FAMILY AND COMMUNITY INFLUENCES ON HEALTH AND SOCIOECONOMIC STATUS:

SIBLING CORRELATIONS OVER THE LIFE COURSE*

Bhashkar Mazumder

Chicago IL

bmazumder@frbchi.org

September 15, 2010

*Prepared for PSID Small Grants Workshop

Abstract

This paper presents new estimates of sibling correlations in health and socioeconomic outcomes over the life course. Sibling correlations provide an omnibus measure of the importance of all family and community influences. I find that sibling correlations in a range of health and socioeconomic outcomes start quite high at birth and remain high over the life course. The sibling correlation in birth weight is estimated to be 0.5. Sibling correlations in test scores during childhood are as high as 0.6. Sibling correlations in adult wages are also around 0.5. Decompositions using techniques developed in Mazumder (2008) offers a convenient and easily understood framework for analyzing the contributions to family inequality in both health and SES. While the evidence in this paper is primarily descriptive in nature it may shed light on some of the sources of disparities over the life course. Preliminary evidence using this approach suggests that there are some limited opportunities for policies to ameliorate the gradients in health and SES. For example, it is possible that improvements in family income when children are between the ages of 6 and 10 may reduce disparities in cognitive and non-cognitive skills.

I. Introduction

There is a growing recognition among economists that health is a critical component of human capital and plays a critical in determining socioeconomic status (Currie, 2009). Recent work has emphasized that key periods of child development may be especially susceptible to environment conditions that affect health capital which in turn may affect socioeconomic success (Cunha and Heckman, 2007). There is, of course, also a long-standing literature in the social sciences that has documented gradients in adult health and mortality by socioeconomic status (SES), suggesting that there are important effects running in the other direction, i.e. from SES to health (e.g. Grossman, 2005). Therefore it would be valuable to gain greater insight into how gradients in both health and socioeconomic status arise over the lifecycle by documenting these patterns.

It is not obvious, however, how to best summarize lifecycle patterns in the formation of these gradients. A potentially useful way to reduce the dimensionality of the issue is to focus on the role of gradients in the health-SES relationship that arise due to differences in family background. This paper attempts to take such an approach by estimating sibling correlations in measures of health and SES over various stages of the life course using variance decomposition models. Sibling correlations provide a straightforward summary measure of the combined effects of family background and community influences on a particular outcome. The measure estimates the fraction of the overall variation that is attributable to differences across families and therefore provides a simple gauge of how much of a gradient exists by a summary measure of family background.

An added virtue of using variance decomposition models to estimate sibling correlations is that it provides a useful way to add covariates into the analysis.[1] This provides a means for assessing the extent to which particular family background characteristics (e.g. family income, parent education) can account

[1] In an earlier paper (Mazumder, 2008) I introduce a simple framework for including a variety of covariates into variance decomposition models using restricted maximum likelihood (REML). In that paper I used the NLSY where I was not able to incorporate the potential effects of early life health as would be possible for the first time with the PSID.

for the sibling correlation. A decomposition of the sibling correlation provides additional useful descriptive that may provide "first order" evidence to guide policymakers.

Acquiring the necessary data for such a rich descriptive analysis is also a daunting challenge. I use data from the PSID which has genealogical data covering 40 years. This enables me to track siblings from the original PSID sample into their middle age. I also take advantage of the detailed questions included in the Child Development Survey (CDS) to track more recent cohorts of siblings.

II. Background and Previous Literature

A long literature in sociology and economics has tried to estimate the importance of family background on children's future SES. Early studies were often missing variables on key family background characteristics which led to the concern that family background might matter more than what the available measures would indicate. As a result researchers began to examine the sibling correlation as an alternative approach to measuring the importance of family background (e.g. Corcoran et al; 1976). Conceptually, the sibling correlation provides a summary statistic that captures *all* of the effects of sharing a common family as well as any other shared factors (e.g. common neighborhoods, school quality).[2] If the similarity in say, general health status between siblings is not much different compared to randomly chosen individuals, then we would expect a small correlation. If, however, a large fraction of the variance in health is due to factors common to growing up in the same family environment then the correlation might be sizable.

A number of studies have exploited the panel dimension of the PSID to estimate sibling correlation in economic outcomes using multiple years of data and have generally estimated correlations of around 0.3 to 0.4 (e.g. Solon et al., 1991; Bjorklund et al, 2002; Solon and Page, 2003). Other studies have utilized the NLS original cohort of young men and found broadly similar results (Altonji and Dunn,

[2] Conversely, many aspects of family background including genetic traits and sibling-specific parental behaviors will not be captured.

1991; Ashenfelter and Zimmerman, 1997; Levine and Mazumder, 2007. Levine and Mazumder (2007) and Mazumder (2008) found larger estimates of around 0.5 when utilizing the NLSY79.

Only a few studies have done any kind of decomposition of sibling correlations. Altonji and Dunn (2000) find evidence of unobserved preferences for work hours among family members using a factor model. Solon Page and Duncan (2000) and Solon and Page (2003) decompose the sibling correlations in schooling and earnings into factors that may be related to neighborhood effects and generally interpret their findings to suggest a small role for neighborhoods. In a recent paper (Mazumder, 2008) I introduce a statistical approach and produce an extensive set of decompositions into factors attributable to a wide range of factors including human capital (education, test scores), physical characteristics (height, weight, BMI), socially deviant behaviors (jail, drug use) and psychological characteristics (Rotter scale, self esteem). As might be expected, I found that human capital can explain 50 percent or more of the brother correlation in wages and earnings, but I also found that non-cognitive measures such as deviant behavior and psychological characteristics can also account for around 20 percent of these correlations. An important omission in that analysis, however, was health either as an outcome or as a covariate for SES. Further, the analysis was not done throughout the life course.

