NBER WORKING PAPER SERIES

DYNAMIC ASPECTS OF CHILDREN'S HEALTH, INTELLECTUAL DEVELOPMENT, AND FAMILY ECONOMIC STATUS

Robert A. Shakotko

Working Paper No. 451

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge MA 02138

February 1980

Research for this paper was supported by a grant from the Robert Wood Johnson Foundation to the National Bureau of Economic Research. I am indebted to Ann Colle for excellent research assistance, and to Anthony Cassese, Michael Grossman, Lee Lillard, and Robert Moffitt for valuable comments and suggestions. The research reported here is part of the NBER's research program in Health Economics. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper 451
February, 1980

Dynamic Aspects of Children's Health, Intellectual Development, and Family Economic Status

ABSTRACT

This paper is an empirical investigation of childhood and adolescent health and cognitive development as determined by family economic variables. The model proposed recognizes that these processes may be jointly dependent, and may in part be determined by common unobserved factors; these factors may also be correleted with the observed family economic variables. A two-factor model is estimated using panel data, and the results indicate that when such factors are taken account of, family income is estimated to have no significant influence on health and cognitive development, but parents' education a strong positive influence.

Robert A. Shakotko
Department of Economics
Columbia University
New York, New York 10027

(212) 280-2449

Natioanl Bureau of Economic Research
15-19 West 4th Street, 8th Floor
Washington Square
New York, New York 10012

(212) 598-3439

The relationships between an individual's adult health and various economic variables have been examined empirically in a number of recent studies (Grossman and Benham, 1974; Grossman, 1976; Lee, 1979; Shakotko, 1979). They stem from the idea that one's state of health is at least partly endogenous, reflecting one's stock of health capital, whose accumulation is determined to some extent by economic and educational factors (Grossman, 1972). In turn, these economic and educational variables may themselves be partly determined by health. Investigators have also recognized that an individual's lifetime economic and physical well-being is due in large measure to conditions present or acquired during childhood. Educational success, for example, which is a significant determinant of adult earnings, may be partly determined by the individual's cognitive development and physical health as a child. Poor adult health, and its consequent economic hardship, may arise in many instances from poor health during childhood or adolescence.[1] Given this link between life-cycle behavior and childhood conditions, it is important to understand the determinants of childhood development.

This paper is an empirical investigation of the determinants of childhood and adolescent health and cognitive development. The main feature of this investigation is that it draws together into one estimable empirical model three classes of postulated determinants of early health and cognitive development, and estimates their contributions simultaneously. In this sense, the model is superior to previous studies which have analysed these determinants on a limited

and piecemeal basis. The model takes explicit account of the possibility that individual-specific unobserved factors jointly determine health and cognitive development, and recognizes that these factors may themselves be related to other observed determinants. The emphasis is on a two-factor model, where all of the health, cognitive development, and background economic variables serve as indicators of the unobserved factors. The model is estimated with panel data for approximately 1400 children.

Three main results emerge from the estimates of this model. First, systematic unobserved factors are found to play a significant role in the determination of both health and cognitive development. These effects are relatively stronger in the equations determining cognitive development. Second, parents' education is found to be a positive contributor to both processes, but family income is observed to have no significant effect. Third, a favorable history of early childhood achievement is found to have a small positive effect on future health, but poor early health has little effect on cognitive development. In fact, there is some evidence that poor health influences cognitive development positively.

The rest of the paper is organized as follows. Section I outlines the basic model both theoretically and empirically. In Section II, there is a brief discussion of the data. In Section III, the estimates of the model are presented. These are compared to estimates of a model with no factor structure in terms of both effects of different variables and predictive power.

I. An Analytic Framework

Three classes of variables which have been widely proposed as determinants of children's development might be labelled "nurture", "nature", and "history." While these are broad classifications, and perhaps at times misleading or inappropriate, they serve well to indicate the dimensions of the analysis. In general, nurture refers to the impact of discretionary (or interventionary) behavior on the part of the household which affects children's development. Economic models of household production have hypothesized that this discretionary behavior is manifested in the form of time and income investments in children.[2] While it is difficult to measure these these investments directly, it is argued that the efficiency and magnitude of the investments are determined to a large extent by parents' education and family income.

The "nature" explanation of childhood development centers on those factors which are predetermined and therefore not controllable by the household. In medical literature, this effect is often labelled predisposition to certain health conditions. In terms of cognitive development, "nature" explanations are commonly associated with arguments for innate abilities and IQ inheritability.[3] It is, however, important to point out that at the margin the distinction between nurture and nature effects is not totally resolved even conceptually. These effects outlined above relate to physical predispositions. But it is not clear, however, where behavioral predispositions which cannot be explained by economic or educational factors should be accounted.

