The Intergenerational Transmission of Health Status: Estimates and Mechanisms

Owen Thompson

This Draft: May 2013. First Draft: August 2012

Abstract

This paper studies the transmission of health across generations using US data on over 100,000 parent-child pairs. I first document that children with a parent who has a particular health problem are typically at least 100 percent times more likely to have the same health problem themselves, and that transmission strength varies by child age and by parent gender. To assess the role of genetic transmission mechanisms in generating these intergenerational associations, I compare the strength of transmission among adopted versus biological children, and find that genetic transmission accounts for approximately 20-30 percent of intergenerational associations in most health conditions. To assess the role of environmental transmission mechanisms, I utilize an extensive set of control variables, and find that while controlling for potential environmental mediators does not substantively reduce intergenerational health transmission estimates, transmission is stronger among low SES families. I conclude that intergenerational persistence in health is an important hindrance to overall socioeconomic mobility, but that interventions which target environmental conditions may be able to mitigate the persistence of health across generations and promote equality of opportunity.

Introduction

One of the most fundamental questions in the social sciences is the extent to which socioeconomic outcomes are transmitted across generations. Because children cannot control the characteristics or choices of their parents, a strong relationship between parent and child socioeconomic outcomes suggests a lack of equal opportunity that many consider problematic. Existing empirical research has focused primarily on measuring and explaining the degree of intergenerational mobility in income and educational attainment, but relatively little is known about the intergenerational transmission of health status. This is an important gap in the literature, both because health status is increasingly recognized as a critical socioeconomic outcome in its own right and because health is a likely proximate cause of more heavily studied outcomes like income and education.

A small but growing empirical literature addresses issues of health mobility. Within developed countries, the health measure that has been most widely studied across generations is birth weight. For example, Currie & Moretti (2007) use vital records from California to show that the children of low birth weight mothers are approximately 50% more likely to be low birth weight themselves, even when extensive controls and sister fixed-effects models are employed.[1] With respect to health outcomes other than birth weight, Coneus & Spiess (2012) use a sample of German children ages 0-3 and document intergenerational associations in anthropomorphic and self-rated health measures; Trannoy et al. (2010) find that parental longevity predicts self-rated health in the next generation in a French sample; and Classen (2010) uses US data from the National Longitudinal Survey of Youth to estimate intergenerational associations in BMI.[2]

Several recent studies have also estimated intergenerational health correlations in developing country contexts, and have typically focused on how various aspects of economic development impact health mobility. Bhalotra & Rawlings (2009) document a correlation between mother's relative height and infant survival in a large sample of microdata from 38 developing countries, and show that this relationship is weaker when mothers have higher incomes, education levels, or better access to public health infrastructure early in the child's life.[3] Kim et al. (2011) study the transmission of several health markers in Indonesia, and find that intergenerational correlations are strong overall, but weaker on the more developed islands of Bali and Java. Finally, Venkataramani (2011) analyzes data from Vietnam and finds strong intergenerational correlations in height.

The present study builds on this existing literature in two important ways. First, I estimate intergenerational health correlations using a more comprehensive set of data sources, health outcomes, and methodological approaches than have previous studies. While previous studies have used a single data source and at most two or three health measures, I use three different nationally representative US data sets and analyze the transmission of seven separate health outcomes which incorporate global

[1] Other important papers on the intergenerational persistence of birth weight include Emanuel et al. (1992), Conley & Bennet (2001), and Royer (2009). All of these studies find positive intergenerational correlations in birth weight, although estimated effect sizes vary considerably across studies.

[2] Additionally, an unpublished 2007 working paper by Akbulut & Kugler (2007) also uses data from the NLSY to estimate intergenerational associations in anthropomorphic measures, asthma and depression.

[3] In a related paper using the same data, Bhalotra & Rawlings (2011) present estimates of intergenerational health associations that use maternal BMI and anemia status in addition to height, and focus on the strength of linkages between mothers and daughters over time.

and anthropomorphic health measures as well as specific conditions. Additionally, I analyze how intergenerational associations vary by child age and by parent and child gender, issues which have not been thoroughly addressed in previous studies.

Second, the present study uses several novel methods to systematically investigate the mechanisms underlying intergenerational associations in health status. An understanding of these mechanisms is critical for designing policies that increase intergenerational mobility, both in terms of health status specifically and socioeconomic outcomes in general. To assess the importance of genetic transmission mechanisms, I estimate and compare health correlations between monozygotic twins, dizygotic twins, and full siblings, and also estimate the strength of health transmission for samples of adopted versus biological children. Both of these methods are well established in the broader intergenerational mobility literature, but have not to my knowledge been applied to the case of health transmission.[4]To assess the importance of environmental transmission mechanisms, I observe the extent to which intergenerational health correlations are reduced when controls for potential non-genetic mediators are added to my baseline specifications. I use this strategy to investigate the importance of socioeconomic status, health care access and utilization, prenatal and early childhood conditions, health behaviors, and cognitive test scores. I also estimate fixed-effects models within schools and small geographic areas to account for unobserved health determinants that are constant within schools or neighborhoods, such as exposure to pollutants or the local availability and prices of healthy foods.

The basic finding is that intergenerational associations in health are qualitatively strong and extraordinarily robust. Across all data sources, health measures and methodological approaches, my baseline estimates indicate that having a parent with a given health condition typically increases the chances that a child will suffer from the same health condition by at least 50%-100%. Additional results from samples of twins, siblings and adoptees suggest that only around 20-40% of these associations can be attributed to genetic mechanisms, and the estimation of models with extensive controls and fixed-effects models indicate that at most only 5-15% of intergenerational correlations can be explained by observed environmental factors or by unobserved factors that are constant within neighborhoods or schools.

The remainder of the paper proceeds in 5 sections. Section 1 describes the data sources and health measures, and presents descriptive statistics. Section 2 reports baseline intergenerational health transmission estimates and shows how these estimates vary by gender and age. Section 3 investigates the importance of genetic health transmission mechanisms, while Section 4 assesses the importance of environmental mechanisms. Section 5 discusses several aspects and implications of the main results and concludes.

1 Data and Health Measures

Data Sources

Data is drawn from the National Health Interview Survey, an annual nationally representative survey

[4]An partial exception is Sacerdote (2007) who uses a sample of adoptees primarily to study the intergenerational transmission of income and education but also presents results for height and BMI.

conducted by the Centers for Disease Control.[5] In its current design, the NHIS randomly selects one adult and one child from each participating household to complete a detailed health questionnaire.[6] The NHIS survey instrument includes questions on respondent's health status, specific health conditions, access and utilization of health services, health behaviors and a variety of other health related topics. The selected adult answered the survey questions directly, while information about the selected child was provided by a "responsible adult family member," usually a parent.[7]

I focus on cases in which the selected adult and child were a parent-child pair and pool the 1998-2011 waves of the survey, which produces a data set with health information on over 100,000 parent-child pairs. Working with such a large sample allows for very precise full-sample estimates, and also has enough power to investigate how health transmission varies by characteristics such as age, gender and socioeconomic background.

To test the replicability and robustness of my results, I also employ data from the National Longitudinal Survey of Youth (NLSY) and the National Longitudinal Survey of Adolescent Health (Add Health). These data sets and described in detail in the appendix.

Health Measures

Health is a complex and multidimensional outcome which is difficult to capture with any single measure. Given this, I utilize seven different health indicators as dependent variables: health related school or work limitations, obesity, asthma, severe allergies or hay fever, chronic headaches or migraines, diabetes, and an indicator for suffering from any of the four chronic conditions just listed (asthma, hay fever, diabetes and chronic headaches). These outcomes were selected primarily because they are all serious health issues that have large impacts on quality of life, and which are also likely to impact human capital accumulation in childhood and labor market success in adulthood. Additionally, they all have at least some prevalence in child and young-adult populations, so that their correlations across generations can be meaningfully studied with available data. Alternative outcomes like mortality, heart disease or cancer are clearly important health outcomes, but are insufficiently common in childhood to study here. Finally, many of these health measures are available in the alternative data sets noted above, allowing me to replicate my baseline NHIS results using independently drawn samples.

Specific definitions of each health measures are as follows. I define a health limitation as having missed 14 or more days of work (for parents) or school (for children) due to illness or injury in the past year. While this definition of a health limitation has the advantage of being relatively clearly defined and objective, but the disadvantage of only applying to parents in the labor force, which reduces the size of the working sample. Parental height and weight were collected in each year of the NHIS, and I use these to calculate BMI and define parental obesity using the standard cutoff of 30. For children, I

[5] To allow pooling of multiple annual NHIS samples, use the integrated data prepared by Minnesota Population Center (2012).

[6] This design was first implemented in 1998. Prior to 1998, only an adult was selected for the detailed survey, and a relatively small amount of information was collected about all other household residents, including children when present.

[7] For current purposes parent refers to biological parents. I drop step-parents and cases where other relatives (e.g. grandmothers) have legally adopted the child. In Section XXX below I also examine health transmission among children adopted by non-relatives, but the baseline estimates use only biological parent-child pairs.

define obesity using the age & gender specific BMI cutoffs provided by the Centers for Disease Control. Unfortunately, child height and weight data was only collected for children over 12 and in the 2008-2011 versions of the survey, resulting is substantially reduced sample sizes for my obesity models. The remaining health measures are all specific conditions: asthma, severe allergies or hay fever, frequent headaches or migraines, and diabetes, as well as an indicator for having one or more of these four chronic conditions. Respondents are asked whether a medical professional has ever told them they have asthma or diabetes, while frequent headaches and hay fever were self-diagnosed.