Bjorklund et al (2008) use the decomposition approach from Mazumder (2008) to estimate which specific characteristics of parents (in addition to parent income) may account for the sibling correlation in income using Swedish data. They find that accounting for parent attitudes and parental involvement can account for some portion of the observed sibling correlation.

III. Methodology

A. Statistical Model

I utilize the following statistical framework.[3] Each outcome is denoted by y_{ijt}, where i indexes families, j indexes siblings and t indexes years (for outcomes observed in multiple years).[4] Outcomes are then modeled as follows:

$$y_{ijt} = \beta X_{ijt} + \varepsilon_{ijt} \quad (1)$$

Here, the vector X_{ijt}, will typically contain age dummies to account for lifecycle effects and a female dummy when both sexes are pooled. For economic outcomes, year effects are included to account for business cycle conditions. The residual, ε_{ijt}, which is purged of these effects is then decomposed as follows:

$$\varepsilon_{ijt} = a_i + u_{ij} + v_{ijt} \qquad (2)$$

The three terms on the right hand side of (2) are treated as random effects that are assumed to be independent of each other.[5] The first term, a_i, is the permanent component that is common to all siblings in family i. The second term, u_{ij}, is the permanent component that is individual-specific. For outcomes for which we use repeated observations on the same individual, v_{ijt}, represents the transitory component that reflects noise due to either temporary shocks or measurement error in the survey.[6] In the case of outcomes that are not repeated, v_{ijt} is omitted. The variance of for example, age-adjusted earnings, ε_{ijt}, then is simply:

$$\sigma^2_{\varepsilon} = \sigma^2_a + \sigma^2_u + \sigma^2_v \qquad (3)$$

[3] The notation here follows Solon et al 1991 and has also been used by Bjorklund et al (2002) and Mazumder (2008) along with various other studies.
[4] Many of the non-economic outcomes considered in this study are only measured in one year or are only taken from one year.
[5] The assumption that a_i and u_{ij} are uncorrelated is purely for analytical convenience and allows one conceptually, to divide the permanent component into a part that is perfectly correlated among siblings, and a part that is perfectly uncorrelated among siblings. Previous studies have found that there is little or no *cross-sectional* correlation in the transitory component.
[6] The model can also be extended to account for serial correlation in the transitory component.

The first term, σ^2_a, captures the variance in the permanent component that is due to differences *between* families while the second term, σ^2_u, captures the variance in the permanent component that is due to differences *within* families. These two components are then used to construct the correlation in permanent outcomes between siblings, ρ.

$$\rho = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_u^2}, \quad (4)$$

Some earlier studies such as Solon at al (1991) and Bjorklund et al (2002) used a two step approach to estimate the variance components in this so called "mixed model" (mixed because it contains both fixed effects, or regressors, and random effects). First they use a regression to estimate (1) and to produce the residuals and then used analysis of variance (ANOVA) formulas on the residuals. Due to the unknown properties of ANOVA when the data is unbalanced (e.g. differing number of siblings per formula), Mazumder (2008) introduced the use Restricted Maximum Likelihood (REML). REML has a number of advantages such as consistency, asymptotic normality, and a known asymptotic sampling dispersion matrix. However, REML requires an assumption that the data are normally distributed. For many of the outcomes to be considered here (e.g. log wages, birth weight) this is not likely to be a major factor. For other outcomes such as education, the assumption of normality may be more suspect. An alternative approach utilized by Altonji and Dunn (1991), Solon et al (2000), Solon and Page (2003) and Bjorklund et al (2009) is to use method of moments. One downside is that judgment calls are required in how exactly to weight families of different sizes. For simplicity we only use REML in this analysis. Standard errors are calculated using the delta method.

In order to decompose the sibling correlation, I will add the additional regressors to the vector X in (1). Their inclusion will reduce the residual variation in the outcome variable and may alter the share of the residual variance due to the family and individual components. For example, the new family component (σ^{2*}_a) may be lower what was found without the inclusion of covariates (σ^2_a). One measure of the contribution of the added covariates is the change in the variance of the family component (σ^2_a - σ^{2*}_a). This provides an upper bound estimate of the *causal* effect because it includes all omitted factors that are also correlated with the included fixed effects. For example, the reduction in σ^2_a due to say the inclusion of parent's education would be comprised of both the direct effects of parent schooling as well as any omitted factors (e.g. inherited perseverance) that also contribute to the outcome variable and are correlated with years of schooling. The change in the variance of the family component divided by the overall variance of the permanent component tells us what fraction of the overall sibling correlation is due to the factor(s) in question. In this analysis I will mainly focus on the change in the sibling correlation when including the covariates.

It is also worth noting that while previous economic studies have taken advantage of repeated observations on individuals to remove the effects of v_{ijt} on the estimated correlation, in principle the same approach can be taken for *non-economic* outcomes. For example, it is well known that test scores can be a quite noisy measure of underlying knowledge or skill. Therefore, multiple observations on test scores can be used to produce error-corrected measures of the sibling correlation.

IV. Data

The analysis uses two distinct groups of cohorts in the PSID to cover sibling correlations over the lifecycle. I use the 1985 to 1997 birth cohorts that were included in the CDS for estimating sibling correlations in birth outcomes, childhood outcomes and adolescent outcomes. I use cohorts born between 1951 and 1968 for measuring the sibling correlations in adult outcomes. The CDS sample is restricted to include only those who were the son or daughter of the household head in 1997. Similarly the adult sample is restricted to those who were the son or daughter of the household head in 1968. The adult sample is further restricted to those who become household heads or wives of a household head. The adult sample is confined to only include those who were members of the nationally representative portion of the sample (SRC). Similar to most of the previous literature I do not restrict either sample to include only biological siblings since the question of interest concerns family background broadly defined. Both samples also include singletons who are useful for calculating the individual component.

