University of Wisconsin-Madison

Institute for Research on Poverty

Discussion Papers

Jacques van der Gaag
Barbara L. Wolfe

ESTIMATING DEMAND FOR MEDICAL CARE: HEALTH AS A CRITICAL FACTOR FOR ADULTS AND CHILDREN

DP #690-82

Estimating Demand for Medical Care: Health as a Critical Factor for Adults and Children

Jacques van der Gaag
The World Bank

Barbara L. Wolfe
University of Wisconsin-Madison

The research reported in this paper was supported in part through funds granted to the Institute for Research on Poverty by the Department of Health and Human Services pursuant to the provisions of Economic Opportunity Act of 1964.

ABSTRACT

In this paper we investigate (1) the measurement of health status for both adults and children, (2) the importance of health status measures in demand analysis for medical care, and (3) attitudinal measures toward health care and their role in demand for medical care. Finally we develop a structural model of demand for medical care, incorporating a number of health status measures as indicators of a comprehensive, one-dimensional health status measure. The model also includes a production-demand relationship which specifies the unobservable health status measure as a function of socioeconomic household and individual characteristics. We find that the choice of health measure in demand analysis should get more attention than is usually the case in the literature, for there are differences in the results of other exogenous variables in response to health status measures. Inclusion of a large set of measures as indicators in a structural model is one way (partially) to deal with this problem.

1. INTRODUCTION

In an era when resources are increasingly directed toward medical care, understanding the factors that influence demand takes on greater importance. Most work in this area (see, for example, Newhouse and Phelps, 1974; Hyman, 1971; Rosett and Huang, 1973; Newhouse, 1981) is directed to understanding the role of personal income and health insurance on demand, with emphasis on the dimensions of insurance. The recent Health Insurance Study conducted by the Rand Corporation focused on measuring responsiveness to various coinsurance rates--the partial payment by the consumer. Other recent work has addressed the value of time, length of wait, and demand for medical care. Equity issues are implicitly or explicitly raised by many of these studies. For example, the Health Insurance Study suggests that persons with low incomes decrease their medical care usage more than higher-income individuals when coinsurance is imposed.

Most medical care is therapeutic rather than preventive;[1] that is, it is for purposes of treating acute and chronic illness. People seek care when they have a health problem. Thus, in the analysis of the demand for medical care, health plays a critical role. It must be included in the demand analysis if we are to get unbiased measures of the role of other factors such as income, insurance, waiting times, and so forth.[2]

The problem is that health is difficult to measure. Generally, one or a number of partial measures—days ill, self-assessed rating of excellent, good, fair, or poor health, functional limitations—is available, but each of these captures only a part of health status and may be influenced by an individual's own expectations. This last characteristic is particularly true for the most commonly used measure, the self-assessment of overall health (see Manning et al., 1981, p. 45, for a review). For example, a handicapped individual may be doing well at the time of the survey and respond that he is in "excellent" or "good" health, yet, from society's point of view, that person might be considered to be in only fair or poor health. It is probably fair to say that the search for an "ideal" health measure is hopeless unless, perhaps, we specify the purpose of this measure in advance. In this paper, we search for a measure of health status suitable to be included in health care demand analysis. An inaccurate measure of health is likely to lead to bias in demand analysis.

The main problem stems from the fact that health itself is influenced by many of the same factors that influence the demand for medical care. Examples of these factors include education (see Edwards and Grossman, 1980; Shakotko, 1980) and income (see Grossman, 1972). Thus, if education--or any other factor exerting these influences--were included in the demand analysis, but health itself were not, the coefficient on education would include the influences of education on health, and health on demand, not simply education on demand. Thus, finding an index (or indexes) to appropriately measure health may be critical to improving analysis of the demand for medical care. It is, however, possible that the omission of health from demand analysis limits us to measuring only the gross effects of variables that influence both demand and health--education and income, for example--but does not create omitted variable bias for the other variables. In this case, demand analysis could proceed without the need for extensive data on health. Analysts would need to realize they are measuring gross effects in certain cases.

Omission of health may, however, result in biased estimates if there is correlation between the omitted term and the additional independent variables.

In this paper we investigate the importance of health as a determinant of the demand for medical care; the influence of "demand-related factors" on health, and the importance of including accurate measure(s) of health in the demand analysis. We do so separately for children and adults--in part because preventive care may play a larger role in medical care demanded for children, and in part because for children the demand is parent-initiated rather than self-initiated.

We begin with a large set of health measures or indicators, some of which may be relevant in the demand for medical care. In Section 2 we explore these measures and attempt to combine them through the use of principal component analyses. In Section 3 we explore the factors that influence health, as measured both by the health factors from the principal component analyses and the separate health measures. In Section 4 we present our demand analyses and explore the differential results as we change the health variables included and alternatively omit health from the analyses. Throughout these sections we also explore the role of attitude on demand for health care--attitude toward quality, convenience, and cost of medical care. Finally, in Section 5 we bring these separate explorations together and extend our analysis to a structural model of the demand for medical care. In this, health is treated as a latent variable, the health measures serve as indicators, and we have equations explaining health and utilization all as part of a Multiple Causes-Multiple Indicators (MIMIC). We present our conclusions in Section 6.

2. PRINCIPAL COMPONENT ANALYSES OF HEALTH AND ATTITUDE VARIABLES

The data used in this study were collected in 1975 by the Rochester Community Child Health Survey. A 1% sample of families with children under 18 years of age in Monroe County (where Rochester, New York, is situated) were interviewed in 1975. Observations on 972 adults and 1191 children aged 1-18 within 514 households are used in this work. The data are rich in health information, both in terms of health status and medical care utilization, and in information regarding the households' attitudes towards the seeking of professional care. To the data we have added provider availability, by matching resident location to physician location and calculating the travel distance to hospitals and health centers. For more detail on the data see Wolfe (1980).

