

Individuals possess three types of capital in adulthood: health, education, and financial. Because much research demonstrates that investments during childhood are crucial to later development, I assume that the amount of education and health human capital in adulthood is proportional to the amount accumulated and preserved during childhood.

Assume children are born to parents of two types—rich or poor. Assume poor parents face credit constraints that prevent them from making all worthwhile investments in the human capital of their children. One constraint, a focal point of this research, is residential location choice. The formation of neighborhoods in this model is assumed to be the byproduct of economic segregation, which emerges because families prefer affluent neighbors because neighbors determine the level of school resources and the quality of peers and role models. Segregation thus sorts children into relatively more homogenous neighborhoods and school districts. Due to residential segregation by income, assume that two types of neighborhoods exist—high- and low-income.

Children's human capital is a function of their parents' human capital, school spending in their community, and the quality of neighborhood conditions and neighbor interactions (increasing in the neighborhood distribution of human capital), which produce human capital externalities. Adult health and economic status are determined by endowments inherited from parents, by own parental (p_i) expenditures, by local public expenditures (s) on amenities such as the quality of public schools, water and air quality, sanitation, pollution and environmental toxins, safety, quantity and quality of health case and social services, and by neighbors' parental expenditures ($p_{n(-i)}$). This last factor arises from behavioral spillovers operating via peer group and role model effects, and the effects of social complementarities.

Assume two periods of life, childhood and adulthood, and that children are born with an initial health stock, H_0. The change in health stock over time is determined by participation in health promoting activities and the influence of these activities on health, and the use of health stock. Following Case and Deaton (2003), the health evolution equation can be specified as:

$$H_{t+1} = \theta m_t + (1-\delta_t)H_t \ , \qquad (1)$$

where m_t is the quantity purchased of medical care or other health promoting activities, θ is the efficiency with which purchases create health, and δ_t is the rate at which health deteriorates at age t. Neighborhood conditions may affect the efficiency of private health investment (θ) as well as the quantity and quality of m_t.

The rate at which health capital depreciates with age in childhood is partly a biological process over which people do not have control, but it is also affected by parental investments and the quality of neighborhood environmental conditions. Parental investments in the child's health (e.g., medical care, nutritional diet, exercise equipment) and favorable neighborhood conditions during childhood produce more healthy adults. Assume health shocks experienced in early life alter the health production function in such a way that reduces the efficiency of health investment and increases the rate at which health deteriorates over time. This is consistent with recent evidence on the long-term effects of early life events (Johnson and Schoeni, 2006; Blackwell et al., 2001; Almond and Chay, 2003).

The rate of depreciation of the health stock increases with age and with the nature and intensity of use. The rate at which health capital depreciates with age in adulthood is partly a biological process, but it is also affected by the extent to which health capital is used in consumption and in work (Case and Deaton, 2003). As emphasized by Muurinen and Le Grand (1985), although all components of capital possessed by individuals—health, education, financial—are unequally distributed, the inequality in inherited health may be less than in other inherited stocks because of its distinctive, genetic component. As a result, the proportional share of health in total available capital is greater for individuals who are born to poorer families. And, because these components of human capital are substitutable, health capital will constitute a more important source of producing income ($y_t(H_t)$) and enjoying leisure (Muurinen and Le Grand, 1985).

The degree of persistence in educational attainment and earnings across generations affects the life course trajectory of health capital depreciation because it affects individual's opportunity sets with respect to adult living and working conditions. For example, in an economically segregated environment with low intergenerational economic mobility, the children of poor, less-educated parents residing in low-income neighborhoods with access to poorer quality schools will be more likely to reach adulthood with less accumulated human capital and will be less likely to qualify for well-paid jobs that do not require manual labor. Thus, they will work disproportionately in physically demanding blue-collar occupations, which will increase the rate of decay of their adult health capital (Muurinen and Le Grand, 1985; Case

and Deaton, 2003). Moreover, due to economic residential segregation, they will be more likely to live in low-income neighborhoods that are not supportive of good health (e.g., neighborhoods with high crime, pollution, poor health care system). Higher stress-related life events that result from these living and work conditions may be further exacerbated by an increase in behaviors such as smoking and binge drinking that, while hazardous in the long-run, relieve day-to-day stress in the short-run.

III. 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. This can arise from the endogeneity of residential location. That is, individuals and families choose where they live based on the characteristics they value (Tiebout, 1956). Although constraints such as racial discrimination and exclusionary zoning may affect residential choice, families that care more about their health and their children's health will be less likely to choose to live in an area with high crime, pollution, or a poor health care system. Because the complex characteristics that influence neighborhood choices are not well measured, we lack convincing evidence on the impact of neighborhoods on individual outcomes.

Moreover, the typical methods used to address endogeneity (e.g., fixed effect approaches) have significant limitations in this context. First, most health outcomes are a product of cumulative exposures to advantaged/disadvantaged environments spanning decades or exhibit long latent periods before problems manifest. Therefore, the connection between current neighborhood and current health may say little about the influence of neighborhoods factors over the life cycle. Because most methods for overcoming endogenous neighborhood choice are based on 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, so most year-to-year variations are noise.

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 the health of children and adults (Katz, Kling, Liebman, 2002; Leventhal and Brooks-Gunn, 2002). This evidence is consistent with the claim that neighborhood factors influence health status, at least in the short-run among poor families.

Among the studies that address endogeneity and self-selection using non-experimental methods, the most common approach is the use of instrumental variable techniques (e.g., Evans et al., 1992; Case and Katz, 1991; and McLanahan, 1996), where the exclusion restrictions are tenuous. An alternative non-experimental approach compares siblings who have been raised in different neighborhoods at different ages because their parents have moved (Aaronson, 1998; Plotnick and Hoffman, 1996). The key assumption is that the family effect is fixed, not time-varying. If, for example, families' preferences change as their children get older, and they become more interested in neighborhoods that are less risky for their children's health, then they might move to better neighborhoods which may in turn lead to better health outcomes for their kids. But if the underlying change in their preferences not only caused them to change neighborhoods, but also to spend more time encouraging their children to practice good health behaviors, then the neighborhood "effect" represents these other factors and not the causal effects of neighborhoods *per se*. Moreover, it is possible that sibling differences may aggravate the endogeneity problem, as has been discussed in the context of the labor market returns to schooling (Griliches, 1979; Bound and Solon, 1999).

Typical neighborhood studies also face the challenge of identifying and measuring relevant factors. The neighborhood qualities that matter may be hard to measure, or they may not be measured in enough spatial detail. This issue is analogous to the finding in the family background literature that sibling correlations in socioeconomic status far exceed what has been explained by any particular measured aspects of the siblings' shared background (Corcoran, Jencks, and Olneck, 1976).

IV. Overview of Empirical Approach

I exploit a unique feature of the PSID and adopt an approach used by Solon et al. (2000) to examine the role of childhood neighborhood factors on educational attainment. Specifically, the initial PSID sample in 1968 was highly clustered with most PSID families having several other sample families living on the same block, who have been subsequently followed over time. I follow the health experiences of those who were children in 1968, and thus who had reached mid-adulthood by 2005. This design allows a comparison of the similarity in childhood to mid-adulthood health between siblings who grew up together, versus unrelated individuals who grew up in the same narrowly defined neighborhood. This approach avoids the difficulty of defining neighborhood quality at the outset, and instead compares sibling correlations with neighbor correlations, placing an upper bound on the neighborhood influence (including effects emanating from school quality) and allowing a comparison of the relative magnitudes of child neighborhood versus family effects. The results are based on the estimation of four-level hierarchical random effects models (neighborhoods, families, individuals, over time) of health status.

The intuition behind this strategy is that if family background and residential community are important determinants of adult health outcomes, there will be a strong correlation between siblings in their health outcomes, as compared to two arbitrarily chosen individuals. Sibling correlations in health outcomes reflect the influence of all family and neighborhood background factors shared by siblings—measured and unmeasured—that may affect health outcomes, such as the socioeconomic status of parents, genetic traits, family structure, and neighborhood and school quality. And, if the neighborhood where the child grew up is important, it will show up as a strong correlation between neighboring children's subsequent health outcomes. The logic of the analytic approach is that if the neighbor correlations are substantial, there is a rationale for the further investigation of which neighborhood characteristics matter.

There are four primary reasons why the approach taken in this paper extends our understanding of neighborhood effects. First, in contrast to the experimental evidence and previous observational studies, the analysis examines effects over a very long time horizon. Second, instead of focusing on

contemporaneous neighborhood effects, I analyze the effects of neighborhood origins, which include indirect effects operating via the economic mobility process as well as cumulative exposure to neighborhood conditions that may vary over the life cycle. Third, I use the census block as the definition of neighborhood, a much smaller geographic area than previous studies utilize. Finally, I use estimates of neighbor correlations as an omnibus measure of the potential effects of neighborhood quality (including unmeasured characteristics), rather than initially focusing the analysis on particular observable neighborhood attributes.

I address endogenous selection of families into neighborhoods in several ways. First, I perform a sensitivity analysis to assess how large the unobservables would have to be, relative to observables, to invalidate the results. Second, I instrument for economic residential segregation with factors that are unlikely to be directly related to health outcomes but that should influence segregation patterns. This strategy posits that, while families may sort endogenously into neighborhoods, parents are constrained in their choices by the set of neighborhoods available in the area where they live. That is, it assumes the metropolitan area (MSA) is the exogenous housing market that parents face in their residential location choices. I isolate exogenous variation in the extent of MSA-level economic residential segregation and exploit this (between-MSA) variation, which results in differential likelihoods that poor families will reside in high poverty neighborhoods, to assess the role of neighborhood poverty.

The innovative research design and unique measures merged on from multiple data sources collected on aspects of neighborhood physical, service and social environments during childhood—including neighborhood poverty and crime, income and education, health insurance, race and residential segregation, school quality, parental expectations for child achievement, health behaviors, housing quality, connectedness to informal sources of support—help illuminate what lies along the "chain of causation" from childhood conditions to adult health outcomes.

V. DATA AND MEASURES

I analyze the restricted, confidential geocoded version of the PSID (1968-2005) with identifiers at the neighborhood block level. I then merge on an array of neighborhood and school information from multiple data sources that prevailed in the 1960s and 1970s when these children were growing up.

The PSID began interviewing a national probability sample of families in 1968 and re-interviewed them each year through 1997, when interviewing became biennial. All persons in PSID families in 1968 have the PSID "gene," which means that they are followed in subsequent waves. When children with the "gene" become adults and leave their parents' homes, they become their own PSID "family unit" and are interviewed in each wave. This sample of "split offs" has been found to be representative (Fitzgerald, Gottschalk and Moffitt, 1998). Moreover, the genealogical design implies that the PSID sample today includes numerous adult sibling groupings who have been members of PSID-interviewed families for nearly four decades.

The PSID used a "cluster sample" when it started to economize on interviewing costs. This design effect is typically a liability in statistical analyses because one has to account for non-independence across individuals within the same cluster. But for our purposes, the clustering provides the unique opportunity to examine health outcomes for adults who were childhood neighbors in 1968. Moreover, because all 1968 children are followed throughout their lives, I can examine the similarity in health status over the life course of both siblings and childhood neighbors.

I define the neighborhood of upbringing as the census block where the child lived in 1968.[1] This is a better definition of neighborhood than the typically-used census tract which consists of roughly 5,000 families. In the original wave of the PSID, a represented census block contains 4 sampled families, on average.[2] The PSID cluster design is discussed in greater detail in Solon et al. (2000).

Measurement of Health. The key childhood and adulthood health outcome is general health status (GHS) based on the question asked of household heads and wives (if present) at each wave between 1984 and 2005: "Would you say your health in general is excellent, very good, good, fair, or poor?" It

was asked of all family members in 1986.[3] 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. (For reviews, see Idler and Benyamini (1997) and Benyamini and Idler (1999).) GHS is frequently used and allows us to compare findings with those from related studies such as Case, Fertig, and Paxson (2005).

In addition, the PSID in 1999 and 2001 asked adults to recall their health in childhood (i.e., ages less than 17) and rate it as excellent, very good, good, fair, or poor. Empirical research findings have supported the validity and reliability of retrospective reports of childhood health conditions (Smith, 2008). Retrospective reports of global/overall childhood health (E/VG/G/F/P) have been shown to be highly correlated with reports of childhood activity-limiting health conditions (Elo, 1998).

In order to scale the GHS categories, I use the health utility-based scale that was developed in the construction of the Health and Activity Limitation index (HALex). (A discussion of various options for treatment of GHS is described in the Appendix). The HALex scores associated with GHS categories are based on the U.S. National Health Interview Survey (NHIS). A multiplicative, multi-attribute health utility model was used to assign scores and quantify the distance between the different GHS categories. The technical details of the scaling procedures are discussed at length elsewhere (Erickson, Wilson, Shannon, 1995; Erickson, 1998). Thus, using a 100-point scale where 100 equals perfect health, the interval health values associated with GHS 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. Consistent with previous research, the skewness and nonlinearity of this scaling is reflected in the fact that the "distances" between excellent health, very good health, and good health are smaller than between fair and poor health. This scaling is currently used by the National Center for Health Statistics to estimate health-related quality of life measures and years of healthy life (*Healthy People* 2000).

I estimate all regression models of health status using the interval regression method. While the HALex approach with interval regressions is superior to alternatives, as described in the appendix, I have

also estimated identical models using poor/fair health as the dependent variable in a multi-level logit model.[4] The substantive conclusions are unchanged.

The sample consists of PSID respondents who were children when the study began and who have been followed into adulthood; they were born between 1949 and 1968 and were between 0 and 18 years old in 1968. I obtain all available information on them for each wave, 1968 to 2005. In 2005, the oldest respondent is 57 and the youngest is 37.[5] (A discussion of sample attrition is presented in the Appendix).

The sample includes males and females and all analyses control for gender, given well-known differences in health status, health behaviors, and labor market outcomes for men and women. Due to the complexity of the health status changes for women during the childbearing years, I exclude self-assessed health status measures of women in the years they were pregnant.

To increase the sample size as well as the number of poor and black families, I 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 sample. I appropriately apply multi-level sample weights at the neighborhood and family levels to produce nationally-representative estimates.[6] The results are robust to the exclusion of the SEO sample (results available upon request).[7]

In addition to detailed measures of family economic resources and socioeconomic status during childhood, I include child health insurance coverage, birth weight, parental and neighborhood-level measures of expectations of child achievement, connectedness to informal sources of support, parental health behaviors (alcohol and smoking), and parental self-reports of neighborhood and housing conditions. The self-reports of housing/neighborhood conditions 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.[8]

Information from the PSID is supplemented with 1970-2000 census tract based measures—particularly, neighborhood poverty rate. I have also merged a set of school quality resource indicators for

1960-1980 (including per-pupil spending, class size) and measures of the extent of racial school segregation.[9]

The sample used to analyze adult health contains 51,082 person-year observations from 4,705 individuals from 1,935 families, 1,428 neighborhoods, and 270 counties. The mean age is 35, with age ranging from 20 to 57, and an average of 11 observations per person. A total of 1,383 families had at least two children, and a total of 357 neighborhoods contained at least two different unrelated families. The sample used to analyze child health contains 2,316 individuals from 1,280 families, 934 neighborhoods, and 210 counties. 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 sample descriptive statistics for all childhood family and neighborhood measures by race.

Descriptive Results. I begin by presenting nationally-representative estimates of the bivariate relationship between childhood-to-midlife health status and socioeconomic status in childhood (i.e., parental education, income, child health insurance coverage), and neighborhood quality in childhood (i.e., poverty and crime, race and residential segregation, and neighborhood housing quality).

The results shown in Figures 1-5 describe the extent of health disparities and how the child socioeconomic gradient in health evolves over the life course. These figures display the proportion of years in poor health as an adult as well as the age pattern of the health index. The age patterns of the conditional expectations are calculated using a Jianqing Fan (1992) locally weighted regression smoother, which allows the data to determine the shape of the function, rather than imposing a functional form. The differences presented are all statistically significant.

These figures reveal several patterns. First, the relationships between the parental income-to-needs ratio and child health and adult health exhibit nonlinearities, with children growing up in poverty experiencing significantly higher rates of problematic health throughout life. As shown in Figure 1, 31 percent of children who grow up in poverty did not possess excellent or very good health during childhood, in contrast to 20 percent among the near-poor and roughly 15 percent among the non-poor. Similarly, among children whose parents did not graduate from high school, 23 percent were not in

excellent or very good health during childhood, while that rate was 15 percent among children of more highly-educated parents.

Furthermore, the socioeconomic gradient in health widens over the life course, as the health deterioration rate is more rapid in adulthood among those who grew up in more disadvantaged child neighborhood and family environments. For example, 23 percent of adulthood years between ages 35 and 55 is spent in fair or poor health among those who grew up in poverty, while those rates are 13, 8, and 6 percent respectively among the near-poor, those whose parental-income-to-needs ratio is 2 to 3, and those growing up in affluent families (Figure 1). As shown in Figure 1, the health status of a 25 year old who grew up in poverty is roughly at the same level of health as a 50 year old who grew up in an affluent family.

This pattern is striking for health status by child neighborhood poverty. As shown in Figure 2a, 27 percent of children who grew up in high poverty neighborhoods lacked excellent or very good health during childhood, compared with 16 percent among children from low poverty neighborhoods. The health status of a 25 year old who grew up in a high poverty neighborhood (i.e., neighborhood poverty rate of thirty percent or higher) is roughly the same as that of a 50 year old who grew up in a low poverty neighborhood.

Johnson (2008) highlights substantial race differences in the incidence and duration of exposure to concentrated poverty over the life course. He documents high rates of immobility from poor neighborhoods over the life course, especially among African-Americans. The average black child spent ¼ of childhood years in high poverty neighborhoods, 1/3 of early-to-mid adulthood years in high poverty neighborhoods, and 15 percent of adulthood years lived in low poverty neighborhoods. This is in stark contrast to those rates for the average white child, who spent just 3 percent of childhood and adulthood years in high poverty neighborhoods, 80 percent in low poverty neighborhoods, and more than half of early-to-mid adulthood years in low poverty neighborhoods. These black-white differences in adulthood exposure to neighborhood poverty are largely accounted for by differences in the likelihood of being born

into a poor neighborhood, and to a lesser extent by differences in rates of upward and downward socioeconomic mobility over the life course (Johnson, 2008).

Building on that work, Figure 2b shows the proportion of adulthood years between ages 35 and 55 spent in fair or poor health by exposure to concentrated neighborhood poverty during childhood and by cumulative exposure up to mid-life. Those who spent their childhood residing in high poverty neighborhoods subsequently experienced one-quarter of their years between ages 35 and 55 in fair or poor health. There appears to be a dose-response in the simple bivariate relationship with duration of exposure to concentrated poverty and the likelihood of problematic health in adulthood. For example, among those who spent less than 20, 50 and 80 percent of their lifetime residing in high poverty neighborhoods, their corresponding proportion of adulthood spent in fair or poor health is roughly 10, 20 and 40 percent, respectively.

These differences by childhood neighborhood and family socioeconomic status likely contribute to the observed racial disparities in health. Thirty percent of blacks did not have excellent or very good health during childhood compared with 15 percent among non-Hispanic whites. As shown in Figure 3, black-white differences in health status widen significantly over the life course. By age 55 the health status of the average African-American is problematic, while the average health status of whites is good or very good (65 versus 85 on the health status index). A quarter of whites report themselves in excellent health well into their 50s; among blacks, the same points are reached before age 40. Figure 3 also shows that blacks who grew up in extremely segregated environments in childhood experienced worse health in childhood and adulthood, relative to blacks in less segregated areas.

Figure 4 presents child and adult health status by childhood neighborhood crime, and neighborhood housing plumbing and insulation problems. About 27 percent of children in high-crime neighborhoods were not in excellent or very good health during childhood compared with 16 percent among those in low crime neighborhoods. Similar patterns of differences exist in childhood health and adulthood health between individuals raised in environments with neighborhood plumbing and insulation problems, relative to individuals raised in environments that did not have these problems.