Inclusion of "history" as a determinant of health and cognitive development reflects the recognition that these processes may be structurally related through time, and that a particular event (an event being the collection of health and cognitive development outcomes at a point in time) may be due in some measure to the occurrence of a prior event. Stated alternatively, even controlling for nurture and nature inputs into developmental processes, observed outcomes may not be temporally independent. For example, some health conditions tend to be persistent, such that poor health at birth or the acquistion of certain conditions in early childhood may explain the presence of these or other conditions later in the individual's life, quite aside from intervening or preceding nature and nurture influences. Also, children who develop intellectual skills early may be able to use these early abilities to acquire greater proficiency in other aspects of cognitive development. It is important to point out that these history effects need not be conceptually restricted to operate in univariate fashions. Indeed, it is conceivable that aspects of health and cognitive development are interdependent, whereby, for example, poor early health could impede subsequent cognitive development and below average levels of cognitive development could interfere with the child's ability to maintain good health.[4]

While it is likely that the child's history of health and cognitive development is itself a function of various nurture and nature variables, controlling for the history of these processes will yield estimates of the structural links between different

aspects of health and cognitive development and also more precise estimates of the effects of discretionary behavior as it is affected by parents' education and family income. These structural links usually cannot be identified in cross-section formulations because of the lack of identifying restrictions provided by the underlying theory. Moreover, modelling these processes whereby past values of health and cognitive development are explicitly controlled for conforms more closely to classical experimental analysis. Typically, in laboratory experiments, assessment of an action's efficacy is made by observing the initial state, administering a measured "dosage", and then observing the final state, all in an effort to determine if and to what extent the procedure had any impact. Different aspects of the initial state and the intervening dosage are then related to the observed final outcome in an attempt to explain the structure of a process. In the absence of a laboratory, the same conceptual framework can be applied to the analysis of children's development by observing a time series for a number of individuals. By treating a particular child's history as a comparative benchmark, one can identify temporal interdependence between processes and perhaps more accurately measure the contributions of other variables to favorable or unfavorable outcomes.

In short, a completely specified model of children's development must be capable of incorporating all of the above as potential sources of observed variation in measures of health and cognitive development, as well as recognizing the possible relationships among the three classes of determinants themselves. It is well-known that

this latter point is a particularly troublesome statistical issue, since "nature" variables are in practice unobserved, and even in principle may not be observable or measurable. In the absence of information on this potential determinant, variance in observed outcomes may be incorrectly ascribed to other observed determinants, especially since it is commonly argued that "nature" is correlated with both history and nurture.

An analogy between this problem and the familiar ability-bias problem in estimates of earnings functions is apparent. If ability is correlated with the observed level of completed schooling, and if both influence earnings, then an estimated earnings function which disregards ability will yield schooling coefficients which are biased upwards. As a solution to this specification problem when ability is not observed, Chamberlain (1977,1978) has proposed joint modelling of equations for all observed variables which may be indicators of an unobserved factor or factors. Given a sufficient set of identifying restrictions, the parameters of such equations, and the distributional parameters of the unobserved factor(s), can be estimated using a random effects procedure.

A linear formulation of such a model may be written

$$(1) \qquad By_i = \Lambda f_i + \varepsilon_i \quad ,$$

where y_i is an nx1 column vector of observed variables including measures of current and past health and cognitive development, parents' education, and family income, where f_i is a kx1 vector of exogenous individual- or family-specific unobserved

factors, and where ε_i is a vector of stochastic disturbances, uncorrelated with f_i. The index i indicates the i-th individual. Denoting the covariance matrices of f_i and ε_i by Σ_f and Σ_ε respectively, and premultiplying (1) by B^{-1} yields

$$y_i = B^{-1}\Lambda f_i + B^{-1}\varepsilon_i$$

and

$$\Sigma_y = B^{-1}\Lambda\Sigma_f\Lambda'(B^{-1})' + B^{-1}\Sigma_\varepsilon(B^{-1})' \quad , \tag{2}$$

where Σ_y is the nxn covariance matrix of y_i. Assuming f_i and ε_i are distributed normally, the log-likelihood function for the i-th individual is

$$L_i(B,\Lambda,\Sigma_f,\Sigma_\varepsilon) = -1/2 \log|\Sigma_y| - 1/2\,(y_i-E(y_i))'\Sigma_y^{-1}(y_i-E(y_i)) \quad , \tag{3}$$

and if the vectors f and ε are uncorrelated across individuals, then the joint likelihood function for a sample of N individuals can be written

$$L(B,\Lambda,\Sigma_f,\Sigma_\varepsilon) = -N/2 \log|\Sigma_y| - N/2\,\mathrm{tr}(S\Sigma_y^{-1}) \quad , \tag{4}$$

where S is the sample covariance matrix of y, and where Σ_y is given by (2). The remainder of the paper presents estimates of B, Λ, Σ_f, and Σ_ε obtained by maximizing (4) under various restrictions.

