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

RUCKER C. JOHNSON*
ROBERT WOOD JOHNSON POSTDOCTORAL RESEARCH FELLOW
UNIVERSITY OF MICHIGAN

ROBERT F. SCHOENI
ASSOCIATE PROFESSOR OF ECONOMICS & PUBLIC POLICY
UNIVERSITY OF MICHIGAN

October 2003

ABSTRACT

Most analyses of health disparities are cross-sectional and do not examine the dynamics of health inequality from a life course perspective. In this paper, we analyze health dynamics and the evolution of health inequality over the life course, and investigate the importance of neighborhood and family background. We use correlations based on nationally representative longitudinal sample of siblings and neighbors to estimate upper bounds on the possible causal effects of family and neighborhood background on health outcomes in early adulthood through mid life. Estimates based on four-level hierarchical random effects models consistently show a higher scope for family background than for neighborhood background.

We find that sibling correlations are large throughout at least the first 30 years of adulthood: the brother correlation in general health status immediately after childhood (age18) is roughly 0.60, and at ages 30-40 remains at 0.55. While the childhood neighbor correlations are much smaller than the sibling correlations, they are substantial among men. Our estimates suggest that disparities in neighborhood background account for between one-fourth and one-third of the variation in health status among men in mid life. In our analyses, we attempt to explain the level and interesting life-cycle patterns of sibling and childhood neighbor correlations in health and the relative importance of neighborhood and family background. Our analysis of the intertemporal character of health inequality provides insights into the extent and nature of feedbacks between disparities in health and disparities in socioeconomic status over the life cycle.

* Please direct correspondence to Rucker Johnson, University of Michigan, 109 Observatory, SPH-II, M2226, Ann Arbor, MI 48197, or email to ruckerj@umich.edu. This research is supported by funding from the Robert Wood Johnson Foundation to Johnson and NIA grant K01AG00670 to Johnson and Schoeni. We thank Gary Solon, Marianne Page, Greg Duncan, and Bobbi Wolfe for comments on an earlier version of this paper.

I. INTRODUCTION

Persistent residential segregation of poor and minority populations has spurred a growing body of literature that investigates the effects of community background on a variety of socioeconomic outcomes. However, the effects of the physical and socioeconomic neighborhood on health outcomes have been relatively unexplored. Analyses of health disparities have focused largely on individual and family-level determinants of health outcomes.

At the same time, the studies that have examined neighborhood differences in health status demonstrate that health outcomes do in fact exhibit a distinctive spatial pattern that mirrors the spatial pattern of physical and socioeconomic disadvantage (e.g., Geronimus et al, 2001; Morenoff, forthcoming; Morenoff and Lynch, 2002; Skinner et al, 2002; Chandra and Skinner, 2003). The similarity of these geographic patterns motivates our investigation into the potential causal effects of neighborhood context on health status.

A primary goal of this paper is to improve our understanding of how and why individual, family, and neighborhood factors produce and reproduce poor health. The principal impact of parents on their children is shaped in the childhood years. In order 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 from a life course perspective. 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? Answers to these questions are at the center of assessing the structure of opportunity in the U.S., and are particularly pertinent for the analysis of inequality in quality of life over the life course. In this paper, we analyze health dynamics and the evolution of health inequality over the life course, and investigate the importance of neighborhood and family background.

The typical analytical approach used in neighborhood studies is to regress individual level outcomes such as education, criminal activity, or health on neighborhood level factors such as census tract level mean income, poverty rates, or rates of single motherhood. But attempts to estimate causal effects of neighborhood context have faced well-documented challenges of endogeneity (Manski, 1993) and of obtaining accurate measures of neighborhood factors. Few

studies have used convincing identification strategies to overcome these challenges, exceptions being experimental evaluations such as Katz, Kling, Liebman (2001) and Leventhal & Brooks-Gunn (2001).

In this paper we take a different approach that largely side-steps these challenges by exploiting a unique feature of the Panel Study of Income Dynamics (PSID). Specifically, the initial PSID sample in 1968 was highly clustered, allowing us to compare the similarity in adulthood health between siblings who grew up together, versus unrelated individuals who grew up in the same narrowly defined neighborhood. We 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. Small neighbor correlations would indicate that community origins can explain only a minor portion of the variation in health outcomes. Large neighbor correlations would leave open the possibility that neighborhoods contribute significantly to inequality in health outcomes, and further analyses of the effects of particular neighborhood characteristics would be warranted.

We find that sibling correlations are large throughout at least the first 30 years of adulthood: the brother correlation in general health status immediately after childhood (age18) is roughly 0.60, and at ages 30-40 remains at 0.55. While the childhood neighbor correlations are much smaller than the sibling correlations, they are substantial. In particular, the male childhood neighbor correlation is roughly 0.30 during most the period between ages 25 to 40, implying that neighborhood factors may in fact play a significant role in determining health status.

The remainder of the paper is organized in the following way. We begin with a discussion of how neighborhood and family background may affect an individual's health trajectory in adulthood. We describe an economic model of health that incorporates the influence of neighborhood factors. The model provides our 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. The data are described in section IV. Sections V and VI discusses the econometric model and estimation methods, respectively. The results are presented in section VII, with concluding statements provided in the final section.

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 clearly 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 lifestyle orientation – eating habits, and exercise and smoking behaviors, for example – across generations may also translate into disparities in adult health.

Similarly, it has been hypothesized that neighborhood background can have direct effects on health. 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 have some influence on childhood health. Health lifestyle orientation may also have a neighborhood component as well, with peer groups and role models within communities or neighborhoods influencing children's opportunities and preferences (Johnson, 2003).

Perhaps equally or more importantly for health dynamics, neighborhood and family background may have indirect effects on health over the life course through their effects on the socioeconomic mobility process. The degree of socioeconomic 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. Since economic status is a major determinant of residential choice, persistence in economic status is likely to lead to persistence in neighborhood quality as well; that is, the lower economic mobility is, the greater the correlation between childhood and adulthood neighborhood characteristics.

Theoretical Framework/Considerations. We briefly describe a simple two-period overlapping generations model of the transmission of health and economic status from parents to children to motivate the empirical analyses that follow. The discussion of the model adopts a simplified version of the basic framework of Becker & Tomes (1986).

Some children have an advantage because they are born into families with favorable genetic attributes, which we will 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 uterine, 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 decisions for himself. 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, with health and education comprising the two forms of human capital. Since much research demonstrates that investments during childhood are crucial to later development, we 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 one of two types of parents—rich or poor. Assume poor parents face credit constraints that prevent them from making worthwhile investments in the human capital of their children. One of these constraints, which will be one of the focal points 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 for a variety of reasons, including their effect on the tax base and the positive role model influences that they produce. Due to residential segregation by income, assume that two types of neighborhoods exist—high- and low-income.

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 water and air quality, sanitation, pollution and environmental toxins, safety, quantity and quality of health case and social services, and public schools, 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 some period of 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 & 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 in the child 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 time in adulthood. 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 claim is consistent with recent empirical evidence on the long-term effects of early life events in the US (Blackwell et al, 2001; Conley, 2000; 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 & Deaton, 2003). As emphasized by Muurinen & 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 from poorer families, and, because these components of human capital are to some extent substitutable, health capital will constitute a more important source of producing income ($y_t(H_t)$) and enjoying leisure (Muurinen & Le Grand, 1985).

The degree of persistence in educational attainment and earnings across generations are determinants of the life course trajectory of health capital depreciation because it affects individual's opportunity sets with respect to living and working conditions in adulthood. For example, in an economically segregated environment with low intergenerational economic mobility, individuals born to poor, less-educated parents residing in low-income neighborhoods will be more likely to reach adulthood with insufficient levels of accumulated human capital to qualify for high-skilled jobs that are well paid and 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 health capital due to the greater intensity of use (Muurinen & Le Grand, 1985; Case & 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 are further exacerbated by an increased need to engage in consumption activities such as smoking and binge drinking that, while hazardous in the long-run, temporarily relieve day-to-day stress in the short-run.

Neighborhood conditions have an indirect effect on health through their effects on health behaviors (Johnson, 2003), as well as its more commonly cited direct effect on health through "weathering," whereby the accumulated stress, lower environmental quality, and limited resources of poorer communities experienced over many years erodes the health of residents in ways that make them more vulnerable to mortality from any given disease (Geronimus, 1992).

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 be placed on that decision. In this context, families and individuals who care more about their health will be less likely to choose to live in an area with high crime, pollution, or a poor health care system. Furthermore, the set of complex and nuisanced characteristics that influence neighborhood choices are not likely to be well measured and accounted for appropriately in econometric models.

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 on-going 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. Evidence from two sites – Boston and New York – demonstrates 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 do in fact influence health status, at least in the short-run among poor families.

Among the studies that have tried to 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 is comparing 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 living in neighborhoods that are less risky for their children's health, then they might move to neighborhoods with less crime or pollution, which may in turn lead to better health outcomes for their kids. But if the underlying change in their preferences towards health outcomes not only caused them to change neighborhoods, but also to spend more time encouraging their children to practice good health behaviors such as eating healthily, exercising, and avoiding high crime areas, then the neighborhood "effect" might actually be representing all of these other factors and not the true causal effects of neighborhoods *per se*. Moreover, it is quite 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 may in fact 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).

Instead of performing another regression analysis focused on particular neighborhood characteristics, in this paper we exploit a unique feature of the PSID and adopt an approach recently used by Solon et al (2000) and Duncan et al (2001) to examine the role of contextual factors on educational attainment, achievement, and delinquency.[1] Specifically, the initial PSID

[1] The analysis of Duncan et al. (2001) focuses on adolescent outcomes using the Add-Health Survey data.

sample in 1968 was highly clustered with most PSID families having several other sample families living on the same block. This survey design allows us to compare the similarity in early 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 and instead compares sibling correlations with neighbor correlations, placing an upper bound on the neighborhood influence and allowing a comparison of the relative sizes of neighborhood versus family effects.

IV. DATA

The PSID began interviewing a national probability sample of families in 1968, and it has re-interviewed the members of those families every year since, with bi-annual interviewing beginning in 1997. Most importantly, when children of the 1968 PSID families became adults and left their parents' homes, these children were interviewed themselves in each year.

The PSID used a "cluster sample" when it started in 1968 in order 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 within a given family are followed throughout their lives, we can examine the similarity in health outcomes over the life-course of both siblings and childhood neighbors.