Figure 5 presents health status over the life course by birth weight and child health insurance status. The gaps widen over the adulthood years among individuals born low weight and those born at normal weight, and to some degree widen between those whose parents possessed private health insurance coverage for their children and those who lacked coverage (see Johnson and Schoeni, 2007). In particular, 27 percent of children born low weight lacked excellent or very good health during childhood compared with 15 percent among those born at normal weight; and by age 55, the gap in the health utility index widens to an average of 70 (among low birth weight individuals) compared to 85 (among normal birth weight individuals).

Of course, families who exhibit different health trajectories are different from one another in a multitude of ways that may also contribute to these differences in their children's adult health status. These bivariate relationships do not necessarily reflect causal relationships, particularly if the propensity for health problems is in part transmitted by nature from one generation to the next. I find significant correlations in health problems across generations. Parents who were in problematic health for the majority of their 50s and 60s were more likely to have children who experienced fair or poor health for a larger share of their adulthood years between ages 35 and 55 (relative to the children of parents who were in good health at these ages).

VI. ECONOMETRIC MODEL & ESTIMATION METHODS

In this section, I present an econometric model that illustrates the connections among sibling correlations, neighbor correlations, and regression analyses of neighborhood effects.[10] I begin by assuming the true model for health status is:

$$H_{sfn} = \alpha' X_{fn} + \beta' Z_n + \varepsilon_{sfn} \qquad (2)$$

where H_{sfn} denotes health status for sibling *s* in family *f* in neighborhood *n*, X_{fn} is the vector that includes all family characteristics (measured and unmeasured) that affect H_{sfn}, Z_n is the vector of all neighborhood characteristics that affect H_{sfn}, and ε_{sfn} is the error term that includes all individual-

specific factors that are not related to X_{fn} or Z_n. Note that for simplicity, at this juncture, I do not incorporate dynamics and potential interactions between family and neighborhood background effects or nonlinearities into the model, but rather assume a linear representation.

Due to the self-selection of advantaged families sorting into advantaged neighborhoods as discussed above, we expect the family background factors, X_{fn}, and the neighborhood background factors, Z_n, to be positively correlated. Because it is difficult to accurately measure every factor in X_{fn} and Z_n, the assumption that ε_{sfn} is uncorrelated with the observable measures of X_{fn} and Z_n will be violated, leading to biased estimates of neighborhood effects (β) and family background effects (α). Using the taxonomy of Manski (1993), it is not possible to distinguish the two types of "social effects" -- "endogenous effects" and "exogenous effects" -- from the nonsocial "correlated effects". Manski also demonstrates that it is not possible to distinguish the two types of social effects from each other. Therefore, the first goal of the analysis is focused on an overall assessment of the relative contributions of individual, childhood family and neighborhood effects on health in childhood and early-to-mid adulthood. I then analyze the relative contribution of a rich array of measured individual, family, neighborhood covariates to the total variation from each component, and test hypotheses about the effects of specific characteristics of childhood families and neighborhoods. The strategy for assessing the importance of contextual effects involves estimating the fraction of variation in health outcomes that lies between families and neighborhoods, to provide an upper bound on the possible effect of these contexts.

As demonstrated in Solon et al. (2000), using the additive model of the effect of family and neighborhood context in equation (2), the population variance of H_{sfn} can be decomposed as:

$$Var(H_{sfn}) = Var(\alpha' X_{fn}) + Var(\beta' Z_n) + 2Cov(\alpha' X_{fn,} \beta' Z_n) + Var(\varepsilon_{sfn}) . \quad (3)$$

Similarly, the covariance in H_{sfn} between siblings s and s' is:

$$Cov(H_{sfn}, H_{s'fn}) = Var(\alpha' X_{fn}) + Var(\beta' Z_n) + 2Cov(\alpha' X_{fn,} \beta' Z_n). \qquad (4)$$

The sibling correlation, $\text{cov}(H_{sfn}, H_{s'fn}) / \text{var}(H_{sfn})$, measures the proportion of the total variation in the health outcome due to factors shared by siblings. From (4) we see that siblings have correlated health outcomes because they have shared family and neighborhood backgrounds, corresponding to the first and second terms of (4), respectively. The sorting of families into neighborhoods is reflected in the third term. The sibling covariance then captures all measured and unmeasured factors shared by siblings that may affect health outcomes, such as the socioeconomic status of parents, genetic traits shared by siblings, family structure, as well as neighborhood effects.

Augmenting the estimation of sibling correlations with the estimation of neighbor correlations enables us to bound the relative importance of family and neighborhood factors. To see this, note the covariance between neighbors is:

$$Cov(H_{sfn}, H_{s'f'n}) = Cov(\alpha' X_{fn}, \alpha' X_{f'h}) + Var(\beta' Z_n) + 2Cov(\alpha' X_{fn,} \beta' Z_n) \qquad (5)$$

The last two terms in (4) and (5) are identical, so we expect the covariance between neighbors to be smaller than the covariance between siblings because siblings share both the same neighborhood and the same family. As Solon et al. (2000) state, if the covariance among neighbors is small relative to the covariance among siblings, the family effects, represented by the first term in (4), must be the main source of the covariance among siblings. Previous studies of sibling correlations do not disentangle family from neighborhood effects, exceptions being Solon et al. (2000, 2001), Raaum, Salvanes, and Sorensen (2002), and Oreopoulos (2003), none of whom examine health outcomes.

The neighbor correlation, $\text{cov}(H_{sfn}, H_{s'f'n}) / \text{var}(H_{sfn})$, measures the proportion of the variation in the health outcome that can be attributed to factors shared by individuals from the same neighborhood. In (5), the neighbor covariance consists of more than the variance in (effect-weighted) neighborhood characteristics given in the second term, and it should be viewed as an upper bound of the neighborhood influence on the covariance in H_{sfn} between neighbors. The first and third terms are both expected to be positive, leading to an upward bias. The first term represents the sorting of similar families into the same

neighborhoods, since neighboring children share similar family characteristics. Similarly, the third term also represents sorting, in that it captures sorting of disadvantaged families into disadvantaged neighborhoods. We see that positive sorting, $Cov(\alpha' X_{fn}, \alpha' X_{f'n}) \geq 0$ and $Cov(\alpha' X_{fn,} \beta' Z_n) \geq 0$, implies that $Var(\beta' Z_n) \leq Cov(H_{sfn}, H_{s'f'n})$.

Access to neighborhood identifiers and family characteristics in the same data enables us to tighten the upper bound on the neighborhood effects and also establish a lower bound on the family effects. First, it follows from (5) that the upper bound on the neighborhood effects can be made tighter by introducing observable family characteristics shared by the neighbors, and by subtracting that as an observable part of the first term of (5).

Following Solon et al. (2000) and Altonji (1988), I estimate the part of $\alpha' X_{fn}$ related to observable childhood family characteristics such as parental income, education, family structure, race, child health insurance coverage, birth weight, parental alcohol and cigarette use, parental expectations for child achievement, and housing quality. Let $\widetilde{X}_{fn}$ denote the observable subset of family characteristics with associated parameters $\hat{\alpha}$ estimated *within* neighborhoods. I then subtract the sorting component arising from the fact that similar families tend to cluster in neighborhoods,

$$Cov_{adj}(H_{sfn}, H_{s'f'n}) = Cov(H_{sfn}, H_{s'f'n}) - Cov(\hat{\alpha}' \widetilde{X}_{fn}, \hat{\alpha}' \widetilde{X}_{f'n}) \ . \qquad (6)$$

While this approach reduces the upper bound, it only captures the direct effect of neighborhoods on health outcomes. Consider the example where neighborhood factors allowed parents to obtain higher paying jobs, which in turn improved health status of children. In this case, the indirect neighborhood effect that works through employment and wages would be attributed to the family component and not the neighborhood component.

The tighter upper bound on neighborhood effects also implies a tighter lower bound on family effects. Specifically, the difference between the sibling correlation and the adjusted neighbor correlation

represents a lower bound of the magnitude of the effect of family background on the health outcome of interest. I refer to this as the "adjusted sibling correlation."

Four-Level Hierarchical Random Effects Interval Regression Model. I decompose both the variance of the level of health and the rate of health depreciation over time into the fraction that lies between neighborhoods, families, and individuals. In order to decompose both the total variation in the health level and the health depreciation rate, I estimate a four-level hierarchical random effects interval regression model. The data are hierarchical because I have multiple observations over time of individuals who are nested within families, which are nested within neighborhoods, and counties. Multilevel modeling techniques can accommodate the hierarchical and unbalanced structure of our data, non-independence of the (sometimes overlapping) pairs of siblings and neighbors, as well as the non-normality of health (Raudenbush and Bryk, 2002).

I begin by estimating the four-level hierarchical random effects model[11] given by

$$H^{*}_{tsfn} = (\beta_{0000} + \beta_{1000} * Age_{t}) + (\eta_{000n}) + (\phi_{00fn}) + (\delta_{0sfn}) + \varepsilon_{tsfn} \qquad (7)$$

I estimate these models separately at four distinct stages of the life cycle: childhood; young adulthood (ages 20-34); ages 35-44; and ages 45-57, in order to gain greater insight into the extent to which childhood family and neighborhood influence the trajectory of health over the life course. These unconditional baseline models also include controls for year of birth and quadratic terms for age (suppressed in the above notation).

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;
$s = 1, 2, \ldots, S_{fn}$ siblings 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 η_{000n}; the family random intercept coefficient is represented by ϕ_{00fn}; the individual random intercept

coefficient is represented by δ_{0sfn}; and ε_{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 var(η_{000n})= σ^2_{0n}, var(ϕ_{00fn})= σ^2_{0fn}, var(δ_{0sfn})= σ^2_{0sfn}, and var(ε_{tsfn}) = σ^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.[12]

Of primary interest is the decomposition of the variance of the level of health over the life course 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 random effect η_{000n}, and siblings are correlated because they share the random effects η_{000n} and ϕ_{00fn}.

We want to evaluate the health correlation between siblings at the same age, and evaluate the health correlation between neighbors at the same age. In this model, the sibling correlation and neighbor correlation in the level of health can be computed, respectively, as:

$$\rho_{\text{sibling,healthlevel}}(\text{age}) = \frac{(\sigma^2_{0n}) + (\sigma^2_{0fn})}{(\sigma^2_{0n}) + (\sigma^2_{0fn}) + (\sigma^2_{0sfn})}$$

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

The sibling correlation is between H^*_{sfn} and $H^*_{s'fn}$, 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.

I then use the estimated sibling and neighbor correlations at the four distinct stages of the life cycle, to construct an age-profile of sibling and neighbor health correlations. The age-profile of sibling and neighbor correlations provides insight into the nature and causes of the evolution of health inequality, and the relative roles of neighborhood and family background.

Health varies with age and gender. Because I did not want the estimates of sibling and neighbor correlations to reflect the influence of either of these two demographic factors, I adjusted for them in the baseline model by including gender and a quadratic specification of age as explanatory variables. Moreover, given that age affects health outcomes and that most same-aged children do not belong to the same family, it is important to control for age in the baseline model. Otherwise, between-family variance could mostly reflect differences between individuals of different ages.

VII. REGRESSION RESULTS

Unadjusted Sibling and Child Neighbor Correlations in Health over the Life Course

The unadjusted sibling and child neighbor correlations of health in childhood through mid-life are presented next. The estimates from the baseline four-level hierarchical random effects model that include only controls for age, year of birth, and gender are presented at four distinct stages of the life cycle (childhood; young adulthood (20-34); ages 35-44; and ages 45-57) in Table 1. The random effects estimates are all significant at each of the childhood neighborhood, family and individual levels. The baseline models measure the overall magnitude of variation at the neighborhood, family, and individual levels over the life course. The sibling and neighbor correlation estimates are based on the decomposition of variance over time into the fraction that lies between neighborhoods, families, and individuals. The age profile of the unadjusted sibling and neighbor correlations calculated from the baseline models are summarized in the first row of Table 4.

Sibling correlations are large throughout at least the first 50 years of life: the correlation in general health status in childhood is 0.56 and remains high at 0.63 through ages 45-57, suggesting that three-fifths of health disparities in adulthood may be attributed to neighborhood and family background influences. To assess the importance of the shared genetic component of health, I contrast full biological sibling correlations versus sibling correlations for step relations and adoptive ties. I find marginal evidence of an effect of "relatedness" on health status beyond living in the same household and neighborhood, as the standard deviation in adult health between half/adoptive ties is about 30 percent

higher than that for full biological siblings. However, small sample sizes of step and adoptive ties prohibit more definitive evidence (see Appendix for further discussion).

Sibling correlations by themselves cannot disentangle how much of the resemblance among siblings in health outcomes is due to the effects of family background and how much to the effects of neighborhood background. While the childhood neighbor correlations are smaller than the sibling correlations, they are substantial through middle-age. The childhood neighbor correlation in child health is 0.30; it increases to 0.43 on average during adulthood. Thus, knowing the adulthood health status of a childhood neighbor predicts nearly one-fifth of the adult health status of another childhood neighbor. By comparing the magnitudes of the sibling and neighbor correlations in adulthood health, the results indicate that at least half of the average sibling correlation in adulthood (0.6) may be attributable to neighborhood effects.

Estimating "Adjusted Neighbor Correlations". I next examine how much of the child neighbor correlations in health can be explained by the fact that families in a neighborhood tend to be similar as opposed to emanating from neighborhood effects *per se*. I estimate "adjusted neighbor correlations", which are net of the similarity arising from childhood neighbors having similar observed family background characteristics. To extract the impact of similar family backgrounds out of the neighbor correlation, I first estimate the following regression; for ease of exposition, here I omit the random effects terms that are included in the estimated model:

$$H^{*}_{tsfn} = \alpha_0 age_{tsfn} + \alpha_1 gender_{sfn} + \alpha_2' X_{\bullet\bullet fn} + \alpha_3' (\overline{X_{\bullet\bullet\bullet n}}) + \varepsilon_{tsfn}, \qquad (8)$$

where $X_{\bullet\bullet fn}$ is a vector of childhood family background characteristics including: average annual family income-to-needs ratio (based on the five-year average as reported in 1967-1972), parental education, parental family structure, race, child health insurance coverage (as reported in 1967-1972), parental annual expenditures on cigarette and alcohol consumption (based on the five-year average in 1967-1972), indicator for low birth weight, parental connectedness to informal sources of help, parental expectations

for child achievement, and housing plumbing and insulation problems. $\overline{X_{\bullet\bullet\bullet n}}$ 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 the within-neighborhood effects of family background characteristics. Using the within-neighborhood estimates of the family background effects of parental income, education, race, family structure, child health insurance coverage, parental health behaviors, birth weight, parental expectations for child achievement, parental 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 of the effects of family background characteristics can be taken as a conservative estimate of $\alpha' X_{fn}$ in equation (2).

I then estimate the between-neighborhood variance in $\hat{\alpha}' X_{fn}$ by estimating a hierarchical random effects model of $\hat{\alpha}' X_{fn}$ on neighborhood-level, family-level, and individual-level random effects. I then subtract the estimate of the between-neighborhood variance in $\hat{\alpha}' X_{fn}$ from the estimate of the overall between-neighborhood variance in H^*_{sfn}. Dividing the resulting quantity by $\hat{Var}(H^*_{sfn})$ yields a tighter upper bound on the proportion of $Var(H^*_{sfn})$ that can be attributed to child neighborhood effects. The estimates of "adjusted neighbor correlation" enable us to ascertain how much of the raw neighbor correlation is due to childhood neighbors having similar (observable) family background characteristics.

From the adjusted neighbor correlation estimates, I find that observable family sorting (controlling for a broad array of family background characteristics described above) does not seem to explain all the resemblance in adulthood health status among individuals who grew up in the same neighborhood. The adjusted neighbor correlation is roughly 10 percent lower than the unadjusted

neighbor correlation, suggesting that differences in neighborhood quality during childhood may account for up to 40 percent of adult health disparities.[13] I show in the next section that child neighbor correlations of this magnitude can imply large effects on subsequent adult health outcomes from changes in children's neighborhood environment.

However, without access to nationally-representative longitudinal data and the ability to identify the permanent component of health, the transitory component would have been captured in resultant point-in-time estimates, significantly diluting the relevant estimated sibling and child neighbor correlations in health over the life course. This result demonstrates the importance of correcting for measurement error, transitory fluctuations and unrepresentative homogenous samples, and parallels those found in the literature on the permanent component of adult earnings (Solon *et al.*, 1991).

MAGNITUDE OF EFFECTS OF CHILDHOOD FAMILY AND NEIGHBORHOOD FACTORS

What do these correlations mean in terms of the absolute size of the effects of family and neighborhood background (including effects emanating from school quality)? Estimates of the neighborhood random components (σ_n) indicate that childhood neighborhood quality has large, significant, and enduring effects on general health status over the life course. From the unconditional hierarchical random effects models and the estimated adjusted neighbor correlation estimates, I calculate how one would expect an individual's adult health status to change given a one standard deviation change in the index of child family environment, and the corresponding predicted change in adult health for a one standard deviation in the index of neighborhood environment. The results suggest that a one standard deviation change in the index of neighborhood environment is equivalent to roughly a 6-, 8-, and 9-point change in the health utility index at ages 20-34, 35-44, 45-57, respectively. This upper bound estimate on the potential scope of child neighborhood/school influences for health trajectories is substantial, as the mean of the index at age 40 is 84.2 and the average year-to-year rate of health deterioration in one's 40s is -0.4 (represented by the annual decline in the index).

I next investigate to what extent observable childhood family-, neighborhood- and school-level characteristics explain the estimated sibling and neighbor correlations at the four stages of the life cycle. Explicitly measuring the magnitude of variation in the effects of unmeasured factors allows an assessment of the importance (quasi-R^2) of the measured variables, *X,* in total variation at each level (e.g., measured vs. unmeasured neighborhood characteristics). In a subset of models, I include measures of the individual's own economic status in adulthood into the four-level hierarchical random effects model to examine the extent to which the resemblance of childhood neighbors' subsequent health in adulthood may be due to the similarity of their economic status in adulthood. These estimates are only suggestive because of endogeneity between contemporaneous health and SES. The results demonstrate what aspects and sources of current adult health disparities are missed using traditional models that focus on contemporaneous socioeconomic factors, without considering earlier life factors.

Parental income and neighborhood poverty are dimensions of childhood families and neighborhoods that are a key focus of the analysis. Growing up in a neighborhood with concentrated poverty may have consequences above and beyond those of growing up in a poor family because of the absence of positive role models, social isolation, weakened social institutions, unrelenting stress, inferior health care accessibility, and other factors.

I control for parental education, parental health status, birth order, whether child was low birth weight, born into a two-parent family, year of birth, and region of birth. I also use measures of parental expectations of children's educational attainment, residential segregation, parental connectedness to informal sources of help, parental aspirations/motivation and long-term planning, parental personality, habits and skills that were collected in the early years of the PSID. These factors may themselves be the product of growing up in a high poverty neighborhood and may represent pathways through which exposure to depressed neighborhood environments during childhood affect health trajectories later in life. However, controlling for this myriad of ways in which children who grow up in high poverty neighborhoods may differ from children who grow up in affluent neighborhood environments allows one

to generate a more conservative estimate of the effect of child neighborhood poverty itself, as well as shed light on the factors that affect adult health status.

Tables 2 and 3 contain the regression results in childhood and adulthood (ages 20-57), respectively, where the models include the raw age-adjusted race gap (column(1)), control for childhood family characteristics (column(2)), and control for childhood neighborhood, school, and family background characteristics (column (3)). Appendix Tables A2-A4 contain the results estimated separated at three distinct stages of adulthood—young adulthood (ages 20-34), ages 35-44, and ages 45-57—in order to examine the lifecycle profile of effects of childhood conditions. To conserve space, I integrate the discussion of the results contained in Tables 2-3 and Appendix Tables A2-A4 and summarize the sequential set of hierarchical random effects models estimated over the life course. The estimated effects of a one standard deviation change in neighborhood or family environment index provide a useful comparison to discuss effect sizes. One must use caution, however, with drawing causal inferences from these coefficient estimates. The estimates summarize the relationships between the health trajectory over the life course with various dimensions of neighborhood and family background.[14] The robustness of the results for causal inference is examined in detail in the final section of the paper.

The specification that includes the childhood family, neighborhood, and school-related factors is shown in column (3) of Table 3 and Appendix Tables A2-A4 (presented separately by lifecycle stage). The childhood school quality factors are included as controls but suppressed in the tables, since the focus of this paper is on family and neighborhood background. The school quality results are presented in detail in Johnson (2009).