Health Factors

As stated earlier, the data contain many health measures. For everyone in the sample, we observe a subjective evaluation of health (HSTAT) given by the respondent, the presence of a handicap (HCAP), whether the individual's activity is limited in any way (LIM), and over the past year the number of days ill (DAYS ILL), the number of days in bed (DAYS BED), and whether or not the family member has been ill (ILL). Table 1 presents the relative frequency distribution of these measures for adults.

Table 1. Proxy Measures of Health Status for Adults

N = 972

Relative Frequency Distribution of HSTAT				Percentage Adults with			Average Number	
1 excellent	2 fair	3 good	4 poor	HCAP	LIM	ILL	DAYS ILL	DAYS BED
48.56	40.33	9.16	1.95	5.04	8.44	59.88	8.63	4.49

We see that although almost half the sample is rated as having excellent health, nearly 60% have been ill during the year. Persons report 4.49 days in bed on average, a figure slightly below that reported on the Health Interview Survey (HIS) of the civilian noninstitutionalized population of the United States for 1980. More days ill than days in bed are reported, as expected. (No comparable data are available in the HIS.) The percentage who report handicaps is similar to those reporting "with limitation in major activity" (5.2%) in the HIS, while the percentage reporting LIM is similar to those reporting "with activity limitations" (8.6%) in the HIS survey.[3]

While it is clear from these data that most adults are relatively healthy, it is not an easy matter to decide which health variables best represent the health status of the adults. Some variables contain overlapping information (DAYS BED, DAYS ILL) while other variables seem to convey conflicting information (HSTAT, ILL).

For children, the picture is even more complicated. First we observe the same health information as for adults. Table 2 presents the relative frequency distribution of these health measures for children.

Table 2. Proxy Measures of Health Status for Children

N = 1191

Relative Frequency Distribution of HSTAT				Percentage Children with			Average Number	
1 excellent	2 good	3 fair	4 poor	HCAP	LIM	ILL	DAYS ILL	DAYS BED
60.03	34.34	4.79	0.84	2.10	5.45	77.29	5.63	2.67

Again, over half the sample is reported to be in excellent health, yet over 75% have been ill during the year. In terms of comparisons to national statistics, these children have fewer days ill and days in bed than those reported in HIS. The handicap percentages are closer: 2.0 for HIS, 2.1 for this sample; 3.8 for activity limitation for HIS, 5.45 for this sample.

In Table 3 we display the incidence of seventeen specific health distortions for children. Taken one by one, the data merely provide frequencies; e.g., nearly a quarter of the sample have allergies other than asthma or hay fever. It is likely, however, that there is overlap; we can expect that the information on breathing problems (24.22%) at least partly contains the same information as that on certain allergies.

Table 3. The Prevalance of Seventeen Health Distortions in the Subsample of Children (percentage)

N = 1191

Asthma	5.30	Seeing	2.10	Diabetes	0.17
Hay fever	8.92	Speaking	7.15	Behavior	7.74
Other Allergy	23.97	Arthritis	0.34	Learning	8.07
Kidneys	1.77	Bronchitis	8.75	Breathing	24.22
Heart	4.71	Epilepsy	1.93	Nose	40.96
Hearing	4.96	Cerebral palsy	0.25		

The main purpose of this paper is to assess which aspects of health are relevant to health care utilization. This task would be greatly simplified if the large number of health measures available could be reduced to a smaller set of independent variables. We construct such a set of health factors by calculating the principal components of the correlation matrices of the health measures.

Tables 4 and 5 show the rotated factor matrices for the health measures of adults and of children.[4] In Table 4 we see that the five variables for adults reduce to two independent components: factor 1, with large loadings on the handicap measures (HANDICAP), and factor 2, with large loadings on those measures usually related to acute illnesses (ACUTE). The total variance explained by both factors is 73%.

Table 4. Rotated Factor Matrix for Adult Health Measures (Varimax Rotation)[a]

	Factor 1 (HANDICAP)	Factor 2 (ACUTE)
HCAP	.958	.032
LIM	.959	.013
ILL	-.046	.512
DAYS ILL	.066	.885
DAYS BED	.074	.864

[a]Only factors with an eigen value exceeding 1.00 are shown.

In Table 5, for children, we find approximately the same two factors as for adults. Factor 1 correlates highly with handicap measures (including cerebral palsy) and factor 3 correlates highly with the acute

Table 5. Rotated Factor Matrix for Children's Health Measures

(Varimax Rotation)[a]

	Factor 1 (HANDICAP)	Factor 2 (RESPIRATORY)	Factor 3 (ACUTE)	Factor 4 (BEHAVIOR)
HCAP	.919	.021	.000	.078
LIM	.910	.017	.011	.112
ILL	.019	-.039	.634	-.002
DAYS ILL	.005	.028	.788	.027
DAYS BED	-.005	.076	.813	.018
ASTHMA	-.003	.520	-.063	.061
HAY FEVER	.063	.700	-.045	-.027
OTHALLGY	.031	.593	.151	-.036
KIDNEYS	-.014	-.033	.035	-.076
HEART	-.023	-.064	-.036	.151
HEARING	.045	.043	-.018	.142
SEEING	.066	-.032	.036	.061
SPEAKING	.058	-.013	.065	.500
ARTH	.098	.004	-.034	-.111
BRONCH	-.074	.256	.186	-.004
EPILEPSY	.190	-.036	-.016	.504
CERPALSY	.585	.072	.014	.134
DIABETES	.022	-.045	-.003	-.050
BEHAVIOR	-.026	.105	-.045	.629
LEARNING	.110	-.062	.047	.688
BREATHING	.002	.534	.070	.059
NOSE	.009	.724	-.030	-.041

[a]There were eight factors with an eigen value exceeding 1.00. Only the first four are shown.

illness proxies (ILL, DAYS ILL, DAYS BED). The other two factors presented in Table 5 are also easily interpretable. Factor 2 scores high on all measures of respiratory diseases (RESPIRATORY), while factor 4 relates to diseases with a large behavioral content (BEHAVIOR). The four factors contain almost 35% of the total variation of the 22 original health measures.