In our analyses, we define the neighborhood of upbringing as the census block where the child lived in 1968.[2] Thus, we are able to use a narrower, compact definition of neighborhood than the vast majority of previous studies of neighborhood effects. Typical studies use census tracts to define neighborhoods, and census tracts, which consist of roughly 5,000 families, are much larger than the neighborhood construct we employ. Although the neighboring families in the PSID sample may or may not have been social neighbors in the sense of interacting closely with each other, they did live in close geographic proximity to each other, and this neighborhood construct should capture important environmental influences. In urban areas, neighboring 1968 families in the PSID may have been a city block or even just part of a block. In rural areas, the

[2] 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.

families were spread further apart, but still were among each other's closest neighbors (Solon et al, 2000). The PSID cluster design is discussed in greater detail in Solon et al (2000).

We use PSID survey responses to the self-assessed health (SAH) question, "how do you rate your health status in general—excellent, very good, good, fair, or poor?" This question was asked of household heads and wives (if present) in each survey between 1984-2001. For a significant share of the individuals in our sample who were children in 1968, 1984 represents roughly the year in which they became heads of households as adults. A detailed discussion of how we construct our measure of health status from these survey responses is presented in Section IV.

We estimate models separately by gender because of well-known differences in health status between men and women. This paper examines men; future work will incorporate women. There are three criteria for inclusion in the sample: (1) the individual was 0-16 years old in 1968 residing in an original PSID sample family, (2) the individual has a valid measure for general health status in adulthood, and (3) the individual has at least one same-sex unrelated childhood neighbor who also meets these two criteria. Because one of our primary interests is in correlations across neighboring families, we restrict the samples to neighborhoods containing children from at least two different families.

In addition, we restrict our analyses to the Survey Research Center (SRC) component of the PSID sample. The Survey of Economic Opportunity (SEO) component, commonly known as the "poverty sample," is problematic for our purposes because two neighboring families could enter that component of the PSID only if they had sufficiently low income. Consequently, the within-neighborhood variation in outcomes in the SEO sample may not be informative about the typical variation among all families within a neighborhood.

The resulting sample for men contains 4,070 person-year observations from 357 individuals from 234 families in 104 neighborhoods. The mean age is about 32, with age ranging from 17 to 47. The sample is about 10% black.

V. ECONOMETRIC MODEL

In this section, we present an econometric model that illustrates the connections among sibling correlations, neighbor correlations, and regression analyses of neighborhood effects.[3] We begin by assuming the true model for health status is:

$$H_{sfn} = \alpha' X_{fn} + \beta' Z_n + \varepsilon_{sfn} \qquad (1)$$

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 illustrative simplicity, at this juncture, we do not attempt to 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 for the reasons discussed in section II, we expect the family background factors, X_{fn}, and the neighborhood background factors, Z_n, to be positively correlated. Because it is difficult to fully and 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 our analysis is focus on an overall assessment of the relative contributions of individual, family, and neighborhood effects on health in early to mid adulthood. We then analyze the relative contribution of a parsimonious set of measured individual, family, and neighborhood covariates to the total variation from each component, and test hypotheses about the effects of specific characteristics of families and neighborhoods.

Our strategy for assessing the importance of contextual effects involves estimating the fraction of variation in health outcomes of interest that lies between families and neighborhoods,

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

to provide an upper bound on the possible effect of these contexts. The intuition motivating the use of this strategy is that if family background and residential community are important determinants of health outcomes, there will be a strong correlation between siblings in their health outcomes, as compared to two arbitrarily chosen individuals. 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.

As demonstrated in Solon et al (2000), using the additive model of the effect of family and neighborhood context in equation (1), the population variance of H_{sfn} can be decomposed as:

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

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

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

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 under consideration due to factors shared by siblings. From (3) we see that siblings have correlated health outcomes because they have shared family and neighborhood backgrounds, corresponding to the first and second terms of (3), 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 have an impact on health outcomes, such as the socioeconomic status of parents, genetic traits shared by siblings, family structure, as well as neighborhood effects stemming from the quality of neighborhood conditions.

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:

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

The last two terms in (3) and (4) 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, which are represented by the first term in (3), 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 (2002), none of whom examine health outcomes.

The neighborhood 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 (4), we notice that the neighborhood covariance consists of more than the variance in neighborhood characteristics given in the second term, and it should therefore be viewed as an upper bound on 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 effect and also establish a lower bound on the family effects. First, it follows from (4) 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 (4). Following Solon et al (2000) and Altonji (1988), we estimate the part of $\alpha' X_{fn}$ related to observable family characteristics such as parental income, education, family structure, and race. Let $\widetilde{X}_{fn}$ denote the observable subset of family characteristics with associated parameters $\hat{a}$ estimated *within* neighborhoods. We can then subtract off the sorting component arising from the fact that similar families tend to cluster in neighborhoods,

$$Cov_{adj}(H_{sfn}, H_{s'f'n}) = Cov(H_{sfn}, H_{s'f'n}) - Cov(\hat{\alpha}' \widetilde{X}_{fn}, \hat{\alpha}' \widetilde{X}_{f'n}) \ . \qquad (5)$$

While this approach helps reduce 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. We call this the "adjusted sibling correlation."

Additional Considerations

There are three reasons why our analytical approach may lead to incorrectly concluding that neighborhood factors are small. First, 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, we estimate the correlation between individuals who were childhood neighbors in 1968, but if 1968 neighborhood is a poor proxy for longer-run neighborhood environment, our estimates of the influence of 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.

We find that the average proportion of childhood spent growing up in the 1968 neighborhood was roughly 2/3 for our sample. To investigate the potential impact of residential mobility further on our findings, we re-estimated all health status correlations on the sample of children who had lived in their 1968 home since at least 1963.[4] The correlations among this sample were very similar to the ones we report here. Solon et al (2000) found that neighbor correlations in education were not sensitive to similar sample restrictions. Therefore, the

[4] We experimented with introducing heteroscedasticity into our multilevel model at the individual level as a function of the percent of childhood years spent growing up in the 1968 neighborhood. We 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.

evidence tends to suggest that residential mobility is not significantly influencing the estimated neighbor correlations.

A second factor that could cause us to underestimate the importance of neighborhood effects follows from the fact that sibling and neighbor correlations alone do not allow for a straightforward examination of subgroups, some of which may be more susceptible to neighborhood factors. For example, neighborhood problems such as violence, lead paint, or pollution, may only have significant impacts on health when they achieve some threshold of incidence. Or certain families, perhaps low-income or single parent families, may be less able to buffer the negative effects of low quality neighborhoods.

Finally, effect sizes that program evaluators commonly view as medium or even large may translate into small neighborhood correlations (Duncan & Raudenbush, 2001; Cain & Watts 1972; Rosenthal & Rubin 1982). As highlighted by Duncan & Raudenbush (2001), a small correlation between neighbors does not rule out a large effect size associated with a measured difference between neighborhoods. To illustrate this point, they consider a standardized mean difference between a set of experimental neighborhoods and an equal number of control neighborhoods, and they show that even a standardized effect size of d=.80, which is commonly viewed as very large, translates into a neighborhood correlation of just 0.14.[5]

VI. ESTIMATION METHODS

Measurement of Health. A principal challenge in analyzing health dynamics and health inequality is that information on health status is only available at an ordinal level. We use PSID survey responses to the self-assessed health (SAH) question, "how do you rate your health status in general—excellent, very good, good, fair, or poor?" SAH has been shown to be closely linked to morbidity reported in surveys or diagnosed through clinical examinations (Larue et al, 1979; Linn et al, 1980; Mays et al, 1992). It is also one of the most powerful predictors of mortality, even when controlling for physician-assessed health status and health-related behaviors, and it is a strong determinant of whether patients choose to use medication and health services. SAH

[5] Using the hypothetical example discussed in Duncan & Raudenbush (2001), assume we have a treatment group of neighborhoods and control group of neighborhoods, and we compute the standardized mean difference, *d,* between these two groups of neighborhoods. The proportion of variance explained by treatment is then *R*-squared, which in this case is $R^2 = d^2/[\, d^2 + 1/p(1-p)]$, where p is the proportion in the treatment group and $(1-p)$ is the proportion in the control group. In this hypothetical example, all between-neighborhood variance is created by the treatment. Thus, *R*-squared is equivalent to the intra-neighborhood correlation. In non-experimental settings, we do not have treatment groups, but rather sets of "high-risk" and "low-risk" neighborhoods.

represents one of the only measures for which we have health information over extended time period in large-scale U.S. longitudinal surveys. It has also been the key health variable used in international comparisons of health inequality in part due to its availability in the largest number of surveys and over the longest period of time (Van Doorslaer et al, 1997).

The multiple-categorical health indicator presents problems for the analysis of health dynamics and the measurement of health inequality. Conventional inequality measures require the outcome to be in the form of either a continuous variable or dichotomous variable. Thus, one must address how to scale categorical responses and subject our indicators to the transformation that best predicts quality of life or life expectancy, thus translating our measures into the metric that reflects the underlying level of health.

A number of previous studies using surveys that contain both the SAH question and questions underlying a health utility index have demonstrated that a change in SAH 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 & Jones, 2003; Humphries & Van Doorslaer, 2000). More generally, this research has shown that health differences between SAH categories increase with lowering SAH categories. In other words, the extent of health deterioration increases with each step in SAH when going down from excellent (5) to poor health (1). Consequently, changes in SAH along the five-point scale have larger implications for morbidity and mortality when health status is poor than when it is good or excellent. Thus, assuming a linear scaling would not be appropriate.

To analyze health inequality in the presence of a multiple-category health indicator, three alternative approaches have previously been employed, each with its own set of advantages and disadvantages. The most common and simplest approach is to dichotomize SAH 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 & Van Doorslaer, 1994).

A second approach is to estimate an ordered logit or ordered probit regression using the SAH 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 & 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).[6] 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 SAH question is a latent, continuous but unobservable health variable with a standard lognormal distribution. This assumption allows "scoring" of the SAH 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 (HUI) (Humphries & 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 imposes the strong assumption of an identical latent health distribution across countries. An additional drawback 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 SAH 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 SAH and health utility index measures,

[6] 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 & van Doorslaer, 2000).

we use the distribution of health utility-based scores across the SAH categories to scale the categorical responses and subject our 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.