Comparing the estimates in column (3) with the estimates in column (2) and the descriptive results shows the bias that occurs when estimating either the direct effects of child neighborhood factors on adult health without controlling for family background characteristics or the direct effects of child family characteristics that omit neighborhood characteristics. Controlling for neighborhood and school characteristics reduces the estimated health effects at ages 35-44 of parental income among those who grew up in poor and middle-class families by between 40-60 percent (as shown in column (2)-(3) of

Appendix Table A3, spline specification coefficient estimates on income-to-needs ratio change from 1.35 to 0.54 when in the range below the poverty line; and change from 2.16 to 1.32 when the income-to-needs ratio is in the range of 1-3). Similarly, all the child neighborhood coefficients decline significantly when family background controls are included (the models that include neighborhood variables without family variables are not shown). However, the estimated effects of various dimensions of neighborhoods remain large and significant with the inclusion of the extensive set of family background factors. Similarly, the effects of various dimensions of family background remain significant with the inclusion of the extensive set of child neighborhood characteristics.

The joint hypothesis that the neighborhood factors are empirically unimportant is clearly rejected; the *F*-statistic yields a p-value less than 0.01. Most of the effect of child neighborhood quality is due to three factors: concentrated neighborhood poverty, high crime, and poor housing quality. I find that blacks who grew up in more segregated neighborhoods and schools had significantly worse health in adulthood, both compared with whites and compared with blacks who grew up in areas where racial neighborhood and school segregation was less extreme.

Gaps in health between blacks and whites are large and exist at all stages in life. As shown in column (2) of Table 2 and column (1) of Table 3 (and Appendix Tables A2-A4), respectively, the general health status (GHS) index in childhood is 2.6 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 by middle-age equivalent to blacks (on average) reaching a level of health deterioration about 20 years prior to their white counterparts. That is, GHS is 9.3761 points lower for black adults at ages 45-57 (column (1) of Appendix Table A4), which is equal to roughly 20 years evaluated at an effect of age of -0.4146.

The raw black-white gap in health status during ages 35-57 is equivalent to about a one standard deviation change in the index of child neighborhood environment. For these birth cohorts, average childhood family and neighborhood environments between blacks and whites differ by as much as one standard deviation of the family/neighborhood environment index.

The estimates in column (4) of Table 2 and column (3) of Table 3 imply that black-white disparities in child and adult health would not exist (or would be small) if it were not for differences in childhood family, neighborhood and school quality factors between the racial groups (e.g., after controlling for both childhood family, and neighborhood and school quality factors, the black-white health gap is completely eliminated in childhood and is reduced by between 80-100 percent during adulthood).

Most prominent among the family background factors is family income, with substantially larger impacts in the lower tail of the distribution highlighting the negative effects of child poverty. For example, the results in column (3) of Appendix Table A4 indicate that a one-unit increase in the family income-to-needs ratio from half of the poverty line to 1.5 times the poverty line translates into a 6.4 point increase in adult GHS at ages 45-57 (0.5*11.5555+0.5*1.3056), which is equivalent to 15 years younger. Parental education, child health insurance coverage, and low birth weight are each strongly associated with adult health. These findings parallel those reported in Johnson and Schoeni (2007) for men.

Most salient among the childhood neighborhood factors is neighborhood poverty. Children who grow up in low poverty neighborhoods have a 1.698 higher child health index, relative to children who did not spend any years in such neighborhoods. An increase in the childhood neighborhood poverty rate from 10 to 20 percent is related to about a 3-point reduction in GHS in middle-age, and growing up in a high poverty neighborhood corresponds with a 9-point lower GHS score at ages 45-57, relative to being raised in a low poverty neighborhood.[15] This latter effect is equivalent to reaching a level of health deterioration roughly 22 years sooner for an individual raised in a high poverty compared to a low poverty neighborhood. For purposes of causal inference, the robustness of this result to alternative thresholds of selection on unobservables is analyzed in the following section.

Other dimensions of childhood neighborhood disadvantage had substantive, independent influences on the health trajectory, including high crime, parental and neighborhood-level average expectations for child achievement, neighborhood connectedness to informal sources of support (which may serve as a proxy for social cohesion), and neighborhood housing problems. These factors have

stronger relationships with health over time, with stronger links to adulthood health than childhood health and stronger links to health in middle-age relative to young adulthood; evidence suggestive that the linkages may be the result of how they influence the socioeconomic mobility process.[16] This age pattern also emerges for the relationship of childhood residential segregation and health among blacks. For example, as shown in column (3) of Appendix Table A4 for health status at ages 45-57, growing up in a high crime neighborhood reduces GHS by 2.4 points; both low parental expectations and neighborhood-level low expectations for child achievement are both independently associated with about a 3 point lower GHS (relative to college-bound expectations); neighborhood housing plumbing and insulation problems are each associated with about a 3.9 lower GHS; and for blacks, a 10-point increase in the black-white dissimilarity index is related to a 5.8 point reduction in GHS (independent of school segregation). Johnson (2008) demonstrates these factors also significantly influence mobility prospects, and explain part of black-white differences in rates of upward mobility from poor families. Taken together, the cumulative set of childhood family, neighborhood and school quality factors account for more than half of the neighborhood-level variance during adulthood (implied quasi-R^2 at the neighborhood level). That these measures account for less of the family-level variance may be the result of the fact that family-level influences include genetic/hereditary risk factors.

Parental health status may influence their offspring's subsequent adult health status due to inherited susceptibility to health problems, lower quality care of sick parents, or common socioeconomic factors across generations. In column (3) of Table 3, I include parental health status measures. The results discussed above are robust with and without the inclusion of these measures and demonstrate a significant intergenerational association of adult health status, independent of childhood factors. The results indicate that mother's health is more strongly associated with their offspring's adult health than is father's health.[17] The adult health status of individuals whose mothers were in fair or poor health throughout their 60s exhibited 4-5 point lower GHS scores, on average, relative to those whose mothers were in good health during their 60s. The magnitudes of the estimated effects of childhood neighborhood and family background factors largely persist with the inclusion of parental health status.

A substantial literature has investigated whether contemporaneous economic factors can account for the racial disparities in adult health. In column (4) of Table 3 (and Appendix Tables A2-A4), I re-examine this issue and find that over eighty percent of the black-white gap in health status remains after accounting for adult socio-economic factors (e.g., -6.5 in column (1) in comparison to -5.4 in column (4) of Table 3). This finding is similar to prior studies as reviewed by Wenzlow, Mullahy, and Wolfe (2004) and found in Johnson and Schoeni (2007).

The final model includes both all childhood family, neighborhood and school quality factors and contemporaneous adult socioeconomic status measures (adult neighborhood poverty rate, educational attainment, adult family income and earnings). As shown in column (5) of Table 3 (and Appendix Tables A2-A4), the racial differences in adult health can be accounted for by childhood family, neighborhood and school quality factors, while contemporaneous economic factors account for relatively little of this gap. Educational attainment was the main adult factor associated with adult health, while contemporaneous adult neighborhood poverty was only weakly related. The coefficient estimates on the childhood family, neighborhood and school quality factors are reduced to some extent with the inclusion of the adult socioeconomic measures, but the childhood factors remain large and significant.

Because there is potential causation running in both directions—from income to health and vice-versa—we, however, cannot disentangle from this analysis how much income affecting health contributes to this overall relationship. While one must use caution with attaching causal inferences from these results, this evidence, taken together with the sensitivity analyses and 3SLS-IV results to follow, provides strong support for the hypothesis that socioeconomic factors during childhood and/or other factors that affect an individual's economic status in adulthood affect an individual's health trajectory in adulthood. Separately identifying the causal pathways through which income affects health and health affects income over the life course has proven to be extremely difficult and is beyond the scope of this paper, but remains an important area for future research (Adda et al., 2003).

Table 4 presents the sibling and child neighbor correlations in health status over the life course as estimated from these hierarchical random effects models, for the unconditional, adjusted, and conditional

model estimates after controlling for childhood factors and adult socioeconomic status. As summarized, the unadjusted child neighbor correlation in adulthood health is about 0.4 and the adjusted neighbor correlation shows that the neighbor correlations were not driven by similarity of family background characteristics, but they reflect the combined influence of neighborhood and school quality effects. As well, after controlling for observable neighborhood, school, and family background factors, the similarity of childhood neighbors' adult health outcomes is less marked, and is estimated at between 0.24 and 0.27. The sibling correlation in adult health is roughly 0.6, and after controlling for the set of observable neighborhood, school, and family background factors, the similarity of siblings' health in adulthood is reduced to roughly 0.4. The broad array of available measures of child family and neighborhood characteristics, which are unique to the PSID, is a tremendous asset.

Sensitivity Analysis of Effects of Childhood Neighborhood Poverty

I conduct a sensitivity analysis to test the robustness of the estimated effects of childhood neighborhood poverty to selection bias due to an omitted variable. The goal 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. I use a novel empirical approach recently proposed by Altonji et al. (2005) and Krauth (2006). This sensitivity analysis allows one to determine the threshold of selection on unobservables, if any, at which neighborhood poverty during childhood 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

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

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 childhood on adult 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 background 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 neighborhood background captured in the PSID are likely to be the most important determinants of adult health. Thus, estimates obtained under the assumption of equal selection will be biased downwards.

Sensitivity of Estimated Effects of Neighborhood Poverty to Selection Bias.

Thus far, I have assumed exogeneity in child neighborhood residence. I now evaluate the robustness of these results to deviations from exogeneity. Table 5 presents the range of estimated coefficients and standard errors on childhood neighborhood poverty as a function of the ratio of selection on unobservables to selection on observables. I find that the effect of child neighborhood poverty on health status later in life remains large and significant even with a reasonably large amount of selection on unobservables. Even if the correlation between neighborhood poverty and unobserved outcome-relevant factors was assumed to be equal to the correlation between neighborhood poverty and observed-relevant factors, this does not eliminate the significant effect of child neighborhood poverty on health status later in life.

Instrumental Variables Analysis

As an alternative approach, following Cutler and Glaeser (1997), I account for the endogeneity of residential location using instruments based on political factors (public finance characteristics) and topographical features of the MSA. I use two measures as instruments for residential segregation by poverty status: (1) the statewide average share of local revenue that came from the state or federal

government in 1962; and (2) the number of rivers that run through an MSA. The logic of using the statewide average local tax burden as an instrument stems from the fact that a greater fraction of funding for local public goods coming from state or federal government sources (i.e., a lower reliance on local taxes) reduce the incentives of households within a city to sort by income, which should lead to less economic residential segregation in an MSA. Secondly, within an MSA that has historically been fragmented into many smaller jurisdictions by topographic barriers (namely, the number of inter- and intra-county rivers flowing through the MSA), it is more likely that significant differences in policies, local public good provision, and income levels persist between jurisdictions.

To motivate the instrumental variables strategy for neighborhood poverty, consider two hypothetical cities: City A, which exhibits high levels of residential segregation by economic status, and City B, which features more integrated residential location by SES. Although a poor family living in City A has a greater chance of living in a neighborhood with concentrated poverty than does one living in City B, children (in similar families) growing up in similar neighborhoods in the two cities face the same opportunities to invest in human/health capital that influence their longer-run health trajectories. The child living in a high poverty neighborhood in City B does not benefit from the fact that the city as a whole is more residentially integrated than City A. Nor is a child who lives in a low poverty neighborhood in City A hindered by the fact that the rest of City A is highly segregated. The effects of living in City A, therefore, are indirect, operating through the neighborhood in which one lives. That is, the neighborhood channels the direction through which broader, city-level forces trickle down to the individual level. (The model assumes MSA-level economic residential segregation has no association with children's subsequent health outcomes except by influencing neighborhood/school and county-level factors). These simplifying assumptions are supported to some degree by the fact that the hierarchical random effects models show that MSA-level effects over and above that of neighborhood/school and county-level factors appear to be relatively small empirically.

I control for the potential endogeneity of racial residential segregation by using an instrument that transcends individual decisions and captures historical patterns of migration by blacks during the Great

Migration out of the South that occurred between World War I and World War II. I instrument county racial residential segregation in 1970 with the net migration of blacks into that county between 1930 and 1940 (as a proportion of the total 1930 county population). The rationale is that as substantial numbers of people migrated to urban areas, the resultant racial composition of the destination neighborhoods depended in large part on the sheer number of new migrants relative to the total population and what share of the new migrants were black. La Ferrara and Mele (2006) use a similar instrumental variables approach in their analysis of the effects of segregation on public school expenditure.

We expect a mass influx of black migrants to increase racial residential segregation when their share of the total population is relatively small, while it may reduce segregation as the black share becomes very large. This historical migration pattern is documented in the work of Massey and Deaton (1993), Margo (1988), Collins (1997), and Vigdor (2002).

The specification of the model underlying the three-stage least squares with instrumental variables (3SLS-IV) estimates is as follows:

$$
\begin{aligned}
H^{*}_{tincm} &= \alpha_1 + \phi NhoodPov_{1970ncm} + \delta Race\,\mathrm{Re}sSeg_{1970c} * black + X_{incm}\beta_1 + \varphi Age_{tincm} + \varepsilon_{1tincm} \\
NhoodPov_{1970ncm} &= \alpha_2 + Pov\,\mathrm{Re}sSeg_{1970m} * FamPov_{1970incm} + X_{incm}\beta_2 + \varepsilon_{2ncm} \\
Pov\,\mathrm{Re}sSeg_{1970m} &= \alpha_3 + IntGovtGrant\%_{1962s} + Rivers_m + Rivers_m^2 + X_{incm}\beta_3 + \varepsilon_{3m} \\
Race\,\mathrm{Re}sSeg_{1970c} &= \alpha_4 + BlackMigrant\%_{1930-1940,c} + X_{inc}\beta_4 + \varepsilon_{4c}
\end{aligned}
\tag{9}
$$

where $NhoodPov_{1970ncm}$ is the neighborhood poverty rate in 1970; $PovResSeg_{1970m}$ is the MSA residential segregation by poverty status dissimilarity index in 1970; $FamPov_{1970incm}$ is an indicator for a child who grew up in a poor family; $RaceResSeg_{1970c}$ is the racial residential segregation dissimilarity index in the childhood county in 1970; $IntGovtGrant\%_{1962s}$ is the state average share of local expenditures funded through intergovernmental grants received from the state and federal government; $Rivers_m$ is the number of inter-county and intra-county rivers flowing through the MSA; $BlackMigrant\%_{1930\text{-}1940c}$ is the net migration rate of blacks to the county during the 1930s (defined as the number of (net) new migrants to the county between 1930-1940 who were black as a proportion of the total 1930 county population); X_{incm} is a vector of child and parental family characteristics; *i* indexes individuals, *n* indexes childhood

neighborhoods, *c* indexes county in childhood, *m* indexes metropolitan area in childhood, *t* indexes age of individual at which adult health outcome is measured, and *s* indexes state of birth. To account for the possible nonlinear effect of the historical net migration rates of blacks on subsequent residential segregation, I include squared and cubic terms of the net migration rate variable.[18] The regression models also include the following 1970 MSA- and county-level controls: ln(population), percent black, share employed in manufacturing, ln(median household income), and region of birth dummy indicators.

3SLS-IV Results of Effects of Childhood Neighborhood Poverty & Residential Segregation

Columns (1) and (2) of Table 6 shows the first-stage results of the MSA poverty dissimilarity index on the two sets of instrumental variables that meet the exclusion restriction, controlling for other MSA characteristics. Both the number of rivers running through the MSA and the share of local revenue that comes from intergovernmental sources are significantly related to economic segregation (as measured by the poverty dissimilarity index), each in the expected direction. Column (3) of Table 6 presents the first-stage results of the county's 1970 black-white dissimilarity index on the black net migration rate in the county during the 1930s, controlling for other county characteristics. The results indicate that inflows of black migrants led to increases in racial residential segregation when their share of the population was relatively small, but the effect diminished as their share increased, and actually led to a reduction in racial segregation when their share became large. In particular, the marginal effect of the net migration rate of blacks increased segregation until black migration rates reached levels around 17 percent after which point black migration significantly decreased racial residential segregation. The F-Statistic for the joint significance of the two sets of instruments (p-value $<.0.01$) are large enough to mitigate concerns regarding weak instrument bias (Staiger and Stock, 1997).

Table 7 displays the results from the 3SLS-IV specifications. In the 3SLS-IV specifications, both the main effect of MSA economic segregation and its interaction term with family poverty status are instrumented for with the two sets of instruments previously discussed. The results indicate that a one standard deviation increase in the MSA poverty dissimilarity index increases the likelihood that a poor family will reside in a high poverty neighborhood by about 25 percentage points and reduces the

likelihood that a non-poor family will live in a high poverty neighborhood by 23 percentage points (column (2)).

As shown in column (1) of Table 7, the 3SLS-IV estimates indicate that growing up in a high poverty neighborhood has substantial negative consequences on adult health. The magnitude and statistical significance of the effects of childhood neighborhood poverty are in line with the hierarchical random effects models presented in Table 3 (and Appendix Tables A2-A4). The results imply that growing up in a high poverty neighborhood accelerates health deterioration and leads to a 10-point reduction in adult health status. The implied effect size of living in a high poverty neighborhood during childhood translates into reaching a level of health deterioration roughly 15 years earlier than individuals who grew up in low poverty neighborhoods. The estimated effects of growing up in concentrated neighborhood poverty are comparable to the effects of a one standard deviation decrease in the index of neighborhood environment as implied by the unconditional hierarchical random effect model estimates. MSA-level economic segregation predicts worse health outcomes for lower-income families because of its influence on the likelihood of living in concentrated neighborhood poverty, and predicts better health outcomes for more affluent families.

The main effect of racial residential segregation and its interaction term with race are instrumented for with the variable capturing black historical migration patterns. As shown in column (4) of Table 7, the 2SLS estimates indicate that, for blacks, growing up in a more segregated community leads to significantly worse adult health—in particular, a 10-point increase in the black-white dissimilarity index results in a 2.2 point reduction in adult health status. Conversely, whites who grew up in more segregated communities had better adult health. Column (5) of Table 7 presents the 3SLS-IV model specification that includes both childhood neighborhood poverty and racial residential segregation. The estimated long-run health impacts of concentrated neighborhood poverty in childhood remain large and significant, and the results show significant negative health consequences for blacks when raised in highly segregated environments in childhood (though the latter estimated effects are smaller in magnitude

after including neighborhood poverty). The estimated effect of residential segregation for whites becomes small and insignificant after the inclusion of (instrumented) childhood neighborhood poverty.

The hierarchical random effects models and the 3SLS-IV models yield a consistent pattern of results. The IV results corroborate the evidence showing child neighborhood quality as a significant determinant of the health trajectory in adulthood, and also indicate negative health consequences for blacks of growing up in highly racially segregated communities. The instruments pass the standard test of overidentifying restrictions, as the p-value on the Sargan statistic for the specification in Table 10 is not significantly different from zero at conventional significance levels. A second test supports the validity of the instruments as they do not have a significant relationship with other key MSA characteristics, such as the poverty rate or poverty rate for blacks, after controlling for residential segregation by poverty status and other MSA characteristics. That is, other than through their relationship to segregation, the instruments do not appear to be related to MSA-level poverty rates. Thus, both tests are consistent with the view that the instruments are not capturing the effects of omitted variables that affect both adult health status and the topographic features of the MSA or historical sources of local revenue.

The final simulation builds on the intergenerational mobility literature and involves estimating the distribution of health status at age 40 by the percentile of the childhood neighborhood background component (which includes effects of school quality). These are based on the results from the hierarchical random effects model estimates of adult health. Table 8 displays the probability that an individual's health status at age 40 lies within the specified percentile ranges as a function of the percentile of their childhood neighborhood background component. The figures are based on the assumption that the neighborhood, family, and individual-level components of adult health are normally distributed. The estimates indicate that a child who grows up in a neighborhood at the 10^{th} percentile of the neighborhood quality distribution has roughly a 0.3 chance of falling in the bottom decile of the adult health distribution and has only a 0.15 chance of rising above the median.

