NBER WORKING PAPER SERIES

SOCIOECONOMIC STATUS AND HEALTH IN CHILDHOOD: A COMMENT ON CHEN, MARTIN, AND MATTHEWS (2006)

Anne Case
Christina Paxson
Tom Vogl

Working Paper 12267
http://www.nber.org/papers/w12267

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2006

Socioeconomic Status and Health in Childhood: A Comment on Chen, Martin and Matthews (2006)
Anne Case, Christina Paxson and Tom Vogl
NBER Working Paper No. 12267
May 2006
JEL No. I1

ABSTRACT

Understanding whether the gradient in children's health becomes steeper with age is an important first step in uncovering the mechanisms that connect economic and health status, and in recommending sensible interventions to protect children's health. To that end, this paper examines why two sets of authors, Chen et al (2006) and Case et al (2002), using data from the same source, reach markedly different conclusions about income-health gradients in childhood. We find that differences can be explained primarily by the inclusion (exclusion) of a handful of younger adults living independently.

Anne Case
367 Wallace Hall
Princeton University
Princeton, NJ 08544
and NBER
accase@princeton.edu

Christina Paxson
316 Wallace Hall
Princeton University
Princeton, NJ 08544
and NBER
cpaxson@princeton.edu

Tom Vogl
Center for Health and Wellbeing
Wallace Hall
Princeton University
Princeton, NJ 08544
tvogl@princeton.edu

Introduction

The relationship between socioeconomic status (SES) and health in childhood is an active and important area of research, one that may improve our understanding of the origins of socioeconomic gradients in adult health, and mechanisms through which the intergenerational transmission of poverty takes place. For these reasons, we read with interest the recent article "Socioeconomic status and health: Do gradients differ within childhood and adolescence?" (Chen, Martin and Matthews, 2006). These authors find that, although assessments of child health differ significantly between richer and poorer households, these differences do not increase as children age. This result stands in sharp contrast to recent findings that the relationship between household income and children's health becomes more pronounced as children age, with no diminution of the gradient in adolescence (Case, Lubotsky and Paxson, 2003). The difference in findings between Chen, Martin and Matthews (CMM) and Case, Lubotsky and Paxson (CLP) are surprising, given that the two sets of authors rely on the same data source – the National Health Interview Survey (NHIS) – and choose the same socioeconomic and health variables for their analyses.

Understanding whether the gradient in children's health becomes steeper with age is an important first step in uncovering the mechanisms that connect economic and health status, and in recommending sensible interventions to protect children's health. To that end, this paper examines why these authors reach such different conclusions about the income-health gradient in childhood. We find these differences can be explained primarily by the inclusion, as if they were dependent children, of younger adults living independently.

Methods

Sample choice

CLP analyzed 10 years of NHIS data for approximately 230,000 children aged 0 to 17, surveyed between 1986 and 1995. CMM analyzed data from the 1994 round of the NHIS, for approximately 29,000 children aged 0 to 18. CMM note that, in most cases, a proxy adult respondent (typically the child's mother) reported on the health of persons under age 19. For simplicity, they refer to these responses as 'parental reports.'

From this description, it is clear that the samples analyzed varied in two important ways. CLP worked with a data set that was an order of magnitude larger. In addition, CLP excluded 18 year olds from their analysis, while CMM did not. CLP noted that they chose not to include 18 year olds because they were concerned about the living arrangements of college-aged individuals, and whether these respondents would report their current incomes or the incomes of the families in which they were raised.

Data from the 1994 NHIS suggest it may be problematic to include 18 year olds in such analyses. In the NHIS, all relationships are recorded relative to a *reference person*, who is the person who rents or owns the housing unit surveyed. While, overall, less than one percent of children in 1994 reported themselves to be the household reference person, two percent of 17 year olds and 17 percent of 18 year olds did so. (In what follows, we will refer to a young adult who is not the household reference person as 'dependent' and one who is the household reference person as 'independent.') Nearly 80 percent of independent 18 year olds have completed at least 12 years of schooling (in contrast to 50 percent of the dependent 18 year olds), suggesting that a non-trivial fraction of 18 year olds living away from their families of origin may be enrolled in college. The current household income of such 18 year olds will be a poor guide to the socioeconomic status of the households in which they were raised, and may be an inaccurate estimate of their access to resources. Indeed, only 2 percent of dependent 18 year olds reported household incomes of less than $5,000, while fully 70 percent of independent 18 year olds did so. Moreover, if the healthiest and highest SES 18 year olds were more likely to have left for college, their systematic absence in their households of origin and their presence in newly formed (apparently low income) households will bias estimates of the impact of household socioeconomic status on health outcomes, at least to some degree.