The full sample size for the CDS cohorts meeting these restrictions is 3,246 individuals in 2177 families. The actual sample sizes, however, will vary depending on the outcome being measured. For some CDS outcomes for which I make use of repeated observations from the same individuals, the actual number of observations used for the estimation will be considerably higher. The full adult sample includes 3265 individuals from 1355 families.

For the CDS sample, I begin the analysis with a set of birth outcomes which includes: birth weight[7] (converted to grams), gestation length (measured in days), health at birth (3 point scale going from better than average to worse than average), poor health at birth (an indicator for worse than average health at birth) and NICU (an indicator if a newborn was admitted to a neonatal intensive care unit). The sample means and other summary statistics for all the samples are shown in Table 1.

I next use our CDS sample to measure childhood health. A useful broad measure of childhood health is health status rated on a 1 to 5 scale (Excellent, Very Good, Good, Fair, Poor) by the parent or care giver. We measure this at three different age ranges (0 to 5, 6 to 10, and 11 to 22). Other measures include whether the child was reported to have had any health limitation, learning disability, speech impairment, emotional problem, allergy, anemia, asthma, developmental delays, diabetes, hyperactivity, height and weight. Although the CDS asks questions concerning an even larger set of health outcomes, the incidence rates were far too low to be meaningful for many of these outcomes (e.g. autism). For some of the outcomes I restricted the age ranges so that the incidence rates are meaningful. We also include age dummies in all the analysis.

[7] Following some previous studies we do not include the birth weight of individuals who were born prematurely.

The next stage of analysis examines a range of childhood educational measures that reflect both resources and outcomes. These include whether the parent expects the child to go to college and an indicator for whether the child has fewer than 10 books. We then examine four specific tests scores from the Woodcock-Johnson tests of achievement: letter-word identification, passage comprehension, calculation and applied problems. We estimate sibling correlations at two age ranges, 6 to 10 and 11 to 15. In future drafts we will make use of the repeated observations and estimate the correlation over the full age range. All of the measures are first converted to z-scores by age and year of test. Finally we make use of the WISC Digit Span Test for short term memory.

We also utilize the Transition to Adulthood (TA) modules given to older cohorts of the CDS in 2005 and 2007 to assess sibling similarities in measures that are of particular relevance for adolescents. These include whether students finished high school (does not include GED) and whether they ever enrolled in a college. We limit those outcomes to those at least 18 years of age. We also examine high school grade point average (GPA) and scores on the SAT math and reading. The TA survey includes an extremely rich and detailed set of questions that assess adolescent attitudes and behaviors. I examine a few behaviors that seemed to be of particular relevance for assessing health or impulse control or have been widely used as an outcome. These include hours of sleep, the frequency of binge eating, the number of alcoholic drinks one drinks when drinking and the frequency of feelings of happiness. I also include some composite measures that summarize answers to questions concerning mental, emotional, social and psychological wellbeing. Finally, I include a set of measures dealing with other relevant health and addictive behaviors: whether an individual ever smoked cigarettes, used diet pills, used amphetamines or used marijuana.

The final set of outcomes utilizes siblings drawn from the original 1968 households. For economic outcomes I focus on annual earnings, wages, hours worked and family income. Future analysis will also examine wealth. Additionally I examine years of completed education, health status and a number of specific health outcomes including: ____________ .

V. Results

Birth Outcomes

Table 2 presents the results for birth outcomes. For birth weight we find a sizable sibling correlation of 0.53 that doesn't change much if we limit the sample to only boys or girls. The standard error is quite small at 0.02. This is similar to a sibling correlation of ___ estimated by _____ using birth records. Like () we also condition on full-term births. A large and growing literature has documented that birth weight is strongly associated with later life health and socioeconomic outcomes. Many researchers have cautioned, however, that care should be taken in how this association is interpreted (e.g. Gluckman and Hanson, 2005). Birth weight is a very rough proxy for fetal health and could reflect a potentially wide variety of underlying mechanisms only some of which may be amenable to policy. Nonetheless, the results here suggest that a reasonably large fraction of health inequality between families is present at the beginning of life.

We find a somewhat smaller correlation in the length of gestation of about 0.37 with some strongly suggestive evidence of a difference by gender. The correlation in gestation length among sisters (0.42) is a bit higher than that for brothers (0.34). Taking the results on birth weight and gestation in combination, suggests that an important source of the variation among families in birth weight is due to intrauterine growth retardation (IUGR). In other words, that birth weight is lower for a given gestation length.

Interestingly we find a reasonably large correlation of about 0.3 among siblings in the parent or caregiver report of relative health at birth on a 1 to 3 scale. On the one hand for such a blunt measure this is quite high, on the other hand this may reflect a systematic bias in reporting by the parent or caregiver.

We find a lower correlation of 0.12 in the incidence of poor health at birth and 0.18 in the probability of being admitted to a NICU. However both of these outcomes are statistically significant.

The columns on the right of Table 2 recalculate the overall sibling correlation (pooling both sexes) but add an additional covariate measuring family background. For example, the sibling correlation in birth weight is reduced to 0.50 from 0.53 when we include the race of the child. This suggests that the sibling correlation is reduced by about 6 percent when I account for race. A much smaller reduction is found when I control for a five year average of family income measured in the years preceding and including birth. Similarly including the years of completed schooling of the household head and wife (if present) does little to affect the estimate. In addition, I also calculate the average health status of the household head and wife over a five years span and find that this also has little effect. Finally, the last column of the table includes all 4 of the sets of covariates simultaneously. This specification also does little to explain birth weight beyond what can be explained by race alone.