VIII. Discussion and Conclusion

This paper provides comprehensive evidence that documents the extent and ways in which childhood family and neighborhood quality factors causally influence later-life health outcomes. I used correlations based on a nationally representative longitudinal sample of siblings and neighbors to estimate bounds on the possible causal effects of family and neighborhood background on general health status in childhood through mid life. Estimates based on four-level hierarchical random effects models show a significant scope for both family background (whether emanating from nature or nurture) and for neighborhood background (including school quality). While the within-family resemblance in adult health is significantly stronger than the within child neighbor resemblance, the child neighbor resemblance is substantial. The estimates indicate that three-fifths of adult health disparities may be attributable to family and neighborhood background, and suggest that disparities in neighborhood background account for between one-third and 40 percent of the variation in health status in mid life.

The neighbor correlations should be interpreted as upper bounds of the scope of neighborhood/school influences on subsequent health trajectories. However, the consistency of the results across the various empirical approaches employed—including hierarchical random effects models and 3SLS-IV models that address potential endogenous selection of families into neighborhoods—bolsters confidence in the findings of the lasting consequences of childhood conditions on adult health.

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 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 growing up in a high poverty neighborhood on adult health. The results reveal that even a large amount of selection on unobservable factors does not eliminate the significant effect of child neighborhood poverty on adult health status.

Childhood neighborhood quality factors play important roles in the intergenerational transmission of health status and influence both contemporaneous and future health outcomes (in part through their influence on the socioeconomic mobility process). I find that growing up in a neighborhood with concentrated poverty substantially increases the likelihood of having problematic health at mid-life, and I document how neighborhood quality influence later-life health in ways that cannot be reduced to the characteristics of the individuals and families themselves.

The analyses pay careful attention to the role of observable characteristics and unobserved heterogeneity. There is no evidence that the results were driven primarily by endogenous selection of families into neighborhoods. Taken together, these results are not consistent with a simple sorting on unobservables story. The fact that the effect of concentrated neighborhood poverty is weaker when the duration of exposure is brief, and that the 3SLS-IV models yielded very similar results, suggests that selection bias is not driving these results. These results are not consistent with the hypothesis that families predisposed to poor outcomes are selecting into poor neighborhoods, but rather are consistent with causal mechanisms such as neighborhood and school resources, or peer contagion in neighborhoods, schools or other locally shared contexts.

The pattern of results points toward true causal effects of childhood neighborhoods. The magnitude of the estimated effects of some dimensions of neighborhood quality are larger than estimates reported in previous research and, taken together, are larger than the impact of increasing parents' income by a comparable amount. Further research on the effects of particular neighborhood characteristics is warranted to identify the causal mechanisms through which concentrated neighborhood poverty effects operate.

Racial differences in adult health can be accounted for by childhood family, neighborhood, and school quality factors, while contemporaneous economic factors account for relatively little of this gap. These results challenge future research to further our understanding of the underlying processes that produce health disparities between different racial, ethnic, and socioeconomic groups. The results indicate that both family background and neighborhood quality during childhood serve as primary

gatekeepers of the intergenerational transmission of adult health status and play a large role in producing racial health disparities.

In order to assess the policy implications of this research, we need a better understanding of the pathways through which families, neighborhoods, and schools affect health. Peer group effects, role model effects, and contextual-complementarity effects each represent distinct influences under the umbrella of neighborhood effects, and each has different policy implications. This paper has focused on quantifying the potential overall magnitude of family, neighborhood, and school effects. Disentangling the causal sources of neighborhood effects is difficult (Manski, 1993; Moffitt, 2001), but investigation into the precise mechanisms of why neighborhoods matter is an important next step for future research.

---

[1] The 1968 addresses were geocoded to census block identifiers using GDT geographic mapping technologies. Census blocks are the smallest level of geographic precision reported by the Census Bureau and represent a narrow definition of neighborhood. Census block identifiers are defined for the entire U.S. in 2000. The average proportion of childhood spent growing up in the 1968 neighborhood was roughly two-thirds for the sample (further discussion provided in the Appendix).

[2] A contribution of this work pertains to measurement and the conceptualization of neighborhood contexts. I analyze the sensitivity of the main results to the level of aggregation, and compare how the results differ if I instead use larger neighborhood constructs for neighborhood groupings (e.g., census block vs. census tract vs. zip code or county).

[3] For a significant share of the respondents who were children in 1968, 1984 represents roughly the year in which they became heads of households as adults.

[4] The key shortcoming of an ordered logit or ordered probit regression is that probit and logit link functions are inadequate to model health due to the significant degree of skewness in the health distribution (i.e., the majority of people report themselves to be in good to excellent health). Van Doorslaer and Jones (2003) assess the validity of using ordered probit regressions to impose cardinality on the ordinal responses comparing it with the McMaster 'Health Utility Index Mark III' (HUI). They conclude "…the ordered probit regression does not allow for any sensible approximation of the true degree of inequality."

[5] The PSID maintains wave-to-wave response rates of 95-98%. Studies have concluded that the PSID sample of heads and wives remains representative of the national sample of adults (Gottschalk et al., 1999; Becketti et al., 1997).

[6] To be eligible for the SEO sample, households had to have income that was below two times the poverty line, which could be problematic for our purposes because two neighboring families could enter that component of the PSID only if they had sufficiently low income. However, due to the significant degree of residential segregation by income, I find evidence that the typical neighbor of a low-income family was also low income; thus, in practice this does not present significant within-neighborhood sample selection bias problems. In particular, in the 1968 SRC sample, the average family with income less than two times the poverty line (in that year) lived in neighborhoods in which neighbors' average income was also among the bottom third of the income distribution. Similarly, using larger national samples geocoded to the census block, Hardman and Ioannides (2005) find that among the poorest 30 percent of households, roughly 75 percent live in neighborhoods in which neighbors' median income is also among the poorest 30 percent of households.

[7] The ability to conduct analyses within families and between neighboring families is a unique feature of the study. Because the study is among the first to report evidence of sibling correlations in health over the life course, I include all neighborhoods to increase the effective sample size for the sibling correlation estimates. Results on the sub-sample of neighborhoods containing children from at least two different families yielded very similar magnitudes of sibling and child neighbor correlations in health outcomes to those reported (results available upon request).

[8] These measures serve as proxies of neighborhood quality as this information was only collected in the 1975 survey and may not reflect the characteristics of the 1968 neighborhood due to residential mobility over the period. However, 1968 families in the PSID tended to move to neighborhoods that had observable neighborhood characteristics that were similar to their previous residential location (Kunz et al., 2001).

[9] The sources for the school data are detailed in Johnson (2009).

[10] This discussion follows Solon et al. (2000).

[11] Maximum-likelihood (ML) estimates based on a numerical integration procedure were computed using aML statistical software (Lillard and Panis, 2003). Some estimates were computed using the gllamm macro in Stata (Rabe-Hesketh et al., 2000). The numerical evaluation of the unconditional-likelihood function uses Gaussian quadrature. I use 10-point quadrature for each level.

[12] Because neighborhoods are nested within counties, I also estimated five-level hierarchical models, where the hierarchical levels represented counties, neighborhoods, families, and individuals over time. This provides a robustness check to ensure that the childhood neighborhood random effects components were not primarily driven by effects operating at higher geographic levels of aggregation (i.e., above the school district level). However, those models did not significantly improve the fit and the between-county random effects component was not statistically significant. This supports the use of the four-level hierarchical model.

[13] Robustness of these baseline results on two dimensions was considered. First, I examined alternative specifications of health status: a) the dichotomous variable poor/fair versus good/very good/excellent, and b) the Health and Activity Limitation Index that attributes scores to combinations of self-assessed health and activity functional limitation categories. The overall patterns of the neighbor and sibling correlations were qualitatively similar for these outcomes and the preferred health status measure. I also sought to identify a health status measure that is largely determined by genetic factors. If such an outcome could be identified, one would not expect it to be correlated among neighbors if in fact correlation was not spurious. Height is largely determined by genetic factors and therefore most likely is not causally influenced by neighborhood characteristics. Re-estimating the models with height as the dependent variable, I find that the neighbor correlation is negligible and statistically insignificant, as expected. This suggests that the substantial neighbor correlations for GHS are not due to spurious sorting of individuals with similar characteristics.

[14] I examined alternative functional forms of the key explanatory variables to best fit the data. As a result, the functional forms vary slightly between the childhood health and adulthood health models.

[15] To facilitate interpretation of marginal effects, I converted the units of county racial residential segregation dissimilarity index so that a 1-unit change represents a 10-point change in the dissimilarity index. Similarly, a one-unit change in the spline specification for neighborhood poverty represents a 10-point change (e.g., change in neighborhood poverty rate from 10% to 20%).

[16] I control for year of birth, as it is important to distinguish these life cycle effects from birth cohort effects. I find the same age pattern emerges when I restrict the sample to the subset of individuals who were born in the 1950s for whom we observe health in one's 30s, 40s, and into the 50s; this robustness check was performed to ensure the estimated age profile of effects was not instead capturing birth cohort effects.

[17] The parental health status measures included are the proportion of years spent when the parent was in their 60s in which they were in fair or poor health (based on self-reports of GHS). These ages correspond to years in life when rates of health deterioration typically begin to accelerate and for which parental health status information in the PSID is most plentiful for this older cohort. Similar results were found when alternatively using parental health status measured in their 50s. Dummy indicators are included in these models for missing information on parental health status.

[18] The net migration rate of blacks between 1930-1940 ranges from -0.12 to 0.27 across the different counties.

Appendix

PSID sample

The sample consists of PSID respondents who were children when the study began and who have been followed into adulthood; they were born between 1949 and 1968 and were between 0 and 18 years old in 1968. I obtain all available information on them for each wave, 1968 to 2005. In 2005, the oldest respondent is 57 and the youngest is 37.

The first wave of PSID interviewing in 1968 included 2,856 families containing 8,710 children 0-18 years old. 167 of these children died by 2005. These individuals are included in the analyses for the years they are observed alive. Any selective attrition with respect to mortality is likely to lead to an understatement of the impact of adverse childhood conditions, if those who suffer premature death disproportionately grow up in the more disadvantaged childhood family and neighborhood environments. I estimated mortality models, but there were too few deaths to precisely estimate any relationships. Of these 8,710 children, 5,628 had at least one valid report of health status in adulthood. Adult GHS is based on reports for PSID heads and wives/"wives" (1984-2005) as well as all family members in 1986. A small minority of respondents lacked valid addresses and were not able to be matched to neighborhoods in the geocode file—these cases were disproportionately located in rural areas. The selection criteria maximize the number of adult person-year observations of adult health and, in the vast majority of cases, child neighbors grew up within eight years of one another.[1] The resultant sample used in the analyses contains 4,705 individuals that came from 1,935 different childhood families, 1,428 neighborhoods, and 270 counties. Data are combined across all waves for each person, and in total there are 51,082 person-year observations, or an average of 11 observations per person, for the analyses of adult health.

While the decline in the initial sample of 46 percent is substantial, it is low given the long period over which these children and their families are followed. For example, among the 17,287 newborns participating in the 1970 British birth cohort sample, 6,454 (37 percent) were not interviewed (i.e., were not in the "observed sample") in 1999/2000 when they were 30 years old. Moreover, studies have concluded that the PSID sample of heads and wives remains representative of the national sample of adults (Fitzgerald, Gottschalk, and Moffitt, 1998a; Becketti et al, 1988), and that the sample of "split offs" is representative (Fitzgerald, Gottschalk and Moffitt, 1998b). The 95-98% wave-to-wave response rate of the PSID makes this possible.

Table A0 contains a summary of the variable definitions and data sources of all key measures used in the analyses, the year(s) of data collection, and the relevant survey questions used to construct

[1] The selection criteria was guided by both sample size considerations as well as the need to ensure the resulting sample comprised children who grew up in neighborhoods during comparable periods (e.g., I did not want to compare adult outcomes of neighboring children who were more than eight years apart, as neighborhood change over the period could cause child neighbor correlations to be downwardly biased).

these measures. Table A1 reports descriptive statistics for the samples used in the models of adult health status both for the full sample and separately by race. The substantial race differences in childhood family and neighborhood characteristics are highlighted in this table.

Income is the total for the family in which the child lives, and it is measured from the five-year average for the years 1967-1972. All dollar values are expressed in 1997 dollars using the CPI-U. The parental income measure is specified as the income-to-needs ratio and I explore nonlinearities in effects at the bottom of the income distribution (child poverty).

Child health insurance coverage is measured through information collected in the first five waves of the PSID (1968-1972) on whether the parent (head of household) had access to private health insurance coverage and if so, whether the entire family was covered. I include an indicator variable defined as lack of private health insurance coverage in childhood years during 1968-1972. Lack of private health insurance may discourage preventive medical care use. For those who lacked private coverage for their children, the data suggest that public health insurance coverage was utilized to some extent, but there were not enough individuals in the sample who persistently lacked public and private insurance during these childhood years to define "no public or private insurance during childhood" as an additional category.

The parental health status measures included are the proportion of years spent when the parent was in their 60s in which they were in fair or poor health (based on self-reports of general health status). These ages correspond to years in life when rates of health deterioration typically begin to accelerate and for which parental health status information in the PSID is most plentiful for this older cohort. Similar results were found when alternatively using parental health status measured in their 50s. Dummy indicators are included in all regression models for missing information on parental health status.

Health Index

A number of previous studies using surveys have demonstrated that a change in GHS from fair to poor represents a much larger degree of health deterioration than a change from excellent to very good or very good to good (e.g., Van Doorslaer and Jones, 2003; Humphries and Van Doorslaer, 2000). More generally, this research has shown that health differences between GHS categories are larger at lower levels of GHS. Thus, assuming a linear scaling would not be appropriate.

To analyze health disparities in the presence of a multiple-category health indicator, three alternative approaches have been used, each with its own set of advantages and disadvantages. The most common and simplest approach is to dichotomize GHS by setting a cut-off point above which individuals are said to be in good health (e.g., excellent/very good/good vs. fair/poor). The disadvantage of this approach is that it does not utilize all of the information on health. Additionally, it uses a somewhat

arbitrary cut-off for the determination of healthy/not-healthy, and the measurement of inequality over time can be sensitive to the choice of cut-off (Wagstaff and Van Doorslaer, 1994).

A second approach is to estimate an ordered logit or ordered probit regression using the GHS categories as the dependent variable, and rescale the predicted underlying latent variable of this model to compute "quality weights" for health between 0 and 1 (Cutler and Richardson, 1997; Groot, 2000). The key shortcoming of this approach is the probit and logit link functions are inadequate to model health due to the significant degree of skewness in the health distribution (i.e., the majority of a general population sample report themselves to be in good to excellent health). Van Doorslaer and Jones (2003) assess the validity of using ordered probit regressions to impose cardinality on the ordinal responses comparing it with a gold standard of using the McMaster 'Health Utility Index Mark III' (HUI).[2] They conclude "…the ordered probit regression does not allow for any sensible approximation of the true degree of inequality."

The third approach, adopted first by Wagstaff and Van Doorslaer (1994), assumes that underlying the categorical empirical distribution of the responses to the GHS question is a latent, continuous but unobservable health variable with a standard lognormal distribution. This assumption allows "scoring" of the GHS categories using the mid-points of the intervals corresponding to the standard lognormal distribution. The lognormal distribution allows for skewness in the underlying distribution of health. The health inequality results obtained using this scaling procedure have been shown to be comparable to those obtained using truly continuous generic measures like the SF36 (Gerdtham et al., 1999) or the Health Utility Index Mark III (Humphries and van Doorslaer, 2000) in Canada, but has not been validated as an appropriate scaling procedure using U.S. data. The disadvantage of this approach is it inappropriately uses OLS on what remains essentially a categorical variable and does not exploit the within-category variation in health. This is particularly problematic for the analysis of health dynamics over a relatively short time horizon. Ignoring within-category variation in health will cause health deterioration estimates to be biased and induce (health) state dependence because within-category variation increases when going down from excellent to poor health.

Several surveys have been undertaken that contain both the GHS question and questions underlying a health utility index. In this paper, we adopt a latent variable approach that combines the advantages of approaches two and three above, but avoids their respective pitfalls. Specifically, utilizing external U.S. data that contain both GHS and health utility index measures, we use the distribution of health utility-based scores across the GHS categories to scale the categorical responses and subject our

[2] The McMaster Health Utility Index can be considered a more objective health measure because the respondents are only asked to classify themselves into eight health dimensions: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain. The Health Utility Index Mark III is capable of describing 972,000 unique health states (Humphries and van Doorslaer, 2000).

indicators to the transformation that best predicts quality of life. This scaling thus translates our measures into the metric that reflects the underlying level of health. Specifically, using a 100-point scale where 100 equals perfect health and zero is equivalent to death, the interval health values associated with GHS are: [95, 100] for excellent, [85, 95) for very good, [70,85) for good, [30,70) for fair, and [1,30) for poor health.

Interval Regression Model. The method assumes that underlying the categorical empirical distribution of the responses to the GHS question is a latent, continuous health variable. I estimate interval regression models using the aforementioned values to scale the thresholds for GHS, where interval regression models are equivalent to probit models with known thresholds.

The measure of health status has categorical outcomes excellent (E), very good (VG), good (G), fair (F), and poor (P). The model can be expressed as

$H_i =$ 1 (E) if $95 \le H_i^* \le 100$ = perfect health
2 (VG) if $85 \le H_i^* < 95$
3 (G) if $70 \le H_i^* < 85$
4 (F) if $30 \le H_i^* < 70$
5 (P) if $1 \le H_i^* < 30$,

where H^* is the continuous latent health variable and is assumed to be a function of socio-economic variables x:

$$H_i^* = x_i\beta + v_i \ , \quad v_i \sim N(0, \sigma_v^2).$$

Given the assumption that the error term is normally distributed, the probability of observing a particular value of y is

$$P_{ij} = P(H_i = j) = \Phi\left(\frac{\mu_U - x_i\beta}{\sigma_v}\right) - \Phi\left(\frac{\mu_L - x_i\beta}{\sigma_v}\right) \ ,$$

where j indexes the categories, $\Phi(\bullet)$ is the standard normal distribution function, and μ represent the threshold values previously discussed. Because the threshold values are known, it is possible to identify the variance of the error term σ_v^2. Because I use the health utility-based values to score the thresholds for GHS, the linear index for the interval regression model is measured on the same scale. This scaling thus translates the measures into the metric that reflects the underlying level of health. With independent observations, the log-likelihood for the interval regression model takes the form:

$$\log L = \sum_i \sum_j H_{ij} \log P_{ij} \quad ,$$

where the H_{ij} are binary variables that are equal to 1 if $H_{ij} = j$. This can be maximized to give estimates of β.

Additional Considerations

Residential mobility. Because siblings typically share similar family environments for longer periods than neighboring children share neighborhood environments, we expect lower correlations for neighbors than for siblings. That is, I estimate the correlation between individuals who were childhood neighbors in 1968, but if 1968 neighborhood is a poor proxy for longer-run childhood neighborhood environment, my estimates of the influence of childhood neighborhoods may be subject to a downward errors-in-variables bias. The potential for measurement error is a serious concern since residential mobility is common in the US, especially among families with younger children. Thus, children sharing a neighborhood at any given point in time may have quite different residential histories. However, Kunz et al. (2001) investigate this issue using the PSID and find a high degree of persistence in the quality of children's neighborhood environments. They estimate the autocorrelations of observed neighborhood characteristics inhabited by the PSID children, and find the autocorrelation between the average of log mean income during the 1970-1980 period and each single year value is at least 0.90 for every year and averages 0.94.

I find that the average proportion of childhood spent growing up in the 1968 neighborhood was roughly two-thirds for the sample. To investigate the potential impact of residential mobility further on the findings, I re-estimated all health status correlations on the sample of children who had lived in their 1968 home since at least 1963.[3] The correlations among this sample were similar to the ones reported in the paper. Solon et al (2000) found that neighbor correlations in education were not sensitive to similar sample restrictions. Therefore, the evidence tends to suggest that residential mobility is not significantly influencing the estimated neighbor correlations.