In order to scale the SAH categories in the PSID, we use the health utility-based scale that was developed in the construction of the Health and Activity Limitation index (HALex). The HALex scores associated with SAH categories are based on the U.S. National Health Interview Survey, which contains a fuller health instrument than utilized in the PSID. A multiplicative, multiattribute health utility model was used to assign scores and quantify the distance between the different SAH 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 and zero is equivalent to death, the interval health values associated with SAH 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 CDC National Center of Health Statistics to estimate health-related quality of life measures and years of health life to evaluate the state/distribution of U.S. health (*Healthy People* 2000).

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

Our 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 we use the health utility-based values to score the thresholds for SAH, the linear index for the interval regression model is measured on the same scale. This scaling thus translates our 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.

Measuring & Decomposing Health Inequality. A primary focus of this paper is to analyze health dynamics and the evolution of health inequality over the life course. Our aim is to decompose the sources of health inequality and its age profile into the part attributable to neighborhood background, family background, and individual heterogeneity. There exists a long-standing literature on inequality and mobility measures developed originally for the measure of income inequality and economic mobility. We will apply these methods in a novel way in the context of health dynamics and health inequality.

We employ two different inequality measures: (1) the variance, and (2) Thiel's general entropy inequality index (mean logarithmic deviation). Both of these inequality measures possess the desired additive decomposability property—i.e., the inequality measure for the total population can be expressed as a sum of the population-weighted average of the inequality measures "within" its subgroups and of the inequality existing "between" them. The decomposability of these two inequality measures is particularly useful in analyzing the

contribution to inequality from hierarchical levels. Our data are hierarchical because we have multiple observations over time of individuals who are nested within families, which are nested within neighborhoods. Thus, the additive multilevel decomposability property enables the inequality index to be expressed as:

$$I_{AGEt}(H_{1t}, H_{2t},\ldots, H_{nt}) = \; I^{B,N} \; + \; I^{B,FN} \; + \; I^{B,SF(N)} \;,$$

where I_{AGEt} denotes total inequality at age t, $I^{B,N}$ represents inequality between neighborhoods, $I^{B,FN}$ represents inequality between families within neighborhoods, and $I^{B,SF(N)}$ represents inequality between siblings. We can then examine the evolution of health inequality by using the estimated I_{AGEt} and its component parts at each age (t=17, 18,…45) to construct an age-profile of health inequality. Both the variance and Thiel's general entropy inequality index (mean logarithmic deviation) also have the additional desirable property as inequality measures of satisfying the Pigou-Dalton condition—i.e., the inequality measure decreases with any health transfer from healthy to less healthy people (Bourguignon, 1979).

Four-Level Hierarchical Random Effects Interval Regression Model. We begin with the variance as our measure of inequality. We 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, we estimate a four-level hierarchical random effects interval regression model. Our data are hierarchical because we have multiple observations over time of individuals who are nested within families, which are nested within neighborhoods. 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 & Bryk, 2002).

We begin by estimating the four-level hierarchical random effects model[7] given by

$$H^{*}_{tsfn} = (\beta_{0000} + \beta_{1000} * Age_t) + (\eta_{000n} + \eta_{100n} * Age_t) + (\phi_{00fn} + \phi_{10fn} * Age_t) + (\delta_{0sfn} + \delta_{1sfn} * Age_t) + \varepsilon_{tsfn}$$

or equivalently,

$$H^{*}_{tsfn} = (\beta_{0000} + \eta_{000n} + \phi_{00fn} + \delta_{0sfn}) + (\beta_{1000} + \eta_{100n} + \phi_{10fn} + \delta_{1sfn}) * Age_t + \varepsilon_{tsfn} \qquad (1)$$

[7] 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 gllamm6 macro in Stata (Rabe-Hesketh et al, 2000). The numerical evaluation of the unconditional-likelihood function uses Gaussian quadrature. We use 10-point quadrature for each level.

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 and random slope coefficient are represented by η_{000n} and η_{100n}, respectively; the family random intercept and random slope coefficient are represented by ϕ_{00fn} and ϕ_{10fn}, respectively; the individual random intercept and random slope coefficient are represented by δ_{0sfn} and δ_{1sfn}, respectively; 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 $\text{var}(\eta_{000n})=\sigma^2_{0n}$, $\text{var}(\eta_{100n})=\sigma^2_{1n}$, $\text{var}(\phi_{00fn})=\sigma^2_{0fn}$, $\text{var}(\phi_{10fn})=\sigma^2_{1fn}$, $\text{var}(\delta_{0sfn})=\sigma^2_{0sfn}$, $\text{var}(\delta_{0sfn})=\sigma^2_{1sfn}$, and $\text{var}(\varepsilon_{tsfn})=\sigma^2_{tsfn}$.

Age_t is the individual's actual age at time *t* minus 17. We center the age variable around age 17 so that the fixed and random intercept terms can be interpreted as the initial health status after childhood (age 17). The fixed and random slope coefficients can thus be interpreted as the subsequent health deterioration rate from young adulthood into mid life (age 45). The average initial health status after childhood (age 17) is captured by β_{0000} and the average subsequent annual health deterioration rate is captured by β_{1000}. We estimate the covariance/correlation between initial health status after childhood (age 17) and subsequent health deterioration rate, given by $\text{cov}(\eta_{000n},\eta_{100n})=\sigma_{n,0,1}$, $\text{cov}(\phi_{00fn},\phi_{10fn})=\sigma_{fn,0,1}$, and $\text{cov}(\delta_{0sfn},\delta_{1sfn})=\sigma_{sfn,0,1}$. We can thus decompose the correlation between initial health status (after childhood) and (subsequent) annual health deterioration rate into its within-family, between-family within-neighborhood, and between-neighborhood components.

Of primary interest is the decomposition of the variance of both the level of health over time and the annual rate of health depreciation 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 effects η_{000n} and η_{100n}, and siblings are correlated because they share the random effects η_{000n} and η_{100n}, and ϕ_{00fn} and ϕ_{10fn}. Here 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:

$$\rho_{\text{sibling, healthlevel}}(\text{age}) = \frac{(\sigma^2_{0n} + 2age\sigma_{n,0,1} + \sigma^2_{1n}age^2) + (\sigma^2_{0fn} + 2age\sigma_{fn,0,1} + \sigma^2_{1fn}age^2)}{(\sigma^2_{0n} + 2age\sigma_{n,0,1} + \sigma^2_{1n}age^2) + (\sigma^2_{0fn} + 2age\sigma_{fn,0,1} + \sigma^2_{1fn}age^2) + (\sigma^2_{0sfn} + 2age\sigma_{sfn,0,1} + \sigma^2_{1sfn}age^2)}$$

$$\rho_{\text{neighbor, healhlevel}}(\text{age}) = \frac{(\sigma^2_{0n} + 2age\sigma_{n,0,1} + \sigma^2_{1n}age^2)}{(\sigma^2_{0n} + 2age\sigma_{n,0,1} + \sigma^2_{1n}age^2) + (\sigma^2_{0fn} + 2age\sigma_{fn,0,1} + \sigma^2_{1fn}age^2) + (\sigma^2_{0sfn} + 2age\sigma_{sfn,0,1} + \sigma^2_{1sfn}age^2)}$$

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.

We can then use the estimated sibling and neighbor correlations at each age t=17, 18,…45, to construct an age-profile of sibling and neighbor health correlations. The age-profile of sibling and neighbor correlations will provide insight into the nature and causes of the evolution of health inequality, and the relative roles of neighborhood and family background. In this model, the sibling correlation and neighbor correlation in the annual health deterioration rate can be computed, respectively, as:

$$\rho_{\text{sibling, deterioration rate}} = \frac{\sigma^2_{1n} + \sigma^2_{1fn}}{\sigma^2_{1n} + \sigma^2_{1fn} + \sigma^2_{1sfn}} \quad ;$$

$$\rho_{\text{neighbor, deterioration rate}} = \frac{\sigma^2_{1n}}{\sigma^2_{1n} + \sigma^2_{1fn} + \sigma^2_{1sfn}} \quad .$$

Health varies with age and gender. Because we did not want our estimates of sibling and neighbor correlations to reflect the influence of either of these two demographic factors, we adjusted for them in our baseline random coefficients model by including age as an explanatory variable at each of the hierarchical levels, and conducting separate analyses by gender, with results for males presented in this paper. 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.

Estimating "Adjusted Neighbor Correlations". We 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, we first estimate the following regression; for ease of exposition, here we omit the random effects terms that are included in the estimated model:

$$H^{*}_{tsfn} = \alpha_1 age_{tsfn} + \alpha_2' X_{\bullet\bullet fn} + \alpha_3' (\overline{X_{\bullet\bullet\bullet n}}) + \varepsilon_{tsfn}, \qquad (7)$$

where $X_{\bullet\bullet fn}$ is a vector of childhood family background characteristics including: log of average annual family income (based on three-year average as reported in 1968-1970), parental education, parental family structure, parent's race, indicator of whether parental family had health insurance in 1968, and parental annual expenditures on cigarette and alcohol consumption in 1968. $\overline{X_{\bullet\bullet\bullet n}}$ is a vector of the 1968 neighborhood-level measures of the same above variables.

Inclusion of family-level and neighborhood-level variables measuring the same concepts enables the vector α_2 of coefficient estimates to capture the within-neighborhood effects of these family background characteristics. Using the within-neighborhood estimates of the family background effects of parental income, education, race, family structure, health insurance coverage, and parental health behaviors 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 these family background characteristics can be taken as a conservative estimate of $\alpha' X_{fn}$ in equation (1).

We then estimate the inter-neighbor 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. We then subtract our estimate of the inter-neighbor variance in $\hat{\alpha}'X_{fn}$ from the estimate of the overall inter-neighbor 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 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. We then investigate to what extent observable childhood neighborhood-level characteristics explain the observed sibling and neighbor correlations and their respective age profiles. In addition to the variables previously discussed, we include self-reports of neighborhood and housing quality collected in the PSID. These measures include self-reports of whether it is a 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 too few police in the neighborhood in which they live.[8]

We will provide estimates of the distinct effects of neighborhood and family level background variables measuring the same concepts—for example, the effects of family SES conditional on neighborhood SES and vise versa. In addition, 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., measures vs. unmeasured neighborhood characteristics). Finally, we 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.

[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, as discussed earlier, 1968 families with children 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).