One cautionary word should be expressed at this point. While the vector structural equation (1) takes explicit account of unobserved variables, and while most discussions of unobserved variables in the context of children's development have centered on genetic factors as

being the chief component of "nature", unobserved variables by their very nature have unobserved names. The primary interest in this paper is not so much in assigning names and interpretations to the estimated factor structure, but rather in estimating the relationships among health, cognitive development, and family background variables once the possibility of common determining factors is taken account of.

II. The Data

The model outlined in the preceding section is estimated using data from Cycles II and III of the U.S. Health Examination Survey (HES). Both sections of the survey collected extensive data on the current health and health histories for a nationally representative sample of children. Cycle II surveyed children aged 6 to 11 in the early 1960s, and Cycle III surveyed adolescents aged 12 to 17 in the late 1960s. In each case, data include results from a physician's examination, scores from different tests, and surveys of the child's parents and school.

Approximately 2200 of the children in Cycle II were re-surveyed in Cycle III. Of these, a final sample of 1434 was selected for analysis. Because of cross-section evidence that the health and cognitive development processes differ significantly for blacks, they were not included in this sample, nor were observations with missing data. It should be noted that there is considerably more age homogeneity in this longitudinal subset than in the overall cross-section surveys: in Cycle III, 99 percent of the adolescents in the longitudinal subset were aged 12 to 15. The re-survey interval

of this subsample averages 42 months, with 70 percent of the sample intervals falling between 36 and 48 months. While it is difficult to assess the effects of even this limited age and interval heterogeneity on the estimates of the model, it is unlikely that any of the main results will be substantially affected. Using the same data, Shakotko, Edwards, and Grossman (1980) found that controlling for these variables did not in most cases contribute significantly to their model's explanatory power. In any case, both measures of cognitive development and one measure of health used in this analysis are age-adjusted.

Specifically, the analysis examines the relationships among thirteen variables. The primary focus is on two measures of cognitive development and two measures of health: (1) the Wechsler Intelligence Scale for Children (WISC); (2) the Wide Range Achievement Test (WRAT); (3) the age- and sex-adjusted periodontal index, which is a measure of oral health (APERI);[6] (4) the diagnosis of a "significant abnormality" by the examining physician (ABN). Each of these four variables is reported in Cycles II and III, and to distinguish the two measures of each variable, the post-script "1" is used to denote a Cycle II value and "2" to denote a Cycle III value. WISC, WRAT, and APERI are all continuous variables, and and ABN is a dichotomous variable taking the value of one if an abnormality was found and zero otherwise.

Other health measures included to control for history effects are dummy variables indicating a parental assessment of poor health for the child during the first year after birth (FYPH) and a dummy

indicating if the child's birth weight was 2000 grams or less (LIGHT).[7] Family background variables used are family income (FAMINC), father's education in years (FATHED), and mother's education (MOTHED). The sample means and standard deviations of these variables are reported in Table 1.

The particular measures of health and cognitive development included in the analysis were chosen to ensure some level of generality while at the same time acknowledging that both health and cognitive development have different components. For example, WISC is designed to measure general cognitive aptitudes, while WRAT measures more specific acquired skills. The abnormal diagnosis indicator (ABN) is an objective measure of the presence of a condition which affects current physical activities, or could affect future activities.[8] The periodontal index (APERI) was used because it is perhaps the most sensitive to differential levels of family health investment, and hence may be best indicative of preventive medical care.

Without doubt, this analysis is restrictive, not only in its consideration of relatively few measures of health and cognitive development, but also in its limited consideration of other observable variables that may affect developmental processes. A more detailed examination of other variables was conducted by Edwards and Grossman (1978) and Shakotko, Edwards, and Grossman (1980); these studies have shown that most other variables contribute relatively little to explaining the variance of the measures analysed here. Moreover, it is likely that consideration of larger and more complex

Table 1 : Sample Statistics

Variable Name	Mean	Standard Deviation
WISC2[a]	104.51	14.00
WRAT2[a]	104.11	13.56
APERI2[b]	-0.138	0.852
AEN2	0.188	0.391
WISC1[a]	103.51	13.92
WRAT1[a]	103.57	12.02
APERI1[b]	-0.055	0.792
AEN1	0.096	0.294
FAMINC[c]	8.060	4.607
FATHED	11.31	3.36
MOTHED	11.22	2.70
FYPH	0.068	0.252
LIGHT	0.008	0.091

[a]Age-adjusted.