Sibling Correlations by Relatedness. I explored the relative impact of shared household environment versus shared genetic unmeasured components using "relatedness" of children in the household. I began by ignoring neighborhoods and investigated whether the "relatedness" of children in the household affects the degree of correlation (heterogeneity) in their health outcomes, i.e., whether more related children have more similar health outcomes. I used a slightly restricted dataset—all children in the household must be of the same "relatedness" and the other children in the household are omitted so that the household correlation structure is clear and simple. That is, each family contains only fully biological siblings, or only half/step or "unrelated" siblings. Only a very small number of children are

[3] I experimented with introducing heteroscedasticity into the multilevel model at the individual level as a function of the percent of childhood years spent growing up in the 1968 neighborhood. I initially thought this would be a good idea because we would expect the within-neighborhood variance to be smaller if most or all individuals grew up in the 1968 neighborhood for their entire childhood. However, upon further reflection, the selection bias issues of who moves outweigh the reduction in errors-in-variables bias, and thus does not justify modeling heteroscedasticity nor keeping only individuals who grew up in the 1968 neighborhood for their entire childhood.

excluded by this restriction, so the effect on estimates should be negligible. Among the 1,257 families containing two or more children, 822 had at least two full biologically-related children.

For this purpose, I estimated a first set of models that included only family- and individual-level unmeasured components, and included only families with two or more children because households with one child contribute no information about family components. First, the model was estimated with the same household component (magnitude of variation) for every household. I then allowed the "household" component to be different for the two types of households with multiple children—fully biological siblings versus half/step siblings and adopted children (i.e., "unrelated" children who grew up in the same household). The latter models allow the between-sibling and between-family random effects variance components to differ by biological relatedness (jointly estimated so covariate effects are the same). I tested whether the magnitude of variation of their household components are different (larger or smaller).

The result is that the degree of heterogeneity is significantly different among the two types of households (at the 10 percent level). As expected, the between-sibling variance is smaller for fully biological siblings versus half-siblings and "unrelated" siblings, reflecting the influence of the genetic component. However, because the subset of fully biological siblings are a more homogenous and advantaged subsample along socioeconomic dimensions, the between-family variance component for full siblings is also smaller than that for half-siblings and unrelated children (results available upon request). As a result, the estimated sibling correlations do not differ significantly between full biological siblings and other siblings (half/adopted).

I then incorporated the significance of both the childhood family and neighborhood components in health status. Similar sibling correlation estimates are found once neighborhood components are introduced. I conclude that there is marginal evidence of an effect of "relatedness" on health status beyond living in the same household and neighborhood, though small samples of step and adoptive ties prohibit more definitive conclusions. The remainder of the analysis does not take into account the relatedness of children in the same household.

REFERENCES

Aaronson, Daniel. 1998. "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes". *The Journal of Human Resources*. Autumn, 33(4): 915-946.

Adda, Jerome, T. Chandola, and M. Marmot. 2003. "Socio-economic Status and Health: Causality and Pathways". *Journal of Econometrics*. 112:57-63.

Almond, Douglas, and Kenneth Y. Chay. 2003. "The Long-Run and Intergenerational Impact of Poor Infant Health: Evidence from Cohorts Born During the Civil Rights Era," mimeo.

Altonji, Joseph, Todd Elder and Christopher Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, Vol. 113(1):151-184.

Becker, G. and N. Tomes. 1986. "Human Capital and the Rise and Fall of Families". *Journal of Labor Economics*. July, 4(3): 1-39.

Blackwell, Debra, M. Hayward, and E. Crimmins. 2001. "Does Childhood Health Affect Chronic Morbidity in Later Life?" *Social Science and Medicine*. 52:1269-84.

Bound, J. and G. Solon. 1999. "Double Trouble: On the Value of Twins-Based Estimation of the Return to Schooling". *Economics of Education Review*. 18:169-82.

Card, David and Jesse Rothstein. 2007. "Racial Segregation and the Black-white Test Score Gap." *Journal of Public Economics* 91(11-12):2158-2184.

Case Anne and L. Katz. 1991. "The Company you Keep: the Effects of Family and Neighborhood on Disadvantaged Families". National Bureau of Economic Research Working Paper no. 3705.

Case, Anne and Angus Deaton. 2003. "Broken Down by Work and Sex: How our Health Declines". NBER Working Paper. Center for Health and Wellbeing, Princeton University.

Case, Anne, Darren Lubotsky and Christian Paxson. 2002. "Economic Status and Health in Childhood: The Origins of the Gradient". *American Economic Review*. December, 92(5): 1308-1334.

Case, Anne, Angela Fertig and Christina Paxson. 2003. "The Development of the Gradient: Earnings and Health Dynamics". Unpublished manuscript, Center for Health and Wellbeing, Princeton University.

Chandra, Amitabh and Jonathan S. Skinner. 2003. "Geography and Racial Health Disparities". Unpublished manuscript. Department of Economics and Center of Evaluative Clinical Sciences, Dartmouth Medical School.

Corcoran, M., R. Gordon, D. Laren, and G. Solon. 1992. "The Association Between Men's Economic Status and their Family and Community Origins". *Journal of Human Resources.* 27: 575-601.

Corcoran, M., C. Jencks, and M. Olneck. 1976. "Intergenerational Transfers of Inequality: The Effects of Family Background on Earnings". *American Economic Review*. 66(2):430-35.

Crane J. 1991. "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping out and Teenage Childbearing". *American Journal of Sociology.* 96: 1226-1259.

Cutler, David and Edward Glaeser. 1997. "Are Ghettos Good or Bad? *Quarterly Journal of Economics* 112: 827-872.

Cutler, David, E. Richardson, and T. Keeler. 1997. "Measuring the Health of the U.S. Population". *Brookings Papers on Economic Activity. Microeconomics*. 1997:217-282.

Deaton, Angus and Christina Paxson. 1998. "Health, Income, and Inequality over the Life Cycle". In *Frontiers in the Economics of Aging*, David Wise, Ed. Chicago: University of Chicago Press. 431-462.

Duncan, Greg, J. Boisjoly, and K. M. Harris. 2001. "Sibling, Peer, Neighbor, and Schoolmate Correlations as Indiciators of Importance of Context for Adolescent Development". *Demography*. August 38 (3):437-447.

Durlauf, Steven. 2004. "Neighborhood Effects," in *The Handbook of Regional and Urban Economics, Volume 4: Cities and Geography*, edited by V. Henderson and J.F. Thisse. Elsevier Science/North Holland.

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

Erickson, Pennifer. 1998. "Evaluation of a Population-based Measure of Quality of Life: the Health and Activity Limitation Index (HALex)". *Quality of Life Research*. 7:101-114.

Erickson, Pennifer, R. Wilson and I. Shannon. 1995. "Years of Healthy Life". *Healthy People 2000: Statistical Notes*.7:1-14.

Evans W.N., W. Oates and R.M. Schwab. 1992. "Measuring peer group effects: a study of teenage behavior". *Journal of Political Economy.* 100:966-91.

Fan, Jianqing. 1992. "Design-Adaptive Nonparametric Regression". *Journal of the American Statistical Association*. 87:998-1004.

Fitzgerald, J., P. Gottschalk and R. Moffitt. 1998. "An analysis of sample attrition in panel data: The Michigan Panel Study of Income Dynamics". *Journal of Human Resources.* 33(2):251-99.

_____. 1998b. "The impact of attrition in the Panel Study of Income Dynamics on intergenerational analysis". *Journal of Human Resources.* 33(2):300-44.

Gerdtham, U. G., M. Johannesson, L. Lundberg and D. Isacson. 1999. "A Note on Validating Wagstaff and van Doorslaer's Health Measure in the Analysis of Inequalities in Health". *Journal of Health Economics*. 18:117-124.

Geronimus, Arline T., John Bound, Timothy A. Waidmann, Cynthia G. Colen and Dianne Steffick. 2001. "Inequality in Life Expectancy, Functional Status, and Active Life Expectancy Across Selected Black and White Populations in the United States". *Demography*. May, 38(2): 227-251.

Griliches, Z. 1979. "Sibling models and data in economics: beginnings of a survey". *Journal of Political Economy*. 87:S37-64.

Groot, W. 2000. "Adaptation and Scale of Reference Bias in Self-Assessments of Quality of Life". *Journal of Health Economics*. 19:403-420.

Hoxby, Caroline M. 2000. "Does Competition Among Public Schools Benefit Students and Taxpayers?" *American Economic Review*, 90(5): 1209–38.

Humphries, K. and E. Van Doorslaer. 2000. "Income-related Inequalities in Health in Canada". *Social Science and Medicine*. 50:663-671.

Jencks, Christopher and Susan E. Mayer. 1990. "The Social Consequences of Growing Up in a Poor Neighborhood," in *Inner-City Poverty in the United States*, edited by L. Lynn and M. McGeary, Washington, D.C.: National Academy Press, 111-186.

Johnson, Rucker C. 2009. "Long-run Impacts of School Desegregation and School Quality on Adult Health." Unpublished manuscript, UC-Berkeley.

_____. 2007. "The Lasting Impacts of Childhood Neighborhood and Family on Adult Economic Status and Poverty over the Life Cycle". Unpublished manuscript, UC-Berkeley.

_____. 2007. "Race Differences in the Incidence and Duration of Exposure to Concentrated Poverty over the Life Course: Upward Mobility or Trapped in the Hood?" Unpublished manuscript, UC-Berkeley.

Johnson, Rucker C. and Robert F. Schoeni. 2007. "The Influence of Early-Life Events on Human Capital, Health Status, and Labor Market Outcomes Over the Life Course". National Poverty Center working paper.

Katz, L., J. Kling and J. Liebman. 2001. "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment". *Quarterly Journal of Economics*. 116(2):607-654.

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

Kremer, Michael. 1997. "How Much Does Sorting Increase Inequality?" *Quarterly Journal of Economics*. Feb., 112(1): 115-139.

Kunz, Jim, Marianne E. Page and Gary Solon. 2001. "Are Single-Year Measures of Neighborhood Characteristics Useful Proxies for Children's Long-Run Neighborhood Environment?" Unpublished manuscript. University of Maryland and Joint Center for Poverty Research.

La Ferrara, Eliana & Mele, Angelo, 2006. "Racial Segregation and Public School Expenditure". C.E.P.R. Discussion paper.

Leventhal, T and J. Brooks-Gunn. 2001. "Moving to Opportunity: What About the Kids?" forthcoming in *Choosing a Better Life: Evaluating the Moving to Opportunity Social Experiment*, eds. J. Goering and J. Ferris. Washington, DC: Urban Institute Press.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem". *The Review of Economic Studies*. 60(3): 531-542.

Margo, R. 1988. "Schooling and the Great Migration". NBER Working Paper #2697.

Massey, D.S. and N.A. Denton. 1993. *American apartheid: Segregation and the making of the underclass.* Harvard University Press, Cambridge, Mass.

Moffitt, R. 2001. "Policy interventions, low-level equilibria, and social interactions". In: Durlauf, S., Young, H.P. (Eds.), *Social Dynamics*. MIT Press, Cambridge.

Morenoff, Jeffrey and John Lynch. 2004. "What makes a place healthy? Neighborhood influences on racial/ethnic disparities in health over the life course", in *Critical Perspectives on Racial and Ethnic Difference in Health in Late Life*, edited by Norman B. Anderson, Randy A. Bulatao, and Barney Cohen, National Research Council, 406-449.

Muurinen, Jaana-Marja and Julian Le Grand. 1985. "The Economic Analysis of Inequalities in Health". *Social Science Medicine*. 20 (10):1029-1035.

Oreopoulos, Philip. 2003. "The Long-Run Consequences of Living in a Poor Neighborhood", *Quarterly Journal of Economics,* Vol. 118(4): 1533-75.

Page, Marianne E. and Gary Solon. 1999. "Correlations between Brothers and Neighboring Boys in their Adult Earnings: The Importance of Being Urban". Unpublished manuscript. University of California at Davis.

_____. 2001. "Correlations between Sisters and Neighboring Girls in Their Subsequent Income as Adults". Unpublished manuscript. University of California at Davis.

Plotnick, D. and S. Hoffman. 1996. "The Effect of Neighborhood Characteristics on Young Adult Outcomes: Alternative Estimates". Institute for Research on Poverty, Discussion Paper 1106-96.

Poterba, James M.. 2003. "Some Observations on Health Status and Economic Status". *Journal of Econometrics*. 112:65-67.

Rabe-Hesketh, Sophia. 2001. "GLLAMM Manual Technical Report 2001/01". Unpublished manuscript. Department of Biostatistics and Computing, Institute of Psychiatry, University of London.

Raaum, Oddbjørn O., K. G. Salvanes and E.O. Sorensen. 2002. "The Neighbourhood is not what it used to be". Unpublished manuscript. Norwegian School of Economics.

Raudenbush, S. and S.W. Bryk. 2002. *Hierarchical Linear Models. Applications and Data Analysis Methods.* Thousand Oaks, CA: Sage Publications.

Rothstein, Jesse. 2007. "Does Competition among Public Schools Benefit Students and Taxpayers?

Comment." *American Economic Review*, 97(5): 2026–2037.

Skinner, Jonathan, D. Staiger, A. Chandra, J. Lee, and M. McClellan. 2002. "Racial Differences in Hospital Quality for the Treatment of Acute Myocardial Infarction: Evidence from the Medicare Population". Unpublished manuscript. Center for Evaluative Clinical Sciences, Dartmouth Medical School.

Solon, Gary, M. Corcoran, R. Gordon, and D. Laren. 1987. "The Effect of Family Background on Economic Status: a Longitudinal Analysis of Sibling Correlations". NBER Working Paper Series. Paper #2282. National Bureau of Economic Research.

Solon, Gary, Marianne Page, and Greg Duncan. 2000. "Correlations between neighboring children in their subsequent educational Attainment". *The Review of Economics and Statistics*. 82(3): 383-392.

Staiger, Douglas and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments". *Econometrica* 65(3): 557-586.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures". *Journal of Political Economy*. 64: 416-424.

Wagstaff, Adam, P. Paci, and E. van Doorslaer. 1991. "On the Measurement of Inequalities in Health". *Social Science Medicine*. 33 (5):545-557.

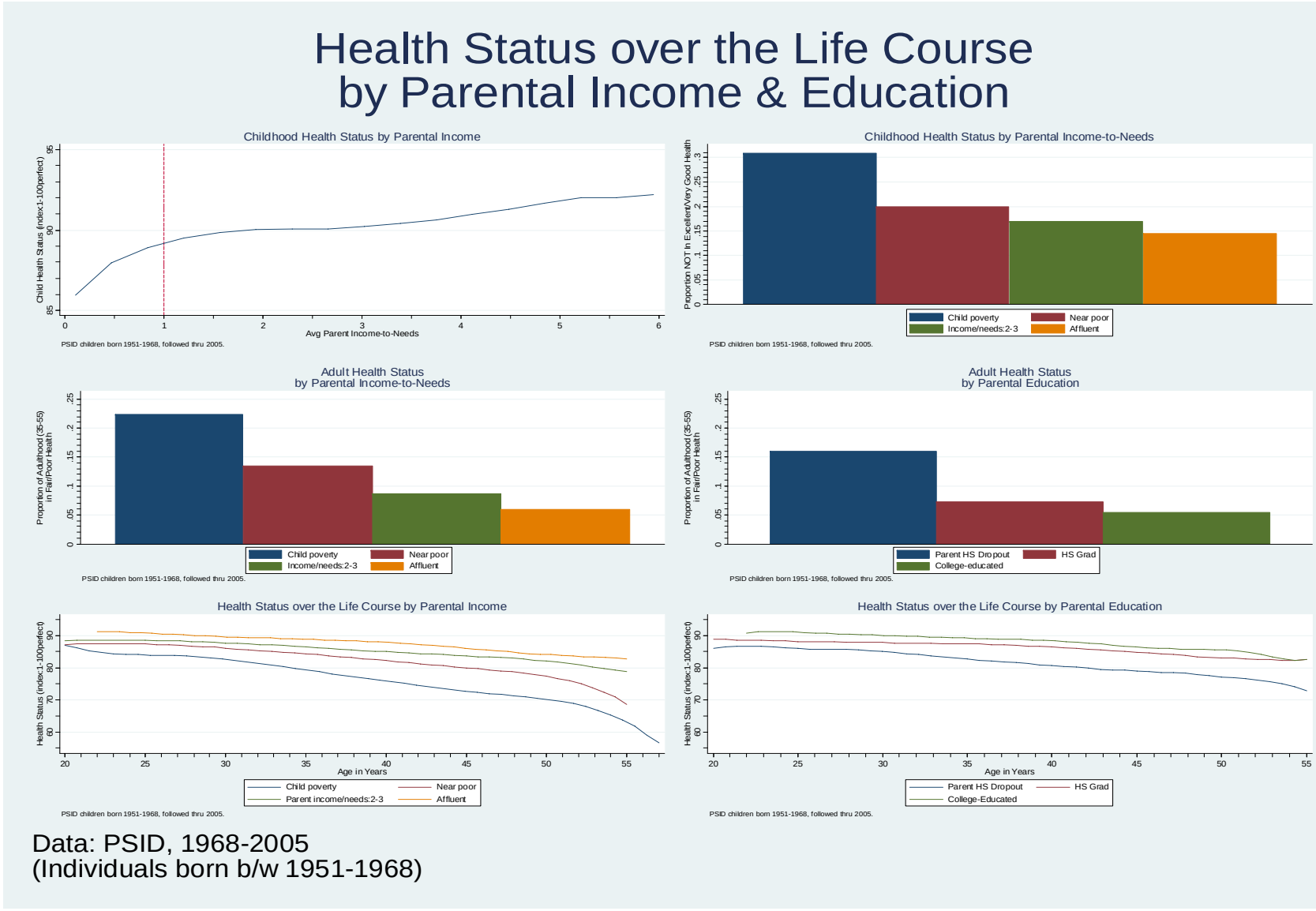
Wagstaff, Adam and Eddy van Doorslaer. 1994. "Measuring Inequalities in Health in the Presence of Multiple-Category Indicators". *Econometrics and Health Economics*. 3:281-291.

Van Doorslaer, Eddy and Andrew Jones. 2003. "Inequalities in Self-Reported Health: Validation of a New Approach to Measurement". *Journal of Health Economics*. 22:61-87.

Van Doorslaer, A. Wagstaff, H. Bleichrodt, et al. 1997. "Income-related Inequalities in Health: Some International Comparisons". *Journal of Health Economics*. 16:93-112.

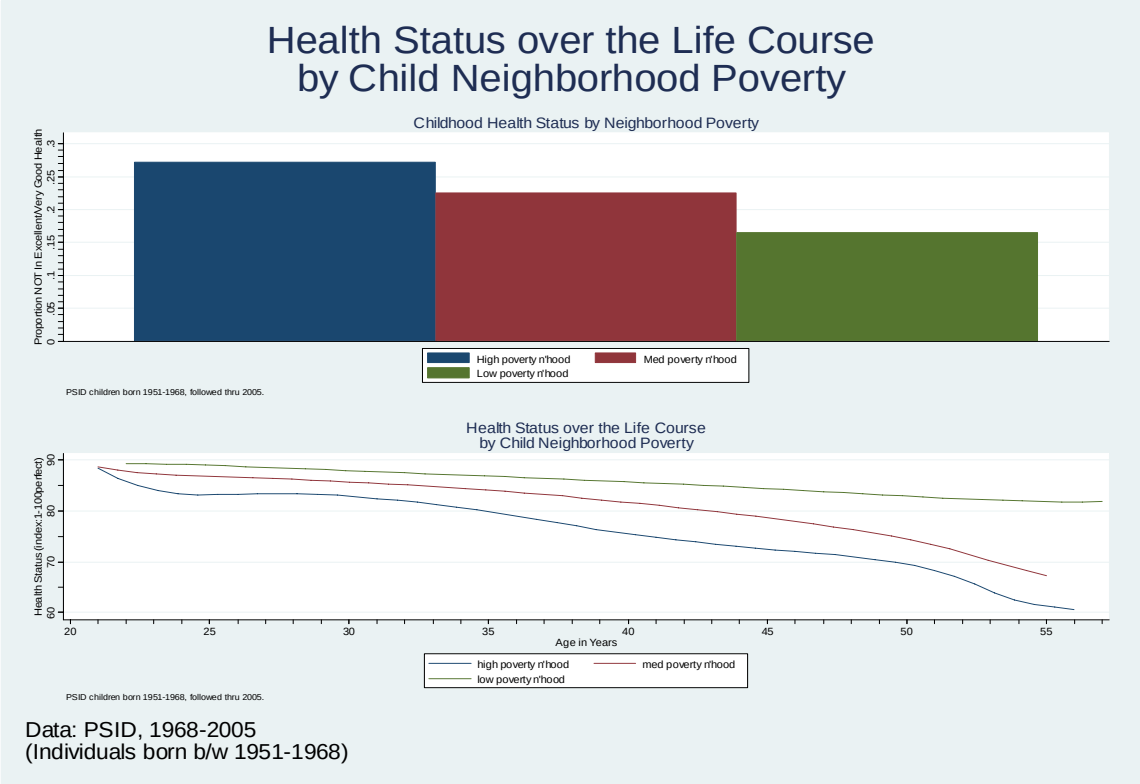
Vigdor, Jacob L., 2002. "The Pursuit of Opportunity: Explaining Selective Black Migration." *Journal of Urban Economics* vol 51(3): 391-417.