It is worth noting that with respect to asthma and diabetes the NHIS does not differentiate between childhood-onset and adult-onset cases in the parent's generation. Child and adult onset asthma are formally the same pathology, but often differ in important ways. For instance, adult asthma is more likely to be chronic, while childhood asthma tends to be sporadic and is more likely to be triggered by viral infections or exposure to allergens. Juvenile and adult onset diabetes are in most cases clinically distinct conditions, referred to as type I diabetes and type II diabetes, respectively.[8] Type I diabetes is more common among children and adolescents, while type II diabetes is more common among adults, although exceptions in either direction do occur. Ideally, I would be able to estimate the relationship between childhood-onset asthma and diabetes among parents and children, and the fact that the two cases are pooled together is a shortcoming that should be kept in mind when interpreting my results.

2 Baseline Results

2.1 Descriptive Statistics and Mean Differences

Table 1 presents the prevalence rates (means) for each of the health measures described above. The first two columns show prevalence rates for parents and children, respectively. Not surprisingly, parents are considerably more likely than children to report most of the health problems studied.[9] The third and fourth columns begin the analysis of intergenerational health transmission by reporting the prevalence rates of each health condition among children with parents who *do* report the same health condition and children with parents who *do not* report the same health condition, respectively. In general, prevalence rates are higher among children of parents who have the specified health condition than they are among children whose parents do not. The fifth column and sixth columns quantify these differences by reporting first the percentage point difference between these two groups of children and then the percent difference. The estimated effect sizes are usually qualitatively large and highly statistically significant. For example, children of parents who reported being asthmatic are themselves 16.6 percentage points (151%) more likely to have asthma in the NHIS, Similar results hold for the other conditions: having a parent with the specified health problem increases prevalence among children by 50% in nearly all cases, and often by 100% or more.

As simple as these mean comparisons are, they contain much of the same information as more complex methods. Specifically, note that the conditional means in columns three and four contain the exact same

[8] Type I diabetes results from an autoimmune attack on the pancreas that renders it incapable of producing insulin, whereas type II diabetes occurs when the pancreas can still produce insulin but does so inadequately for the bodies needs, and is closely associated with obesity.

[9] The exceptions is asthma, which is well known to have high prevalence among children.

information as transition probability matrices, which are commonly used to study intergenerational mobility at multiple points of the income or education distribution. Also note that the percentage point differences in prevalence rates given in column 5 are exactly equal to the coefficient on parental health when it is regressed onto child health with no additional controls.

Table A1 of the web appendix replicates the results in Table 1 using data from the NLSY and Add Health. While the results are not identical, which as discussed in the appendix is likely due to differences in health definitions and survey procedures, there are qualitatively large and statistically significant intergenerational health associations in all three data sets. The percent increase in the prevalence of a given health measure among the children of parents who have the same health condition rarely falls below 50 percent and is frequently 100 percent or more.

2.2 Multivariate Results

Table 2 builds on the simple mean comparisons of Table 1 by estimating the following baseline regression specification:

$health_1 = \alpha + \beta\, health_0 + X^{'}\gamma + \varepsilon$

where the 1 subscript refers to the child's generation, the 0 subscript refers to parent's generation, and X is a vector of demographic control variables that includes indicators for the age, race and gender of both parents and children, as well as the survey year to account for any secular health trends. This equation is estimated as a linear probability model using OLS, but the results are very similar if non-linear estimation methods like logit or probit regression are used instead. The coefficient of interest is β, which estimates the percentage point change in the likelihood of a child reporting each health problem which is associated with having a parent who reports the same health problem. Results from this equation are shown in Table 2. To help ease the interpretation of the coefficient's magnitudes and facilitate comparisons across measures and samples, Table 2 also converts the percentage point effect given by β into percentage terms using the prevalence rates from column 4 of Table 1 (β/ prevalence).

The regression based estimates of intergenerational health transmission are similar in magnitude to the conditional mean results from Table 1. In most cases, a given health problem in the parent's generation increases the probability of observing the same problem in the child's generation by 1-20 percentage points, which translates into percent increases of approximately 50-300%. For instance, children in the NHIS with clinically obese parents are 14 percentage points (157%) more likely to be obese themselves.

2.3 Life Cycle Biases and Correlation Coefficients

The risk of health problems are generally not static across the life cycle, but instead tend to increase with age. Since the data sources used here observe the health of the first generation in adulthood and the health of the second generation in childhood, prevalence rates are expected to be higher among parents than children (as was confirmed in Table 1). For binary variables like those studied here, these

higher prevalence rates imply higher variances for health outcomes in the parent's generation as well.[10] Differences in the variance of health outcomes could arise due to factors other than upward sloping age-incidence profiles as well, for example public health measures implemented between generations or secular trends in other health determinants.

These issues are in many ways analogous to the life cycle issues that are common in studies of intergenerational income elasticities, since both the mean and variance of earnings are unstable over the life cycle. The preferred method of life cycle adjustments is to observe parents and children at the same age or range of ages, but when this is not possible a common technique is to calculate the intergenerational correlation coefficient, which is simply $\beta \times \frac{\sigma_1}{\sigma_0}$, where$\sigma$ denotes the standard deviation of the outcome in the specified generation (see Black & Devereux 2010). This adjustment standardizes the regression coefficient, so that the resulting number can be interpreted as the *standard deviation* change in child health associated with a one *standard deviation* change in parental health. Intergenerational correlation coefficients are reported in Table 2, and are typically 20% to 60% larger than the regression coefficients, reflecting the higher variances of the health measures in the parental generation.

To the extent that the processes generating intergenerational associations when children are observed in childhood continue in a similar fashion into adulthood, the intergenerational correlation coefficient can be interpreted as an estimate of what the regression coefficient would be if the children in these samples were observed at the same point in the life cycle as their parents.[11] But such an interpretation should only be made cautiously, because the nature of intergenerational transmission may change as children age. For instance, parental health may be a more important determinant of health in the subsequent generation when children are still living at home, and less important as they transition to adulthood. Alternatively, health determination in childhood may have a larger genetic component that fades in adulthood, or vice-versa. The main point though is simply to recognize that intergenerational effects tend to be larger after factoring out differences in cross-sectional dispersion between the two generations, even if these differences cannot be given a well-defined interpretation without additional information, and the next section explicitly analyzes how transmission strength changes as children age.

Another method of gaining insight into how intergenerational health transmission estimates are impacted by life cycle issues is to examine how these estimates change as children age, and in the next subsection I use the the very large sample sizes of the NHIS to do this with considerable precision.

2.4 Child Ages at the time of Observation

While none of the data sets used in this study contain health measures for parents and children observed at the same point in the life cycle, the NHIS does contain health measures for children of varying ages, and an important question is whether the strength of intergenerational associations change as children

[10] More precisely, the fact that prevalence rates in the parent's generation are closer to 1/2 implies higher variance. This is because the variance of a binary variable is given by p(1-p), where p is the proportion of the population with the attribute, and variance is therefore maximized at p=1/2.

[11] The only existing paper that compares parents and children at the same point in the life cycle is Classen's (2010) study of BMI transmission. His estimates are somewhat stronger than those presented here, although the two sets of results are not directly comparable because Classen primarily analyzes continuous BMI as opposed obesity indicators.

age. The regression results in Table 2 contain controls for child (and parent) ages, but the reported β effectively averages the strength of intergenerational transmission for each health outcome across all child age groups, and is uninformative with respect to potential heterogeneity by child age.

Figure 1 addresses this issue by presenting age-prevalence profiles for each health problem separately for children with a parent reporting the same health problem and for children without a parent reporting the same health problem. The profiles are estimated using partially-linear regressions with a bandwidth of 2 years, and parametrically control for race, gender and survey year but do not make any functional form assumptions with respect to child age. [12] The strength of intergenerational transmission at each age is estimated as the vertical distance between the two profiles.

All of profiles in Figure 1 are upward sloping, which confirms that the health problems studied here do in fact become more common as children grow older. Also, for all the health problems except headaches and diabetes, there are substantial gaps between children with and without a parent reporting the same health problem even at the earliest observed age, which in most cases is infancy. Most importantly, the figure shows that with the exception of health limitations, the age-health profile of children with a parent who reports a given health problem is steeper than the profile for children whose parent does not, indicating that (in level terms) the strength of intergenerational health transmission grows as children age.[13] For example, children with a parent who has had asthma are about 6 percentage points more likely to have asthma themselves in infancy (10% prevalence versus 3%), but by age 17 the difference grows to approximately 21 percentage points (34% versus 13%). Qualitatively similar results are present for self-rated health, hay fever, headaches and diabetes.