Variable construction

CLP and CMM also differ in their construction of the household income variable. For all years from 1986 to 1995, the NHIS reports information on total household income for 27 income categories, coded from 0 to 26—corresponding to $1000 intervals for incomes between $0 and $20,000, and to $5,000 intervals for incomes between $20,000 and $50,000. The highest income category is "$50,000 or more," so all household incomes above $50,000 are top coded.

CLP assigned incomes to these income categories using data from the 1986-1995 March Current Population Surveys (CPS). Specifically, they calculated, for each income category in each year, the mean total household income in the CPS for households whose head's education matched that of the reference person in the household and whose income fell into that income category. (See the Appendix to CLP for details.) CLP's specifications generally used the logarithm of household income as an explanatory variable. Implicit in this specification is the assumption that, in wealthy and poor households alike, a one-percent increase in household income has the same marginal effect on the health of children of a given age. However, CLP note that their results are robust to changes in functional form. In particular, their finding that the gradient in children's health becomes steeper with each year of age continues to hold when household income is used in place of the logarithm of household income.

CMM chose to measure household income as the categorical income variable published in the NHIS. That is, their income variable consists of integer values from 0 to 26. Implicit in this choice is an assumption that a $1000 increase in income in poorer households (i.e., a one-unit change in the NHIS's categorical income variable for households with incomes below $20,000) has the same marginal effect on children's health as does a $5,000 increase in income in wealthier households (i.e., a one-unit change for households with incomes above $20,000 annually). This specification has the very odd implication that the effect of an additional dollar is five times as large for households with incomes of $19,999 than it is for households with incomes of $20,000.

Data analysis

We re-analyzed the NHIS data using STATA/SE 9.2. Following CMM, we first ran logistic regressions using 1994 data only, in which an indicator that a child is reported in fair or poor health is regressed on the income measure used by CMM, age, an income-age interaction, and controls for race (white, black, other) and sex. The analysis is repeated, restricting the sample to dependent children (that is, removing 232 children – primarily

17 and 18 year olds – reported to be the household reference person). We then estimate this model using 10 years of NHIS data from 1986-95. Finally, we repeat these analyses using the logarithm of income in place of the CMM measure. All results were weighted using sampling weights. In all cases, we report odds ratios (ORs) so that our results can be directly compared to those of CMM.

Results

Table 1 presents the results of logistic regressions using the CMM categorical income measure in columns 1 to 4, and using the CLP log income measure in columns 5 to 7. Column 1 reproduces CMM's estimate of the impact of household income on the log odds of reporting a child in fair or poor health, and column 2 produces results reported in CMM that they find no significant age-income interaction term. The lack of a significant age-income interaction term is also seen in column 5, where the categorical income measure has been replaced by log income. However, once the 232 independent young adults are removed from the sample (column 6), the age-income interaction using log income becomes statistically significant ($P<.05$). For both income specifications using 10 years of NHIS data (columns 4 and 7), once young adults living independently are removed from the sample, we find age-income interactions that are statistically significant (P's<.001). Furthermore, the estimates imply substantial changes in the relationship between income and health as children become older. The results in column 7 suggest that a change in household income from the 75th to the 25th percentile of the 1986-95 income distribution (from about \$37,000 to about \$13,000) at age 3 increases the probability of being in fair or poor health from 1.7 percent to 2.7 percent. The same change in household income at age 15 increases the probability from 1.8 percent to 3.3 percent.

Table 1 suggests that the inclusion of a small number of young adults living on their own changes the estimated age-income interaction term dramatically. This can also be seen in Figure 1, which plots the coefficients on log income from logistic regressions run separately for each age, from 0 to 18, for all children observed in the NHIS from 1986 to 1995. The lightly shaded diamonds represent the odds ratios for log income from age-specific regressions run for all children, while the darkly shaded circles correspond to the odds ratios for dependent children only. These show convincingly that income is increasingly protective with age when the newly-independent children are excluded. Note that, from age 0 to age 13, there are no differences in the estimated coefficients for the two samples—dependents only and all children—because no children below age 14 are reported to be living on their own. The low incomes of college-aged young adults living independently have a very large effect on the estimated relationship between income and health status in the NHIS for 17 and 18 year olds. Similar results are found using ordered probits of health status, measured on a five-point scale.

Discussion

The differences in the age patterns observed for income-health gradients between CMM and CLP are not driven by CMM's use of a categorical income variable. Both the CLP and the CMM age-income interactions terms are insignificant when young adults living independently are included in the analysis. Both become highly significant and imply large changes in the gradient with age when the analysis is restricted to dependent children and 10 years of NHIS data are used.

Neither are the differences between the CLP and CMM results due to the much larger sample size used by CLP. The age-income interaction term becomes significant using CMM's income measure, but restricting the analysis to dependent children, if one combines only two years of NHIS data – 1994 and 1995, for example ($P<.001$). Rather, the differences between the CLP and CMM results are driven largely by CMM's inclusion of independent 18 year olds. We believe it is inappropriate to include these young adults in an analysis of the impact of socioeconomic status on health in childhood and adolescence, as the "socioeconomic status" markers available for these independent college-aged individuals are very possibly a poor reflection of the socioeconomic status they enjoyed for much (or most) of their lives.