Figure 1.



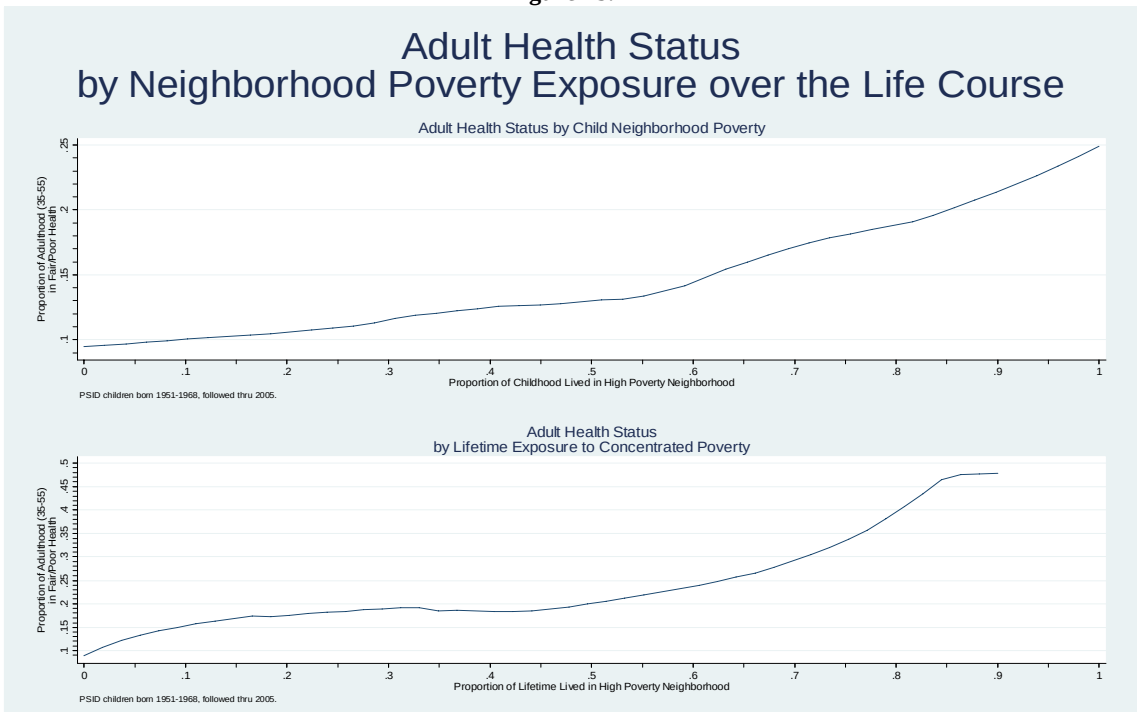
Data: PSID, 1968-2005
(Individuals born b/w 1951-1968)

Figure 2a.



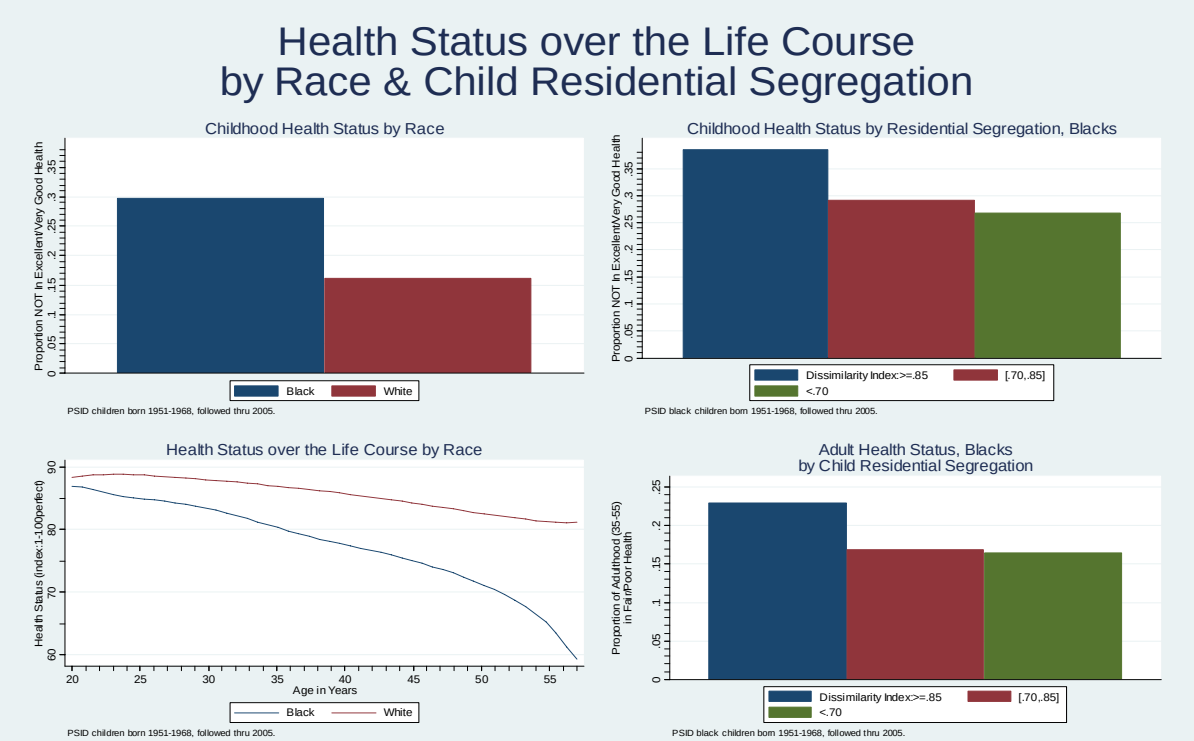
Data: PSID, 1968-2005
(Individuals born b/w 1951-1968)

Figure 2b.



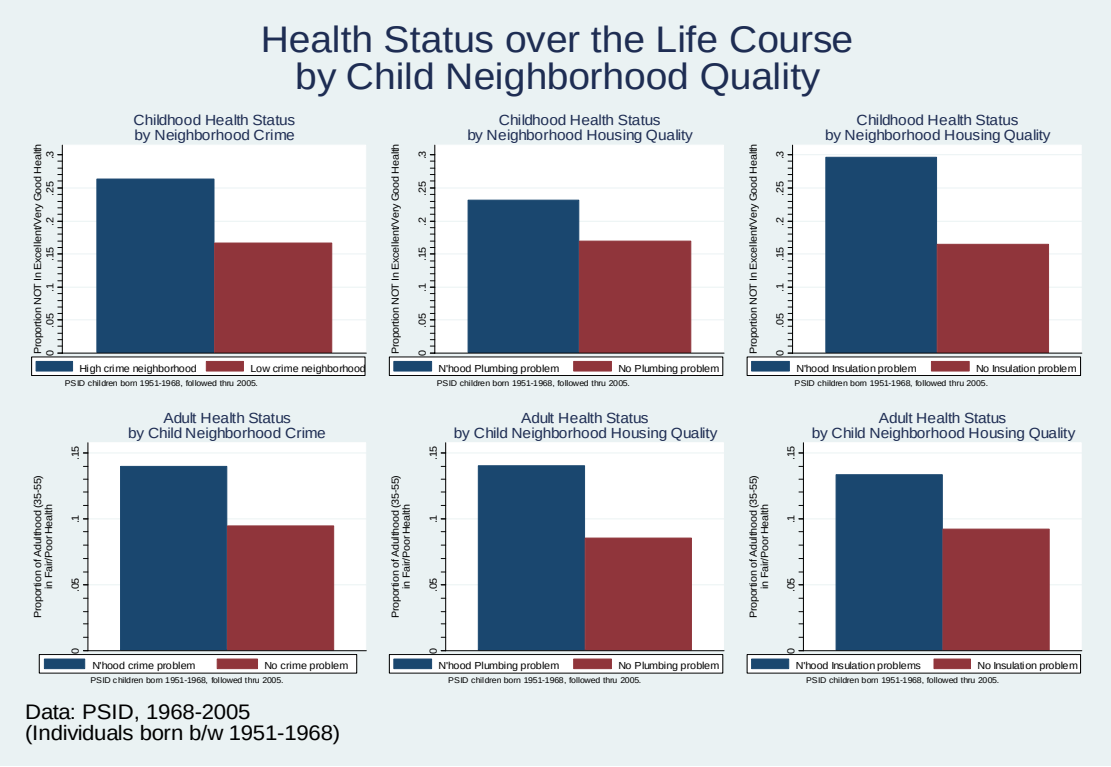
Data: PSID, 1968-2005
(Individuals born b/w 1951-1968)

Figure 3.



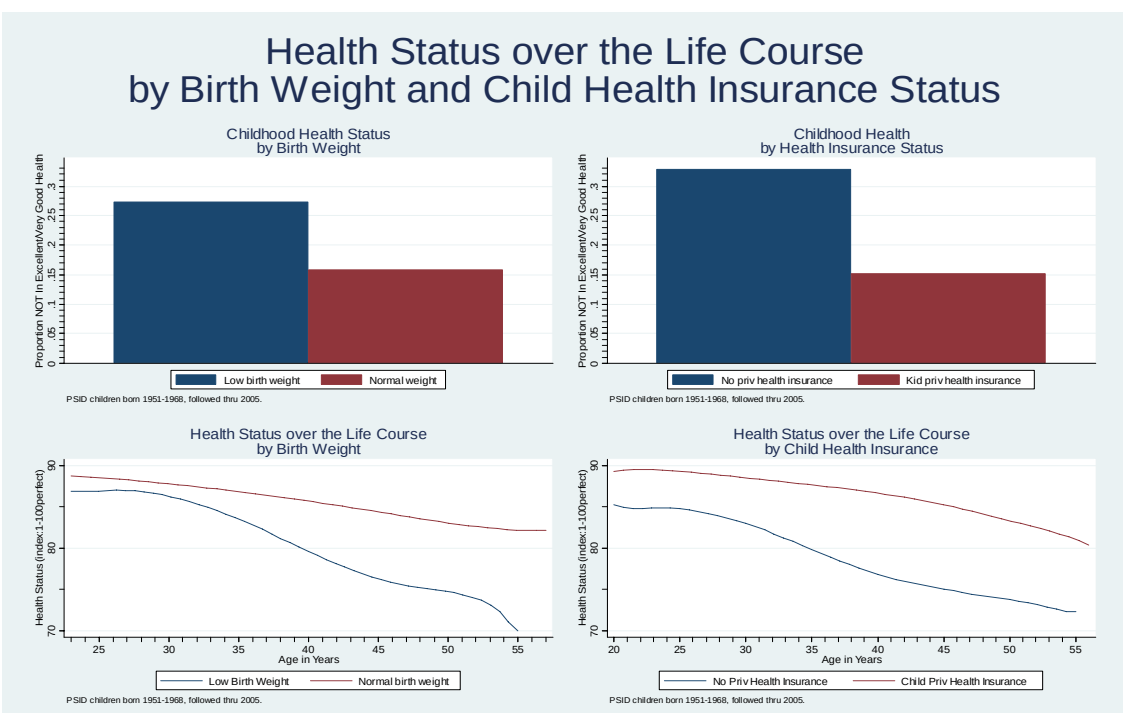
Data: PSID, 1968-2005
(Individuals born b/w 1951-1968)

Figure 4.



Data: PSID, 1968-2005
(Individuals born b/w 1951-1968)

Figure 5.



Data: PSID, 1968-2005
(Individuals born b/w 1951-1968)

Table 1. Health over the Life Course: Importance of Child Neighborhood & Family Background

(Dependent variable: general health status) Hierarchical Random Effects Interval Regression Model: 100pt-scale, 100=perfect health					
	Childhood	All Adulthood yrs	Ages 20-34	Ages 35-44	Ages 45-57
	(1)	(2)	(3)	(4)	(5)
Constant	90.8154***	86.7975***	87.1879***	84.1733***	79.2558***
	(0.0846)	(0.0797)	(0.0756)	(0.1019)	(0.2123)
Age - 30		-0.2145***	-0.1970***		
		(0.0021)	(0.0040)		
Age - 40				-0.3501***	
				(0.0034)	
Age - 50					-0.4146***
					(0.0093)
Female	-1.6545***	-0.5559***	-0.8757***	-0.5812***	-0.6245***
	(0.0865)	(0.0744)	(0.0760)	(0.0995)	(0.1625)
Random Effects, Unmeasured (Std Dev)					
Child Neighborhood component	5.7535***	7.3531***	6.5621***	8.2610***	9.5961***
	(0.1267)	(0.0777)	(0.0763)	(0.1025)	(0.1817)
Child Family component	5.3088***	4.0843***	3.1935***	4.5187***	7.0197***
	(0.1235)	(0.0870)	(0.1029)	(0.1315)	(0.2124)
Individual component	6.9342***	7.4047***	7.3883***	9.0133***	9.0402***
	(0.0420)	(0.0313)	(0.0335)	(0.0437)	(0.0799)
Transitory error component		6.3373***	5.1087***	4.5655***	5.9360***
		(0.0055)	(0.0065)	(0.0067)	(0.0190)
Log-likelihood	-144056.62	-2452537.9	-1209527	-906291.68	-263015.27
Number of counties	210	270	270	270	270
Number of neighborhoods	934	1,428	1,388	1,224	711
Number of families	1,280	1,935	1,868	1,652	923
Number of individuals	2,316	4,705	4,405	3,483	1,507
Number of person-year observations		51,082	27,349	19,256	4,477

*** 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 column (2) includes controls for $(age-30)^2$ and $(age-30)^3$ (coefficients suppressed to conserve space).

Table 2. Race & SES Differences in Child Health: Importance of Neighborhood & Family Background

(Dependent variable: general health status in childhood)

3-Level Hierarchical Random Effects Interval Regression Model: 100pt-scale, 100=perfect health

	Uncond'l model	Raw race gap	Controls for Fam bckgrd	Controls for Child Nhood + School + Fam
Childhood factors	(1)	(2)	(3)	(4)
Black		-2.6329***	-0.6334***	1.0447***
Non-Hispanic White (reference category)		(0.2160)	(0.2408)	(0.3246)
Family income-to needs ratio (avg during 1967-1972), spline:				
Income-to-needs ratio*ratio is <1			7.1305***	8.8800***
			(0.7724)	(0.8210)
Income-to-needs ratio* ratio is >=1			0.4271***	0.3605***
			(0.0405)	(0.0414)
Parental head's education:				
High school dropout			-0.3239**	-0.2862*
High school graduate (reference category)			(0.1590)	(0.1652)
College-educated			-0.0308	-0.1198
			(0.1607)	(0.1634)
No Private Child HI coverage, 1968-1972			-1.6476***	-1.3550***
			(0.2349)	(0.2375)
Low birth weight			-2.2641***	-2.1867***
			(0.1976)	(0.1967)
Mother unmarried at child's birth			-0.3647	-0.2214
			(0.2322)	(0.2329)
Parent smoked cigarettes at some point, 1968-1972			-1.1010***	-1.0828***
			(0.1411)	(0.1416)
Parental annual alcohol expenditures (in $100's),			-0.0310***	-0.0323***
5-year average 1968-1972			(0.0107)	(0.0109)
Child Neighborhood factors				
%of childhood yrs lived in low poverty neighborhood				1.6998***
				(0.2184)
High crime neighborhood				-1.3155***
				(0.1589)
Residential segregation dissimilarity$_{county}$, 1970				0.0753
				(0.0762)
Residential segregation dissimilarity index*Black				-0.3421
				(0.5368)
Parental low expectations for child achievement				-1.1001***
College-bound expectations (ref categeory)				(0.2614)
Neighborhood low expectations for child achievement				-0.3132
				(0.2601)
N'hood connectedness to informal sources of help				0.2996***
				(0.0564)
Neighborhood plumbing problems				-0.3093
				(0.2655)
Neighborhood housing insulation problems				-2.3002***
				(0.2608)
Random Effects, Unmeasured (Std Dev)				
Childhood Neighborhood component	5.7535***	5.4938***	5.0989***	4.8133***
	(0.1267)	(0.1352)	(0.1468)	(0.1557)
Childhood Family component	5.3088***	5.4387***	5.4517***	5.5460***
	(0.1235)	(0.1250)	(0.1279)	(0.1284)
Individual component	6.9342***	6.9378***	6.9369***	6.8838***
	(0.0420)	(0.0421)	(0.0421)	(0.0420)
Log-likelihood	-144056.62	-143969.29	-143573.02	-143220.03
Number of counties	210	210	210	210
Number of neighborhoods	934	934	934	934
Number of families	1,280	1,280	1,280	1,280
Number of individuals	2,316	2,316	2,316	2,316

*** p<0.01, ** p<0.05, * p<0.10

Note: All models include a constant and controls for year of birth, gender, and columns (3) and (4) control for birth order and include indices intended to capture parental aspirations/motivation and long-term planning horizon (rate of time preference proxy); and column (4) includes dummy indicators for expectations of child achievement that were in between "low" and "college-bound" expectations and also includes the following controls for child school quality: school segregation dissimilarity index interacted with race, school district per-pupil spending, and class size (coefficients supressed to conserve space).