We will use the two factors obtained for adults and the four factors for children in the analyses that follow, and interpret them as suggested above.

Attitude or Taste Factors

A large number of measures of attitudes toward seeking medical care are available in the data. The attitude part of the survey includes questions on the importance of having guaranteed access to a doctor (Guaranteed Access) and the importance of having convenient office hours (Convenient Hours). The replies take on values from 1 = very important to 3 = not important. Another set of questions relate to the attention received while seeing a doctor, including: does the doctor spend enough time with you (MD Time)? The responses range from 1 = not enough time to 3 = enough time. Finally, we have questions rating the health care received (Quality of Care). The responses range from 1 = excellent to 4 = poor. Table 6 presents the relative frequencies of these attitude variables.

Again, there is so much information that it is nearly impossible to characterize attitudes. Guaranteed access to a doctor seems very important, but comprehensive services do not. People believe M.D.s do not listen enough but do give enough time, and so on. Many of these measures

Table 6. Average Values of Attitude Variables for Families

Importance of				Medical Attention		Quality of Care	
1. Guaranteed Access	1.08	7. Reasonable Fees	1.38	13. MD Careful	1.16	18. Quality of Care	1.43
2. Convenient Hours	1.47	8. Fast Appointments	1.35	14. MD Concerned	1.34	19. Satisfied	1.13
3. Convenient Location	1.70	9. Short Office Wait	1.56	15. MD Listens	1.16	20. Relative Care	1.50
4. Recommended by Friend	2.25	10. Friendliness of Staff	1.47	16. MD Time	2.78	21. Find MD	2.38
5. 24 Hr Emerg. Care	1.23	11. Type of Patients	2.81	17. MD Info	2.54		
6. Comprehensive Services	1.92	12. All See 1 MD	2.07				
Values:	1 (very important) to 3 (not important)			Values:	1 (not enough) to 3 (enough)	Values:	1 (excellent) to 4 (poor)

no doubt overlap and represent the same underlying concerns. In order to gain insight into these concerns, we calculated the principal components of this set of attitude variables. The results are presented in Table 7. We see that the variables reduce to 3 independent components. Factor 1 has high loadings on factors related to M.D. or medical attention and to other factors generally related to the quality of care. Factor 2 has heavy loadings on cost factors with an emphasis on the cost of time. Factor 3 stresses convenience, showing heavy loadings on a convenient location, one M.D. for the family, and comprehensive services. In the analysis to follow we use these three generally interpretable attitude factors to represent family tastes toward medical care.

We thus have assembled a unique data set which includes socioeconomic, individual, and family characteristics, data on health care utilization, matched availability data, and constructed independent factors to measure health characteristics and attitudes.

3. WHAT FACTORS AFFECT HEALTH AND ATTITUDE?

As pointed out in the Introduction, it is likely that some or all of the health and attitude measures are systematically related to a number of socioeconomic variables that enter the demand equations. Thus, if the demand equations are estimated without the health and attitude variables, some of the coefficients obtained are likely to be biased. On the other hand, if the health measures are included, the coefficients of the socioeconomic variables show only partial effects on health care utilization, and should be interpreted as such.

Table 7

Rotated Factor Matrix for Attitude Variables (Varimax Rotation)[a]

Variable	Factor 1 (Quality)	Factor 2 (Cost)	Factor 3 (Convenience)
Guaranteed Access	-.096	.071	-.208
Convenient Hours	.103	.516	.079
Convenient Location	-.033	.329	.443
Recommended by Friend	.033	.035	.118
24 Hr Emerg. Care	.051	.039	.214
Comprehensive Services	-.018	.210	.658
Reasonable Fees	-.007	.579	.268
Fast Appointments	-.029	.608	-.016
Short Office Wait	-.018	.756	.011
Friendliness of Staff	.150	.542	.022
Type of Patients	-.129	.101	-.018
All See 1 MD	-.045	-.033	.739
MD Careful	.698	.062	.013
MD Concerned	.744	.099	.077
MD Listens	.785	-.031	.031
MD Time	-.702	.089	-.025
MD Info	-.652	-.062	.074
Quality of Care	.722	-.019	-.159
Satisfied	.579	.039	-.032
Relative Care	.187	-.011	-.179
Find MD	.068	.006	.058

[a]There are six factors with eigen values exceeding 1.00. Only 3, those easily interpretable, are shown.

The magnitude of this potential problem is an empirical question that often is not addressed due to lack of data. In this section, we will assess to what extent health (H) and attitude (T) variables are systematically related to various socioeconomic variables. We will estimate equations of the form

H = H (individual characteristics, family characteristics)

T = T (family characteristics)

where H represents a health measure and T represents a taste factor.

In Table 8 we present the health equations for adults. As health measures we use both the two health factors and the six separate health variables. As explanatory variables we include individual characteristics: age, sex (FEMALE), race (NONWHITE), education, employment status (WORKPART, WORKFULL), and occupation (as measured by a commonly used occupation status scale, the Bogue Scale). Family characteristics include marital status (MARRIED), family size (FAMSIZE), family income (FAMINC) and the median income of the census tract where the family lives (MEDINC). The last variable may be viewed as a better proxy for economic status ("permanent income") than annual family income. (The means and standard deviations for these variables are in Appendix A).