2.5 Transmission Strength by Parent and Child Gender

In addition to age, we might expect the strength of intergenerational health transmission to vary by parent or child gender. One consideration is that mothers often engage in more child rearing activities than fathers, and certainly have more control over the prenatal conditions experienced by their children. Since parenting behaviors are potentially effected by parental health and may also impact child health, gender asymmetries in parenting responsibilities could lead to differing strengths of health transmission by parental gender. The strength of health transmission could also vary by child gender if, for instance, children of one gender are less vulnerable to a given health problem or are treated differently by their parents in ways which impact health. Finally, we might expect the strength of health transmission to be different for parent-child pairs of the same gender, such as fathers and sons, (concordant pairs) than those of differing genders, such as mothers and sons (discordant pairs). This could arise either because the nature of the relationships between parent child pairs of the same gender are different, or because the overall prevalence rates of certain conditions vary by gender.[14]

[12] Obesity is excluded because child BMI data was only collected for children over 12 and in the 2008-2011 versions of the NHIS, so that both the available ages and the within-age sample sizes are very restricted.

[13] Coneus & Spiess (2011) also find increasing intergenerational transmission strength by age for some of their health outcomes, but their child sample is restricted to children ages 0-3.

[14] For example consider the transmission of urinary tract infections (UTIs). Since this condition is far more common among females than males, any intergenerational transmission will occur primarily in concordant mother-daughter pairs, and be close to zero for other parent-child gender permutations. While all of the conditions studied here are reasonable common in both genders, conceptually similar processes could still be in operation.

Since the NHIS chose an adult and a child respondent randomly from each household, it includes large numbers of each possible parent-child gender combination, allowing for the investigation of how transmission differs by gender.[15] The first two rows of Table 3 report results from regression specifications identical to the baseline models from Table 2, but estimated separately for mothers and fathers, respectively. The results indicate that maternal health has a substantially larger effect on child health than paternal health does, as for all seven health measures the coefficient on parental health is larger for the mother-child sample than the father-child sample. In most cases the differences are qualitatively large. For instance having a mother with asthma increases the probability of a child having asthma by .191, but the analogous effect from having a father with asthma is only .120, or 35% weaker.

The second two rows of Table 3 shows results from a similar exercise splitting the sample by child gender. In contrast to the results for parental gender, the strength of intergenerational health transmission is broadly similar for daughters and sons, and no clear pattern of gender differences are seen. The final two rows of Table 3 show results where the sample is split into concordant parent-child pairs and discordant parent-child pairs. Similarly to child gender, the strength of intergenerational health transmission does not differ substantially between parent-child pairs of the same gender and parent-child pairs of opposite genders, as the coefficients in the final two rows of Table 3 are all approximately equal.[16]

While documenting the existence of intergenerational health associations and how they vary by basic demographic characteristics is an important initial exercise, the ability to design policies which promote equal opportunity in health requires a sound understanding of the mechanisms underlying health transmission. Intergenerational associations could be the result of genetic traits being passed from parents to children, the result of environmental factors that effect both parent and child health, or of some (possibly interactive) combination of the two. The remainder of the paper attempts to assess the relative importance of genetic and environmental transmission mechanisms.

3 Genetic Mechanisms

Comparing Biological and Adopted Children

Health conditions manifestly have a substantial genetic component, and if the intergenerational associations documented above are due primarily to genetic traits shared by parents and children it may be more difficult to design policies that reduce health persistence. This is not to say that effective policy interventions would precluded by strong genetic transmission, because the ultimate consequences

[15] Unfortunately, the other two surveys used here contain little or no reliable health information for fathers, and so are not suitable for analyzing gender based heterogeneity.

[16] Results from models estimated on sub-samples composed of each possible parent and child gender combination (mother-daughter, mother-son, father-daughter, father-son) are quite similar to those shown here, and are omitted for brevity. Also, to test whether the presented results were sensitive to family structure, I re-estimated all models using the the following subsamples: 1) families with both biological parents present 2) families with two parental figures present, even if one of the parents was a step-parent or adoptive parent and 3) only families with a single, biological parent. In all cases, the results were similar to those in Table 3, where maternal health is consistently a stronger predictor of child health than paternal health, but no clear patterns emerge with respect to child gender or parent-child gender concordance.

of genetic predispositions can often be modified by environmental factors. For instance, Goldberger (1979) famously points out that geneticly driven inequality in eyesight can be greatly mitigated by providing access to eyeglasses. Still, it would be useful to quantify the approximate extent to which intergenerational health transmission operates via genetic channels.

To do so, I compare the strength of health transmission among biological versus adopted children, a technique which has been widely used in the broader literature on intergenerational mobility, but has not to my knowledge been applied to the case of health persistence. While the NHIS did not intentionally seek out adopted children for inclusion in the survey, the very large size of the combined annual samples led to the incidental inclusion of over 2,000 children who had been adopted by non-relatives, a sample size which is fairly large relative to those used in much of the existing literature[17] The appeal of estimating intergenerational models using this sample is that because adoptive parents are not biologically related to their adopted children, it severs the most direct link between parental and child genes. Under a set of fairly strong but not unreasonable assumptions, which I discuss in detail below, the reduction in intergenerational health associations among adopted children relative to biological children can be interpreted as the proportion of intergenerational health transmission that is due to genetic factors.

To quantify the differences in health transmission strength among adopted versus biological children, I estimate models of the following form and report the results in Table 4:

$$health_1 = \alpha + \beta_1 health_0 + \beta_2 adopt + \beta_3 (adopt \times health_0) + X^{'}\gamma + \varepsilon$$

The term of interest in the interaction, β_3, which estimates the percentage point difference in transmission strength between adopted and biological children. Table 4 also reports the percent decline in health transmission among adoptees which, as noted, can under certain assumptions be interpreted as the proportion of intergenerational health transmission that is due to genetic factors As expected, all of the interaction terms are negative, indicating that intergenerational correlations are weaker among adoptees. However, the magnitude of these reductions are modest, typically falling on the 20-30 percent range, and in four of the seven cases are statistically insignificant.[18] The largest reduction is for obesity, at 60 percent, but reductions for other health conditions do not exceed 40 percent.

Although the results in Table 4 suggest that genetics play an important role in the transmission of health across generations, the relatively small magnitudes of the reductions may be surprising. As an external check of whether these estimates are reasonable, I estimate correlations in the same health measures within a sample of monozygotic (MZ) twins available in the Add Health data (described in the appendix). Because MZ twins share 100 percent of their genes, the correlation in health outcomes within MZ twin pairs provides a credible upper bound on the proportions of variance in the outcomes that can be attributed to genetic factors.[19] Intuitively, if a trait were 100 percent genetically determined, then that trait would be present in either both members of a MZ twin pair or neither member,

[17] For example Das & Sjogren (2002) work with a sample of 126 adoptees; Plug (2004) uses information on 610 adoptees; and Sacerdote (2002) uses two samples of adoptees, one with a size of 300 and the other with a size of 128.

[18] The lack of statistical significance is partially a function of the relatively small number of adoptees in the sample, but is also due to the modest magnitudes of the estimated interaction terms.

[19] A vast interdisciplinary literature attempts to estimate the heritability of a wide range of traits by comparing the strength of correlations in monozygotic versus dizygotic twins, as well as other types of relatives. The validity of these heritability point estimates generated using these "quantitative genetic" models rests on a number of very strong assumptions, including additive separability of genetic and environmental influences, no associative mating, and equal

and the trait would therefore be perfectly correlated within MZ twin pairs. Observed correlations are upper bounds, however, because MZ twins typically share a great deal of environmental influences in addition to sharing all of their genes, and for this reason the within MZ twin correlations probably substantially exceed the true proportion of variance due to genes.

Correlations among MZ twins are reported for all available health measures in the final row of Table 4,[20] and are broadly in line with the results derived by comparing biological and adopted children. The correlation is greatest for obesity at approximately .6, while the other conditions have within twin correlations in the .2-.4 range. Because they are derived using different methodology, a separate, smaller sample, and different definitions of health problems, it is not appropriate to directly compare these upper bound estimates to their analogs from the NHIS sample of adoptees. However, the fact that correlations in these health outcomes are typically positive but well below .5 broadly corroborates the adoptee based results in suggesting that genes are an important contributor to the documented persistence of health across generations, but that their role is far from dominant or deterministic.

Assumptions Required for Valid Comparisons of Adopted and Biological Children

While studying the relative strength of health transmission among adoptees surely reduces the genetic linkages between parents and children, there are two central sets of that must hold for this approach to be fully valid (See Black & Devereux 2010 and Björklund, Lindahl & Plug 2004). First, it is necessary to assume that the pairing of adoptive parents and children is done randomly with respect to any genes that effect the traits under study, in this case health status. Second, we must assume that adopted and biological children are treated similarly by their respective parents, and that any unobserved parental or child characteristics that could impact child health are similar for adoptive and biological parent-child pairs.

These two sets of assumptions can be viewed as corresponding to the internal and external validity of comparisons between adopted and biological children, respectively. Violations of the first assumption reduce the internal validity of the results, since without this assumption the association between parent and child health among adoptees cannot be seen as free of genetic influences. Violations of the second set of assumptions reduce the external validity of the results, because the nature of the parent-child relationship among adoptees or the characteristics of the adoptive parents and children will not be a good representation of the general population.

While directly testing the validity of these assumptions is not possible in my data, the results of previous studies conducted with adoption data sets more well suited this purpose may provide some insight. Regarding the assumption that the assignment of adoptive children to families is as good as random, most existing adoption studies of intergenerational mobility have no information on the process by which adoptees are assigned to families, and therefore leave the issue unaddressed or interpret their estimates of environmental contributions as upper bounds (Das & Sjogren 2002; Plug & Vijverberg 2003; Plug 2004). Two exceptions are Björklund, Lindahl & Plug (2004), who use extraordinarily rich administrative data from Sweden which contains information on the characteristics of both biological

degrees of environmental similarity among identical and fraternal twins (see the thorough discussion of these issues by Sacerdote forthcoming and Black & Devereux 2010).