Table 3. Race & SES Differences in Adult Health (Age 20-57): Importance of Neighborhood & Family Background

(Dependent variable: general health status in adulthood)

4-Level Hierarchical Random Effects Interval Regression Model: 100pt-scale, 100=perfect health

	Raw race gap	Controls for Fam bckgrd	Controls for Child Nhood + School + Fam	Only Adult Nhood + SES	Child bckgrd + Adult SES
Childhood factors	(1)	(2)	(3)	(4)	(5)
Black	-6.5151***	-2.6961***	-0.4129*	-5.3799***	-0.9870***
Non-Hispanic white (reference category)	(0.1874)	(0.1943)	(0.2428)	(0.1703)	(0.2334)
Family income-to needs ratio (avg during 1967-1972), spline:					
Income-to-needs ratio*ratio is <1		3.2709***	3.7694***		2.8009***
		(0.6388)	(0.6395)		(0.6083)
Income-to-needs ratio* ratio is 1 to 3		2.0318***	1.2958***		0.7896***
		(0.1018)	(0.1022)		(0.0981)
Income-to-needs ratio* ratio is >3		0.4609***	0.3531***		0.2996***
		(0.0437)	(0.0424)		(0.0407)
Parent head's education:					
High school dropout		-2.3852***	-1.6221***		-1.2208***
High school graduate (reference category)		(0.1341)	(0.1324)		(0.1271)
College-educated		0.7565***	0.4134***		-0.0650
		(0.1413)	(0.1385)		(0.1345)
No Private Child HI coverage, 1968-1972		-1.5283***	-1.2287***		-1.1720***
		(0.1811)	(0.1772)		(0.1706)
Low birth weight		-2.3272***	-2.1575***		-1.7826***
		(0.1741)	(0.1718)		(0.1672)
Mother unmarried at child's birth		-1.9316***	-1.9897***		-1.9096***
		(0.1817)	(0.1233)		(0.1185)
Parent smoked cigarettes at some point, 1968-1972		-0.6190***	-0.4184***		-0.2774**
		(0.1199)	(0.1162)		(0.1117)
Parental annual alcohol expenditures (in $100's),		-0.0516***	-0.0329***		-0.0316***
5-year average 1968-1972		(0.0085)	(0.0083)		(0.0079)
Child Neighborhood factors					
Neighborhood poverty rate (1970), spline:					
Low poverty neighborhood (ref category)					
Medium poverty neighborhood			-3.4528***		-3.0306***
			(0.2164)		(0.2036)
(Neighborhood poverty rate - 20)* rate 10 to 30%			-2.5216***		-2.0164***
			(0.3076)		(0.2933)
High poverty neighborhood			-3.6740***		-3.1042***
			(0.3414)		(0.3231)
High crime neighborhood			-0.8406***		-0.6532***
			(0.1316)		(0.1239)
Residential segregation dissimilarity$_{county}$, 1970			-0.1763***		-0.1417**
			(0.0636)		(0.0607)
Residential segregation dissimilarity index*Black			-1.2828***		-1.0038***
			(0.4017)		(0.3822)
Parental low expectations for child achievement			-2.0113***		-1.7452***
College-bound expectations (ref category)			(0.1993)		(0.1854)
N'hood low expectations for child achievement			-1.9121***		-1.2766***
			(0.1860)		(0.1695)
N'hood connectedness to informal sources of help			0.6308***		0.6505***
			(0.0421)		(0.0397)
Neighborhood plumbing problems			-1.9631***		-1.7327***
			(0.2074)		(0.1971)
Neighborhood housing insulation problems			-1.7496***		-1.6664***
			(0.2059)		(0.1952)
Parental health status					
Proportion of 60s mother in fair/poor health			-3.8133***		-3.4163***
			(0.1766)		(0.1699)
Proportion of 60s father in fair/poor health			-0.5050**		-0.1590
			(0.2001)		(0.1917)

Table 3 (cont'd). Race & SES Differences in Adult Health (Age 20-57): Importance of Neighborhood & Family Background

(Dependent variable: general health status in adulthood)

4-Level Hierarchical Random Effects Interval Regression Model: 100pt-scale, 100=perfect health

	Raw race gap	Controls for Fam bckgrd	Controls for Child Nhood + School + Fam	Only Adult Nhood + SES	Child bckgrd + Adult SES
	(1)	(2)	(3)	(4)	(5)
Adulthood SES					
Neighborhood poverty rate, spline:					
Low poverty neighborhood (reference category)					
Medium poverty neighborhood				-0.2418***	-0.2339***
				(0.0336)	(0.0326)
(Neighborhood poverty rate - 20)* rate is 10 to 30%				-0.2990***	-0.2524***
				(0.0372)	(0.0343)
High poverty neighborhood				-0.2215***	-0.1640***
				(0.0592)	(0.0592)
Educational attainment:					
High school dropout				-4.1767***	-3.1097***
High school graduate (reference category)				(0.1365)	(0.1375)
Some college				2.0647***	1.4020***
				(0.0940)	(0.0947)
College graduate or higher				3.8057***	2.5659***
				(0.1035)	(0.1089)
Family income-to needs ratio, spline:					
Income-to-needs ratio*ratio is <2				0.1735***	0.1719***
				(0.0233)	(0.0233)
Income-to-needs ratio* ratio is 2 to 4				0.2586***	0.1954***
				(0.0152)	(0.0143)
Income-to-needs ratio* ratio is >4				0.0204***	0.0211***
				(0.0018)	(0.0018)
No annual earnings				-3.9736***	-4.0089***
				(0.0538)	(0.0537)
No annual earnings*Female				2.3221***	2.3612***
				(0.0646)	(0.0645)
Random Effects, Unmeasured (Std Dev)					
Childhood Neighborhood component	6.7102***	5.6001***	5.1556***	5.7841***	4.8134***
	(0.0825)	(0.0894)	(0.0950)	(0.0801)	(0.0939)
Childhood Family component	4.2435***	4.1678***	3.8833***	3.6063***	3.6853***
	(0.0884)	(0 .0903)	(0.0982)	(0.0919)	(0.0978)
Individual component	7.4303***	7.4708***	7.4985***	7.3274***	7.3330***
	(0.0316)	(0.0320)	(0.0323)	(0.0317)	(0.0319)
Transitory error component	6.3373***	6.3377***	6.3379***	6.3054***	6.3069***
	(0.0055)	(0.0055)	(0.0055)	(0.0055)	(0.0055)
Log-likelihood	-2451929	-2450393.3	-2449445.5	-2445640.7	-2444400
Number of counties	270	270	270	270	270
Number of neighborhoods	1,428	1,428	1,428	1,428	1,428
Number of families	1,935	1,935	1,935	1,935	1,935
Number of individuals	4,705	4,705	4,705	4,705	4,705
Number of person-year observations	51,082	51,082	51,082	51,082	51,082

*** p<0.01, ** p<0.05, * p<0.10

Note: All models include a constant and controls for age, age squared, age cubed, gender, year of birth, and columns (2)-(3) and (5) include controls for region of birth, birth order and indices intended to capture parental aspirations/motivation and long-term planning horizon (rate of time preference proxy); and columns (3) and (5) include dummy indicators for expectations of child achievement that were in between "low" and "college-bound" expectations and also include the following controls fo child school quality: school segregation dissimilarity index interacted with race, school district per-pupil spending, and class size (coefficients supressed to conserve space). To facilitate interpretation of marginal effects, I converted the units of county racial residential segregation dissimilarity index so that a 1-unit change represents a 10-point change in the dissimilarity index. Similarly, a one-unit change in the spline specification for neighborhood poverty represents a 10-point change (e.g., change in neighborhood poverty rate from 10% to 20%).

Table 4. Sibling and Child Neighbor Correlations in Health Status over the Life Course

	Childhood		All Adulthood yrs		Age 20-34		Age 35-44		Age 45-57	
	Sibling Correlation	**Child Neighbor Correlation**	**Sibling Correlation**	**Child Neighbor Correlation**	**Sibling Correlation**	**Child Neighbor Correlation**	**Sibling Correlation**	**Child Neighbor Correlation**	**Sibling Correlation**	**Child Neighbor Correlation**
Unconditional	0.5604	0.3027	0.5634	0.4306	0.4938	0.3993	0.5218	0.4017	0.6337	0.4128
	(0.0068)	(0.0119)	(0.0046)	(0.0068)	(0.0055)	(0.0072)	(0.0056)	(0.0079)	(0.0072)	(0.0133)
Adjusted (net of residential sorting of HHs w/similar family bckgrd)	--	0.2750	--	0.3991						
Conditional, control for child family/neighborhood/school factors	0.5323	0.2287	0.4256	0.2715	0.3623	0.2621	0.3810	0.2484	0.4700	0.2411
	(0.0072)	(0.0139)	(0.0057)	(0.0089)	(0.0065)	(0.0090)	(0.0070)	(0.0112)	(0.0104)	(0.0163)
Conditional, control for child family/neighborhood/school + adult SES	--	--	0.4060	0.2559	0.3355	0.2445	0.3530	0.2273	0.4419	0.1956
			(0.0059)	(0.0090)	(0.0067)	(0.0089)	(0.0073)	(0.0113)	(0.0111)	(0.0174)

Table 5. Estimated Effect of Child Neighborhood Poverty Rate on Adult Health for a Proportional Correlation Model with Varying Values of the Relative Correlation

Relative Correlation	Estimated Effect of 10-percentage point increase in Childhood Neighborhood Poverty Rate
0 (exogeneity)	-4.1774***
	(0.0972)
0.2	-4.4727***
	(0.1787)
0.4	-5.9051***
	(0.5777)
0.8	-8.2675***
	(0.5708)
1	-10.1697***
	(0.2658)

Table 6. First-stage Estimates of Effects of Instrumental Variables on Economic & Racial Residential Segregation

	Dependent variable:		
	1970 poverty status dissimilarity index$_{MSA}$		**1970 black-white dissimilarity index$_{county}$**
	(1)	(2)	(3)
Instrumental variables:			
Share of Local Revenue from Intergovt Grants$_{1962}$	-0.1160***		
	(0.0359)		
# of Intercounty Rivers		0.0408***	
		(0.0106)	
# of Intercounty Rivers squared		-0.0053	
		(0.0045)	
# of Intracounty Rivers		-0.0199*	
		(0.0128)	
# of Intracounty Rivers squared		-0.0048	
		(0.0046)	
Black net migration rate$_{1930\text{-}1940,\ county}$			0.0861***
			(0.0328)
Black net migration rate squared$_{1930\text{-}1940}$			0.0752*
			(0.0406)
Black net migration rate cubed$_{1930\text{-}1940}$			-0.0381**
			(0.0178)
Number of MSAs	87		---
Number of counties	---		140

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

Notes: The regression models in columns (1)-(2) of the extent of residential segregation of the poor include the following 1970 MSA-level controls: ln(MSA population), MSA %black, share of MSA employed in manufacturing, and MSA ln(median household income). To facilitate interpretation of coefficients in column (2), I centered the river variables around their respective overall MSA-level means and converted the variables into standard deviation units. The regression model in column (3) of the extent of black-white residential segregation at the county-level include the following 1970 county-level controls: ln(population), %black, and ln(median household income), and region dummy indicators. To facilitate interpretation of marginal effects, I converted the units of the black net migration rate so that a 1-unit change represents a 10-percentage point change in the black net migration rate (range: (-.12,.27)).

Table 7. Three-Stage Least Squares Instrumental Variable Estimates of Effects of Child Neighborhood Poverty and Racial Residential Segregation on Adult Health

	Dependent variables:				
	3SLS-IV Model			2SLS	3SLS-IV
	(1)	(2)	(3)	(4)	(5)
	Adult Health Status (100pt-scale, 100=perfect health)	Probability (High Poverty Child Neighborhood)	Probability (Medium Poverty Child Neighborhood)	Adult Health Status (100pt-scale, 100=perfect health)	
High Poverty Child Neighborhood	-10.3865***				-12.3569***
(ref cat: Low Poverty Neighborhood)	(1.4273)				(1.4184)
Medium Poverty Child Neighborhood	-2.5368***				-2.4509**
	(0.9243)				(1.1438)
1970 Racial Segregation Dissimilarity $Index_{county}$*Black				-2.2215***	-1.2957**
				(0.5808)	(0.5929)
1970 Racial Segregation Dissimilarity $Index_{county}$*White				2.0612**	-0.3921
				(0.8569)	(0.8675)
1970 Poverty Dissimilarity $Index_{MSA}$ (instrumented)		0.2473***	-0.1947***		
		(0.0111)	(0.0137)		
1970 Poverty Dissimilarity $Index_{MSA}$*					
Non-Poor Child Family		-0.2223***	0.0901***		
		(0.0165)	(0.0205)		
Person-year observations		37,965		34,600	30,827
Number of individuals		3,702		3,321	2,966
Number of families		1,384		1,183	1,059
Number of neighborhoods		1,009		899	797
Number of counties		142		140	107
Number of MSAs		87		76	69

Robust Standard errors in parentheses (clustered on county)
*** p<0.01, ** p<0.05, * p<0.10

Notes: The number of rivers flowing through the MSA and the (state average) 1962 share of local revenue from intergovernmental grants are used as instruments for the extent of 1970 MSA economic residential segregation as measured by the MSA segregation poverty dissimilarity index; the black net migration rate in the county during the 1930s is used as an instrument for the extent of 1970 racial residential segregation in the county. An F-test of the joint significance of the excluded instruments is highly significant with a p-value of <.01. The first-stage IV estimates are presented in Table 6. The regression models include the following 1970 MSA- and county-level controls: ln(population), %black, share employed in manufacturing, ln(median household income), and region of birth dummy indicators. All models include the same set of child and family background controls as in Table 3. The model in columns (1)-(3) is restricted to the cities for which data is available on the number of rivers; the models in columns (4)-(5) are restricted to cities for which there is a sizeable black population since that is when segregation indices are most meaningful. To facilitate interpretation of marginal effects, the units of the dissimilarity indices have been multipled by 10, so tha

Table 8. Distribution of Health Status at Age 40 by Percentile of Childhood Neighborhood Background Component

	Distribution of Adult Health Status Attainment (proportion falling within specified percentile range)						
Child Neighborhood Percentile	0-10	10-20	20-40	40-50	50-60	60-80	80-100
10	0.27	0.21	0.28	0.09	0.06	0.07	0.02
20	0.17	0.18	0.30	0.11	0.09	0.12	0.04
40	0.07	0.12	0.26	0.13	0.12	0.20	0.10
60	0.03	0.07	0.20	0.12	0.13	0.26	0.19
80	0.01	0.03	0.12	0.09	0.11	0.30	0.35

Data Appendix Table A0.

Measures	Data Source	Year(s) collected	Survey Question	Definition
General Health Status	PSID	Adulthood:1984-2005; Childhood (retrospective): 1999/2001	"Would you say your health in general is excellent, very good, good, fair, or poor?"	--
Parental Health Status	PSID	Measured during parent's ages 50s and 60s (1984-2005).	"Would you say your health in general is excellent, very good, good, fair, or poor?"	Proportion of years when parent was in 50s and 60s in which they were in fair/poor health
Neighborhood Poverty Rate	1970-2000 Census	Child neighborhood: 1970 Census; Adult neighborhood: 1980-2000 (linearly interpolate for non-census years)	PSID respondent's residential location (1968-2005) matched to decennial census tract info	low poverty neighborhood (<10% poor); medium poverty neighborhood (10-30%); high poverty neighborhood (>30%)
Childhood Racial Residential Segregation	1970 Census	1970 Census	Black-white dissimilarity $index_{county}$: b_{it} & w_{it} = # of black & white individuals in neighborhood i at time t; B_t & W_t = total # black & white individuals in county.	$\frac{1}{2} * \sum_{i=1}^{n} \left\| \frac{b_{it}}{B_t} - \frac{W_{it}}{W_t} \right\|$
Childhood Economic Residential Segregation	1970 Census	1970 Census	Poverty status dissimilarity $index_{MSA}$: p_{it} & r_{it} = # of poor & non-poor families in neighborhood i at time t; P_t & R_t = total # poor & non-poor families in MSA.	$\frac{1}{2} * \sum_{i=1}^{n} \left\| \frac{p_{it}}{P_t} - \frac{r_{it}}{R_t} \right\|$
Childhood Neighborhood/Housing Quality	PSID	1975	Parental self-reports: whether there exist plumbing or insulation problems, or burglary, robbery, assault, drug use problems, or too few police in neighborhood in which they live.	High crime neighborhood=avg response among all PSID households who live in same neighborhood report major crime-related problems; housing insulation/plumbing problems=avg response among all PSID households who live in same neighborhood report insulation/plumbing problems.
Parental/neighborhood Expectations for Child Achievement	PSID	1968-1972	Parental self-reports: "How much education do you think your children will have when they stop going to school? What do you really think will happen?"	low expectations=may not finish high school; college-bound expectations (ref. cat). Neighborhood-level measures obtained by computing avg reponse among all PSID HHs who live in same neighborhood.
Parental/neighborhood Connectedness to informal sources of support	PSID	1968-1972	Index (0-9) of Connectedness to Potential Sources of Help (constructed from survey responses): Attends church once a month or more; # of neighbors known by name; Has relatives within walking distance; Goes to organizations once a month or more (PTA mtg).	Neighborhood-level measures obtained by computing avg index score based on responses among all PSID HHs who live in same neighborhood.
Child School quality	Office of Civil Rights (OCR) School data; Common Core data of NCES; Census of Governments	1962-1982	PSID respondent's residential location during school-age years matched to school resource data	School district per-pupil spending; avg class size; school segregation
Black net Migration $Rate_{county,\ 1930\text{-}1940}$	Census	1930-1940	--	Net migration of blacks b/w 1930-1940 as a proportion of total 1930 county population
Share of Local Revenue from Intergov't $Grants_{1962}$	Census of Governments	1962	--	Computed state avg
# of $Rivers_{MSA}$	Compiled by Jesse Rothstein	--	--	--

Table A1. Descriptive Statistics by Race

	All (N=4,705)	Black (N=2,213)	White (N=2,413)
Adult Health Status:			
Excellent	0.26	0.20	0.30
Very Good	0.35	0.29	0.39
Good	0.29	0.36	0.24
Fair	0.09	0.13	0.05
Poor	0.02	0.03	0.01
Age (range: 20-57)	34.8	34.8	34.8
Year born (range: 1950-1970)	1959	1959	1959
Female	0.50	0.55	0.50
Childhood family variables:			
Income-to-needs ratio (5-yr avg, 1968-1972):			
<1 (child poverty)	0.12	0.43	0.06
1-3	0.55	0.48	0.56
>3	0.34	0.09	0.38
Parent's (head's) education:			
High school dropout	0.41	0.74	0.35
High school graduate	0.31	0.20	0.33
College-educated	0.28	0.05	0.32
Born into two-parent family	0.80	0.49	0.85
Low birth weight (<5.5 pounds)	0.07	0.09	0.06
No private child health insurance, 1968-1972	0.10	0.24	0.08
Parental health behaviors (1997 $):			
Smoked cigarettes at some point, 1968-1972	0.73	0.80	0.72
Alcohol consumption (5-yr avg, 1968-1972)	$421	$299	$437
Parental health status:			
Proportion of 60s mother in fair/poor health	0.32	0.64	0.27
Proportion of 60s father in fair/poor health	0.33	0.66	0.31
Childhood neighborhood variables:			
Neighborhood poverty:			
High poverty neighborhood (>30%)	0.05	0.24	0.01
Medium poverty neighborhood (10-30%)	0.18	0.40	0.14
Low poverty neighborhood (<10%)	0.78	0.36	0.85
Residential segregation dissimilarity index$_{county}$	0.70	0.71	0.70
High crime neighborhood	0.16	0.26	0.15
N'hood low expectations for child achievement	0.17	0.29	0.15
N'hood college-bound expectations	0.72	0.58	0.74
N'hood connectedness to informal sources of help	6.09	5.82	6.14
Neighborhood plumbing problems	0.14	0.24	0.12
Neighborhood housing insulation problems	0.14	0.18	0.14

Note: All descriptive statistics are sample weighted to produce nationally-representative estimates of means. Black-white differences in all childhood family and neighborhood factors are statistically significant.