Though the R^2 for the HANDICAP factor is low, .036, we find significant coefficients for AGE, FEMALE, EDUCATION, WORKFULL, MEDINC. Note that the effects of WORKFULL and MEDINC may be a case of reversed causation. We do not interpret the equations presented in Table 8 as "health production functions." We merely assess the extent of systematic relationships among health measures and socioeconomic variables.

Table 8. Determinants of Adults' Health

	Factor 1 (HANDICAP)	Factor 2 (ACUTE)	HSTAT (excellent=1, poor=4)	HCAP	LIM	DAYS ILL	DAYS BED	ILL
AGE	.011 (2.79)	-.003 (.71)	.012 (4.01)	.002 (2.08)	.005 (3.32)	.009 (.08)	- .018 (.29)	-.003 (1.33)
FEMALE	-.179 (2.79)	.156 (1.56)	.004 (.06)	-.028 (1.18)	-.088 (2.19)	2.29 (.84)	1.51 (.91)	.111 (1.81)
EDUCATION	-.027 (1.83)	.015 (1.06)	-.014 (1.24)	-.005 (1.47)	-.011 (1.92)	.223 (.56)	.018 (.08)	.016 (1.83)
FAMSIZE	-.030 (1.12)	-.001 (.03)	.007 (.35)	-.004 (.71)	-.017 (1.63)	.524 (.73)	- .222 (.51)	-.011 (.67)
NONWHITE	-.077 (.62)	-.082 (.66)	.209 (2.13)	-.030 (1.02)	-.013 (.26)	-2.47 (.72)	.949 (.46)	-.113 (1.49)
MARRIED	-.217 (1.41)	.112 (.73)	-.205 (1.69)	-.032 (.88)	-.105 (1.70)	-1.66 (.39)	-1.73 (.68)	.200 (2.13)
WORKFULL	-.231 (1.80)	.079 (.62)	.019 (.19)	-.041 (1.33)	-.110 (2.13)	-1.94 (.55)	2.11 (.99)	.080 (1.02)
WORKPART	-.128 (.85)	.037 (.24)	.005 (.04)	-.031 (.86)	-.048 (.80)	-2.01 (.48)	2.81 (1.12)	-.013 (.14)
OCCUPATION	.001 (.37)	-.001 (.33)	-.002 (1.12)	.000 (.39)	.000 (.33)	- .007 (.10)	- .026 (.62)	.000 (.03)
FAMINC	-.006 (1.10)	-.012 (2.05)	-.010 (2.12)	-.001 (.97)	-.003 (1.34)	- .373 (2.34)	- .116 (1.21)	-.004 (1.15)
MEDINC	-.025 (1.92)	.016 (1.26)	-.021 (2.04)	-.006 (2.05)	-.009 (1.66)	.349 (.99)	.343 (1.62)	-.001 (.14)
CONSTANT	.886 (2.84)	-.366 (1.18)	1.97 (8.03)	.228 (3.08)	.464 (3.72)	1.83 (.21)	2.42 (.47)	.216 (1.14)
R^2	.051	.016	.103	.016	.063	.017	.012	.027

N = 755

t-statistics in parentheses

With respect to the HANDICAP factor, these relationships seem to be of some importance. But with respect to the ACUTE factor, we find only one significant coefficient: adults in higher income families show a lower score (are "healthier) on the ACUTE factor.

The equation explaining HSTAT shows that older individuals judge themselves to be in relatively poorer health. So do nonwhites. Married adults and adults living in higher income families, on the other hand, give themselves high health scores (i.e., low scores on HSTAT).

The five final columns of Table 8 show slight effects of socio-economic variables on the separate health measures, but the overall picture is mixed. For instance, MEDINC shows a negative coefficient for HCAP and LIM and is not significant for DAYS ILL and ILL, but shows a significant positive coefficient for DAYS BED. All socioeconomic variables included (except occupation), however, show an impact on one or more of the health measures for adults.

In our analysis for children, the variables included in the regression explaining the health measures are similar to those for adults except that more variables are now family variables. MARRIED refers to the marital status of the head of household.

The employment and occupation variables are included for both parents. A few individual variables are also added: a dummy variable which indicates if the child was born while the mother was less than 20 years old (LMAGE) and birth order (BIRTHORD).

Table 9 presents the estimation results for the four health factors and for HSTAT. We also regressed all individual health measures against the socioeconomic variables, but the estimates did not yield any additional information. We therefore do not present these results.

Table 9

Determinants of Children's Health

	Factor 1 (HANDICAP)	Factor 2 (RESPIRATORY)	Factor 3 (ACUTE)	Factor 4 (BEHAVIOR)	HSTAT
AGE	.001 (.07)	.006 (.81)	-.010 (1.19)	-.003 (.29)	- .011 (2.22)
FEMALE	-.076 (1.14)	-.057 (1.00)	-.024 (.38)	-.287 (4.42)	.013 (.35)
FAMINC	.015 (2.09)	-.003 (.44)	-.004 (.51)	.002 (.34)	- .005 (1.35
MEDINC	.012 (.86)	-.003 (.27)	.012 (.89)	-.002 (.15)	- .019 (2.46)
BIRTHORD	.006 (.13)	-.082 (2.19)	-.007 (.18)	-.086 (2.01)	- .011 (.48)
LMAGE	-.100 (.55)	-.179 (1.15)	-.16 (.09)	-.067 (.38)	- .017 (.18)
FAMSIZE	.018 (.51)	.025 (.82)	-.092 (2.67)	.035 (1.01)	- .003 (.15)
NONWHITE	.060 (.48)	-.029 (.27)	-.143 (1.18)	-.150 (1.22)	.041 (.60)
MARRIED	-.525 (2.50)	.040 (.22)	-.065 (.32)	.190 (.92)	- .003 (.03)
FFULL	.295 (2.07)	.022 (.18)	.001 (.01)	-.131 (.94)	.082 (1.07)
FPART	.208 (.53)	.633 (1.88)	-.461 (1.22)	-.444 (1.15)	.109 (.51)
FOCC	.005 (1.40)	.000 (.14)	-.003 (.81)	-.007 (2.12)	- .002 (1.08)
MFULL	.113 (.71)	.185 (1.34)	-.038 (.25)	-.195 (1.25)	.228 (2.64)
MPART	.039 (.25)	.062 (.46)	.049 (.32)	-.073 (.47)	.223 (2.61)
MOCC	-.002 (.70)	-.001 (.58)	-.000 (.07)	.002 (.85)	- .003 (2.10)
MEDUC	-.046 (2.88)	.027 (1.91)	.033 (2.15)	-.026 (1.67)	- .005 (.53)
CONSTANT	.110 (.40)	-.405 (1.72)	.286 (1.08)	.832 (3.10)	1.93 (13.05)
R^2	.036	.026	.038	.046	.059