Looking across the other outcomes, the results of the decomposition are broadly similar. However, for gestation, the inclusion of family background factors has literally no impact whatsoever. This suggests that the observed family background characteristics may in fact, play a greater role in explaining differences in IUGR than in explaining birth weight, which I will explore in future work. For health at birth, poor health at birth and NICU, the observed family background characteristics only explain between 3 and 4 percent of the sibling correlation

Childhood Health Outcomes

In Table 3 we show the results for a set of health outcomes measured during childhood. We start by estimating sibling correlations in the general health status of children. As discussed earlier this is reported by their parents or caregivers on a 1 to 5 scale. We find that correlations vary a bit depending upon the age at which it is measured ranging from around 0.35 to 0.45. By gender, the estimates are a bit noisy suggesting some caution in interpreting the estimates. With that caveat in mind there is suggestive evidence that the sibling correlation may rise with age for boys but not girls. The correlation in having a health limitation is relatively low at 0.16 but this may reflect the low incidence and relative bluntness of this outcome. The overall sibling correlation in having a learning disability is about 0.3 but the point estimate for boys is quite high at 0.50 compared to girls, 0.17. This is somewhat consistent with the pattern for development delays shown later in the table. Among the other specific health outcomes shown, correlations range from a low of 0.14 for diabetes to 0.29 for allergies. One striking finding that deserves further examination is that the correlation in diabetes among sisters is extremely high at 0.58 with a standard error of just 0.03.

I also estimate the sibling correlation in two physical characteristics, height and weight. It should be noted that as with all the estimates, I control for a full set of age dummies. The overall correlation in height is 0.38 with a brother correlation of 0.49 and a sister correlation of 0.35. The brother correlation is virtually identical to the 0.492 correlation reported by Mazumder (2008) using a sample of adults in the NLSY. The sister correlation, however, is a bit lower than the 0.467 reported by Mazumder (2008). For weight, the estimated correlation is about 0.35 with a higher correlation among sisters, 0.38 than brothers 0.20. Mazumder (2008) reported correlations of 0.33 for brothers and 0.29 for sisters. In future work it may be useful to see how the sex specific correlations change over childhood.

The decomposition of the sibling correlation in general health outcomes during childhood by observable family characteristics appears to bear more fruit than the analogous exercise for birth outcomes. I find that nearly 36 percent of the sibling correlation in health status measured between the ages of 6 and 10 can be explained by the observed covariates. Both family income and parent health status appear to account for large portions of the sibling correlation. For health status measured both earlier and later life, however, these characteristics can account for about 15 percent of the sibling correlation. The finding that family income around the time child enters school may matter for health may offer some hope that for example, income support policy may play a role. For the more specific health outcomes shown in Table 3, however, it is far less clear that family income can play much of a role.

Childhood Educational Outcomes

Table 4 presents estimates for sibling correlations related to educational measures. The first measure is whether parents expect their child to attend college. As might be expected, the sibling correlation is quite large at about 0.75. Nevertheless, it suggests that some parents may distinguish between their children's likelihood of academic success. This suggests that this variable may be useful as an explanatory variable in models of actual academic achievement later in life which I will explore in subsequent drafts. Similarly, the correlation in having less than 10 books is perhaps unsurprisingly high at 0.61. This measure may also be potentially more informative as a covariate in later outcomes.

The remaining measures in Table 4 consist of the 4 Woodcock-Johnson tests and the WISC digit span test. The sibling correlations in these outcomes are measured using three different samples. The first two samples limit the age range to either 6 to 10, or 11 to 15. Due to the timing of the CDS interviews this approach will by construction contain only one observation per sibling. The third approach is to allow for repeated observations on siblings and to allow the σ^2_v term to remove any transitory variation, or noise, in the data under the assumption that the object of interest is the overall performance during all ages of childhood. This approach is analogous to trying to capture "permanent income".

For the letter word score I estimate the sibling correlation to be 0.41 using just 6 to 10 year olds. This climbs to 0.49 when I examine 11 to 15 year olds. Finally, when using repeated scores of siblings the estimated sibling correlation rises to 0.62. This affirms the notion that a large fraction of the overall variance in test scores is due to variation within individual students. Roughly similar patterns are found with the other Woodcock Johnson tests across the three specifications, though the estimates are generally smaller. Interestingly, in nearly every case the estimates of the correlation are larger for brothers than for sisters. Mazumder (2008) found a sibling correlation of 0.62 in military test scores using the NLSY but found nearly identical estimates by gender.

The decompositions for test scores suggest that family background and parental characteristics can explain somewhere between 20 and 40 percent of the sibling correlation. For many of the outcomes family income appears to be the predominant factor. With respect to parents' expectations of a child going to college and the number of books owned, the decompositions can explain 5 percent and 9 percent of the sibling correlation, respectively with family income being the key factor.

Young Adult Outcomes

In Table 5 I estimate sibling correlations in measures from the TA modules fielded in 2005 and 2007 to older cohorts. The sibling correlation in completing high school is estimated to be 0.36, with a much higher estimate among sisters (0.50) than brothers (0.23). I obtain a roughly similar estimate for the pooled sample for attending college (0.35). In contrast Mazumder (2008) produced estimates of the sibling correlation in years of completed schooling of 0.6 and Solon et al (2002) reported estimates in the 0.5 range.

For high school grade point average the sibling correlation is estimated to be 0.32 with the correlation estimated to be nearly three times higher for sisters (0.60) than for brothers (0.21). For a relatively small sample, SAT scores on math and reading were asked. These estimates are quite noisy, especially for math where the correlation is only 0.22. For reading, however, the estimates are fairly stable at just over 0.5.

The next set of estimates primarily concern behaviors and attitudes that have been the focus of increasing attention among economists who have begun to emphasize the importance of non-cognitive skills. The sibling correlation in hours of sleep is quite high at 0.6. There are mixed signals provided by two measures that might proxy for impulse control, the frequency of binge eating (0.28) and the typical number of alcoholic drinks consumed when drinking (0.70).