Table A2. Race & SES Differences in Adult Health (Age 20-34): Importance of Neighborhood & Family Background

(Dependent variable: general health status in adulthood)

4-Level Hierarchical Random Effects Interval Regression Model: 100pt-scale, 100=perfect health

	Raw race gap	Controls for Fam bckgrd	Controls for Child Nhood + School + Fam	Only Adult Nhood + SES	Child bckgrd + Adult SES
Childhood factors	(1)	(2)	(3)	(4)	(5)
Black	-5.6700***	-2.1689***	-0.8603***	-4.5235***	-1.3590***
Non-Hispanic white (reference category)	(0.1770)	(0.1862)	(0.2386)	(0.1624)	(0.2324)
Family income-to needs ratio (avg during 1967-1972), spline:					
Income-to-needs ratio*ratio is <1		2.8124***	3.7843***		2.6498***
		(0.6124)	(0.6155)		(0.5876)
Income-to-needs ratio* ratio is 1 to 3		1.7510***	1.1829***		0.7212***
		(0.0968)	(0.0979)		(0.0950)
Income-to-needs ratio* ratio is >3		0.2827***	0.1955***		0.1110***
		(0.0417)	(0.0408)		(0.0392)
Parent head's education:					
High school dropout		-1.7584***	-1.1252***		-0.7667***
High school graduate (reference category)		(0.1274)	(0.1266)		(0.1219)
College-educated		1.2176***	0.8895***		0.3350***
		(0.1337)	(0.1319)		(0.1289)
No Private Child HI coverage, 1968-1972		-0.9511***	-0.6389***		-0.5329***
		(0.1814)	(0.1780)		(0.1718)
Low birth weight		-2.2059***	-1.9914***		-1.6425***
		(0.1734)	(0.1713)		(0.1660)
Mother unmarried at child's birth		-2.4532***	-2.1149***		-2.0195***
		(0.1819)	(0.1234)		(0.1188)
Parent smoked cigarettes at some point, 1968-1972		-0.4078***	-0.2893***		-0.0940
		(0.1137)	(0.1110)		(0.1070)
Parental annual alcohol expenditures (in $100's),		-0.0640***	-0.0464***		-0.0449***
5-year average 1968-1972		(0.0080)	(0.0079)		(0.0076)
Child Neighborhood factors					
Neighborhood poverty rate (1970), spline:					
Low poverty neighborhood (ref category)					
Medium poverty neighborhood			-2.9701***		-2.7080***
			(0.2060)		(0.1992)
(Neighborhood poverty rate - 20)* rate 10 to 30%			-2.8578***		-2.4885***
			(0.2946)		(0.2889)
High poverty neighborhood			-3.4090***		-2.9885***
			(0.3265)		(0.3179)
High crime neighborhood			-0.8582***		-0.7807***
			(0.1249)		(0.1207)
Residential segregation dissimilarity$_{county}$, 1970			-0.1118*		-0.0780
			(0.0617)		(0.0608)
Residential segregation dissimilarity index*Black			-0.2239		-0.1996
			(0.3803)		(0.3733)
Parental low expectations for child achievement			-1.4597***		-1.2277***
College-bound expectations (ref category)			(0.1904)		(0.1776)
N'hood low expectations for child achievement			-1.9565***		-1.4887***
			(0.1786)		(0.1669)
N'hood connectedness to informal sources of help			0.4660***		0.4971***
			(0.0403)		(0.0389)
Neighborhood plumbing problems			-1.2812***		-1.0818***
			(0.1976)		(0.1931)
Neighborhood housing insulation problems			-2.0040***		-1.9420***
			(0.1950)		(0.1904)
Parental health status					
Proportion of 60s mother in fair/poor health			-2.8869***		-2.4400***
			(0.1720)		(0.1657)
Proportion of 60s father in fair/poor health			-0.5604***		-0.2625
			(0.1919)		(0.1837)

Table A2 (cont'd). Race & SES Differences in Adult Health (Age 20-34): Importance of Neighborhood & Family Background

(Dependent variable: general health status in adulthood)

4-Level Hierarchical Random Effects Interval Regression Model: 100pt-scale, 100=perfect health

	Raw race gap	Controls for Fam bckgrd	Controls for Child Nhood + School + Fam	Only Adult Nhood + SES	Child bckgrd + Adult SES
	(1)	(2)	(3)	(4)	(5)
Adulthood SES					
Neighborhood poverty rate, spline:					
Low poverty neighborhood (reference category)					
Medium poverty neighborhood				-0.4143***	-0.3114***
				(0.0389)	(0.0378)
(Neighborhood poverty rate - 20)* rate is 10 to 30%				-0.0494	-0.1681***
				(0.0426)	(0.0396)
High poverty neighborhood				-0.2809***	-0.2377***
				(0.0679)	(0.0681)
Educational attainment:					
High school dropout				-3.3268***	-2.2308***
High school graduate (reference category)				(0.1371)	(0.1377)
Some college				2.0944***	1.5068***
				(0.0958)	(0.0955)
College graduate or higher				3.8286***	2.7564***
				(0.1036)	(0.1089)
Family income-to needs ratio, spline:					
Income-to-needs ratio*ratio is <2				0.4113***	0.3918***
				(0.0325)	(0.0325)
Income-to-needs ratio* ratio is 2 to 4				0.1694***	0.0678***
				(0.0191)	(0.0175)
Income-to-needs ratio* ratio is >4				0.0361***	0.0483***
				(0.0062)	(0.0061)
No annual earnings				-2.4431***	-2.4951***
				(0.0791)	(0.0790)
No annual earnings*Female				1.7545***	1.8036***
				(0.0904)	(0.0903)
Random Effects, Unmeasured (Std Dev)					
Childhood Neighborhood component	6.0789***	5.0887***	4.7997***	5.2453***	5.0191***
	(0.0801)	(0.0881)	(0.0917)	(0.0786)	(0.0789)
Childhood Family component	3.3038***	3.2982***	2.9672***	2.6210***	2.4161***
	(0.1038)	(0.1063)	(0.1177)	(0.1165)	(0.1267)
Individual component	7.4103***	7.4489***	7.4868***	7.3707***	7.2481***
	(0.0337)	(0.0342)	(0.0345)	(0.0341)	(0.0337)
Transitory error component	5.1087***	5.1095***	5.1097***	5.1038***	5.1124***
	(0.0065)	(0.0065)	(0.0065)	(0.0065)	(0.0065)
Log-likelihood	-1209018.2	-1207719.8	-1207040.6	-1206338.8	-1205580.9
Number of counties	270	270	270	270	270
Number of neighborhoods	1,388	1,388	1,388	1,388	1,388
Number of families	1,868	1,868	1,868	1,868	1,868
Number of individuals	4,405	4,405	4,405	4,405	4,405
Number of person-year observations	27,349	27,349	27,349	27,349	27,349

*** p<0.01, ** p<0.05, * p<0.10

Note: All models include a constant and controls for age, age squared, age cubed, gender, year of birth, and columns (2)-(3) and (5) include controls for region of birth, birth order and indices intended to capture parental aspirations/motivation and long-term planning horizon (rate of time preference proxy); and columns (3) and (5) include dummy indicators for expectations of child achievement that were in between "low" and "college-bound" expectations and also include the following controls for child school quality: school segregation dissimilarity index interacted with race, school district per-pupil spending, and class size (coefficients supressed to conserve space). To facilitate interpretation of marginal effects, I converted the units of county racial residential segregation dissimilarity index so that a 1-unit change represents a 10-point change in the dissimilarity index. Similarly, a one-unit change in the spline specification for neighborhood poverty represents a 10-point change (e.g., change in neighborhood poverty rate from 10% to 20%).

Table A3. Race & SES Differences in Adult Health (Age 35-44): Importance of Neighborhood & Family Background

(Dependent variable: general health status in adulthood)

4-Level Hierarchical Random Effects Interval Regression Model: 100pt-scale, 100=perfect health

	Raw race gap	Controls for Fam bckgrd	Controls for Child Nhood + School + Fam	Only Adult Nhood + SES	Child bckgrd + Adult SES
Childhood factors	(1)	(2)	(3)	(4)	(5)
Black	-7.5944***	-3.5474***	-1.2838***	-6.0671***	-1.7749***
Non-Hispanic white (reference category)	(0.2343)	(0.2509)	(0.3167)	(0.2172)	(0.3056)
Family income-to needs ratio (avg during 1967-1972), spline:					
Income-to-needs ratio*ratio is <1		1.0803	0.5695		-0.4143
		(0.8122)	(0.8198)		(0.7840)
Income-to-needs ratio* ratio is 1 to 3		2.1715***	1.3528***		0.7071***
		(0.1285)	(0.1295)		(0.1250)
Income-to-needs ratio* ratio is >3		0.4874***	0.3927***		0.3412***
		(0.0529)	(0.0515)		(0.0498)
Parent head's education:					
High school dropout		-3.0718***	-2.1536***		-1.7340***
High school graduate (reference category)		(0.1708)	(0.1698)		(0.1640)
College-educated		0.3131*	-0.0998		-0.6482***
		(0.1747)	(0.1715)		(0.1682)
No Private Child HI coverage, 1968-1972		-2.4679***	-2.1382***		-2.2884***
		(0.2324)	(0.2281)		(0.2213)
Low birth weight		-2.6888***	-2.5553***		-2.2269***
		(0.2289)	(0.2263)		(0.2224)
Mother unmarried at child's birth		-0.5561**	-1.0540***		-0.9947***
		(0.2437)	(0.1621)		(0.1569)
Parent smoked cigarettes at some point, 1968-1972		-0.4214***	-0.2470*		-0.1182
		(0.1497)	(0.1456)		(0.1407)
Parental annual alcohol expenditures (in $100's),		-0.0322***	-0.0149		-0.0160+
5-year average 1968-1972		(0.0106)	(0.0103)		(0.0099)
Child Neighborhood factors					
Neighborhood poverty rate (1970), spline:					
Low poverty neighborhood (ref category)					
Medium poverty neighborhood			-2.9678***		-2.7520***
			(0.2878)		(0.2729)
(Neighborhood poverty rate - 20)* rate 10 to 30%			-2.4171***		-1.7805***
			(0.4131)		(0.3963)
High poverty neighborhood			-4.6574***		-4.2327***
			(0.4307)		(0.4105)
High crime neighborhood			-0.8284***		-0.6350***
			(0.1644)		(0.1558)
Residential segregation dissimilarity$_{county}$, 1970			-0.1314*		-0.0985+
			(0.0791)		(0.0757)
Residenital segregation dissimilarity index*Black			-1.0686**		-0.6735+
			(0.5029)		(0.4795)
Parental low expectations for child achievement			-3.4689***		-2.9394***
College-bound expectations (ref category)			(0.2506)		(0.2353)
N'hood low expectations for child achievement			-1.5833***		-0.7920***
			(0.2360)		(0.2181)
N'hood connectedness to informal sources of help			0.6187***		0.6376***
			(0.0536)		(0.0508)
Neighborhood plumbing problems			-3.2017***		-2.9015***
			(0.2688)		(0.2565)
Neighborhood housing insulation problems			0.2417		0.2846
			(0.2623)		(0.2507)
Parental health status					
Proportion of 60s mother in fair/poor health			-4.2607***		-3.8468***
			(0.2149)		(0.2081)
Proportion of 60s father in fair/poor health			-0.2057		0.2484
			(0.2434)		(0.2347)

Table A3 (cont'd). Race & SES Differences in Adult Health (Age 35-44): Importance of Neighborhood & Family Background

(Dependent variable: general health status in adulthood)

4-Level Hierarchical Random Effects Interval Regression Model: 100pt-scale, 100=perfect health

	Raw race gap	Controls for Fam bckgrd	Controls for Child Nhood + School + Fam	Only Adult Nhood + SES	Child bckgrd + Adult SES
	(1)	(2)	(3)	(4)	(5)
Adulthood SES					
Neighborhood poverty rate, spline:					
Low poverty neighborhood (reference category)					
Medium poverty neighborhood				-0.5851***	-0.5106***
				(0.0530)	(0.0519)
(Neighborhood poverty rate - 20)* rate is 10 to 30%				-0.4859***	-0.5134***
				(0.0579)	(0.0536)
High poverty neighborhood				-0.8355***	-0.7760***
				(0.0871)	(0.0870)
Educational attainment:					
High school dropout				-5.6038***	-4.1135***
High school graduate (reference category)				(0.1956)	(0.1980)
Some college				2.1554***	1.4623***
				(0.1282)	(0.1296)
College graduate or higher				4.2833***	2.9901***
				(0.1358)	(0.1444)
Family income-to needs ratio, spline:					
Income-to-needs ratio*ratio is <2				0.0769***	0.0827***
				(0.0241)	(0.0241)
Income-to-needs ratio* ratio is 2 to 4				0.3316***	0.2279***
				(0.0202)	(0.0188)
Income-to-needs ratio* ratio is >4				0.0052***	0.0067***
				(0.0016)	(0.0016)
No annual earnings				-1.7694***	-1.7985***
				(0.0655)	(0.0654)
No annual earnings*Female				0.0739	0.0789
				(0.0812)	(0.0812)
Random Effects, Unmeasured (Std Dev)					
Childhood Neighborhood component	7.4218***	6.5014***	5.8108***	6.6132***	5.3912***
	(0.1125)	(0.1227)	(0.1424)	(0.1113)	(0.1438)
Childhood Family component	4.7685***	4.4706***	4.2445***	3.8623***	4.0085***
	(0.1334)	(0.1419)	(0.1613)	(0.1475)	(0.1650)
Individual component	9.0546***	9.1212***	9.1730***	9.0562***	9.0946***
	(0.0443)	(0.0450)	(0.0458)	(0.0453)	(0.0460)
Transitory error component	4.5654***	4.5656***	4.5660***	4.5507***	4.5520***
	(0.0067)	(0.0067)	(0.0067)	(0.0067)	(0.0067)
Log-likelihood	-905730.28	-904639.42	-903825	-903145.42	-902082.4
Number of counties	270	270	270	270	270
Number of neighborhoods	1,224	1,224	1,224	1,224	1,224
Number of families	1,652	1,652	1,652	1,652	1,652
Number of individuals	3,483	3,483	3,483	3,483	3,483
Number of person-year observations	19,256	19,256	19,256	19,256	19,256

*** p<0.01, ** p<0.05, * p<0.10

Note: All models include a constant and controls for age, age squared, age cubed, gender, year of birth, and columns (2)-(3) and (5) include controls for region of birth, birth order and indices intended to capture parental aspirations/motivation and long-term planning horizon (rate of time preference proxy); and columns (3) and (5) include dummy indicators for expectations of child achievement that were in between "low" and "college-bound" expectations and also include the following controls for child school quality: school segregation dissimilarity index interacted with race, school district per-pupil spending, and class size (coefficients supressed to conserve space). To facilitate interpretation of marginal effects, I converted the units of county racial residential segregation dissimilarity index so that a 1-unit change represents a 10-point change in the dissimilarity index. Similarly, a one-unit change in the spline specification for neighborhood poverty represents a 10-point change (e.g., change in neighborhood poverty rate from 10% to 20%).

Table A4. Race & SES Differences in Adult Health (Age 45-57): Importance of Neighborhood & Family Background

(Dependent variable: general health status in adulthood)

4-Level Hierarchical Random Effects Interval Regression Model: 100pt-scale, 100=perfect health

	Raw race gap	Controls for Fam bckgrd	Controls for Child Nhood + School + Fam	Only Adult Nhood + SES	Child bckgrd + Adult SES
Childhood factors	(1)	(2)	(3)	(4)	(5)
Black	-9.3761***	-1.8004***	2.5708***	-7.4029***	1.7027***
Non-Hispanic white (reference category)	(0.3682)	(0.4088)	(0.5459)	(0.3353)	(0.5175)
Family income-to needs ratio (avg during 1967-1972), spline:					
Income-to-needs ratio*ratio is <1		7.4561***	11.5555***		8.1044***
		(1.3051)	(1.3045)		(1.2285)
Income-to-needs ratio* ratio is 1 to 3		2.8014***	1.3056***		0.5899***
		(0.2178)	(0.2133)		(0.2038)
Income-to-needs ratio* ratio is >3		0.2541***	0.2590***		0.2581***
		(0.0788)	(0.0747)		(0.0706)
Parent head's education:					
High school dropout		-2.9483***	-1.0620***		-0.6546**
High school graduate (reference category)		(0.2753)	(0.2687)		(0.2552)
College-educated		-0.6693**	-1.2428***		-1.3725***
		(0.2927)	(0.2824)		(0.2723)
No Private Child HI coverage, 1968-1972		-2.8687***	-2.7205***		-2.2384***
		(0.3695)	(0.3567)		(0.3405)
Low birth weight		-5.9354***	-5.2730***		-5.0336***
		(0.3874)	(0.3775)		(0.3656)
Mother unmarried at child's birth		-5.1901***	-4.4944***		-4.0112***
		(0.4206)	(0.2622)		(0.2502)
Parent smoked cigarettes at some point, 1968-1972		-1.7447***	-1.5204***		-1.2435***
		(0.2459)	(0.2321)		(0.2217)
Parental annual alcohol expenditures (in $100's),		0.0386*	0.0219		0.0449**
5-year average 1968-1972		(0.0202)	(0.0192)		(0.0183)
Child Neighborhood factors					
Neighborhood poverty rate (1970), spline:					
Low poverty neighborhood (ref category)					
Medium poverty neighborhood			-4.4725***		-3.8270***
			(0.4914)		(0.4586)
(Neighborhood poverty rate - 20)* rate 10 to 30%			-1.9016***		-1.0406*
			(0.6694)		(0.6267)
High poverty neighborhood			-9.0301***		-7.4461***
			(0.6909)		(0.6505)
High crime neighborhood			-2.4227***		-2.0692***
			(0.2542)		(0.2361)
Residential segregation dissimilarity$_{county}$, 1970			-0.7183***		-0.5755***
			(0.1165)		(0.1090)
Residential segregation dissimilarity index*Black			-5.0609***		-5.0653***
			(0.7402)		(0.6902)
Parental low expectations for child achievement			-2.8210***		-1.5079***
College-bound expectations (ref category)			(0.3876)		(0.3619)
N'hood low expectations for child achievement			-3.7517***		-3.8552***
			(0.3772)		(0.3498)
N'hood connectedness to informal sources of help			1.2945***		1.0532***
			(0.0883)		(0.0821)
Neighborhood plumbing problems			-3.8848***		-3.5148***
			(0.4483)		(0.4202)
Neighborhood housing insulation problems			-3.8576***		-3.3694***
			(0.4226)		(0.3971)
Parental health status					
Proportion of 60s mother in fair/poor health			-5.9573***		-5.4055***
			(0.3269)		(0.3128)
Proportion of 60s father in fair/poor health			-1.6620***		-0.9990***
			(0.3965)		(0.3757)

Table A4 (cont'd). Race & SES Differences in Adult Health (Age 45-57): Importance of Neighborhood & Family Background

(Dependent variable: general health status in adulthood)

4-Level Hierarchical Random Effects Interval Regression Model: 100pt-scale, 100=perfect health

	Raw race gap	Controls for Fam bckgrd	Controls for Child Nhood + School + Fam	Only Adult Nhood + SES	Child bckgrd + Adult SES
	(1)	(2)	(3)	(4)	(5)
Adulthood SES					
Neighborhood poverty rate, spline:					
Low poverty neighborhood (reference category)					
Medium poverty neighborhood				-0.5861***	-0.4792***
				(0.1608)	(0.1519)
(Neighborhood poverty rate - 20)* rate is 10 to 30%				-0.8739***	-0.8344***
				(0.1761)	(0.1547)
High poverty neighborhood				-0.2956	-0.2697
				(0.2917)	(0.2898)
Educational attainment:					
High school dropout				-8.2261***	-5.0665***
High school graduate (reference category)				(0.3632)	(0.3678)
Some college				1.7639***	0.7456***
				(0.2123)	(0.2137)
College graduate or higher				2.7371***	1.1087***
				(0.2179)	(0.2284)
Family income-to needs ratio, spline:					
Income-to-needs ratio*ratio is <2				0.9383***	0.8960***
				(0.1054)	(0.1051)
Income-to-needs ratio* ratio is 2 to 4				0.5437***	0.4596***
				(0.0642)	(0.0620)
Income-to-needs ratio* ratio is >4				0.0325***	0.0259***
				(0.0056)	(0.0054)
No annual earnings				-3.6917***	-3.5894***
				(0.1620)	(0.1612)
No annual earnings*Female				-0.4301**	-0.4159*
				(0.2142)	(0.2130)
Random Effects, Unmeasured (Std Dev)					
Childhood Neighborhood component	8.4842***	7.2789***	6.2632***	7.0043***	5.3556***
	(0.2064)	(0.2224)	(0.2236)	(0.2136)	(0.2471)
Childhood Family component	7.4453***	6.9222***	6.1012***	6.4127***	6.0092***
	(0.2138)	(0.2213)	(0.2371)	(0.2264)	(0.2366)
Individual component	9.0662***	9.1865***	9.2853***	9.0792***	9.0468***
	(0.0804)	(0.0819)	(0.0846)	(0.0848)	(0.0850)
Transitory error component	5.9363***	5.9347***	5.9363***	5.9258***	5.9225***
	(0.0190)	(0.0190)	(0.0190)	(0.0192)	(0.0191)
Log-likelihood	-262694.41	-261895.85	-260993.93	-261211.04	-259988.06
Number of counties	270	270	270	270	270
Number of neighborhoods	711	711	711	711	711
Number of families	923	923	923	923	923
Number of individuals	1,507	1,507	1,507	1,507	1,507
Number of person-year observations	4,477	4,477	4,477	4,477	4,477

*** p<0.01, ** p<0.05, * p<0.10

Note: All models include a constant and controls for age, age squared, age cubed, gender, year of birth, and columns (2)-(3) and (5) include controls for region of birth, birth order and indices intended to capture parental aspirations/motivation and long-term planning horizon (rate of time preference proxy); and columns (3) and (5) include dummy indicators for expectations of child achievement that were in between "low" and "college-bound" expectations and also include the following controls for child school quality: school segregation dissimilarity index interacted with race, school district per-pupil spending, and class size (coefficients supressed to conserve space). To facilitate interpretation of marginal effects, I converted the units of county racial residential segregation dissimilarity index so that a 1-unit change represents a 10-point change in the dissimilarity index. Similarly, a one-unit change in the spline specification for neighborhood poverty represents a 10-point change (e.g., change in neighborhood poverty rate from 10% to 20%).