[b]Age- and sex-adjusted.

[c]In thousands of dollars; measured in Cycle II.

models, while conceptually more desireable, would add substantially to computational difficulty.

III. Estimates

The basic linear model (1) which is estimated consists of thirteen equations. Each of the Cycle III measures of health and cognitive development (WISC2, WRAT2, ABN2, APERI2) is written as a linear function of the health and cognitive development measures in Cycle II (WISC1, WRAT1, ABN1, APERI1), family income (FAMINC), and parents' education (FATHED, MOTHED). The four measures from Cycle II are written as linear functions of family income, parents' education, and initial health (FYPH, LIGHT). To complete the model, family income is presumed to depend on the education of each of the parents, and "first-year poor health" (FYPH) on low birth weight. Parents' education and low birth weight are not presumed to depend on any observable variable. It is apparent that this linear specification corresponds to restrictions on the elements of the matrix B in (1).

To this specification of the structural links between observable variables is added different specifications of an unobserved factor structure. Estimates of two models are reported here. In the first, the factor space is presumed to be two-dimensional, with one factor operating in the health equations (ABN2, APERI2, ABN1, APERI2, FYPH, LIGHT) and the other operating in the cognitive development equations (WISC2, WRAT2, WISC1, WRAT1), each with equal effects in the equations determining income and parents' education (FAMINC, FATHED, MOTHED). The factors are normalized to have unit variance, and the

correlation between the two factors is a parameter to be estimated. The second specification presumes that observed variables alone determine y (i.e. that $\Lambda=0$ in (1)). In the discussion following, this is referred to as the zero-factor model. In both the two-factor and zero-factor specifications, the unsystematic stochastic effects in each equation (the elements of the vector ε in (1)) are presumed to be uncorrelated, except in the case of the two education equations.[9] This constrains the covariance matrix Σ_ε to be nearly diagonal, with the one exception noted above. It can be verified that these implied restrictions are sufficient to identify all the unrestricted parameters of (4).

To a large extent, the empirical focus of this paper is exploratory, so that the zero-factor model serves as a comparative benchmark. Two empirical questions underlie the use of a model with unobserved factors: (1) can a model which takes account of such factors better explain the observed pattern of health and cognitive development? and (2) how do the estimates of such a model differ from those based solely on observed determinants?

The two specifications were estimated by maximizing (4) with respect to the unknown and unconstrained elements of B, Λ, Σ_f, and Σ_ε. Since B is constrained to be triangular and Σ_ε nearly diagonal, the estimates of the zero-factor model are equivalent to those obtained by estimating each structural equation in (1) by OLS. This makes these estimates generally comparable to those reported in Edwards and Grossman (1978) and Shakotko, Edwards, and Grossman (1980), and it is easily verified that the estimates are very similar.

One should point out that the presence of dichotomous variables in the vector y violates the normality assumption implicit in the proposed likelihood function (4), so that maximization of (4) is a quasi-likelihood procedure. A correct likelihood specification to take account of the dichotomous variables, or a two-stage procedure such as that suggested by Heckman (1978a), would be methodologically superior, but the computational difficulties for this thirteen equation model make such techniques impractical.

Tables 2 and 3 present estimates of the matrix B for the two-factor and zero-factor models. Table 4 presents the estimated matrix of factor loadings in the two-factor model. The remainder of this section summarizes the main results.

(a) Interactions Between Health and Cognitive Development

It was argued in Section I that the use of longitudinal data and the inclusion of Cycle II data in the Cycle III equations is one way to assess the structural dependencies between measures of health and cognitive development. The underlying question is whether each temporal event (an event being a particular outcome for WISC, WRAT, APERI, ABN) depends on previous events, or whether it is the result of an independent drawing from a particular distribution whose parameters depend on some set of observed and unobserved background variables.[10] In terms of the estimates reported in Table 2, the coefficient on each variable's own lagged value is a measure of univariate dependence, or structural persistence, in each process. The coefficients on

TABLE 2

Structural Coefficients in Cycle III Equations
For (a) Two Factor Model
For (b) Zero Factor Model