Two of the composite measures produced in the TA survey, mental health and social well being have quite high sibling correlations of 0.43 and 0.51, respectively. In contrast composite measures of emotional well being and psychological well being have lower correlations of around 0.2. These smaller estimates are consistent with sibling correlations of around 0.25 for self-esteem found by Mazumder (2008). Finally, I estimate sibling correlations in the use of three drugs –two of which are illegal-- diet pills, amphetamines and marijuana. The sibling correlations in all three measures are similar ranging from between 0.1 and 0.2. Interestingly the correlation in use among sisters is consistently higher and ranges from 0.2 to 0.3. Mazumder (2008) also found a higher correlation in illegal drug use among sisters than brothers.

On the whole, the decompositions with these outcomes do not appear to be all that revealing as the estimates vary considerably across outcomes. One exception is college attendance which appears to be strongly affected by both aprent education and family income.

Adult Outcomes

Table 6 estimates sibling correlations in adult outcomes using siblings drawn from the original 1968 families. For years of education we find results largely in line with previous results reported in Mazumder (2008) and Solon et al (2002) of sibling correlations in the 0.5 to 0.6 range. We also confirm Mazumder’s (2008) finding of a sibling correlation in log earnings and log wages of around 0.5. This adds further confirmation that the previous findings of sibling correlations in earnings and wages of around 0.4 using earlier PSID samples may have been somewhat too low due to the younger age of the samples. With respect to health outcomes such as self reported health and disability I find correlations that are centered between 0.2 and 0.3. In future drafts of this paper, I will add more detailed health outcomes (e.g. diabetes, heart disease). Further, I will decompose the sibling correlation using early life health characteristics (e.g. born low birth weight) and family background characteristics (e.g. parent income and parent education).

VI. Conclusion

This study provides a set of new descriptive facts concerning disparities in health and socioeconomic status that emerge over the life course due to differing family and community influences by using sibling correlations. I find that sibling correlations in a range of health and socioeconomic outcomes start quite high at birth and remain high over the life course. The sibling correlation in birth weight and adult wages are both estimated to be about 0.5. The sibling correlation in measures of cognitive skill during childhood are even higher. Decompositions using techniques developed in Mazumder (2008) offers a convenient and easily understood framework for analyzing the contributions to family inequality in both health and SES. While the evidence in this paper is primarily descriptive in nature it may shed light on some of the sources of disparities over the life course. Preliminary evidence using this decomposition approach suggests that there are some limited opportunities for policies to ameliorate the gradients in health and SES. For example, it is possible that improvements in family income when children are between the ages of 6 and 10 may reduce disparities in cognitive and non-cognitive skills.

VII. References

Altonji JG, Dunn TA (1991) Relationships among the Family Incomes and Labor Market Outcomes of Relatives. In: Ehrenberg RG. (ed) *Research in Labor Economics*, Volume 12, JAI Press, Greenwich, Conn. 269-310

Altonji JG, Dunn TA (2000) An Intergenerational Model of Wages, Hours and Earnings. *Journal of Human Resources* 35(2):221-257

Ashenfelter O, Zimmerman D (1997) Estimates of the Returns to Schooling from Sibling Data: Fathers, Sons and Brothers. *Review of Economics and Statistics* 79(1):1-9

Bjorklund A, Eriksson T, Jantti M, Raaum O, Osterbacka E (2002). Brother Correlations in Earnings in Denmark, Finland, Norway and Sweden Compared to the United States. *Journal of Population Economics* 15(4):757-772

Bjorklund, Anders and Markus Jantti and Matthew Lindquist (2009). "Family background and income during the rise of the welfare state: trends in brother correlations for Swedish men born 1932-1968." *Journal of Public Economics*, 93 (5-6), 671-680

Cunha, Flavio and James J. Heckman (2007). "The Technology of Skill Formation." *American Economic Review*, 97(2):31-47.

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

Currie, Janet (2009). "Healthy, Wealthy and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Development," *Journal of Economic Literature* 47(1): 87-122.

Grossman, Michael (2005). "Education and nonmarket outcomes," *National Bureau of Economic Review*, working paper, No. 11582, August

Johnson and Schoeni (2007). "Early-Life Origins of Adult Disease: The Significance of Poor Infant Health and Childhood Poverty." Manuscript. Goldman School of Public Policy, UC Berkeley.

Johnson, Rucker (2009). "Health Dynamics and the Evolution of Health Inequality Over the Life Course: The Importance of Neighborhood and Family Background." Manuscript. Goldman School of Public Policy, UC Berkeley.

Levine and Mazumder (2007). "The Growing Importance of Family: Evidence from Brother's Earnings." *Industrial Relations* Vol 46(1) p. 7-21

Lleras-Muney, A., 2005, "The Relationship between education and adult mortality in the United States," *Review of Economic Studies*, Vol. 71, No. 1, pp. 189-221

Mazumder (2008). "Sibling Similarities and Economic Inequality in the US" *Journal of Population Economics*, 21 p685-701.

Mazumder (2008b). "Does education improve health? A reexamination of the evidence from compulsory schooling laws." *Economic Perspectives*, Smith, James (2009). “Reconstructing Childhood Health Histories” *Demography* Vol.46 (2), May 2009, pp. 387-403.