VII. RESULTS

The unadjusted sibling and neighbor correlations of health in adulthood are presented first. We then examine how much of the neighbor correlations can be explained by the fact that families in a neighborhood tend to be similar as opposed to emanating from neighborhood effects *per se*. We then attempt to explain the life-cycle pattern of sibling and childhood neighbor correlations and explore potential mechanisms that underlie the relative roles of neighborhood and family background on the health trajectory over the life course.

The estimates from the baseline four-level hierarchical random effects model that includes controls only for age are presented in Table 1. Here, age is entered as a linear function—a quadratic specification for age did not significantly improve the fit of the model. As shown in Table 1, the random effects intercept and age-slope estimates are all significant at each of the neighborhood, family and individual levels. The baseline random effects model enables the measurement of the overall magnitude of variation at the neighborhood, family, and individual levels over time. 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 estimated unadjusted sibling and neighbor correlations calculated from the baseline model are presented in Figure 1.

SIBLING AND NEIGHBOR CORRELATIONS

We find that brother correlations are large throughout at least the first 30-45 years of life: the correlation in general health status immediately after childhood (age 18) is roughly 0.60, and at ages 30-45 remains at roughly 0.55. To assess the importance of the shared genetic component of health, we attempted to contrast biological sibling correlations versus sibling correlations for step relations and adoptive ties. We do not find any significant differences between full biological siblings and other siblings (half/adopted), but very small sample sizes of step and adoptive ties prohibit any definitive evidence on this issue in the PSID.

Sibling correlations by themselves cannot disentangle how much of the resemblance among siblings in their health outcomes is due to the effects of family background and how much is due to the effects of neighborhood background. Augmenting the sibling correlation estimates with corresponding neighbor correlation estimates reveals family background is the most important determinant of general health status in adulthood. While the childhood neighbor

correlations are much smaller than the brother correlations, they are substantial throughout at least the first 30-45 years of life. In particular, the male childhood neighbor correlation immediately after childhood (age18) is roughly 0.20. By the time men are in their 30s and 40s, the correlation in general health status between childhood neighbors increases, reaching a peak of more than 0.35 between the ages of 30 and 35. Furthermore, by comparing the magnitudes of the sibling and neighbor correlations in adulthood health among males, the results indicate that roughly half of the average brother correlation in adulthood (0.55) may be attributable to neighborhood effects.

From the adjusted neighbor correlation estimates, we 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 males who grew up in the same neighborhood. The adjusted male neighbor correlation is roughly 10% lower than the unadjusted neighbor correlation.

Sensitivity Analyses

Robustness of the results on two dimensions was considered. First, we examined alternative specifications of health status. Specifically, models were estimated using: 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 brother correlations were qualitatively similar for these outcomes and our preferred health status measure.

We also sought to identify a health status measure available in the data 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 not causally influenced by neighborhood characteristics. Re-estimating our models with height as the dependent variable, we find that the neighbor correlation is zero, as expected. This suggests some evidence that the substantial neighbor correlations that are being estimated for the measure based on SAH are not due to spurious sorting of individuals with similar characteristics.

MAGNITUDE OF EFFECTS OF FAMILY AND NEIGHBORHOOD BACKGROUND

What do these correlation estimates mean in terms of the absolute size of the effects of family and neighborhood background? Estimates of the neighborhood random components (σ_n) indicate that neighborhood quality has very large and significant effects on general health status for males. To provide further insight and facilitate interpretation of our results, we present graphically in Figure 2 the predicted health trajectories in adulthood for the average man from an average family background, who grows up in : a) a neighborhood of average quality, b) a neighborhood that is one standard deviation below average neighborhood quality, and c) a neighborhood that is one standard deviation above average neighborhood quality, respectively.

Similarly, in Figure 3, we simulate the predicted health trajectories in adulthood for the average man from an average neighborhood background, who grows up in: a) a family of average background, b) a family that is one standard deviation below average family background, and c) a family that is one standard deviation above average family background, respectively. The graphical representation of the results highlights the dramatic impact of neighborhood and family background on health in adulthood.[9]

These effects undoubtedly contribute to current racial health disparities, given the well-known racial differences that exist in their family and neighborhood backgrounds. For example, Deaton & Paxson (1999) find that a quarter of white men report themselves in excellent health until their late 50s; among blacks, the same points are reached before age 40 among males. The sample we employ in this paper do not contain enough blacks to conduct a systematic analysis of the evolution of racial health disparities, but we plan to address this issue in future work.

EXPLAINING THE LEVEL AND AGE-PROFILE OF SIBLING & NEIGHBOR CORRELATIONS

We next estimate a series of models building toward a full model specification that includes a complete set of observable 1968 family-level and neighborhood-level background characteristics to attempt to identify determinants of health dynamics and the evolution of health inequality. In order to try to explain the level and age-profile of sibling and neighbor correlations, we estimate the following series of models that include 1968 family-level and neighborhood-level versions of each variable:

[9] In future work, we plan to also explore potential interactions between family and neighborhood effects—e.g., families who lack social and economic resources may be more vulnerable to poor neighborhood conditions (Currie & Hyson, 1999).

(1) baseline model
(2) (1)+race+family structure
(3) (2)+parental education
(4) (3)+parental income
(5) (4)+health insurance
(6) (5)+parental health behaviors
(7) (6)+neighborhood and housing quality measures

The inclusion of both family-level and neighborhood-level versions of each variable enables us to disentangle family-level and neighborhood contextual effects. One must use some caution, however, with drawing causal inferences from these coefficient estimates. The estimates are intended instead to summarize the relationships between the health trajectory in adulthood with various measures of neighborhood and family background.

The results are presented in Table 2. Note, across these model specifications, both random intercept and random slope coefficient terms at every level are only included if they significantly improve the fit of the model. From the parsimonious specification of Models (2), as shown in columns (2) of Table 2, we find that growing up in female-headed households and in predominantly black neighborhoods negatively affects health in adulthood. While the magnitude of the effect of growing up in black neighborhoods is sizable, it is not statistically significant likely because of the small numbers of blacks in our sample. The estimated average annual health deterioration rate is roughly 0.25 in adulthood, ages 18-45. Thus, based on the estimates in Model 2, the same level of health deterioration would be reached 25 years earlier among individuals who grew up in female-headed households relative to individuals who grew up with both parents. Similarly, the same level of health deterioration would be reached 16 years sooner among individuals who grew up in black neighborhoods relative to individuals who grew up in white neighborhoods.

Models (4)-(6) reveal that the negative effects of growing up in predominantly black neighborhoods and growing up in neighborhoods with a large proportion of female-headed households can be fully explained by the effects of lower average neighborhood income. The effects of both mean neighborhood income and family income during childhood have significant enduring effects on health in adulthood. In particular, the effect sizes imply that the same level of health deterioration would be reached 20 years sooner among individuals who grew up in neighborhoods with fifty percent lower mean incomes. Similarly, the same level of health deterioration would be reached 8 years sooner among individuals who grew up in families with

fifty percent lower incomes. We find that the educational attainment of the parental head has a marginally significant positive effect on health, net of the effects of family income. In the fuller model specifications, some of the significance of individual effects are reduced due to the significant degree of multicollinearity between the variables measuring SES background.

From Models (5) and (6), we see that health insurance coverage during childhood is associated with better health in adulthood. The neighborhood mean health insurance coverage rates are also large and significant. We interpret the latter coefficient as reflecting correlation between neighborhood mean health insurance coverage rates and unmeasured causal factors. While the coefficients on family- and neighborhood-level alcohol and cigarette consumption (parental health behaviors) are not significant, most of these effects are in the expected directions.

We next estimate a model (Model 7) that adds a vector of housing and neighborhood quality measures. These results are presented in the last two columns of Table 3. Among the housing and neighborhood quality measures, the neighborhood average rate of cockroach/rat problems, and prevalence of burglary, robbery, assault, drug use, or problems related to too few police in the neighborhood in which children grew up were associated with the largest lasting impacts on health in adulthood. We also see that the inclusion of these neighborhood quality measures reduces the significance of the effects of neighborhood mean income (Model 7b).

The full model that includes the complete set of observable neighborhood and family background characteristics explains a significant degree of the level and age-profile of sibling and neighbor correlations. In the specification of Model (6), the magnitude of the variance of the neighborhood-level components have been reduced by an average of 67 percent in adulthood by the addition of the neighborhood and family background measures (implied quasi-R^2 at the neighborhood level). As well, the magnitude of the variance of the family-level components have been reduced by an average of 36 percent in adulthood by the addition of the neighborhood and family background measures.

Finally, we estimate models that include the individual's own economic status in adulthood, as measured by labor earnings in adulthood. If inclusion of individual own labor earnings significantly reduces neighborhood random components and the associated conditional neighbor correlation is significantly smaller, this would provide (suggestive) evidence that childhood neighbors exhibit a significant degree of correlation in adulthood economic status that

is contributing to the observed significant neighbor correlations of health in adulthood. Solon and Page (2002) estimate the raw male childhood neighbor correlations in earnings in adulthood to be roughly 0.20.

Indeed, as shown in Table 4, we find that the magnitude of the variance of the neighborhood-level components is reduced dramatically by the inclusion of individual own earnings in adulthood—namely, the variance of the neighborhood-level components is reduced by an average of 90 percent in adulthood by the inclusion of individual own earnings in adulthood (implied quasi-R^2 at the neighborhood level). As well, we find that the magnitude of the variance of the family-level components is reduced by an average of 27 percent in adulthood by the inclusion of individual own earnings in adulthood.

In Figure 4, we graph the conditional neighbor correlations that adjust for the variables in Models (7b) and (8) along with the unadjusted neighbor correlations; and in Figure 5 we graph the conditional sibling correlations that adjust for the variables included in Models (7b) and (8) along with the unadjusted sibling correlations. The graphs containing the conditional sibling and neighbor correlations reveal that similarity in childhood neighbors' subsequent economic status in adulthood can account for the lion share of the resemblance in childhood neighbors' subsequent health outcomes in adulthood. As well, after controlling for observable neighborhood background variables, the similarity of childhood neighbors' health outcomes is less marked. Recall that the adjusted neighbor correlations showed that the neighbor correlations were not driven by similarity of family background characteristics, but seem to reflect neighborhood effects. After controlling for observable neighborhood and family background variables, the similarity of brothers' health in adulthood is also significantly reduced. The broad array of available measures of family and neighborhood background, that are in many ways unique to the PSID, is a tremendous asset to our analyses.