Explanatory Variables		WISC2	WRAT2	APERI2	ABN2
WISC1	(a)	0.124 (0.41)	-0.111 (0.63)	-0.001 (0.28)	0.002 (1.13)
	(b)	0.623 (34.2)	0.183 (11.4)	-0.003 (1.73)	-0.001 (0.62)
WRAT1	(a)	-0.094 (0.50)	0.565 (5.04)	-0.005 (1.82)	-0.001 (0.30)
	(b)	0.215 (10.6)	0.746 (42.0)	-0.006 (3.40)	-0.003 (2.87)
APERI1	(a)	0.640 (2.01)	0.318 (1.20)	0.232 (6.88)	-0.042 (1.42)
	(b)	0.316 (1.11)	0.127 (0.51)	0.269 (10.3)	0.009 (0.72)
ABN1	(a)	-0.345 (0.42)	0.491 (0.71)	-0.097 (1.20)	0.052 (0.83)
	(b)	-0.996 (1.32)	0.108 (0.16)	-0.027 (0.39)	0.148 (4.26)
FAMINC	(a)	-0.014 (0.04)	0.032 (0.14)	0.000 (0.06)	0.008 (1.12)
	(b)	0.101 (1.73)	0.100 (1.94)	-0.002 (0.45)	0.004 (1.64)

(continued on next page)

TABLE 2 (concluded)

Explanatory Variables		WISC2	WRAT2	APERI2	ABN2
FATHED	(a)	0.906 (2.77)	0.543 (2.68)	-0.011 (1.25)	-0.009 (1.39)
	(b)	0.223 (2.90)	0.141 (2.08)	-0.008 (1.15)	-0.004 (1.16)
MOTHED	(a)	0.653 (1.36)	0.526 (1.81)	-0.023 (2.00)	0.004 (0.45)
	(b)	0.142 (1.58)	0.225 (2.84)	-0.023 (2.84)	0.004 (0.93)
Residual Variance	(a)	43.55	45.20	0.578	0.119
	(b)	70.80	54.68	0.594	0.149

Note: Each equation is reported in the form $y = X\hat{\beta} + \hat{\varepsilon}$. Asymptotic t-statistics are reported in parentheses below each coefficient estimate.

TABLE 3
Structural Coefficients in Cycle II, Income, and Initial Health Equations
For (a) Two Factor Model
For (b) Zero Factor Model

Explanatory Variables		WISC1	WRAT1	APERI1	ABN1	FAMINC	FYPH
FAMINC	(a)	-0.062	-0.030	-0.004	0.001	-	-
		(0.17)	(0.11)	(0.42)	(0.24)	-	-
	(b)	0.220	0.184	-0.009	-0.001	-	-
		(2.61)	(2.42)	(1.67)	(0.37)	-	-
FATHED	(a)	0.934	0.739	-0.019	0.002	0.528	-
		(3.69)	(3.66)	(2.01)	(0.71)	(12.8)	-
	(b)	0.946	0.748	-0.019	0.002	0.539	-
		(8.91)	(7.84)	(2.76)	(0.91)	(17.9)	-
MOTHED	(a)	0.914	0.519	-0.016	0.000	0.351	-
		(1.97)	(1.49)	(1.27)	(0.08)	(5.11)	-
	(b)	1.328	0.833	-0.024	-0.002	0.383	-
		(10.8)	(7.51)	(3.09)	(0.57)	(10.3)	-
FYPH	(a)	2.426	-0.247	-0.023	0.047	-	-
		(2.43)	(0.20)	(0.21)	(1.26)	-	-
	(b)	0.305	-1.854	0.124	0.085	-	-
		(0.24)	(1.62)	(1.52)	(2.75)	-	-
LIGHT	(a)	-5.046	-8.100	-0.042	-0.031	-	0.087
		(1.94)	(3.10)	(0.18)	(0.35)	-	(1.16)
	(b)	-8.602	-10.80	-0.013	-0.023	-	0.099
		(2.43)	(3.40)	(0.06)	(0.27)	-	(1.36)
Residual Variance	(a)	52.85	64.99	0.570	0.084	14.29	0.061
	(b)	148.30	119.79	0.604	0.086	14.55	0.064

Note: Each equation is reported in the form $y = X\hat{\beta} + \hat{\varepsilon}$. Asymptotic t-statistics are reported in parentheses below each coefficient estimate.

other lagged variables measure dependence between processes.