Solon G, Corcoran M, Gordon R, Laren D (1991) A Longitudinal Analysis of Sibling Correlations in Economic Status. *Journal of Human Resources* 26(3):509-534

Solon G (1999) Intergenerational Mobility in the Labor Market. In: Ashenfalter OC, Card D (eds.) *Handbook of Labor Economics*, Volume 3A. Elsevier, Amsterdam, North Holland

Solon G, Page ME, Duncan GJ (2000) Correlations Between Neighboring Children in their Subsequent Educational Attainment. *Review of Economics and Statistics* 82(3):383-392

Table 1: Summary Statisics

	Mean	S.D.	N, Ind'l	N, Fam
Birth Outcomes				
Birth Weight	3304.7	638.4	2995	2046
Length of Gestation	282.9	13.3	3145	2125
Health at Birth	1.8	0.6	3223	2166
Poor Health at Birth	0.09	0.28	3223	2166
Admitted to NICU	0.12	0.33	3215	2164
Childhood Health Outcomes				
Health Status (0-5)	1.7	0.8	1037	
Health Status (6 - 10)	1.7	0.8	2105	
Health Status (11 -22)	1.7	0.8	4220	
Ever Had Health Limitation	0.12	0.32	3246	
Learning Disability (8 -14)	0.07	0.26	1449	
Speech Impairment	0.11	0.31	3246	
Emotional Problem	0.24	0.43	3246	
Allergy	0.29	0.45	3246	
Anemia	0.10	0.31	3246	
Asthma	0.19	0.40	3245	
Developmental Delays	0.09	0.28	3246	
Diabetes	0.01	0.08	3245	
Hyperactivity	0.10	0.30	3245	
Height (1997)	46.6	11.0	2979	
Weight (1997)	59.5	33.5	2722	
Childhood Educational Outcomes				
Expected to go to College	0.74	0.44	3179	
Own fewer than 10 Books	0.12	0.32	2717	
Letter Word Score (11-15)	3E-10	1.0	2367	
Passage Comprehension (11-15)	2E-10	1.0	2360	
Calculation (11-15)	1E-09	1.0	520	
Applied Problems (11-15)	7E-10	1.0	2364	
Digit Span Test (11-15)	-7E-10	1.0	2244	

Table 1: Summary Statisics

	Mean	S.D.	N, Ind'l	N, Fam
Adolescent Outcomes				
At least High School	0.83	0.38	1154	
Any College	0.69	0.46	1154	
High School GPA	0.80	0.13	854	
SAT Math	561.9	127.7	280	
SAT Reading	566.8	119.0	264	
Hours of Sleep	7.44	1.69	714	
Eating Binge Freq.	2.73	1.54	713	
Number of Drinks	1.77	1.76	716	
Happiness Freq.	4.98	0.95	716	
Mental Health	13.47	2.52	714	
Emotional Well Being	5.02	0.93	716	
Social Well Being	3.43	1.30	716	
Psych. Well Being	5.07	0.93	716	
Ever Smoke	0.39	0.49	1153	
Diet Pills	0.17	0.37	1155	
Amphetamines	0.13	0.34	1155	
Marijuana	0.47	0.50	1155	
Adult Outcomes				
Log Earnings				
Log Wages				
Hours Worked				
Log Family Income				
Education				
Health Status				
Disability				

Table 2: Sibling Correlation in Birth Outcomes

				All Sibs Controlling for ….				
Outcome	All Sibs	Brothers	Sisters	Race	Family Income	Parent Education	Parent Health	All Factors
Birth Weight	0.528	0.507	0.522	0.500	0.519	0.522	0.512	0.499
	(0.019)	(0.034)	(0.031)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
	2995	1529	1466	2995	2995	2995	2995	2995
Gestation	0.377	0.348	0.424	0.377	0.376	0.377	0.377	0.376
	(0.009)	(0.040)	(0.036)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
	3145	1601	1544	3145	3145	3145	3145	3145
Health at Birth	0.312	0.321	0.309	0.311	0.309	0.304	0.305	0.298
	(0.023)	(0.040)	(0.039)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
	3223	1644	1579	3223	3223	3223	3223	3223
Pr Health at Birth	0.121	0.184	0.081	0.121	0.120	0.118	0.121	0.117
	(0.027)	(0.046)	(0.044)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
	3223	1644	1579	3223	3223	3223	3223	3223
NICU	0.184	0.194	0.192	0.184	0.180	0.180	0.180	0.177
	(0.027)	(0.052)	(0.046)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
	3215	1640	1575	3215	3215	3215	3215	3215

Table 3: Sibling Correlation in Children's Health Outcomes

Outcome	All Sibs	Brothers	Sisters	All Sibs Controlling for …. Race	Family Income	Parent Education	Parent Health	All Factors
Health Status	0.438	0.352	0.534	0.418	0.397	0.408	0.380	0.372
(ages 0 to 5)	(0.047)	(0.106)	(0.078)	(0.049)	(0.052)	(0.052)	(0.051)	(0.052)
	1037	542	495	1037	1037	1037	1037	1037
Health Status	0.345	0.407	0.297	0.313	0.290	0.317	0.236	0.220
(ages 6 to 10)	(0.029)	(0.047)	(0.059)	(0.030)	(0.031)	(0.030)	(0.033)	(0.034)
	2105	1076	1029	2105	2105	2105	2105	2105
Health Status	0.368	0.428	0.371	0.360	0.335	0.350	0.319	0.308
(ages 11 to 22)	(0.016)	(0.023)	(0.025)	(0.016)	(0.017)	(0.016)	(0.017)	(0.017)
	4220	2144	2076	4220	4220	3223	3223	3223
Health Limitation	0.162	0.099	0.077	0.160	0.154	0.161	0.147	0.141
	(0.027)	(0.058)	(0.054)	(0.028)	(0.028)	(0.027)	(0.028)	(0.029)
	3246	1656	1590	3246	3246	3246	3246	3246
Learning Disab.	0.315	0.501	0.168	0.297	0.315	0.308	0.322	0.294
(ages 8 to 14)	(0.042)	(0.066)	(0.087)	(0.044)	(0.042)	(0.043)	(0.042)	(0.045)
	1449	728	721	1449	1449	1449	1449	1449
Speech Impair.	0.116	0.170	0.241	0.112	0.115	0.117	0.108	0.101
	(0.027)	(0.053)	NA	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
	3246	1656	1590	3246	3246	3246	3246	3246
Emotional Prob.	0.275	0.216	0.372	0.262	0.271	0.274	0.273	0.258
	(0.024)	(0.051)	(0.038)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
	3246	1656	1590	3246	3246	3246	3246	3246
Allergy	0.291	0.372	0.184	0.287	0.289	0.289	0.287	0.284
	(0.024)	(0.040)	(0.051)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
	3246	1656	1590	3246	3246	3246	3246	3246
Anemia	0.267	0.282	0.382	0.260	0.264	0.261	0.260	0.255
	(0.023)	(0.047)	(0.039)	(0.024)	(0.023)	(0.024)	(0.024)	(0.024)
	3246	1656	1590	3246	3246	3246	3246	3246
Asthma	0.190	0.283	0.128	0.188	0.185	0.187	0.187	0.184
	(0.026)	(0.042)	(0.047)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
	3245	1656	1589	3245	3245	3245	3245	3245