[20] Diabetes is excluded because it is relatively rare in younger populations, and as a result there was no within-twin diabetes variation in the relatively small Add Health MZ twin sample.

and adoptive parents, as well as the siblings of adopted children, and Sacerdote (2007), who uses a sample of Korean American adoptees for whom the procedural and institutional details of the placement process are well documented and known to be quasi-random.

These studies provide mixed evidence on the validity of the random assignment assumption. To test whether intergenerational correlations in income among adoptees are an artifact of the non-random placement of adopted children, Björklund, Lindahl & Plug (2004) regress child income onto the income of their adoptive parents, and then add direct controls for the income of the biological parents to the model and observe the extent to which the coefficient on adoptive parent's income declines.[21] The authors point out that if genetic resemblances between adopted and biological parents are inflating the estimated transmission coefficients for adoptive parent-child pairs, then in their specification these biases will be absorbed by the controls for the biological parent's characteristics. The authors find no evidence of such bias, and conclude that "the impact of adoptive parents' characteristics on those of their children are quite insensitive to the inclusion of the birth parent's characteristics." These results suggest that while important in principle, in practice nonrandom placement of adoptees may not be especially problematic. On the other hand, Sacerdote (2007) compares the relative strength of intergenerational correlations in income and education between adopted and biological children using a sample of adoptees who are known to be randomly assigned to adoptive parents, the intergenerational correlations among adoptive children are smaller than those found in some previous studies that did not use randomly assigned adoptees. He concludes that "the most natural explanation for this difference is that there is a strong positive selection of adoptees in families." This suggests that non-random assignment of adoptees may be problematic in the NHIS data used for this study.

Turning to the assumptions that parents treat biological and adoptive children similarly and that the unobserved characteristics of adoptive parents and children are similar to those of their biological counterparts, Björklund, Lindahl & Plug (2004) is the only existing study which attempts to explicitly test these assumptions. Their study exploits the richness of the Swedish administrative they are using data to compare the strength of intergenerational correlations across the following three types of children:

1) Adopted children with siblings who are the biological children of their adoptive parents and adopted children with siblings who are the adoptive children of their adoptive parents

2) Biological children from families with an adopted child and biological children from families without and adopted child

3) Biological children from families where a child was placed *into* adoption ("adopted out") and biological children from families where a child was not placed into adoption

Björklund, Lindahl & Plug (2004) note that equal transmission strength in the first comparison would indicate that parents treat adoptive and biological children comparably; that equal transmission strength in the second comparison would indicate that the relevant unobserved characteristics of adoptive parents are similar to those of parents who do not adopt; and that equal transmission strength in the third comparison would indicate that the unobserved characteristics of adoptive children are

[21] The authors conduct a similar analysis with education as well, and the results are comparable to their income estimates.

similar to those of biological children. In all three comparisons, the authors are unable to reject a hypothesis of equal transmission strength, leading them to conclude that "there is no evidence of differential treatment effects among adoptees and own-birth children" and that "adoptees and adoptive parents are [not] different from other children and their parents." Of course, this study uses Swedish data and examines the intergenerational transmission of income and education, not health, so it is not assured that their findings extend to the current study. Still, their results are suggestive that the adoptee results presented in Table 5 are a useful representation of the broader population.

One strategy for mitigating remaining concerns about both sets of assumption's validity is to add controls for the income and education levels of parents to the simple adoptee specification. Recall that for the first assumption to hold, the matching of adoptive parents and children need not be strictly random, but rather must be random with respect to genes that effect health status. While it would be impractical for adoption agencies or governments to take the genetic-based health of prospective adoptive parents directly into account when placing adoptive children, it is plausible that parents with favorable genetic health profiles may have higher education levels or more financial resources to help navigate the often complex and expensive adoption process, leading to the adoption of healthier children. If potential genetic matching does indeed occur through these channels, controlling for parental socioeconomic background should reduce the resulting bias. That is, even if adoptees are not in general randomly assigned to families with respect to genetically driven health status, assignment may be as good as random conditional on the income and education of the adoptive parents. Adding these controls may also help to make the second set of assumptions described above more plausible. Adoptive parents in my sample do tend to come from higher socioeconomic backgrounds than biological parents (for instance only 7.2% of adoptees in the NHIS came from families with income below the poverty line, compared to 14.7% of non-adoptees), and the addition of these controls will help to account for these general differences in observed SES.

Table A4 of the appendix shows results from models that include controls for the household income to needs ratio and the highest grade completed by the responding parent. The reduction in transmission strength among adoptees relative to biological children are modestly larger in these models than in the baseline models without SES controls, but the overall reductions are still relatively modest and again do not exceed 50% except in the case of obesity.

Despite the outlined efforts to increase the plausibility of the identifying assumptions, it remains possible that the assignment of adopted children to parents is non-random in germane ways or that adoptive and biological parents and children differ along unobserved dimensions that effect health transmission. If this is the case, then the estimated intergenerational health associations within the adoptive sample should be viewed as upper bound estimates of non-genetic influences, and extrapolation to non-adoptive populations should only be done tentatively. Still, it should be noted that the intergenerational correlations among adoptees remain practically large in most cases, and as such it seems reasonable to conclude that these associations are not exclusively due to bias caused by violations of the identifying assumptions, and that non-genetic factors are an important source of intergenerational health transmission. Given this, the next section investigates a number of possible non-genetic transmission mechanisms.

4 Environmental Mechanisms

The results from Table 4 indicate that non-genetic mechanisms likely play an important role in health transmission. Previous research has incorporated environmental factors into the analysis of health determination and persistence using two distinct ways. First is to test whether the strength of intergenerational health transmission varies depending on economic circumstances, in particular SES. Second is to observe the extent to which health transmission is reduced when controls for potential mediating factors are added to the baseline specification. In this section I implement these two approaches in turn.

Transmission Strength by SES

It is reasonable to expect that health transmission may vary by SES. While having a parent with a given health condition may predispose a child towards developing similar health problems, these predispositions are perhaps less likely to be realized among children who experience favorable environmental conditions associated with high SES. Several recent studies have investigated this issue in a developing county context (Bhalotra & Rawlings 2009, 2011; KIim et al. 2011; Venkataramani 2011), and have generally found that health transmission weakens when general economic development increases, but it remains unknown whether similar heterogeneity exists in developed countries, where basic public health infrastructures are typically well established and abject poverty is far less prevalent.

To test for heterogeneity in transmission strength by SES, I estimate models of the following form:

$$health_1 = \alpha + \beta_1 health_0 + \beta_2 SES + \beta_3 (SES \times health_0) + X^{'}\gamma + \varepsilon.$$

The sign and magnitude of the interaction term β_3 indicates the manner in which transmission of the specified health condition varies by SES, if at all. I measure SES using binary variables that indicate if the child is from a household with an income to needs ratio above the federal poverty level and whether the child's responding parent completed college. The use of binary SES measures is simply to aid in the straightforward interpretation of the magnitudes of the interaction terms, and analogous results using continuous SES measures are comparable to those shown below and are available upon request.

Results from estimating these specifications are presented in Table 5, and indicate that health transmission is indeed weaker among high SES families. For both SES measures, the interaction terms are negative for 6 of the 7 health conditions, and in most cases the implied reductions are non trivial, ranging from approximately 10 to 35 percent.

Controlling for Potential Environmental Mediators

Another approach to studying the role environmental factors is to control for potential mediators and observe the extent to which intergenerational associations decline. The basic underlying idea in this case is that intergenerational health correlations may arise because many environmental factors which impact health are experienced mutually by parents and children, because poor health reduces parent's ability to engage in parenting behaviors that encourage child health, or because parents with unhealthy lifestyles transmit them to children via modeling or the creation of household rules and norms. To the extent that the baseline intergenerational associations documented in Section

2 are due to observable environmental mediators, controlling for those mediators will reduce those intergenerational associations. [22]

Perhaps the most basic environmental factor that could help account for correlations in health across generations is simply socioeconomic status. A strong relationship between health and SES is one of the most persistent and well documented relationships in the public health and health economics literatures (Meer, Miller & Rosen Smith 1999; Ettner 1996), and has been shown to emerge very early in life (Berg, Lindeboom & Portrati 2006; Case, Fertig & Paxson 2005; Case, Lubtosky & Paxson 2002;). Since parents and children live in the same household, both are subject to the same general socioeconomic conditions and any health effects that those conditions may have, making SES an obvious potential transmission mechanism. I measure SES using indicators for the highest grade completed by the responding parent and with the family income to needs ratio.

Another potentially important mediator is access and utilization of health care services. To varying degrees, all of the health outcomes studied here are preventable and treatable. Because parents and children are often covered by the same insurance plans and because parents often make decisions about how frequently to access health services for both themselves and thier children, this raises the possibility that the intergenerational correlations reflect patterns of health care access and utilization common to parents and children. To measure access to and utilization of health care, I use indicators of whether both the parent and child were uninsured at the time of the interview, as well as an indicator of whether parents reported that they delayed care for themselves or their child due to cost in the last year, and whether the child has had a routine checkup in the past year.