The estimated coefficients for the two-factor model (denoted by rows (a) in Table 2) indicate significant qualitative differences in the structural persistence of each of the four processes. WISC2 and ABN2 do not depend significantly on WISC1 and ABN1 respectively, while WRAT2 and APERI2 exhibit substantial dependence on WRAT1 and APERI1. It is not hard to rationalize at least this latter result. If WRAT measures acquired skills, then it is not likely that these skills will disappear over a three- or four-year interval, so that current WRAT scores will reflect to a large extent past accumulation of skills, quite aside from whether this accumulation is related to other variables. Similarly, one might imagine that oral health is also persistent: decay or periodontal disease tends not to appear or disappear in a fairly short interval. WISC, on the other hand, purports to measure aptitudes or innate abilities. It is reasonable to argue that the score on one test should not depend on the score in a previous test, but that both scores reflect independent measurements of underlying individual-specific variables.[11] It is interesting and puzzling, however, that there is not significant persistence in the ABN equation. The same argument used for the periodontal index might apply in the case of a diagnosed abnormality. It is noteworthy that this lack of persistence is also evident in the raw data, where the correlation between ABN1 and ABN2 is only 0.11, so that the small coefficient on ABN1 may not be a statistically artifact.

The point estimates for the two-factor model indicate less structural persistence than corresponding estimates for the zero-

factor model, also reported in Table 2 (and denoted by rows (b)). The implication is that the observed serial correlation in health and cognitive development is better explained by persistent underlying factors than persistence in the actual processes. This is examined in greater detail in section (c) below.

In general, the estimated cross-effects between processes are small and not significant. This is found in both the two-factor and zero-factor models. Two exceptions should be noted. First, WISC and WRAT contribute positively to APERI2 in both models, although only WRAT is significant in the APERI2 equation at the 90 percent level in the two-factor model. The effect is not very large, however: expressing the point estimate of -0.005 in terms of standard deviations, a one standard deviation increase in WRAT1 results in a 0.07 standard deviation improvement in APERI2.

The second exception is in the WISC equation, where high values of APERI1 (i.e. poor oral health) are associated with high WISC scores. Even though the magnitude of this effect is small (a one standard deviation change in APERI1 results in a 0.04 standard deviation change in WISC2), the coefficient is significant, and the direction of the effect is contrary to what one might expect from a structural relationship. The coefficient may, however, reflect a substitution effect in a heterogeneous population, in the sense that some families may choose to make time and income investments directed toward health, and other families investments directed toward cognitive development of children.

(b) Income and Education Effects

In all the equations reported in Tables 2 and 3, family income is not a significant determining factor in the two-factor model. This is in contrast to the estimates of the zero-factor model, where income has a significant positive effect on both WISC and WRAT. It should be noted, moreover, that in the two-factor model, the point estimates of these and other income effects are substantially smaller than those in the zero-factor model, so that the insignificance does not arise totally from larger standard errors.

On the other hand, the two-factor model confirms many of the education effects in the zero-factor model, and in some cases, the point estimates are larger. In the WISC2 and WRAT2 equations, the effect of a marginal year of father's education increases from 0.223 to 0.906 and 0.141 to 0.543 respectively. Mother's education is less significant in these equations, but nevertheless, the point estimates show this same pattern of increase. The education estimates are generally significant and of the same order of magnitude in the WISC1 and WRAT1 equations.

In the health equations, parents' education is significant in both APERI equations, although at marginal confidence levels. This is not the case in the ABN equations, where neither the coefficients nor the confidence levels are large.

Two other equations linking observed variables were also estimated jointly with the Cycles II and III health and cognitive development equations. In the first, family income was supposed to

depend on parents' education. It is found that a marginal year of father's education contributes on average $528 and mother's education $351. Evaluated at the mean level of income, these estimates imply rates of return of 6.5 and 4.5 percent respectively, which are comparable to other estimates of earnings functions. The second equation supposes that "first-year poor health" depends to some extent on low birth weight. A positive effect is found in both models, but in neither case is it significant.

(c) Factor Loadings

A casual inspection of Table 4, which shows the estimated factor loadings in the two-factor model, suggests that a large part of the variance in health and cognitive development in children can be explained by common unobserved factors. This is especially evident in the cognitive development equations, where the positive point estimates (reported as Factor 1 in Table 4) indicate a factor contributing positively to WISC and WRAT. Squaring the ratio of the factor loading to the standard deviation of the corresponding dependent variable yields the proportion of variance explained by the factor. In the case of WISC, this proportion is about 53 percent in both Cycles II and III. In the case of WRAT, the proportion is 40 percent in Cycle II, and declines to 19 percent in Cycle III. This is perhaps indicative of the growing relative influence of other factors on WRAT in the course of the child's development.