Table 3: Sibling Correlation in Children's Health Outcomes

				All Sibs Controlling for				
Outcome	All Sibs	Brothers	Sisters	Race	Family Income	Parent Education	Parent Health	All Factors
Dev. Delays	0.157	0.255	0.147	0.152	0.154	0.158	0.153	0.139
	(0.027)	(0.050)	(0.045)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
	3246	1656	1590	3246	3246	3246	3246	3246
Diabetes	0.136	0.200	0.584	0.136	0.136	0.135	0.136	0.138
	(0.027)	NA	(0.032)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
	3245	1655	1590	3246	3246	3245	3245	3245
Hyperactivity	0.172	0.201	0.092	0.167	0.166	0.172	0.161	0.155
	(0.026)	(0.048)	(0.042)	(0.027)	(0.027)	(0.026)	(0.027)	(0.027)
	3245	1656	1589	3245	3245	3245	3245	3245
Height	0.375	0.491	0.345	0.374	0.375	0.376	0.373	0.374
	(0.022)	(0.034)	(0.041)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
	2979	1514	1465	2979	2979	2979	2979	2979
Weight	0.346	0.204	0.379	0.339	0.345	0.346	0.345	0.341
	(0.027)	(0.075)	(0.053)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
	2722	1404	1318	2722	2722	2722	2722	2722

Table 4: Sibling Correlation in Children's Educational Outcomes

				All Sibs Controlling for ….				
Outcome	All Sibs	Brothers	Sisters	Race	Family Income	Parent Education	Parent Health	All Factors
Expected to	0.749	0.780	0.733	0.743	0.712	0.727	0.732	0.710
go to College	(0.011)	(0.018)	(0.021)	(0.011)	(0.013)	(0.012)	(0.012)	(0.013)
	3179	1623	1556	3179	3179	3179	3179	3179
Own fewer	0.608	0.591	0.653	0.563	0.554	0.577	0.568	0.550
than 10 Books	(0.017)	(0.031)	(0.028)	(0.019)	(0.019)	(0.018)	(0.019)	(0.019)
	2717	1396	1321	2717	2717	2717	2717	2717
Letter Word Score	0.411	0.453	0.325	0.380	0.326	0.357	0.366	0.317
(Age 6 -10)	(0.028)	(0.049)	(0.052)	(0.029)	(0.031)	(0.029)	(0.029)	(0.031)
	1888	962	926	1888	1888	1888	1888	1888
Letter Word Score	0.485	0.536	0.470	0.475	0.390	0.420	0.410	0.374
(Age 11 -15)	(0.023)	(0.039)	(0.041)	(0.023)	(0.026)	(0.025)	(0.026)	(0.027)
	2367	1200	1167	2367	2367	2367	2367	2367
Letter Word Score	0.618	0.667	0.557	0.614	0.510	0.557	0.553	0.501
(All ages, repeated)	(0.026)	(0.049)	(0.048)	(0.027)	(0.030)	(0.029)	(0.029)	(0.031)
	4255	2162	2093	4255	4255	4255	4255	4255
Passage	0.315	0.275	0.296	0.281	0.223	0.244	0.244	0.196
Comprehension	(0.035)	(0.074)	(0.067)	(0.037)	(0.039)	(0.038)	(0.039)	(0.040)
(Age 6 -10)	1731	882	849	1731	1731	1731	1731	1731
Passage	0.422	0.494	0.325	0.405	0.327	0.337	0.335	0.286
Comprehension	(0.024)	(0.040)	(0.047)	(0.025)	(0.027)	(0.027)	(0.027)	(0.029)
(Age 11 -15)	2360	1195	1165	2360	2360	2360	2360	2360
Calculation	0.351	0.320	0.188	0.338	0.303	0.274	0.304	0.266
(Age 6 -10)	(0.062)	(0.108)	(0.126)	(0.062)	(0.066)	(0.069)	(0.068)	(0.070)
	770	381	389	770	770	770	770	770
Calculation	0.153	0.679	0.090	0.110	0.168	0.102	0.068	0.077
(Age 11 -15)	(0.118)	(0.158)	(0.149)	(0.121)	(0.105)	(0.118)	(0.137)	(0.127)
	520	265	255	520	520	520	520	520
Applied Problems	0.368	0.465	0.315	0.309	0.266	0.298	0.304	0.235
(Age 6 -10)	(0.030)	(0.051)	(0.059)	(0.032)	(0.034)	(0.032)	(0.033)	(0.035)
	1883	960	923	1883	1883	1883	1883	1883