Health behaviors are a third potentially important set of non-genetic mediators. Behavioral choices related to diet, exercise, smoking, drinking and personal safety are important health determinants among both adults and children, and these behavioral choices have been shown to persist across generations (Loureiro, Sanz-De-Galdeano & Vuri 2010; Wickrama, Conger & Wallace et. al 1999), making them a plausible health transmission mechanism. Intergenerational links in health behaviors can be quite direct, for example when parents and children eat meals jointly or engage in physical activity together. Even in the absence of such directly shared experiences, parents with better health behaviors may transmit them to their children through modeling or by creating and enforcing rules and norms around healthy lifestyle choices. The NHIS only collects information on parent's health behaviors, and I utilize controls for parent's smoking status (current daily smoker, current occasional smoker, former smoker, and never been a smoker), and the number of days in the past month that they engaged in 10 or more minutes of moderate exercise and vigorous exercise.

Table 6 reports the coefficients on parental health in baseline models that control only for demographic characteristics, and then in models that include all of the controls described above. the final row of Table 5 also reports the percent decline in transmission strength that results from adding the controls. Surprisingly, these reductions are quite small, averaging around 5 percent and never exceeding 15 percent. One possible explanation for these modest reductions is that the controls available in the

[22] While I am not aware of any studies that systemically investigate the effect of adding controls to an intergenerational health specification, this approach has been used in the broader intergenerational mobility literature (for example Blanden, Gregg & Macmillan 2007), and to investigate mediators of other important health determinants (see, for example, Cutler and Lleras-Muney 2010).

NHIS are not sufficiently detailed to capture the relevant environmental conditions. In the appendix, I assess this possibility by using NLSY and Add Health data to control for different and arguably more comprehensive sets of environmental controls. For instance I am able to control for permanent income; an expanded set of health behavior measures that includes for instance diet and seat belt usage and that are measured for both generations; scores on cognitive achievement tests in both generations; prenatal and early infancy factors such as access to prenatal care, smoking and drinking behaviors during pregnancy, and breast feeding; and neighborhood fixed effects that account for unobservables that are constant at the neighborhood level, such as exposure to environmental toxins and quality of school lunches. Even when these more comprehensive sets of controls are employed, the associated reductions in intergenerational transmission rarely exceed 10 or 15 percent.

5 Conclusion

This study aimed to make two substantive contributions to the growing literature on intergenerational health transmission. The first contribution was to provide a more detailed and comprehensive set of baseline transmission estimates than previous studies. To do so, intergenerational models were estimated for seven separate health measures and transmission strength was shown to vary by child age and parental gender. These baseline results were replicated using two independently drawn samples.

The paper's second contribution was to investigate the mechanisms underlying these strong intergenerational health associations. To assess the role of genetic transmission, I compared the strength of health transmission among biological versus adopted children. This exercise indicated that shared genes are an important but relatively modest mechanisms of health transmission, accounting for approximately 20-30 percent of the baseline correlations. The plausibility of these relatively small estimates was confirmed by examining the strength of correlations among identical twins, which provides a reliable upper bound on the proportion of variance in a given trait that is due to genetic factors. With respect to environmental mechanisms, I showed that transmission strength is lower among high SES families, and that adding large sets of controls to my baseline specifications did not meaningfully reduce intergenerational transmission.

Appendix

Supplementary Data Sources

The National Longitudinal Survey of Adolescent Health (Add Health)

Add Health is a longitudinal study that follows an initial sample of 20,745 adolescents who were in grades 7-12 during the 1994-1995 school year. Respondents have been interviewed in 4 waves, with the most recent wave occurring in 2008. Most of the survey's attention is focused on the adolescents themselves, but the first wave of the study included an extensive questionnaire filled out by a resident parental figure of the adolescent. This parental survey included a section where several health problems were listed, and the responding parental figure was asked whether the adolescent's biological mother and biological father had each condition. In over 85% of cases, the responding parental figure was the adolescent's biological mother, and I restrict my analysis to these cases so that parental health is in effect measured using self-reports.[23] In the third wave of the survey, the adolescents were asked whether they suffered from a similar list of health problems. I use these two sets of questions to construct comparable health measures for mothers and children. With these restrictions, my working Add Health sample contains around 10,000 parent-child pairs.

Another unique feature of the Add Health data, which was utilized in Section 3 above, is that within the main survey frame Add Health collected data on an overlapping "genetic subsample" specifically designed to study genetic influences using behavioral genetic methods. To construct this subsample, students who identified themselves as twins in a preliminary questionnaire entered the survey with certainty, resulting in the inclusion of over 150 monozygotic twin pairs.

The National Longitudinal Survey of Youth (NLSY), Mother-Child Files

The NLSY began in 1979 with a nationally representative sample of 12,686 individuals between the ages of 14 and 21. Participants were eligible to be interviewed annually until 1994 and biennially thereafter, with the most recent wave occurring in 2010. Some basic health information was collected in all survey waves, and as NLSY cohorts started turning 40 years old (beginning in the 1998 wave of the survey) they were asked to complete a "40 and older health module" that collected information on a large range of health measures. Starting in 1986, an additional biannual survey of all biological children of female NLSY respondents began, and collected a wide variety of health information about children, thus allowing for the construction of an unusually rich intergenerational health data set. Of the 6,283 original female NLSY respondents 4,929 gave birth to a total of 11,495 children who participated in the supplemental survey, and my working NLSY sample contains 6,000-8,000 children born to approximately 4,000 mothers.

The NLSY survey instrument is very detailed, containing, for example, information on maternal health behaviors during pregnancy and both maternal and child cognitive test scores. I also take advantage of the fact that the NLSY is a long-running longitudinal survey to calculate permanent income, to

[23] Biological fathers comprised about 4% of the responding parental figures, while step mothers, adoptive mothers, and grandmothers each comprised about 2%. But irrespective of who filled out the parental survey, only questions regarding the health issues of biological mothers and fathers were asked. This design implies that only data on biological mothers is reliably available for most of the sample, which is why I restrict the sample accordingly.

construct detailed insurance coverage histories for parents and children, and to observe child health conditions at multiple points in time.

Health Measures in Add Health and the NLSY

Only the NLSY contains information on health limitations in both generations. In each wave, NLSY adults were asked if they were (or would be) limited in the type or amount of work they could perform due to health problems, and children were asked if they had a health condition that limited school attendance. I used these questions to create a variable equal to one if the respondent reported such a limitation in any survey wave, and zero otherwise.[24]

Both Add Health and the NLSY contain a measure of obesity in both generations. The NLSY contains the variables needed to calculate body mass index (BMI) for both parents and children, and Add Health contains this information for children.[25] As in the NHIS data, I define obesity using the standard BMI cutoff of 30 for adults and the age & gender specific BMI cutoffs provided by the Centers for Disease Control for children. In the NLSY, I determine obesity at each survey wave, and classify a child as obese if they were obese at any point. In Add Health, BMI data is not available for parents, but mothers reported whether they considered themselves to be currently obese, and I use this information to measure maternal obesity.

With respect to the remaining health measures, which are all specific conditions, information is collected as follows: In Add Health, mothers self-report the current presence of asthma, severe allergies or hay fever, migraine headaches, and diabetes without reference to formal diagnoses. Children in Add Health were asked if they had ever been diagnosed with asthma or diabetes, and whether they had taken a prescription medication for headaches and for allergies or hay fever in the past year. In the NLSY, parents and children each self-report the presence of asthma, migraine headaches, and diabetes at each wave, without reference to formal diagnoses.

Baseline Models using Add Health and the NLSY

Tables A1 and A2 replicate the exercises from Tables 1 and 2 of the main body of the paper using Add Health and NLSY data. The health transmission estimates are broadly similar to those generated using the larger NHIS sample in that having a parent with a particular health problem greatly increases the probability that a child will develop the same health problem. The NHIS estimates are in some cases non-trivially larger that those from the other surveys. These differences are probably due to the fact that the NHIS is the only one of the three surveys where all child health information is reported by parents. Parents with a given health problem may be more cognizant of that problem, and therefore more likely to perceive its presence in their child or to have it diagnosed by a medical professional. While these differences are a reminder that the data collection procedures of specific surveys can have a non-trivial impact on the estimates and reinforce the importance of replication, the broad consistency of transmission estimates across independent samples is reassuring.

[24] The results are again not sensitive to alternative approaches of combining information from multiple NLSY waves, for example using the proportion of survey waves in which respondents reported health limitations.

[25] In the NHIS, child BMI data was only collected for children over 12 and in the 2008-2010 versions of the survey, leading to a substantially reduced sample sizes for the obesity models.

Adding Controls to the Baseline Add Health and NLSY Models

The set of available control variables in the NHIS is somewhat limited relative to the other two data sets, and below I describe a number of additional controls that are availbale in Add Helath and the NLSY and then perform an exercise similar to that in Table 6 of the main paper, in which the controls are added and the extent to which intergenerational associations are reduced is observed.

With respect to SES measures, Add Health contains information on the highest grade completed by the adolescent's mother and a poverty indicator equal to one if gross family household income was under the poverty line for the reporting year or if the family reported receiving AFDC payments. In the NLSY, I include a variable measuring the highest grade completed by the child's mother, and take advantage of the NLSY's longitudinal design to create a measure of permanent net household income, calculated as the average net household income reported by the mother between ages 22 and 38. The use of permanent income is a potentially important data improvement, as year-to-year income is known to often be highly transitory.