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. On the other hand, while one must use caution with attaching causal inferences from these results, this evidence taken together is consistent with (provides strong support for) the hypothesis that family income 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 beyond the scope of the present paper, but remains an important area for future research (Adda et al., 2003).

Characterizing the Dynamics of Health Inequality. The usual indices of inequality are derived from observations on health, income, wealth, etc., corresponding to a particular point in time. As we have demonstrated in the case of health, cross-sectional health inequality measures by themselves do not accurately reflect the differences between individuals, since the true situation depends to large extent on how the relative positions of individuals vary over time. A class of mobility measures have been developed in the income inequality literature that considers the relationship between the long-run and the short-run distributions of income (Shorrocks, 1978). We apply these methods in a novel way to health. Health mobility may be defined as a measure of movements over the life course within the health distribution.

The basic concept of mobility adopted here was introduced by Shorrocks (1978). A comparison of inequality in the distribution of long-run health, measured over 29 years, age t = 17, 18,…, 45, with some representative value of inequality in the short-run (single year) distributions reveals the degree of mobility, stability, or equalization over time. Measurement of inequality in long-run health is of interest in itself since such measurements can be free of short-run transitory effects. We estimate Shorrocks' rigidity index R which relates intra-distributional mobility to permanent inequality. The rigidity index is defined as

$$R \equiv \frac{I\left(\sum_{t=17}^{45} H_t^*\right)}{\sum_{t=17}^{45} w_t I\left(H_t^*\right)} \quad ,$$

where $I(\cdot)$ is Thiel's entropy inequality index (mean logarithmic deviation), H^* is a vector of health measures, and t = 17, 18,…,45, denotes age. Cross-section (annual) inequality is weighted using shares of health (quality of life) at age t in total health (quality of life) in the 29-year period. This index, then measures the proportion by which inequality for quality of life measured over roughly a 30-year period is lower than a weighted average of cross-section inequalities.

In order to construct an age-profile of health inequality using this inequality measure requires balanced panel data. Thus, we use each individual's estimated random intercept and

slope coefficients to predict the level of health for each individual at each age t=17, 18,…, 45. We use these predictions to create a "simulated" balanced panel of health measures for each individual between ages 17 and 45.

Additionally, the shape of the sequence containing the rigidity values as the time horizon is lengthened one additional period (i.e., the rigidity profile) helps examine the nature of health changes—i.e., transitory vs. permanent health changes.[10] The graphical representation of Shorrocks' rigidity index applied to health is presented in Figure 6. These results highlight the significant degree of persistence in health inequality over a 30-year period of adulthood health.

Shorrocks' rigidity index R is particularly useful for our analysis of the determinants of health dynamics and the evolution of health inequality because it can be additively decomposed into a between and a within groups components, which in turn, help reveal the importance of neighborhood and family background. More precisely, R can be decomposed into between (R^B) and within (R^W) rigidity, weighted by the share of between groups (S^B) and within group (S^W) inequality in total cross-section inequality:

$$R = \frac{I^B\left(\sum_{t=17}^{45} H_t^*\right)}{\sum_{t=17}^{45} w_t I^B\left(H_t^*\right)} \frac{\sum_{t=17}^{45} w_t I^B\left(H_t^*\right)}{\sum_{t=17}^{45} w_t I\left(H_t^*\right)} + \frac{I^W\left(\sum_{t=17}^{45} H_t^*\right)}{\sum_{t=17}^{45} w_i I^W\left(H_i^*\right)} \frac{\sum_{t=17}^{45} w_t I^W\left(H_t^*\right)}{\sum_{t=17}^{45} w_t I\left(H_i^*\right)} = R^B S^B + R^W S^W$$

In future work, we plan to decompose Shorrocks' rigidity index into its between and within groups components to further our understanding of the importance of neighborhood and family background.

CONCLUDING REMARKS AND DIRECTION FOR FUTURE RESEARCH

In this paper, we have used correlations based on nationally representative longitudinal sample of siblings and neighbors to estimate upper bounds on the possible causal effects of family and neighborhood background on general health status in early adulthood through mid life. Estimates based on four-level hierarchical random effects models consistently show a higher scope for family background (whether emanating from nature or nurture) than for neighborhood background. However, our estimates suggest that disparities in neighborhood background factors account for between one-fourth and one-third of the variation in health status

[10] We wish to thank Philippe Van Kerm for sharing programs that assisted us in estimating and decomposing Shorrocks' rigidity index.

among males in mid life. While the neighborhood correlations must be strictly interpreted as upper bounds, the estimates suggest that childhood neighborhood factors play an important role in the intergenerational transmission of health status and influence both contemporaneous and future health outcomes.

The evidence presented in this paper challenges future research to further our understanding of the underlying processes that produce health disparities between different racial, ethnic, and socioeconomic groups. Future theoretical models and empirical analyses must incorporate and carefully model how the timing of neighborhood exposures intersects with the trajectory of health outcomes and how neighborhood effects may vary over the life-course.

This evidence suggests further research on the effects of particular neighborhood characteristics is warranted. In order to assess the policy implications of this research, we need a better understanding of the pathways through which neighborhoods and families 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. We have focused on quantifying the potential overall magnitude of family and neighborhood effects. Disentangling the causal sources of neighborhood effects is extremely difficult (Manski 1993; Moffitt 1998), but the decomposition and investigation into the mechanisms of why neighborhoods matter are an important next step and area for future research.

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.

Barsky, Robert B., F. T. Juster, M. S. Kimball and M. D. Shapiro. 1997. "Preference Paramaters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study". *The Quarterly Journal of Economics*. May, 112(2): 1937-1996.

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.

Bourguignon, Francois. 1979. "Decomposable Income Inequality Measures". *Econometrica*. 47 (4):901-920.

Cain, G. and H. Watts. 1972. "Problems in Making Policy Inferences from the Coleman Report". *American Sociological Review*. 35(2): 228-252.

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.

Conley, D. and N. Bennett. 2000. "Is Biology Destiny? Birth Weight and Life Chances". *American SociologicalReview*. 65:458-67.

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.

Currie, Janet and Rosemary Hyson. 1999. "Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birthweight". *The American Economic Review*. May, 89(2): 245-250.

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.

Dow, William H., J. Holmes, T. Philipson, and X. Sala-i-Martin.1995. "Disease Complementarities and the Evaluation of Public Health Interventions". Working Paper no. 5216. National Bureau of Economic Research.

Dow, William H., T. Philipson, X. Sala-i-Martin. 1999. "Longevity Complementarities under Competing Risks". *The American Economic Review*. Dec., 89(5): 1358-1371.

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.

Duncan, Greg J. and Stephen W. Raudenbush. 1999. "Assessing the Effects of Context in Studies of Child and Youth Development". *Educational Psychologist*. 34(1): 29-41.

_____. 2001. “Neighborhoods and Adolescent Development: How Can We Determine the Links”. Unpublished manuscript. Northwestern University.

Durlauf, Steven N. 2000. “The Memberships Theory of Poverty: The Role of Group Affiliations in Determining Socioeconomic Outcomes”. Unpublished manuscript. Department of Economics, University of Wisconsin.

_____. 2001. “A Framework for the Study of Individual Behavior and Social Interactions”. Unpublished manuscript. Department of Economics, University of Wisconsin.

_____. 2002. “Groups, Social Influences and Inequality: A Memberships Theory Perspective on Poverty Traps”. Unpublished manuscript. Department of Economics, University of Wisconsin.

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.

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.

Ganz, Michael Lee. 2001. “Health Threats and Health Practices”. Unpublished manuscript. Department of Maternal and Child Health, Harvard School of Public Health.

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. 1996. “Black/White Differences in the Relationship of Maternal Age to Birthweight: A Population-Based Test of the Weathering Hypothesis”. *Social Science Medicine*. 42(4): 589-597.

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.

Hayward, Mark D. and Bridget K. Gorman. "The Long Arm of Childhood: The Influence of Early Life Social Conditions on Men's Mortality". Unpublished manuscript. Population Research Institute, Penn State University.

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

Johnson, Rucker C. "Risk Preference Formation and Health Behaviors in Adolescence and Adulthood: The Importance of Neighborhood and Family Background". Unpublished manuscript, University of Michigan

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.

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.

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.

Lillard, Dean and Richard Burkhauser. 2003. "Income Inequality and Health: a Longitudinal Study". Unpublished manuscript. Prepared for Joint EPUNet-2003 and BHPS-2003, Colchester, United Kindom.

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

Maasoumi, Esfandiar and Sourushe Zandvakili. 1986. "A Class of Generalized Measures of Mobility with Applications". *Economic Letters*. 22:97-102.

Moffitt, R. 1998. "Policy Interventions, Low-Level Equilibria and Social Interactions". Department of Economics, Johns Hopkins University.

Morenoff, Jeffrey D. Forthcoming. "Neighborhood Mechanisms and the Spatial Dynamics of Birthweight." American Journal of Sociology.

Morenoff, Jeff, and John Lynch. 2002. "What makes a place healthy? Neighborhood influences on racial/ethnic disparities in health over the life course." Paper prepared for a Workshop on Ethnic Disparities in Aging Health, Washington, DC.

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

Oreopoulos, Phillip. 2002. "The Long-Run Consequences of Living in a Poor Neighborhood". Unpublished manuscript.

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.

Roberts E.M. 1999. "Socioeconomic position and health: The independent contribution of community socioeconomic context". *Annual Review of Sociology*. 25:489-516.

Rosenthal, R. and D.R. Rubin. 1982. "Comparing Effect Sizes of Independent Studies". *Psychoglogical Bulletin*. 92:500-504.

Shorrocks, A. F. 1978. "Income Inequality and Income Mobility". *Journal of Economic Theory*. 19:376-393.

_____. 1978. "The Measurement of Mobility". *Econometrica*. September 46 (5):1013-1024.

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.

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

Timberlake, Jeffrey M. 2002. "Racial and Ethnic Inequality in Childhood Exposure to Neighborhood Poverty and Affluence". Unpublished manuscript. Department of Sociology, University of Chicago.

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

Yen I.H. and S. L. Syme. 1999. "The social environment and health: A discussion of the epidemiologic literature". *Annual Review of Public Health*. 20:287-308.