The second common factor was constrained to operate in the health

TABLE 4
Factor Loadings[a] and Residual Variances: Two Factor Model

Dependent Variable	Factor 1	Factor 2	Residual Variance
WISC2	10.180 (3.67)	- -	43.55
WRAT2	6.003 (3.64)	- -	45.20
APERI2	- -	-0.140 (2.81)	0.578
ABN2	- -	-0.192 (3.09)	0.119
WISC1	10.019 (6.15)	- -	52.85
WRAT1	7.592 (6.44)	- -	64.99
APERI1	- -	-0.194 (2.69)	0.570
ABN1	- -	-0.050 (2.23)	0.084
FAMINC1	0.314 (0.82)	0.314 (0.82)	14.29
FATHED	0.343 (1.35)	0.343 (1.35)	10.87 (covar.=5.12)
MOTHED	0.343 (1.35)	0.343 (1.35)	7.02
FYPH	- -	-0.050 (3.28)	0.061
LIGHT	- -	-0.002 (0.50)	0.008

[a]Factors are normalized to have unit variances. The correlation between the two factors was estimated at 0.397 (std. err. = 0.100). Asymptotic t-statistics are reported in parentheses below each coefficient.

equations for Cycles II and III, and in the equations determining FYPH and LIGHT (Factor 2 in Table 4). It enters with a negative sign in all health equations, indicating a common positive contribution to health. The estimated loadings are significantly different from zero in all except the LIGHT equation. The explanatory contribution of this health factor is not nearly so dramatic as that of the cognitive development factor discussed above. In the APERI equations, 5 percent of the Cycle II variance can be attributed to the factor, and less than 3 percent in Cycle III, again indicative of growing environmental influence. In the ABN equations, this proportion is 3 percent in Cycle II, but grows to 25 percent in Cycle III. Approximately 4 percent of the FYPH variance was explained by the factor.

Both factors were presumed to be reflected equally in family income and parents' education. Being positive factors in the sense that they contribute to higher levels of WISC and WRAT and better health, one might expect that if these factors are generationally related, they should also be positively related to income and education levels. The estimates confirm this expectation, although for neither income nor education are the estimated loadings significantly different from zero. This insignificance may reflect more of an attenuation in the factor communalities between parents and children than the absence of common factors in these "parents'" equations. Despite the insignificance, the point estimate for education is fairly large. One positive standard deviation in either factor is associated with 0.34 additional years of schooling for each parent.

Besides the factor loadings, the correlation between the two factors was estimated. This correlation is estimated to be 0.397, with a standard error of 0.100, indicating a significant positive relationship between the health and cognitive development factors.

One final issue should be addressed, namely that of the goodness-of-fit of each of the models. Under the assumption of normality, $-2L(max)$ is distributed as $\chi^2(q)$, where $q=n(n+1)/2 - p$ and p is the number of parameters to be estimated, and where n is the number of equations. For the two-factor model, $\chi^2(10)$ was estimated to be 18.29, which falls within the 90 percent confidence region; this indicates that the null hypothesis of a two-factor model, together with effects from observed variables, cannot be rejected at the 90 percent level in explaining the observed relationships between health, cognitive development, and economic variables. As a comparison, the χ^2-statistic for the zero-factor model was over 100, indicating that reliance on observable variables alone is not sufficient to satisfactorily explain the covariance structure of these variables.

It should be noted that a similar one-factor model was also estimated; the estimates, which are not reported here, were very close to those of the two-factor model, and indeed many of the standard errors were reduced. The χ^2-statistic for this model (with 14 d.f.) was estimated at 29.94, which does not fall within the 99 percent confidence region. In any case, it suggests that the two-factor estimates, and particularly those which relate observed variables, are fairly robust with respect to specification of the factor structure.

IV. Conclusions and Implications

This analysis has shown that the introduction of unobserved factors into an empirical model explaining chaildren's health and cognitive development adds considerably more explanatory power to such models. More notably, some findings from previous empirical models are confirmed and strengthened, and other findings rejected. The significant results from the policy-makers point of view center on the effects of parents' education and family income. Parents' education is a significant positive contributor to cognitive development and oral health. No significant relationship was found between parents' education and diagnosed abnormalities. Contrary to other studies, income was found to play an insignificant role in both the health and cognitive development processes.

The implication is that an incomes' policy will not have a net impact on children's development. What cannot be distinguished here, of course, is whether increased income has both beneficial and detrimental effects which may be offsetting. As such, an incomes' policy combined with other programs may yield net benefits.