N = 999

t-statistics in parentheses

From the estimation results presented in Table 9 we can make some general observations. In addition to the age, sex, and birth order of the child, the variables for family income, family size, and mother's education seem to be related to one or more of the five health measures. But the direction of their effect depends on the particular measure employed. Mother's education, for instance, shows a negative effect on the HANDICAP factor, a positive effect on the RESPIRATORY and ACUTE factor, and a negative effect on the BEHAVIOR factor. No significant relationship between HSTAT and MEDUC is found. Thus, general conclusions like "mother's education has a positive effect on children's health" cannot be drawn from this analysis. The important point is that when health measures are used in the analysis for the demand for health care, one should be aware that these measures are related to the socioeconomic variables that are themselves included as explanatory variables in the demand analyses. Moreover, since some socioeconomic variables usually employed in demand analysis do have a positive (or negative) effect on some health measures, the estimation results of the demand analysis may depend on the choice of the health measure used.

In the equations analyzing the determinants of the attitude or taste factors, we include a similar set of variables. For these, since the unit of observation is the family, all variables are family variables. They include both parents' labor force participation and occupation, age of the head, whether or not they own the home in which they reside, race, marital status, family size, and family and tract median income measures. As can be seen in Table 10, for only one factor, CONVENIENCE, do these variables have much impact. For this factor, family income, mother's

Table 10

Equations "Explaining" Taste Factors (Households as Unit of Observation)

	Factor 1 "Quality"	Factor 2 "Cost"	Factor 3 "Convenience"
FAMINC (10,000's)	.107 (1.11)	.115 (1.20)	.206 (2.28)
MEDINC (10,000's)	-.013 (.061)	-.284 (1.48)	.197 (1.09)
FAMSIZE	.010 (.274)	-.008 (.201)	-.015 (.420)
NONWHITE	-.081 (.448)	-.312 (1.75)	-.406 (2.41)
MARRIED	-.232 (.809)	-.277 (.977)	-.327 (1.22)
FFULL	.216 (1.09)	.017 (.089)	-.068 (.371)
FPART	.203 (.375)	-.644 (1.21)	-.738 (1.47)
FOCC	-.002 (.373)	.005 (1.06)	.003 (.731)
MFULL	-.179 (.865)	-.172 (.845)	.024 (.126)
MPART	-.107 (.528)	-.089 (.445)	.101 (.532)
MOCC	.001 (.182)	-.001 (.185)	.001 (.395)
MEDUC	.009 (.426)	.026 (1.24)	.040 (2.04)
AGE HEAD	-.006 (1.05)	-.008 (1.51)	-.021 (4.05)
OWN HOME	.027 (.180)	-.164 (1.09)	.267 (1.88)
CONSTANT	.054 (.136)	.447 (1.14)	-.278 (.751)
R^2	.01	.04	.15

N = 514

education, and homeownership all have positive effects, while race (being nonwhite) and age of head both have negative effects.

In order not to complicate the analysis too much, we will, in the next section, always include these taste factors in the demand analysis. Thus, we should bear in mind that if a significant impact of one of the taste factors (especially factor 3) on utilization is found, the coefficients for income, race, and mother's education show only partial effects, "holding taste constant."

In general, the analysis of health-care utilization is hampered by the fact that no generally acceptable unidimensional health measure exists. As shown above, principal component analyses or factor analytical techniques can successfully be employed to reduce the sometimes large number of correlated measures into a smaller set of independent ones. But this approach is quite mechanical and still does not yield one unidimensional measure.

It is probably fair to say that one unidimensional measure of health status, representing all facets of health, and usable for a variety of purposes, simply does not exist. However, in Section 5 we will show how a single, comprehensive health measure can be obtained, once the purpose of that measure is specified. But first we will present an analysis of the demand for medical care, including taste factors and using the health factors derived in the previous section as proxy measures for health. We will provide comparisons with results obtained when the longer list of the health measures is used and when the health measures are completely omitted.

4. HEALTH CARE UTILIZATION

For our demand analysis, in addition to modeling the determinants of the total number of provider visits, we distinguish four categories: visits to emergency rooms (HOSPERVS), visits to hospital outpatient clinics (HOSPOPVS), visits to health centers or clinics (HCORCLVS) and physician visits at office or home (OFFHMVS). The explanatory variables include family variables such as income, race, marital status, attitude, insurance coverage by type and family size. For adults they also include labor force participation, age, sex, and health variables. For children they include age, sex, family characteristics, and health variables. Finally, availability is measured by the distance to the nearest hospital (HOSP), the distance to the nearest HMO[5] or non-HMO clinic (HMO, XHMO) and the number of doctors per population (ALL indicates all physicians, for adults; GPPED indicates general practitioners and pediatricians, for children) is also included.

Table 11 presents the adult health care equations for adults. The health variables included are the two health factors constructed in Section 2, plus the subjective measure HSTAT.