Table 1.
Four-Level Hierarchical Random Effects Interval Regression Model

	Baseline Model
constant	93.4705 ***
	(1.0993)
age_17	-0.2379 ***
	(0.0661)
sig_n0	4.7263 ***
	(1.5018)
sig_n1	0.2566 ***
	(0.0740)
nrho	-0.0353
	(0.4768)
sig_f0	5.9921 ***
	(1.2255)
sig_f1	0.4284 ***
	(0.0606)
frho	-0.7914 ***
	(0.1028)
sig_i0	6.0361 ***
	(0.9637)
sig_i1	0.3914 ***
	(0.0276)
irho	-0.3656 **
	(0.1458)
sig_eps	5.5982 ***
	(0.0279)
ln-L	-5160.32

NOTE: Asymptotic standard errors in parentheses;
Significance: '*'=10%; '**'=5%; '***'=1%.

Random Effect Terms:
sig_n0= neighborhood random intercept component
sig_n1= neighborhood random slope/age component
nrho=correlation between neighborhood random intercept and slope terms

sig_f0= family random intercept component
sig_f1= family random slope/age component
frho=correlation between family random intercept and slope terms

sig_i0= individual random intercept component
sig_i1= individual random slope/age component
irho=correlation between individual random intercept and slope terms
sig_eps=transitory error component