Applied Problems	0.490	0.566	0.403	0.462	0.366	0.391	0.397	0.330
(Age 11 -15)	(0.022)	(0.036)	(0.043)	(0.023)	(0.027)	(0.026)	(0.026)	(0.028)
	2364	1194	1165	2364	2364	2364	2364	2364
Digit Span Test	0.324	0.352	0.312	0.302	0.297	0.291	0.281	0.258
(Age 6 -10)	(0.031)	(0.056)	(0.057)	(0.032)	(0.032)	(0.033)	(0.033)	(0.035)
	1978	1003	975	1978	1978	1978	1978	1978
Digit Span Test	0.410	0.418	0.359	0.411	0.372	0.384	0.377	0.359
(Age 11 -15)	(0.026)	(0.043)	(0.049)	(0.026)	(0.027)	(0.027)	(0.027)	(0.028)
	2244	1116	1128	2244	2244	2244	2244	2244

Table 5: Sibling Correlation in Young Adult Outcomes

				All Sibs Controlling for ….			
Outcome	All Sibs	Brothers	Sisters	Family Income	Parent Education	Parent Health	All Factors
At least High School	0.359	0.234	0.499	0.297	0.317	0.327	0.292
(18 and older)	(0.048)	(0.143)	(0.065)	(0.051)	(0.051)	(0.050)	(0.051)
	1154	550	604	1154	1154	1154	1154
Any College	0.351	0.328	0.351	0.233	0.227	0.273	0.192
(18 and older)	(0.050)	(0.123)	(0.090)	(0.059)	(0.063)	(0.056)	(0.063)
	1154	550	604	1154	1154	1154	1154
High School GPA	0.323	0.212	0.596	0.299	0.298	0.355	0.325
	(0.064)	(0.161)	(0.079)	(0.066)	(0.064)	(0.060)	(0.063)
	854	392	428	854	854	854	3223
SAT Math	0.229	0.121	0.051	0.200	0.161	0.181	0.141
	(0.130)	(0.288)	(0.447)	(0.137)	(0.142)	(0.141)	(0.147)
	280	123	157	280	280	280	280
SAT Reading	0.538	0.504	0.525	0.532	0.529	0.587	0.555
	(0.092)	(0.169)	(0.182)	(0.094)	(0.091)	(0.087)	(0.092)
	264	115	149	264	264	264	264
Hours of Sleep	0.601	--	--	0.585	0.582	0.559	0.540
	(0.189)	--	--	(0.191)	(0.192)	(0.191)	(0.196)
	1782	--	--	1782	1782	1782	1782
Eating Binge Freq.	0.284	--	--	0.278	0.257	0.258	0.263
	(0.108)	--	--	(0.109)	(0.113)	(0.112)	(0.112)
	1784	--	--	1784	1784	1784	1784
Number of Drinks	0.700	--	--	0.634	0.665	0.635	0.647
	(0.074)	--	--	(0.084)	(0.081)	(0.081)	(0.086)
	1788	--	--	1788	1788	1788	3246
Happiness Freq.	0.123	--	--	0.107	0.049	0.017	0.011
	(0.133)	--	--	(0.138)	(0.150)	(0.154)	(0.158)
	1787	--	--	1787	1787	1787	1787
Mental Health	0.432	--	--	0.423	0.393	0.409	0.388
	(0.078)	--	--	(0.079)	(0.083)	(0.081)	(0.083)
	1775	--	--	1775	1775	1775	1775

Table 5: Sibling Correlation in Young Adult Outcomes

				All Sibs Controlling for ….			
Outcome	All Sibs	Brothers	Sisters	Family Income	Parent Education	Parent Health	All Factors
Emotional Well Being	0.192	--	--	0.186	0.145	0.138	0.128
	(0.109)	--	--	(0.111)	(0.118)	(0.119)	(0.122)
	1788	--	--	1788	1775	1775	1775
Social Well Being	0.505	--	--	0.494	0.448	0.479	0.438
	(0.086)	--	--	(0.089)	(0.093)	(0.090)	(0.095)
	1788	--	--	1788	1788	1788	1788
Psych. Well Being	0.240	--	--	0.235	0.228	0.219	0.224
	(0.113)	--	--	(0.114)	(0.115)	(0.117)	(0.116)
	1788	--	--	1788	1788	1788	1788
Ever Smoke	0.162	0.019	0.362	0.164	0.161	0.169	0.171
	(0.062)	(0.189)	(0.085)	(0.061)	(0.061)	(0.061)	(0.061)
	1153	551	602	1153	1153	1153	1153
Diet Pills	0.181	0.107	0.290	0.182	0.181	0.181	0.181
	(0.056)	(0.120)	(0.090)	(0.056)	(0.056)	(0.056)	(0.056)
	1155	551	604	1155	1155	1155	1155
Amphetamines	0.132	0.189	0.242	0.111	0.105	0.109	0.088
	(0.059)	(0.152)	(0.084)	(0.061)	(0.062)	(0.062)	(0.064)
	1155	551	604	1155	1155	1155	1155
Marijuana	0.197	0.102	0.227	0.185	0.193	0.185	0.198
	(0.060)	(0.182)	(0.108)	(0.062)	(0.061)	(0.062)	(0.061)
	1155	551	1590	1155	1155	1155	1155

Table 6: Sibling Correlation in Adult Outcomes

Outcome	Brothers	Sisters
Education	0.665	0.527
	(0.031)	(0.036)
	671	776
Log Earnings	0.505	--
	(0.038)	--
	10201	--
Log Wages	0.500	0.406
	(0.038)	(0.045)
	7700	7161
Log Annual Hours	0.260	--
	(0.073)	--
	10643	--
Log Family Income	0.461	0.452
	(0.039)	(0.035)
	11366	11886
Health Status	0.145	0.381
(age >=40)	(0.101)	(0.100)
	1036	1113
Disability	0.221	0.288
	(0.048)	(0.046)
	7077	7443
Disability	0.291	0.188
(age >=40)	(0.090)	(0.122)
	1036	1112