With respect to the health measures, we find that a high score on the HANDICAP factor (FACT1) is only significant with respect to hospital outpatient visits. The ACUTE health factor (FACT2) shows a significant effect on all but one of the measures of health care utilization. Visits to a health clinic or health center are the exception. The subjective health evaluation measure, HSTAT, seems to be a strong predictor for health care utilization, except for hospital outpatient care.

Table 11

Health Care Utilization Equations for Adults

	HOSPERVS	HOSPOPVS	HCORCLVS	OFFHMVS	TOTAL
Constant	.333 (2.23)**	.440 (1.66)*	.965 (2.70)**	-1.082 (1.52)	.591 (.72)
GT55	-.014 (.13)	-.319 (1.53)	.534 (1.92)*	.270 (.48)	.410 (.64)
FEMALE	-.020 (.55)	-.011 (.16)	.033 (.36)	.995 (5.53)**	.992 (4.70)
FAMSIZE	-.011 (1.02)	-.002 (.12)	-.038 (1.42)	-.145 (2.70)**	-.189 (3.09)
NONWHITE	-.095 (1.86)**	.199 (2.08)**	.364 (2.80)**	-.498 (1.94)**	-.040 (.13)
MARRIED	.003 (.04)	-.200 (1.61)*	.304 (1.82)*	.187 (.55)	.251 (.65)
FAMINC	-.000 (.05)	-.005 (1.09)	-.006 (1.12)	.025 (2.22)**	.014 (1.08)
MEDINC	-.000 (.09)	.014 (1.43)	-.006 (.47)	.059 (2.30)**	.079 (2.61)
ATTIT1	-.006 (.30)	-.072 (1.81)*	-.009 (.17)	-.122 (1.12)	-.204 (1.66)
ATTIT2	-.021 (1.14)	-.021 (.59)	-.077 (1.64)*	-.019 (.21)	-.129 (1.20)
ATTIT3	-.003 (.20)	-.050 (1.69)*	-.049 (1.24)	.060 (.75)	-.053 (.59)
MCAID	.173 (1.94)**	.313 (1.88)*	.907 (4.07)**	-.152 (.34)	1.231 (2.40)
PRIVINS	-.209 (2.68)**	-.067 (.46)	-.658 (3.39)**	.475 (1.21)	-.447 (1.01)
HMOINS	-.026 (.52)	.059 (.64)	.870 (6.99)**	-.128 (.51)	.785 (2.75)
WORKFULL	.019 (.50)	-.126 (1.77)*	-.030 (.31)	.013 (.07)	-.130 (.60)
WORKPART	-.046 (1.10)	.015 (.20)	-.202 (1.93)*	-.166 (.79)	-.399 (1.67)
HOSP	-.000 (.11)	.003 (.69)			-.015 (1.03)
ALL	.370 (.09)			-.468 (.23)	-.008 (.30)
HMO			-.003 (.63)		.013 (1.31)
XHMO			-.004 (.75)		-.011 (.84)
FACT1	.001 (.05)	.185 (6.60)**	.030 (.80)	.071 (.94)	.200 (3.37)
FACT2	.069 (4.54)**	.223 (7.90)**	.052 (1.38)	.757 (9.92)**	1.093 (12.63)
HSTAT	.052 (2.64)**	.008 (.21)	.121 (2.47)**	.649 (6.57)**	.845 (7.5)
R^2	.070	.174	.220	.272	.355

*Significant at 10% level.

**Significant at 5% level.

All three measures are important in explaining the total number of visits. Clearly the HSTAT measure contains information that is not contained in the two more objective health factors.

With respect to the other explanatory variables, we find only a few significant coefficients for HOSPERVS. Individuals with Medicaid insurance have more visits to a hospital emergency room than privately insured individuals. There are also slight racial differences. The overall explanatory power of this equation is low, $R^2 = .070$.

For HOSPOPVS we find significant racial differences: nonwhites seek care more often in a hospital outpatient clinic than whites. Individuals with Medicaid coverage also show more visits to an outpatient clinic. Being married and being employed full-time reduces the number of these visits. We note, finally, that high scores on the "Quality" and "Convenience" (ATTIT1, ATTIT3) factors show a negative impact on HOSPOPVS. Apparently this type of health service does not stand in high esteem for the quality conscious.

Our regression results explain 22% of the variation in HCORCLVS and 27% of the variation in OFFHMVS. Nonwhites with Medicaid coverage or HMO insurance show a relatively high number of visits to health centers or clinics. Whites from high income families and "rich" neighborhoods, and with private health insurance show more visits to the physician's private office.

These racial and income-related differences are less pronounced for the total number of visits. The variables NONWHITE and FAMINC show no significant effect, but median income in the neighborhood is positively related to overall utilization. Adults with Medicaid coverage or HMO insurance show a higher number of visits than do the privately insured. The total number of visits of adults scoring high on ATTIT1 ("Quality") is slightly below average, but the other attitude factors show no effect.

We also find two familiar results: women show higher utilization rates than men, and individuals living in large families show a lower number of visits than members of small families. We finally note that our availability measures do not show any significant impact on utilization. The measurement errors inherent in the way we constructed these variables might have caused this result. Or the differences in availability in the relatively small area from which we obtained the data are simply so small that no effect on utilization can be observed.

The above results appear to be somewhat sensitive to the use of alternative variables "to control for health." Table 12 gives some selected regression results for the case where no health variables are included (column 1), only the two health factors (column 2), only HSTAT (column 3) and, finally, in column 4, the two health factors plus HSTAT (as in Table 11). The regression coefficients of the variables not included in the table appear to be not sensitive to the changes in health variables.