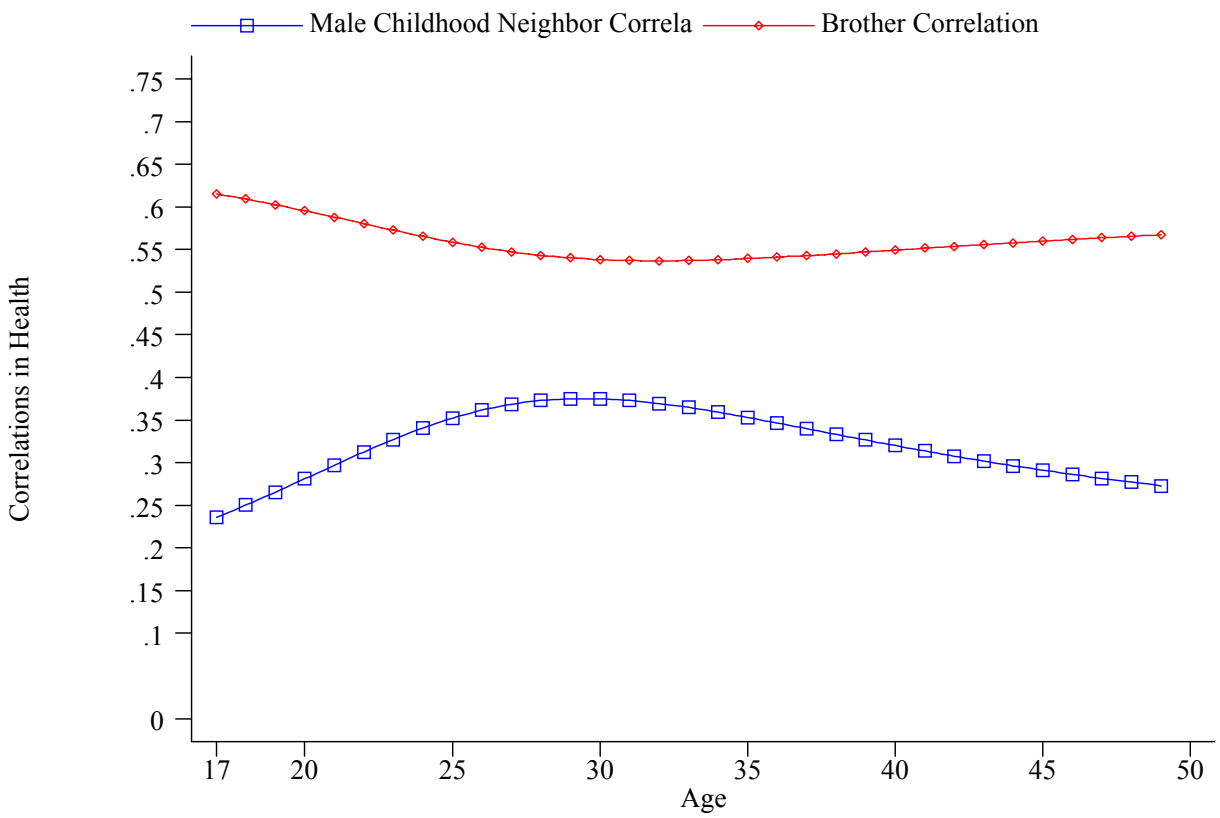


Figure 1. Male Sibling and Neighbor Correlations, by Age

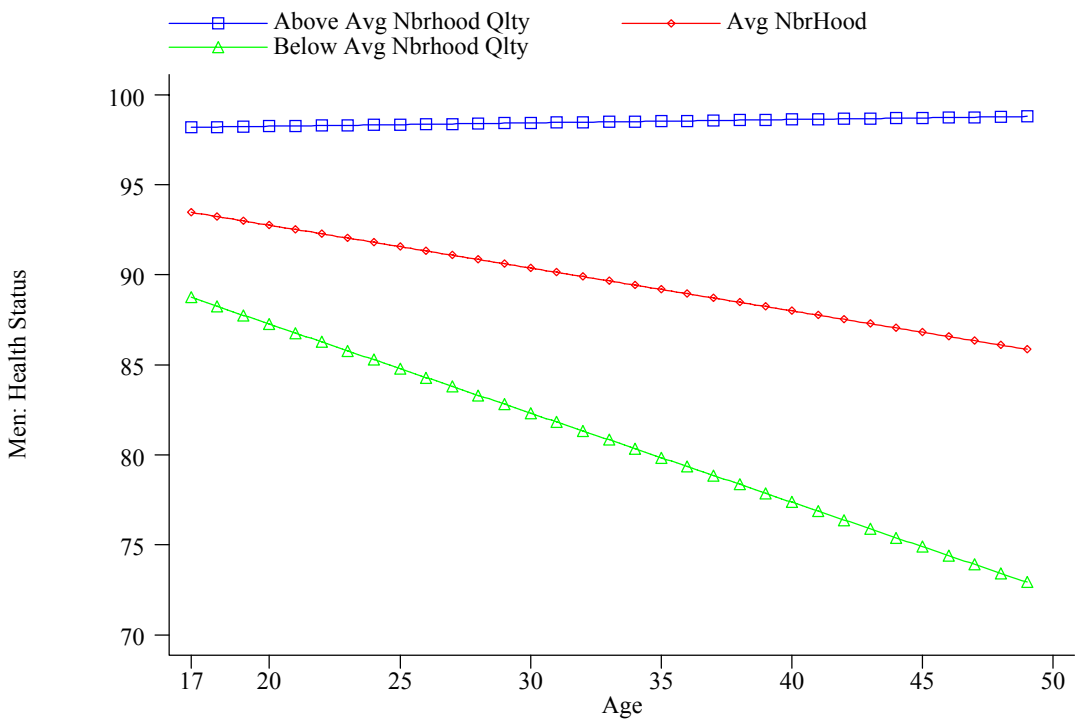


Figure 2. Health Status for Men in Various Levels of Neighborhood Quality

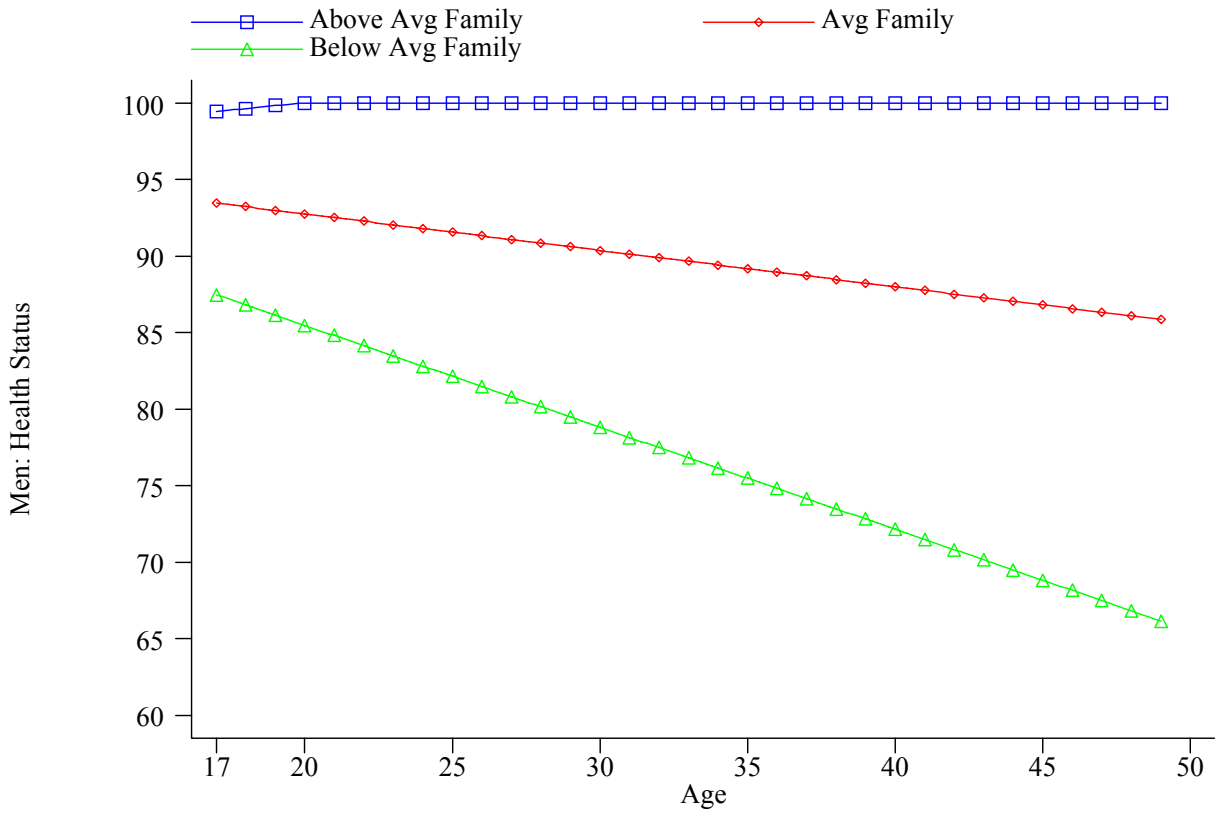


Figure 3. Health Status for Men in Various Levels of Family Quality

Table 2. Family & Neighborhood Background Effects
(Asymptotic standard errors in parentheses)
Significance: '*'=10%; '**'=5%; '***'=1%.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	93.4705 ***	93.6322 ***	93.3179 ***	94.0744 ***	93.9044 ***	94.1122 ***
	(1.0993)	(1.2928)	(1.2172)	(1.3736)	(1.3022)	(1.9702)
Age – 17	-0.2379 ***	-0.3097 ***	-0.2844 ***	-0.2910 ***	-0.2822 ***	-0.2984 ***
	(0.0661)	(0.0581)	(0.0613)	(0.0566)	(0.0646)	(0.0630)
Black		0.6038	0.2704	-1.1048	0.3140	3.0805
		(9.0842)	(9.0266)	(11.5395)	(10.3661)	(10.8190)
HMBlack		-4.0155	-2.3368	0.0040	3.3218	0.5817
		(11.4059)	(9.6550)	(12.0748)	(11.7219)	(13.6839)
Married		6.4339 ***	8.5111 **	6.3437 **	9.0181 ***	8.5991 **
		(2.1840)	(3.4994)	(2.7467)	(3.1937)	(3.6047)
HMMarry		5.0703	-0.3037	-14.1075 *	-9.4065	-13.6557 *
		(5.2793)	(8.3309)	(7.1137)	(7.5584)	(7.7658)
ParHdEd			0.6302	0.7406 *	1.0067 *	0.7124
			(0.4857)	(0.4436)	(0.5348)	(0.5454)
HMHdEd			-0.6441	-1.6378	-1.6249	-1.1434
			(0.9630)	(0.9905)	(1.0857)	(1.1601)
ParWfEd			-0.6611	-0.6726	-1.0064	-0.8691
			(0.6355)	(0.4967)	(0.6267)	(0.6626)
HMWfEd			2.5792 *	2.2762	1.5035	1.6324
			(1.5303)	(1.4931)	(1.3378)	(1.4336)
FamInc				2.1399	1.3239	2.0771
				(1.6277)	(1.7172)	(1.9669)
HMFInc				6.3150 **	6.1455 *	5.9079 *
				(2.8394)	(3.4554)	(3.3951)
HlthIns					1.3701	2.2222
					(2.2021)	(2.1123)
HMHthIns					11.7115 ***	10.8684 ***
					(3.7527)	(3.7397)
Alcohol						-0.0602
						(0.1184)
HMDrink						-0.0432
						(0.1923)
CigSmoke						0.0422
						(0.0967)
HMSmoke						-0.0282
						(0.2089)
sig_n0	4.7263 ***	4.1136 ***	4.9954 ***	4.5398 ***	3.7489 ***	3.7786 ***
	(1.5018)	(0.4917)	(0.5596)	(0.6908)	(0.6526)	(0.8427)
sig_n1	0.2566 ***					
	(0.0740)					
Nrho	-0.0353					
	(0.4768)					
sig_f0	5.9921 ***	5.5364 ***	5.2676 ***	5.1325 ***	5.4654 ***	5.8292 ***
	(1.2255)	(0.9519)	(1.0341)	(1.1998)	(1.1771)	(1.1634)
sig_f1	0.4284 ***	0.3953 ***	0.3975 ***	0.4382 ***	0.3972 ***	0.3998 ***
	(0.0606)	(0.0428)	(0.0458)	(0.0458)	(0.0521)	(0.0510)
Frho	-0.7914 ***	-0.7680 ***	-0.8599 ***	-0.9320 ***	-0.9322 ***	-0.9046 ***
	(0.1028)	(0.0905)	(0.0793)	(0.0683)	(0.0588)	(0.0796)
sig_i0	6.0361 ***	5.5382 ***	5.9713 ***	6.1053 ***	5.8840 ***	5.7293 ***
	(0.9637)	(0.8996)	(0.7710)	(0.7532)	(0.8833)	(0.9073)
sig_i1	0.3914 ***	0.4036 ***	0.3900 ***	0.3880 ***	0.4062 ***	0.3887 ***
	(0.0276)	(0.0244)	(0.0290)	(0.0276)	(0.0315)	(0.0272)
Irho	-0.3656 **	-0.4189 ***	-0.4137 ***	-0.3942 ***	-0.3269 **	-0.3955 ***
	(0.1458)	(0.1027)	(0.1057)	(0.1057)	(0.1271)	(0.1202)
sig_eps	5.5982 ***	5.4616 ***	5.4607 ***	5.3740 ***	5.4243 ***	5.3642 ***
	(0.0279)	(0.0277)	(0.0292)	(0.0253)	(0.0267)	(0.0291)
ln-L	-5160.32	-5149.75	-5103.24	-5052.06	-5041.82	-5042.21

Variable Definitions:
Age – 17 = age is centered around 17, and all other explanatory variables are centered around their grand means or set at their modal value for categorical variables, so that the grand intercept term can be interpreted as the health level at the beginning of adulthood/end of childhood for an individual with average neighborhood and family background characteristics.

1968 Family-level and Neighborhood-level measures of the same concept are entered in the model.

Black=parental head is black
HMblack=neighborhood %Black
Married=parental family includes both head and wife
HMMarry=neighborhood % of household heads that are married

Parental Family SES:
ParHdEd=education level of parental head
HMHdEd=neighborhood mean average education of household head
ParWfEd=education level of wife, if present
HMWfEd=neighborhood mean average education of wife, if present
Faminc=log(average annual family income), where 3-year average based on reports in 1968-1970 IWs
HMFInc=log(neighborhood mean average income), based on 3-year average
HlthIns=parental family had health insurance
HMHthIns=neighborhood % of households with health insurance

Parental Health Behaviors:
Alcohol=annual expenditures of 1968 parental family on alcohol
HMDrink=neighborhood mean annual family expenditure on alcohol
CigSmoke= annual expenditures of 1968 parental family on cigarettes
HMSmoke=neighborhood mean annual family expenditure on cigarettes

Neighborhood Quality:
HMRoach=neighborhood rate of rat/cockroach problems
HMCrime=mean neighborhood prevalence of burglary, robbery, assault, drug use,too few police

Random Effect Terms:
sig_n0= neighborhood random intercept component
sig_n1= neighborhood random slope/age component
nrho=correlation between neighborhood random intercept and slope terms
sig_f0= family random intercept component
sig_f1= family random slope/age component
frho=correlation between family random intercept and slope terms
sig_i0= individual random intercept component
sig_i1= individual random slope/age component
irho=correlation between individual random intercept and slope terms
sig_eps=transitory error component

Table 3. Family & Neighborhood Background Effects

	Model 1	Model 7a	Model 7b
Constant	93.4705 ***	95.8679 ***	94.8254 ***
	(1.0993)	(0.8712)	(1.7309)
Age - 17	-0.2379 ***	-0.2721 ***	-0.2970 ***
	(0.0661)	(0.0605)	(0.0704)
HMRoach		-8.9402 ***	-4.9749
		(1.8342)	(4.3957)
HMCrime		-19.7755 ***	-6.1097
		(2.8397)	(4.6733)
Black			2.9315
			(10.2621)
HMblack			1.5572
			(12.2217)
Married			7.8470 **
			(3.5908)
HMmarry			-14.3404 *
			(7.3942)
ParHdEd			0.7310
			(0.6122)
HMHdEd			-1.5148
			(1.0572)
ParWfEd			-0.7839
			(0.6639)
HMWfEd			2.4110 *
			(1.2937)
FamInc			1.7411
			(2.0500)
HMFInc			4.9596
			(3.3359)
HlthIns			2.3059
			(2.3783)
HMHthIns			11.5153 ***
			(4.2257)
Alcohol			-0.0559
			(0.1228)
HMDrink			-0.1352
			(0.1720)
CigSmoke			0.0852
			(0.1008)
HMSmoke			-0.0278
			(0.2207)
sig_n0	4.7263 ***	6.0630 ***	3.9393 ***
	(1.5018)	(0.4056)	(0.6711)
sig_n1	0.2566 ***		
	(0.0740)		
Nrho	-0.0353		
	(0.4768)		
sig_f0	5.9921 ***	4.2504 ***	5.5126 ***
	(1.2255)	(1.0172)	(1.2046)
sig_f1	0.4284 ***	0.3547 ***	0.3862 ***
	(0.0606)	(0.0479)	(0.0509)
Frho	-0.7914 ***	-0.9742 ***	-0.8866 ***
	(0.1028)	(0.0254)	(0.0839)
sig_i0	6.0361 ***	6.2942 ***	5.9274 ***
	(0.9637)	(0.8544)	(1.0102)
sig_i1	0.3914 ***	0.4263 ***	0.3947 ***
	(0.0276)	(0.0296)	(0.0319)
Irho	-0.3656 **	-0.3166 ***	-0.4214 ***
	(0.1458)	(0.1161)	(0.1184)
sig_eps	5.5982 ***	5.6339 ***	5.4188 ***
	(0.0279)	(0.0268)	(0.0292)
ln-L	-5160.32	-5150.10	-5041.00

Table 4. Controlling for Own Labor Income in Adulthood

	Model 1	Model 8
Constant	93.4705 ***	94.2077 ***
	(1.0993)	(0.8269)
Age – 17	-0.2379 ***	-0.3034 ***
	(0.0661)	(0.0631)
Ln(OwnLaborInc)		0.6154 **
		(0.2440)
Sig_n0	4.7263 ***	2.0859 **
	(1.5018)	(1.0414)
Sig_n1	0.2566 ***	
	(0.0740)	
Nrho	-0.0353	
	(0.4768)	
Sig_f0	5.9921 ***	5.0146 ***
	(1.2255)	(0.7340)
Sig_f1	0.4284 ***	0.3924 ***
	(0.0606)	(0.0506)
Frho	-0.7914 ***	-0.8471 ***
	(0.1028)	(0.0858)
Sig_i0	6.0361 ***	5.1909 ***
	(0.9637)	(0.7527)
Sig_i1	0.3914 ***	0.3606 ***
	(0.0276)	(0.0275)
Irho	-0.3656 **	-0.3659 ***
	(0.1458)	(0.1091)
Sig_eps	5.5982 ***	5.2803 ***
	(0.0279)	(0.0367)
ln-L	-5160.32	-4653.03

NOTE: Asymptotic standard errors in parentheses;
Significance: '*'=10%; '**'=5%; '***'=1%.