To measures health care access and utilization in the Add health data I use a dummy indicating whether the child was insured at the time that health outcome data was collected and an indicator of whether parents reported that it was "somewhat hard" or "very hard" for them to get medical care for their family. In the NLSY, I again make use of the longitudinal design and create variables measuring the percentage of survey waves in which the mother and the child were uninsured, as well as an indicator for respondents (both parents and children) who were covered in all survey waves. Finally, in both Add Health and the NLSY I include indicators of whether children were reported to have had a routine checkup in the past year.

For health behaviors controls, both Add Health and the NLSY have the advantage of containing information about both parents and children, and also offer a generally more comprehensive set of behavioral variables than the NHIS. For parents in the Add Health data, these include indicators of whether the parent currently smokes, drinks alcohol 3 or more days a week, had 5 or more alcoholic drinks at once in the past month, and whether they report always wearing a set belt. For children in Add Health, they include indicators of whether children have ever tried a cigarette, whether they have ever had an alcoholic drink, whether they have had 5 or more alcoholic drinks at once in the past year, the number of times they played sports or exercised in the past week, the number of hours they watch television in a typical week, whether they ate any vegetables yesterday, whether they ate any cookies, doughnuts or cake yesterday, and whether they report always wearing a seat belt. In the NLSY, these include indicators of whether mothers have ever been smokers, drank alcohol 15 or more days in the past month, drank 5 or more alcoholic drinks at once in the past month, and the proportion of survey waves in which they reported engaging in light exercise once a month and strenuous exercise once a month. Finally, for children in the NLSY I include indicators of whether each respondent has ever been a smoker, ever drank alcohol, and the frequency with which they reported exercising.

Recent research has shown that scores on cognitive achievement tests are a consistent predictor of health outcomes in both adults and children (Auld & Sidhu 2005; Gottfredson & Deary 2004), and cognitive ability is also strongly correlated across generations, so that some of the documented intergenerational correlations in health may reflect intergenerational transmission of cognitive skills. Both

Add Health and the NLSY contain cognitive test scores for children, and the NLSY contains cognitive test score results for parents as well. Specifically, Add Health reports each child's score on the Peabody Picture Vocabulary Test (PPVT), while the NLSY reports each child's score on the Peabody Individual Achievement Tests (PIAT) in mathematics, reading recognition and reading comprehension and each parent's score on the Armed Forces Qualification Test (AFQT). All scores were converted to percentiles using nationally representative norming samples of similarly aged individuals.

Non-genetic process which occur while children are still in-utero or very early in life may be important mechanisms for transmitting health across generations as well. The degree to which parents and children are subject to similar environmental influences is particularly strong during this period, as children in-utero are directly impacted by their mother's actions and conditions with respect to nutritional intake, prenatal care, exposure to toxins and other important health determinants. In early infancy, children who are breastfeeding continue to be directly effected by their mother's health decisions, and all infants have limited agency in determining the environments they are exposed to. Only the NLSY contains detailed information on prenatal and early infancy conditions. I make use of variables indicating whether the mother reported receiving professional prenatal care, smoking cigarettes and drinking alcohol during pregnancy, each child's gestational age at birth (in weeks), whether each child was breastfed, and whether each child received routine well-baby care during their first year of life, as well as each child's birth weight, which has been widely used as a general proxy for prenatal conditions (Almond & Currie 2010).

Some potentially germane transmission mechanisms are unobserved in my data sets but are common to neighborhoods. Examples of neighborhood factors may include local pollution levels, proximity to health care services, the availability and prices of healthy foods, and the accessibility of parks and other recreational spaces. One way that these unobserved factors can be accounted for is by estimating fixed-effects models within small geographic areas. The Add Health data is well suited for this purpose, as it contains pseudo geocodes which can be used to identify families living in the same census block group.[26] I use these codes to include a neighborhood fixed effect.

Tables A3 shows the baseline Add Health and NLSY regression specifications which include only demographic controls, then adds the controls described above and calculates the percent change in the estimated transmission strength. Adding all of the described covariates leads to at most modest reductions in the estimated strength of intergenerational health transmission, and in many cases transmission conditional on these extensive controls is nominally stronger than in the baseline case.[27]

[26] Geocode information is also available in the NLSY, but the smallest reported geographical areas are counties, which are typically much larger than block groups and of limited use for present purposes.

[27] The models in Table A3 perform casewise deletion of observations with missing values, and in the most saturated models this approach leads to nontrivial reductions in samples sizes. To ensure that any changes in transmission estimates (or lack thereof) were not due to non-random differences in the sample composition, I re-estimated the baseline models using only observations that did not contain missing values for the full set of covariates, and the results were very similar to those reported here.

References

Akbulut, M., & Kugler, A. (2007). Intergenerational transmission of health status in the US among natives and immigrants. Mimeo.

Almond, D., & Currie, J. M. (2010). Human capital development before age five. National Bureau of Economic Research Working Paper.

Auld, C., & Sidhu, N. (2005). Schooling, cognitive ability and health. Health Economics, 14.

Barker, D. (1997). Maternal nutrition, fetal nutrition and diseases later in life. Nutrition, 13(9).

Barnes KC. (1999). Gene-environment and gene-gene interaction studies in the molecular genetic analysis of asthma and atopy. Clinical and Experimental Allergy : Journal of the British Society for Allergy and Clinical Immunology, 29.

Berg, G. J. V. D., Lindeboom, M., & Portrati, F. (2006). Economic conditions early in life and individual mortality. American Economic Review, 96(1).

Bhalotra, S., Rawlings, S. B. (2011). Intergenerational persistence in health in developing countries: The penalty of gender inequality? Journal of Public Economics, 95(3-4).

Bhalotra, S., & Rawlings, S. (2012). Gradients of the intergenerational transmission of health in developing countries. Review of Economics & Statistics, 94(2).

Björklund, A., Lindahl, M., & Plug, E. (2004). Intergenerational effects in Sweden: What can we learn from adoption data? IZA Working Paper.

Björklund, A., Jäntti, M., & Solon, G. (2007). Nature and nurture in the intergenerational transmission of socioeconomic status: Evidence from Swedish children and their biological and rearing parents. B.E.Journal of Economic Analysis & Policy: Advances, 7(2).

Björklund, A., Lindahl, M., & Plug, E. (2006). The origins of intergenerational associations: Lessons from Swedish adoption data. Quarterly Journal of Economics, 121(3).

Black, S. E., Devereux, P. J. (2010). Recent developments in intergenerational mobility. National Bureau of Economic Research Working Paper.

Blanden, J., Gregg, P., & Macmillan, L. (2007). Accounting for intergenerational income persistence: Noncognitive skills, ability and education. The Economic Journal, 117(519).

Case, A., Fertig, A., & Paxson, C. (2005). The lasting impact of childhood health and circumstance. Journal of Health Economics, 24(2).

Case, A., & Lubtosky, D. :. P., C. (2002). Economic status and health in childhood: The origins of the gradient. American Economic Review, 92(5).

Christakis NA, & Fowler JH. (2007). The spread of obesity in a large social network over 32 years. The New England Journal of Medicine, 357(4).

Clark, A. E., & Lohéac, Y. (2007). "It wasn't me, it was them!" social influence in risky behavior by adolescents. Journal of Health Economics, 26(4).

Classen T.J. (2010). Measures of the intergenerational transmission of body mass index between mothers and their children in the United States, 1981-2004. Economics and Human Biology, 8(1).

Coneus, K., & Spiess, C. K. (2012). The intergenerational transmission of health in early childhood-evidence from the German socio-economic panel study. Economics and Human Biology, 10(1).

Conley, D., & Bennett, N. G. (2000). Is biology destiny? birth weight and life chances. American Sociological Review, 65(3).

Currie, J., & Moretti, E. (2007). Biology as destiny? short- and long-run determinants of intergenerational transmission of birth weight. Journal of Labor Economics, 25(2).

Cutler, D. M., & Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. Journal of Health Economics, 29(1).

Das, M., & Sjogren, T. (2002). The inter-generational link in income mobility: Evidence from adoptions. Economics Letters, 75(1).

Devlin B, Daniels M, & Roeder K. (1997). The heritability of IQ. Nature, 388(6641).

Dickens WT, & Flynn JR. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. Psychological Review, 108(2).

Duncan, G. J., Boisjoly, J., & Harris, K. M. (2001). Sibling, peer, neighbor, and schoolmate correlations as indicators of the importance of context for adolescent development. Demography, 38(3).

Emanuel I, Filakti H, Alberman E, & Evans SJ. (1992). Intergenerational studies of human birthweight from the 1958 birth cohort: Evidence for a multigenerational effect. British Journal of Obstetrics and Gynaecology, 99(1).

Ettner SL. (1996). New evidence on the relationship between income and health. Journal of Health Economics, 15(1).

Godfrey, K., & Barker, D. (2001). Fetal programing and adult health. Public Health and Nutrition, 4.

Goldberger, A. S. (1979). Heritability. Economica, 46(184).

Gottfredson, L., & Deary, I. (2004). Intelligence predicts health and longevity, but why? Current Directions in Psychological Science, 13.

Herrnstein, R. J., & Murray, C. (1994). The bell curve: Intelligence and class structure in American life. New York, New York: First Free Press.

Kim, Y. (2011). Intergenerational correlations of health among older adults empirical evidence from Indonesia. Rand Corporation Working Paper.