From Table 12 we learn that it does matter whether or not one controls for differences in health status. For instance, no income effect and no significant racial differences are measured for OFFHMVS if no health variables are included, but both variables show a significant effect in column 4, when the two health factors and HSTAT are added to the equation.[6]

The choice of the health variables is also relevant. If only HSTAT is included, we find no significant racial differences for HOSPOPVS but a

Table 12

Selected Regression Results for Adults, Using Various Health Measures

		(1) No Health Measures	(2) 2 Health Factors	(3) HSTAT	(4) 2 Health Factors + HSTAT
HOSPERVS	FAMINC	-.001 (.55)	-.000 (.12)	-.001 (.35)	-.000 (.05)
	NONWHITE	-.092 (1.76)*	-.085 (1.66)*	-.103 (2.00)**	-.095 (1.86)*
	ATTIT1	.002 (.10)	-.001 (.03)	-.007 (.32)	-.006 (.30)
	ATTIT2	-.024 (1.27)	-.025 (1.33)	-.020 (1.06)	-.021 (1.14)
	ATTIT3	-.010 (.64)	-.009 (.60)	-.001 (.07)	-.003 (.20)
HOSPOPVS	FAMINC	-.009 (.39)	-.005 (1.10)	-.008 (1.73)*	-.005 (1.09)
	NONWHITE	.157 (1.54)	.200 (2.10)**	.138 (1.36)	.199 (2.08)**
	ATTIT1	-.048 (1.14)	-.071 (1.80)*	-.063 (1.50)	-.072 (1.81)*
	ATTIT2	-.016 (.43)	-.021 (.60)	-.010 (.25)	-.021 (.59)
	ATTIT3	-.057 (1.82)*	-.051 (1.75)*	-.042 (1.34)	-.050 (1.69)*
HCORCLVS	FAMINC	-.008 (1.41)	-.007 (1.19)	-.007 (1.24)	-.006 (1.12)
	NONWHITE	.377 (2.90)**	.339 (2.99)**	.352 (2.72)**	.364 (2.80)**
	ATTIT1	.010 (.19)	.004 (.07)	-.008 (.15)	-.009 (.17)
	ATTIT2	-.084 (1.78)*	-.085 (1.81)*	-.075 (1.60)	-.077 (1.64)*
	ATTIT3	-.064 (1.63)*	-.063 (1.59)	-.048 (1.20)	-.049 (1.24)
OFFHMVS	FAMINC	.011 (.92)	.023 (1.97)**	.017 (1.44)	.025 (2.22)**
	NONWHITE	-.453 (1.58)	-.373 (1.41)	-.597 (2.19)**	-.498 (2.30)**
	ATTIT1	-.015 (.12)	-.051 (.46)	-.126 (1.10)	-.122 (1.12)
	ATTIT2	-.055 (.52)	-.059 (.61)	-.003 (.03)	-.019 (.21)
	ATTIT3	-.030 (.34)	-.019 (.23)	.082 (.97)	.060 (.75)
TOTAL	FAMINC	-.007 (.44)	.011 (.82)	.001 (.10)	.014 (1.08)
	NONWHITE	.012 (.03)	.132 (.43)	-.203 (.62)	-.040 (.13)
	ATTIT1	-.049 (.34)	-.115 (.91)	-.201 (1.47)	-.204 (1.66)*
	ATTIT2	-.174 (1.37)	-.184 (1.65)*	-.097 (.82)	-.129 (1.20)
	ATTIT3	-.165 (1.55)	-.147 (1.58)	-.022 (.22)	-.053 (.59)

*Significant at 10% level.

**Significant at 5% level.

significant income effect. If only the two health factors are included, we find just the opposite. A similar type of reversal--although in the opposite direction--appears for OFFHMVS.

The effect of the attitude variables is also sensitive to whether or not health variables are included. The results suggest, not surprisingly, that one's attitude toward seeking professional medical care is not independent of one's health status.

The estimation results for children are presented in Table 13. The four health factors (HANDICAP, RESPIRATORY, ACUTE, and BEHAVIOR) are included in the regressions, together with HSTAT.

The HANDICAP factor (FACT1) does not show any significant impact on utilization, while the BEHAVIOR factor (FACT4) shows a positive effect on hospital outpatient visits only. The other two health factors show the expected positive impact on utilization almost everywhere.

As was the case for adults, HSTAT seems to contain information about the children's health that is not contained in the four health factors included in the regression. With the exception of HOSPERVS, HSTAT is significantly positively related to all forms of health-care utilization.

A similar result was obtained for adults--i.e., HSTAT contains information not included in the other health variables. Given the large amount of other health information contained in the health factors included (especially for children), these results are surprising. In fact, they cast serious doubt on the use of HSTAT in health-care utilization equations, unless HSTAT is collected at the beginning of the period under investigation. Otherwise there is the obvious danger that a score on the HSTAT scale is influenced by previous health care utilization patterns. This seems to be the case here. Manning et al. (1981) show the