Random Effect Terms:
sig_n0= neighborhood random intercept component
sig_n1= neighborhood random slope/age component
nrho=correlation between neighborhood random intercept and slope terms

sig_f0= family random intercept component
sig_f1= family random slope/age component
frho=correlation between family random intercept and slope terms

sig_i0= individual random intercept component
sig_i1= individual random slope/age component
irho=correlation between individual random intercept and slope terms
sig_eps=transitory error component

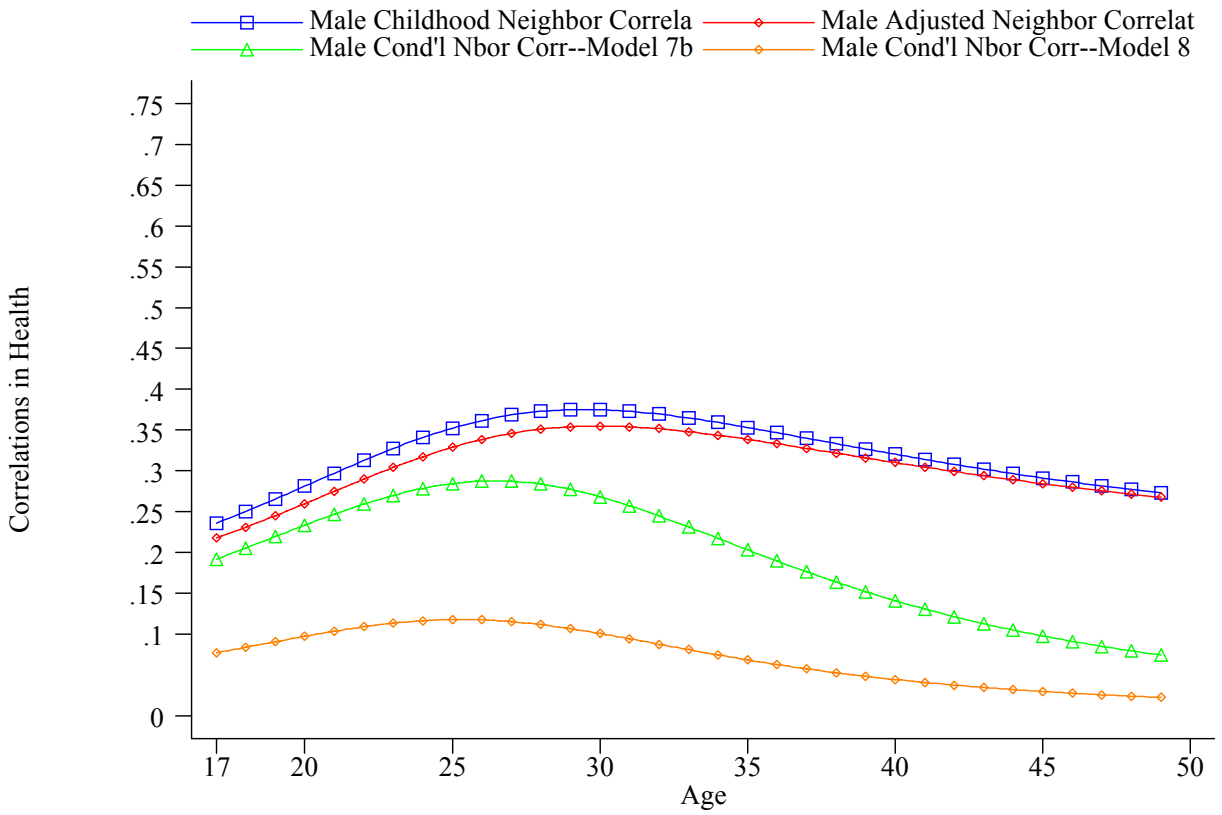


Figure 4. Neighbor Correlations Adjusting for Various Observed Factors

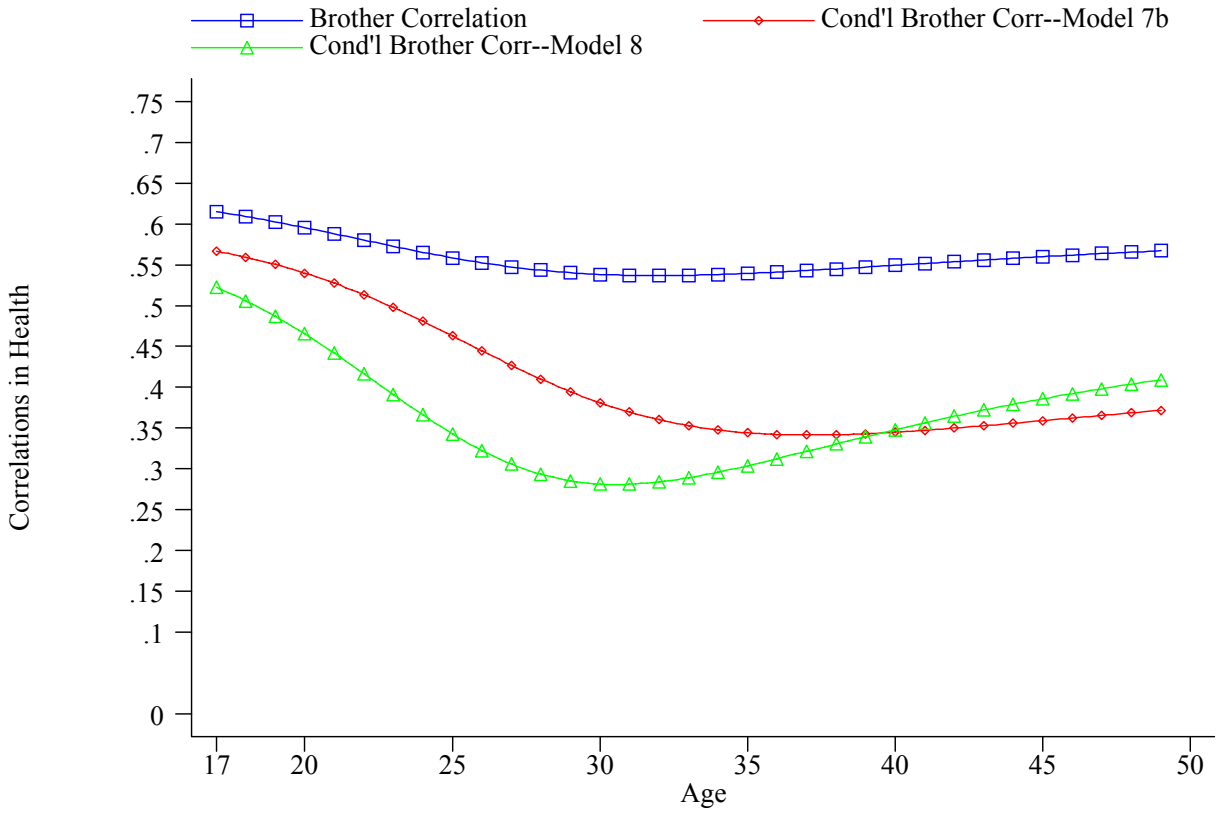
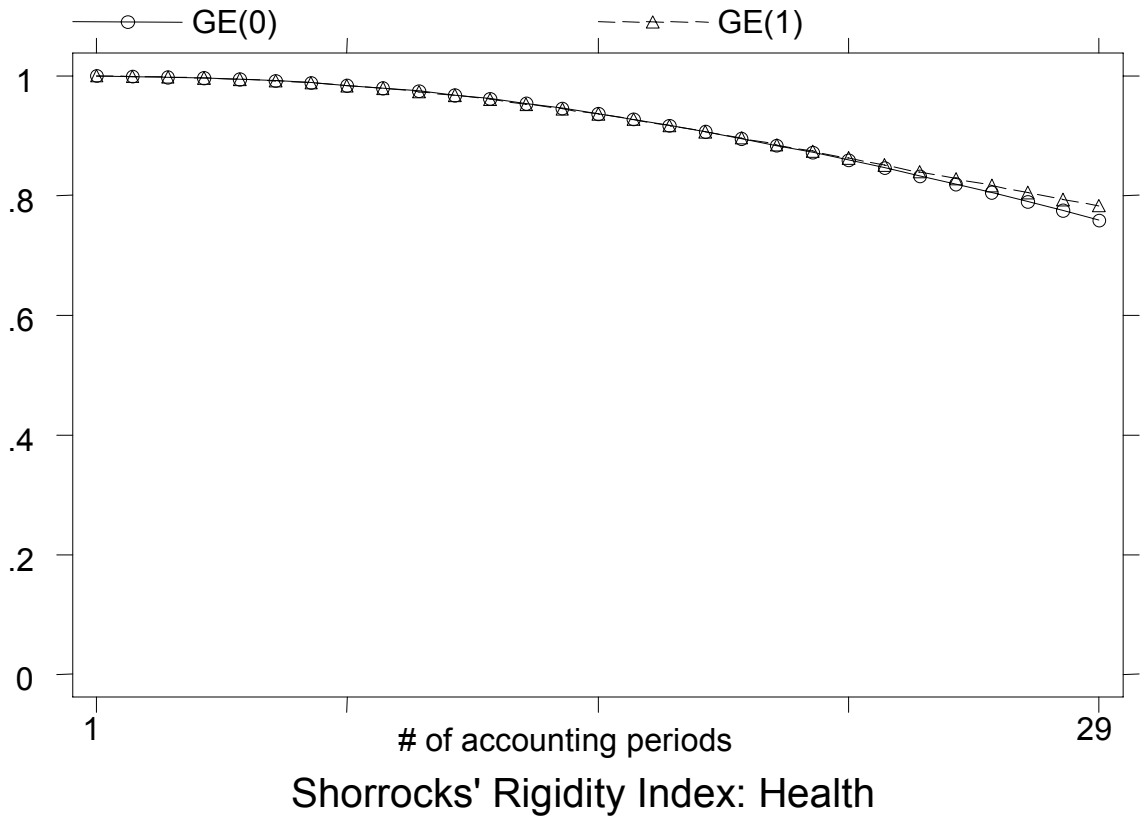


Figure 5. Brother Correlations Adjusting for Various Observed Factors

Shorrocks [1978]'s Rigidity/Immobility Index applied to Health:
using health variable that lies in interval [0,100] and is increasing in health



Shorrocks [1978]'s rigidity index applied to Health for Men:

Using GE(0) inequality index : 0.24075
Using GE(1) inequality index : 0.21729
Thiel's generalized entropy inequality index (mean logarithmic deviation)