Kleeberger, S. R., & Peden, D. (2005). Gene-environment interactions in asthma and other respiratory diseases. Annual Review of Medicine, 56(1).

Loehlin, J. C. (2012). The differential heritability of personality item clusters. Behavior Genetics, 42(3).

Loureiro M.L., Sanz-De-Galdeano A., & Vuri D. (2010). Smoking habits: Like father, like son, like mother, like daughter? Oxford Bulletin of Economics and Statistics, 72(6).

Meer J, Miller DL, & Rosen HS. (2003). Exploring the health-wealth nexus. Journal of Health Economics, 22(5).

Minnesota Population Center (2012). Integrated health interview series: Version 5.0. Minneapolis: University of Minnesota.

Oreopoulos, P. (2003). The long-run consequences of living in a poor neighborhood. The Quarterly Journal of Economics, 118(4).

Plug, E. (2004). Estimating the effect of mother's schooling on children's schooling using a sample of adoptees. American Economic Review, 94(1).

Plug, E., & Vijverberg, W. (2003). Schooling, family background, and adoption: Is it nature or is it nurture? Journal of Political Economy, 111(3).

Qi, L., & Cho, Y. A. (2008). Gene-environment interaction and obesity. Nutrition Reviews, 66(12), 684-694.

Royer, H. (2009). Separated at girth: US twin estimates of the effects of birth weight. American Economic Journal: Applied Economics, 1(1).

Sacerdote, B. (Forthcoming). Nature and nurture effects on children's outcomes: What have we learned from studies of twins and adoptees? Handbook of Social Economics, North Holland, Amsterdam

Sacerdote, B. (2002). The nature and nurture of economic outcomes. American Economic Review, 92(2).

Sacerdote, B. (2007). How large are the effects from changes in family environment? A study of Korean American adoptees. Quarterly Journal of Economics, 122(1).

Smith, J. P. (1999). Healthy bodies and thick wallets: The dual relation between health and economic status. The Journal of Economic Perspectives, 13(2).

Trannoy A., Tubeuf S., Jusot F., & Devaux M. (2010). Inequality of opportunities in health in France: A first pass. Health Economics, 19(8).

Tsuang, M. (2000). Schizophrenia: Genes and environment. Biological Psychiatry, 47(3).

Turkheimer E, Haley A, Waldron M, D'Onofrio B, & Gottesman II. (2003). Socioeconomic status modifies heritability of IQ in young children. Psychological Science, 14(6).

Venkataramani AS. (2011). The intergenerational transmission of height: Evidence from rural Vietnam. Health Economics, 20(12).

Wickrama, K. A. S., Conger, R. D., Wallace, L. E., & Elder, G. H. (1999). The intergenerational transmission of health-risk behaviors: Adolescent lifestyles and gender moderating effects. Journal of Health and Social Behavior, 40(3).

Table 1: Prevalence of Health Outcomes Among Parents and Children

Condition	(1) Parents	(2) All Children	(3) Children of Parents With Condition	(4) Children of Parents Without Condition	(5) Percentage Point Difference	(6) Percent Difference
Health Limitation	0.051	0.044	0.097	0.036	6.1	169%
Obesity	0.28	0.135	0.229	0.089	14.0	157%
Asthma	0.106	0.126	0.281	0.108	17.3	160%
Hay Fever	0.092	0.104	0.362	0.078	28.4	364%
Headaches	0.201	0.063	0.129	0.046	8.3	180%
Diabetes	0.04	0.002	0.007	0.002	0.5	250%
Any Condition	0.344	0.23	0.34	0.172		

Notes: All means are calculated using sampling weights. The percentage-point and percent differences in the final two columns refer to the differences in prevalence rates among children with parents who *do* report each health condition and children with parents who *do not* report the same health condition.

Table 2: The Effect of Parental Health on Child Health, Baseline Estimates

	Health Limitation	Obesity	Asthma	Chronic Headaches	Diabetes	Hay Fever	Any Chronic Condition
Parent has Condition (β)	0.060***	0.136***	0.170***	0.084***	0.005***	0.276***	0.161***
	(0.006)	(0.010)	(0.005)	(0.003)	(0.001)	(0.005)	(0.003)
Percent Increase (β/Prevalence)	**166%**	**152%**	**159%**	**184%**	**328%**	**355%**	**94%**
Correlation Coefficient ($\beta \times \sigma_1/\sigma_0$)	0.071	0.182	0.155	0.14	0.026	0.259	0.183
Observations	64,446	9,417	115,821	94,171	115,928	115,180	116,012

Table 3: Intergenerational Transmission by Parent and Child Gender

	Health Limitation	Obesity	Asthma	Chronic Headaches	Diabetes	Hay Fever	Any Chronic Condition
Mothers	0.073***	0.140***	0.191***	0.087***	0.007***	0.316***	0.177***
	(0.007)	(0.012)	(0.006)	(0.003)	(0.002)	(0.007)	(0.004)
Fathers	0.035***	0.126***	0.120***	0.073***	0.002	0.199***	0.128***
	(0.010)	(0.015)	(0.008)	(0.006)	(0.002)	(0.009)	(0.005)
Daughters	0.062***	0.131***	0.154***	0.086***	0.006***	0.258***	0.154***
	(0.008)	(0.013)	(0.006)	(0.004)	(0.002)	(0.008)	(0.004)
Sons	0.059***	0.140***	0.188***	0.083***	0.005**	0.293***	0.170***
	(0.009)	(0.014)	(0.007)	(0.004)	(0.002)	(0.008)	(0.004)
Same Sex Parent-Child Pairs	0.060***	0.143***	0.159***	0.089***	0.006***	0.276***	0.156***
	(0.008)	(0.013)	(0.006)	(0.004)	(0.002)	(0.008)	(0.004)
Opposite Sex Parent-Child Pairs	0.062***	0.128***	0.183***	0.080***	0.004**	0.277***	0.168***
	(0.009)	(0.014)	(0.007)	(0.004)	(0.002)	(0.008)	(0.004)

Notes: All models include indicators of the age and race of both parents and children, and are estimated using sampling weights. Robust standard errors are in parenthesis. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Table 4: Intergenerational Transmission among Biological and Adopted Children

	Health Limitation	Obesity	Asthma	Chronic Headaches	Diabetes	Hay Fever	Any Chronic Condition
Adopted	-0.007	0.016	0.031***	0.007	0.000	0.001	0.042***
	(0.006)	(0.026)	(0.009)	(0.006)	(0.001)	(0.008)	(0.013)
Parent has Condition	0.060***	0.135***	0.170***	0.084***	0.005***	0.276***	0.161***
	(0.006)	(0.010)	(0.005)	(0.003)	(0.001)	(0.005)	(0.003)
Adopted x Parent has Condition	-0.009	-0.081*	-0.048	-0.024	-0.002	-0.054*	-0.053**
	(0.032)	(0.048)	(0.031)	(0.018)	(0.006)	(0.032)	(0.021)
Percent Decline among Adoptees	**-15%**	**-60%**	**-28%**	**-29%**	**-39%**	**-20%**	**-33%**
Correlation among MZ Twins (Add Health)	0.149	0.583	0.302	0.457	-	0.132	0.315

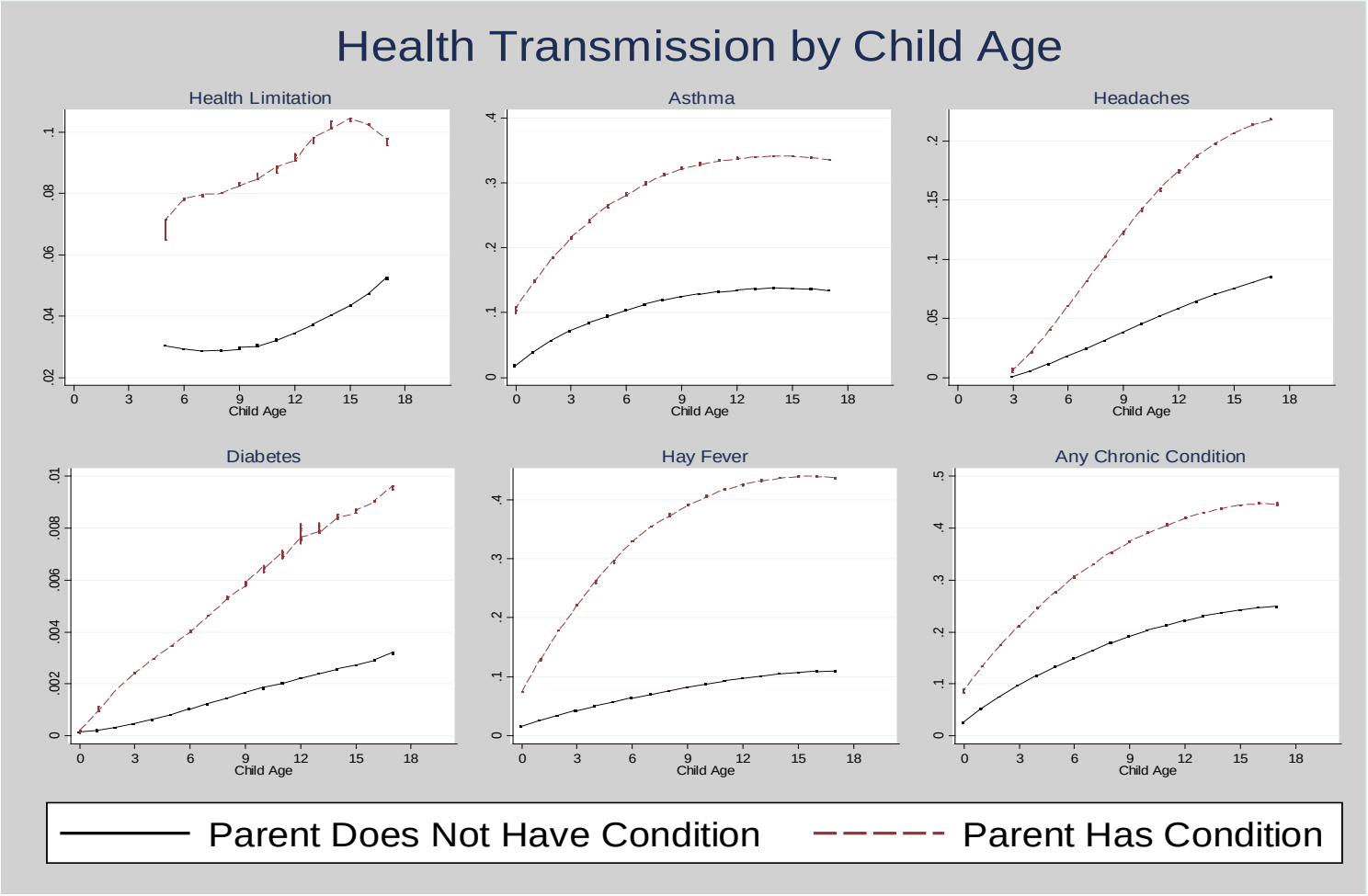
Table 5: Intergenerational Transmission Strength by SES