While the model considered here is limited in its consideration of explanatory and explained variables, and while the statistical methodology is second-best in that it makes imperfect normality assumptions, the general techniques seem worthy of future investigation with other data and other measures of health.

Footnotes

1. Grossman (1976) surveys the relationships among earnings, schooling, health, and measured intelligence of adults and children that have been extensively discussed in recent literature.

2. See, for example, Becker and Tomes (1976) and Inman (1976).

3. See Kamin (1974) for a survey of the issues and arguments that relate to the inheritance of mental faculties.

4. A more detailed discussion of the hypothetical links between health and cognitive development can be found in Shakotko, Edwards, and Grossman (1980).

5. This model is a member of the class of structural equation models proposed by Joreskog (1973), who discusses more general forms of such models and aspects of the identification problem. A computer program to estimate models of this sort, and which was used for this analysis, is described by Joreskog and Sorbom (1978).

6. Low values of APERI indicate better oral health.

7. A dummy variable was used instead of a continuous measure of birth weight since it is likely that the relationship between birth weight and poor subsequent health is non-linear.

8. Of course, the dichotomous nature of ABN precludes measurement of the severity of these conditions. A more rigorous specification of health would use a combination of ABN and a severity index or else an indicator vector of different health conditions.

9. Quite aside from any considerations in this model, a theory of matching would predict that spouses' education levels are positively correlated. To purge this effect from the model, a non-zero covariance between the stochastic effects in the education equations was permitted. An estimate of this covariance is reported in Table 4.

10. This is in essence the issue at stake in the analysis of state dependence that has been used by Heckman (1978b) to study the time profile of labor force participation.

11. This independence might not be preserved if individuals "learned" how to take and score well on tests in the process of re-testing.

References

Becker, Gary S., and Tomes, Nigel, 1976. "Child Endowments and the Quantity and Quality of Children." Journal of Political Economy 84 (4) : S143-S162.

Chamberlain, Gary, 1977. "Education, Income, and Ability Revisited." Journal of Econometrics 5 (2) : 241-257.

Chamberlain, Gary, 1978. "Omitted Variable Bias in Panel Data: Estimating the Returns to Schooling." In The Econometrics of Panel Data, Annales de l'insee 30-31 : 49-82.

Edwards, Linda N., and Grossman, Michael, 1978. "Children's Health and the Family." National Bureau of Economic Research, Working Paper No. 256.

Grossman, Michael, 1972. The Demand for Health: A Theoretical and Empirical Investigation. New York: Columbia University Press for the National Bureau of Economic Research.

Grossman, Michael, 1976. "The Correlation Between Health and Schooling." In Household Production and Consumption, edited by N. E. Terleckyj. New York: Columbia University Press for the National Bureau of Economic Research.

Grossman, Michael, and Benham, Lee, 1974. "Health, Hours, and Wages." In The Economics of Health and Medical Care, edited by Mark Perlman. London: Macmillan.

Heckman, James J., 1978a. "Dummy Endogenous Variables in a Simultaneous Equation System." Econometrica 46 (4) : 931-959.

Heckman, James J., 1978b. "New Evidence on the Dynamics of Female Labor Supply." University of Chicago, mimeo.

Inman, Robert P., 1976. "The Family Provision of Children's Health: An Economic Analysis." In The Role of Health Insurance in the Health Services Sector, edited by R. Rosett. New York: Columbia University Press for the National Bureau of Economic Research.

Joreskog, K., 1973. "A General Method for Estimating a Linear Structural Equation System." In Structural Equation Models in the Social Sciences, edited by A. S. Goldberger and O. D. Duncan. New York: Seminar Press.

Joreskog, K., and Sorbom, D., 1978. LISREL-IV. Estimation of Linear Structural Equation Systems by Maximum Likelihood Methods: A FORTRAN-IV Program. Chicago: International Educational Services.

Kamin, Leon J., 1974. The Science and Politics of IQ. New York: John Wiley and Sons.

Lee, Lung-Fei, 1979. "Health and Wage: A Simultaneous Equation Approach." University of Minnesota, mimeo.

Shakotko, Robert A., 1979. "State Dependence Between Assets and Health." Columbia University, Department of Economics, Discussion Paper No. 78-7901.

Shakotko, Robert A., Edwards, Linda N., and Grossman, Michael, 1980. "An Exploration of the Dynamic Relationship Between Health and Cognitive Development in Childhood." National Bureau of Economic Research, mimeo.