Alternative measures of socioeconomic status in a child's household of origin face the same problem. CMM construct and use a variable measuring the years of education of a parent (or responsible adult) in each child's household. However, creation of this variable is made problematic by the 17 and 18 year olds who report themselves as the household reference person. For these young adults, the education of the "responsible adult" defaults to their own education (in most cases, 12 years of completed schooling). This adds measurement error to a variable intended to capture long-term effects of parents' education, and may bias all estimated coefficients.

Young adults are an important group, deserving of attention. However, we will learn more by following them longitudinally (out of their households of origin and into adulthood – see, for example, Case, Fertig and Paxson, 2005) than we will learn looking at them in cross-sectional studies, such as the NHIS.

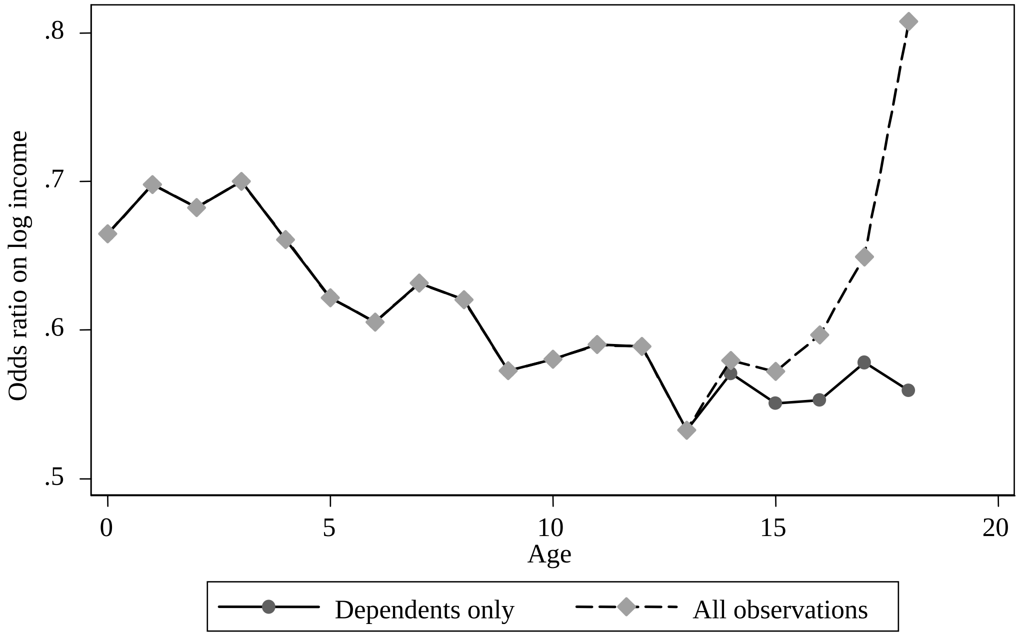
References

Case, A., Fertig, A., & Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of Health Economics*, 24, 365-389.

Case, A., Lubotsky, D., & Paxson, C. (2002). Economic status and health in childhood: The origins of the gradient. *American Economic Review*, 92, 1308-1334.

Chen, E., Martin, A. D., & Matthews, K. A. (2006). Socioeconomic status and health: Do gradients differ within childhood and adolescence? *Social Science & Medicine*, 62, 2161-2170.

Figure 1: Age patterns in the fair/poor health gradient, NHIS 1986-1995



Note: The dependent variable equals 1 if the child is in fair/poor health, 0 otherwise. Odds ratios are adjusted for race, sex, and survey year. All specifications include survey weights.

Table 1
Exploring income-age interactions using data from the National Health Interview Survey

	Chen et al income measure				Log income		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income	0.935 (0.004)**	0.935 (0.008)**	0.941 (0.008)**	0.940 (0.003)**	0.636 (0.029)**	0.683 (0.034)**	0.682 (0.012)**
Age × Income		1.000 (0.001)	0.999 (0.008)	0.998 (0.0003)**	1.004 (0.004)	0.989 (0.005)*	0.987 (0.002)**
Sample	All	All	Dependents	Dependents	All	Dependents	Dependents
Years	1994	1994	1994	1986-95	1994	1994	1986-95
Observations	27833	27833	27601	274623	27833	27601	274623

Note: Odds ratios from logistic regression models, robust standard errors in parentheses. The dependent variable equals 1 if the child is in fair or poor health, 0 otherwise. Log income is calculated using the mean of each income bracket in the March Current Population Surveys. All models include controls for age, race, and sex, and are weighted using survey weights. * significant at 5%. ** significant at 1%.