	Health Limitation	Obesity	Asthma	Chronic Headaches	Diabetes	Hay Fever	Any Chronic Condition
Income Above Poverty Line	-0.037***	-0.034**	-0.022***	-0.020***	-0.000	0.012***	-0.013***
	(0.004)	(0.015)	(0.004)	(0.003)	(0.000)	(0.003)	(0.005)
Parent has Condition	0.059***	0.150***	0.208***	0.108***	0.004	0.313***	0.175***
	(0.017)	(0.025)	(0.011)	(0.007)	(0.003)	(0.017)	(0.007)
Above Povery Line x Parent has Condition	-0.005	-0.010	-0.050***	-0.030***	0.001	-0.041**	-0.019**
	(0.019)	(0.027)	(0.013)	(0.008)	(0.003)	(0.018)	(0.008)
Percent Change among Families not in Poverty	**-8.8%**	**-7.0%**	**-24.2%**	**-28.1%**	**22.5%**	**-13.0%**	**-10.6%**
Parent College Grad	-0.017***	-0.055***	-0.014***	-0.013***	-0.000	0.015***	-0.002
	(0.002)	(0.008)	(0.003)	(0.002)	(0.000)	(0.002)	(0.004)
Parent has Condition	(0.007)	(0.011)	(0.005)	(0.003)	(0.002)	(0.007)	(0.003)
Parent College Grad x Parent has Condition	-0.007	-0.017	-0.041***	-0.034***	0.001	-0.034***	-0.011
	(0.014)	(0.021)	(0.011)	(0.006)	(0.004)	(0.011)	(0.007)
Percent Change among College Educated Families	**-11.7%**	**-12.8%**	**-22.8%**	**-37.6%**	**28.6%**	**-12.0%**	**-6.7%**

Table 6: Intergenerational Transmission Estimates with Environmental Controls Added

	Health Limitation	Obesity	Asthma	Chronic Headaches	Diabetes	Hay Fever	Any Chronic Condition
Baseline with only Demographic Controls	0.054***	0.143***	0.167***	0.084***	0.005***	0.278***	0.159***
	(0.006)	(0.010)	(0.005)	(0.003)	(0.002)	(0.006)	(0.003)
With SES, Health Care Access, and Health Behavior Controls Added	0.047***	0.138***	0.161***	0.076***	0.005***	0.274***	0.151***
	(0.006)	(0.010)	(0.005)	(0.003)	(0.002)	(0.006)	(0.003)
Percent Decline After Adding Controls	**-13.6%**	**-3.3%**	**-3.7%**	**-9.5%**	**-0.2%**	**-1.4%**	**-4.9%**

Figure 1: Health Transmission by Child Age (NHIS)



Profiles are estimated using partially-linear regressions with a bandwidth of 2 years, and parametrically control for race and gender. See Section 1.2 of the text for exact definitions of health conditions. NHIS sampling weights are applied.

Table A1: Prevalence of Health Outcomes Among Parents and Children in Add Health and the NLSY

Add Health

Condition	(1) Parents	(2) All Children	(3) Children of Parents With Condition	(4) Children of Parents Without Condition	(5) Percentage Point Difference	(6) Percent Difference
Health Limitation			--			
Obesity	0.19	0.225	0.386	0.185	20.1	109%
Asthma	0.09	0.168	0.302	0.155	14.7	95%
Headaches	0.28	0.058	0.075	0.051	2.4	47%
Diabetes	0.039	0.009	0.032	0.008	2.4	300%
Hay Fever	0.425	0.152	0.172	0.138	3.4	25%

NLSY

Condition	Parents	All Children	Children of Parents With Condition	Children of Parents Without Condition	Percentage Point Difference	Percent Difference
Health Limitation	0.47	0.163	0.214	0.119	9.5	80%
Obesity	0.262	0.328	0.445	0.302	14.3	47%
Asthma	0.096	0.084	0.219	0.075	14.4	192%
Headaches	0.166	0.014	0.027	0.011	1.6	145%
Diabetes	0.055	0.006	0.023	0.006	1.7	283%
Hay Fever			--			

Notes: All means are calculated using sampling weights. The percentage-point and percent differences in the final two columns refer to the differences in prevalence rates among children with parents who *do* report each health condition and children with parents who *do not* report the same health condition.

Table A2: The Effect of Parental Health on Child Health, Baseline Estimates in Add Health and the NLSY

			Add Health			
	Health Limitation	Obesity	Asthma	Headaches	Diabetes	Hay Fever
Parent has Condition (β)		0.197***	0.151***	0.022***	0.026**	0.033***
		(0.016)	(0.020)	(0.007)	(0.011)	(0.010)
Percent Increase (β/Prevalence)	Not Available	107%	97%	43%	329%	24%
Correlation Coefficient ($\beta \times \sigma_1/\sigma_0$)		0.187	0.113	0.043	0.053	0.046
Observations		9,889	10,435	10,437	10,395	10,421
			NLSY			
	Health Limitation	Obesity	Asthma	Headaches	Diabetes	Hay Fever
Parent has Condition (β)	0.079***	0.139***	0.140***	0.015***	0.017**	
	(0.009)	(0.015)	(0.017)	(0.006)	(0.008)	
Percent Increase (β/Prevalence)	63%	46%	186%	132%	306%	Not Available
Correlation Coefficient ($\beta \times \sigma_1/\sigma_0$)	0.102	0.134	0.140	0.050	0.049	
Observations	8,587	6,936	8,389	7,353	7,363	

Table A3: Intergenerational Transmission Estimates with Environmental Controls Added in Add Health and the NLSY

	Add Health					
	Health Limitation	**Obesity**	**Asthma**	**Chronic Headaches**	**Diabetes**	**Hay Fever**
Baseline with only Demographic Controls	-	0.197*** (0.016)	0.151*** (0.020)	0.033*** (0.010)	0.022*** (0.007)	0.026** (0.011)
With Additional Controls Added	-	0.182*** (0.021)	0.202*** (0.028)	0.048*** (0.014)	0.017* (0.010)	0.030* (0.018)
Percent Decline After Adding Controls	**-**	**-3.3%**	**-3.7%**	**-9.5%**	**-0.2%**	**-1.4%**
	NLSY					
Baseline with only Demographic Controls	0.079*** (0.009)	0.139*** (0.015)	0.140*** (0.017)	0.015*** (0.006)	0.017** (0.008)	-
With Additional Controls Added	0.068*** (0.012)	0.122*** (0.017)	0.121*** (0.019)	0.022*** (0.008)	0.019* (0.010)	-
Percent Decline After Adding Controls	**-13.9%**	**-11.9%**	**-13.4%**	**47.4%**	**6.1%**	**-**

Table A4: Intergenerational Transmission among Biological and Adopted Children with SES controls

	Health Limitation	Obesity	Asthma	Chronic Headaches	Diabetes	Hay Fever	Any Chronic Condition
Adopted	-0.007	0.016	0.031***	0.007	0.000	0.001	0.042***
	(0.006)	(0.026)	(0.009)	(0.006)	(0.001)	(0.008)	(0.013)
Parent has Condition	0.060***	0.135***	0.170***	0.084***	0.005***	0.276***	0.161***
	(0.006)	(0.010)	(0.005)	(0.003)	(0.001)	(0.005)	(0.003)
Adopted x Parent has Condition	-0.009	-0.081*	-0.048	-0.024	-0.002	-0.054*	-0.053**
	(0.032)	(0.048)	(0.031)	(0.018)	(0.006)	(0.032)	(0.021)
Percent Decline among Adoptees	**-15%**	**-60%**	**-28%**	**-29%**	**-39%**	**-20%**	**-33%**