Table 13

Health Care Utilization Equations for Children

	HOSPERVS	HOSPOPVS	HCORCLVS	OFFHMVS	TOTAL
CONSTANT	.111 (.87)	-.312 (1.44)	1.081 (3.70)**	-.795 (1.52)	.121 (.20)
LT6	.049 (1.52)	.056 (1.00)	.193 (2.50)**	.824 (6.16)**	1.130 (7.24)**
12-17	.063 (2.04)**	.018 (.33)	-.051 (.72)	.150 (1.18)	.184 (1.23
FEMALE	-.067 (2.61)**	.072 (1.61)*	.078 (1.31)	.050 (.47)	.140 (1.12
FAMSIZE	.013 (1.26)	.017 (.96)	-.046 (1.95)**	-.135 (3.20)**	-.149 (3.00)**
NONWHITE	.042 (.83)	.370 (4.24)**	.251 (2.10)**	-.593 (2.82)**	.087 (.35)
MARRIED	-.113 (2.05)**	.016 (.168)	.154 (1.19)	.217 (.94)	.236 (.87)
FAMINC	-.002 (.68)	.001 (.13)	.005 (.83)	.004 (.36)	.008 (.66)
MEDINC	.001 (.18)	.005 (.57)	-.028 (2.18)**	.078 (3.44)**	.059 (2.16)**
ATTIT1	-.006 (.24)	.013 (.33)	.065 (1.22)	-.210 (2.18)**	-.128 (1.14
ATTIT2	-.021 (1.06)	.016 (.47)	-.085 (1.85)*	.082 (.99)	-.011 (.12)
ATTIT3	.005 (.29)	.004 (.13)	-.084 (2.07)**	.257 (3.59)**	.172 (2.04)**
MCAID	.157 (1.96)**	.121 (.88)	.405 (2.18)**	-.438 (1.32)	.265 (.68)
PRIVINS	.126 (1.87)*	-.033 (.28)	-.401 (2.57)**	.132 (.47)	-.157 (.48)
HMOINS	-.066 (1.34)	-.054 (.63)	.947 (8.30)*	-.556 (2.72)**	.257 (1.08
MFULL	.013 (.36)	-.047 (.76)	-.075 (.91)	-.088 (.59)	-.198 (1.14
MPART	-.014 (.43)	.088 (1.51)	-.170 (2.17)**	.403 (2.89)**	.301 (1.85)*
HOSP	-.003 (1.45)	-.004 (1.02)			-.022 (1.70)*
GPPED	.157 (.91)			-1.81 (.03)	-.463 (.55)
HMO					
XHMO			.000 (.06)		.001 (.04)
FACT1	-.015 (1.19)	.017 (.82)	.036 (1.27)	.011 (.21)	.051 (.85)
FACT2	.036 (2.49)**	-.036 (1.45)	.117 (3.50)**	.187 (3.14)**	.305 (4.40)**
FACT3	.063 (4.80)**	.087 (3.88)**	.106 (3.51)**	.477 (8.80)**	.738 (11.67)**
FACT4	-.006 (.50)	.052 (2.39)**	.038 (1.27)	-.039 (.73)	.045 (.73)
HSTAT	.012 (.50)	.119 (2.89)**	.180 (3.23)**	.558 (5.59)**	.873 (7.50)**
R^2	.053	.063	.221	.275	.280

*Significant at 10% level.

**Significant at 5% level.

inconsistency of results obtained using a "postdicted" HSTAT measure. Our results underscore this problem.

With respect to the other explanatory variables, Medicaid coverage is again one of the variables that shows a significant effect on HOSPERVS. We also observe slight age and sex differences, while children from intact households (i.e., the mother is married) show slightly lower utilization rates. The overall explanatory power is low: $R^2 = .053$.

In addition to slight sex differences, we see a significant, and relatively large, influence of NONWHITE on hospital outpatient visits. But again, we are not very successful in explaining outpatient utilization differences: $R^2 = .063$.

Children living in low-income neighborhoods, who are nonwhite, and who have either Medicaid or HMO insurance show a relatively high number of visits to a health center or clinic. High utilization of private offices is observed for white children with neither HMO nor Medicaid insurance, living in "richer" neighborhoods. The private insurance variable is positive, as expected, but not significant.[7]

The race and insurance-related differences are not observed for total utilization, but children living in high-income neighborhoods show a slightly higher overall utilization rate than their less well-off counterparts.

Children from families who score high on the Convenience scale (ATTIT3) show fewer visits to a health center, but see the private physician more often. A surprising result is that emphasis on Quality (ATTIT1) is negatively related to the number of private physician visits.

We note finally that children whose mothers work part-time have more private visits, less health center visits, and relatively high overall utilization. Children from large families show, as usual, somewhat lower utilization rates.

In Table 14 we show some selected regression results, explaining health care utilization using alternative health-control variables.[8] The results are more stable than for adults.[9] The effect of HSTAT on observed racial differences are not sensitive to the health information included. The positive effect of ATTIT1 on health center visits disappears as soon as some health information is included, but becomes significantly negative if all health information is added to the OFFHMVS equation.

The effect of MPART on utilization seems to be slightly overestimated when the subjective health measure HSTAT is not included in the equations. When HSTAT is included, the positive effect on HOSPOPVS becomes nonsignificant, the negative effect on HCORCLVS increases in absolute value, and the positive effect on OFFHMVS decreases.

The negative effect of FAMSIZE on HCORCLVS and OFFHMVS is more sensitive to the objective health measures, and becomes less pronounced (but remains significantly negative) when these measures are included.

This section, then, confirms once again that health is an important determinant of health care utilization. More important, we show that the inclusion or exclusion of certain health variables affects the coefficients on variables which themselves affect health. Consequently, the casual choice of one or more health measures from an ad hoc list "to control for health," as is often the case in the literature is not without consequences for the measured impact of other variables.
Finally, we note that the use of several health measures makes it hard to generalize about the role of health in determining demand.

Table 14

Selected Regression Results for Children, Using Alternative Health Measures

		(1) No Health Measures	(2) 2 Health Factors	(3) HSTAT	(4) 2 Health Factors + HSTAT
HOSPOPVS	FEMALE	.056 (1.26)	.078 (1.74)*	.054 (1.23)	.072 (1.61)
	FAMSIZE	.008 (.47)	.018 (1.00)	.010 (.55)	.017 (.96)
	NONWHITE	.365 (4.12)**	.378 (4.31)**	.358 (4.07)**	.370 (4.24)**
	MPART	.101 (1.71)*	.096 (1.66)*	.088 (1.51)	.088 (1.51)
	ATTIT1	.039 (1.96)**	.024 (.62)	.019 (.47)	.013 (.33)
	ATTIT2	.014 (.42)	.017 (.50)	.013 (.39)	.016 (.47)
	ATTIT3	.002 (.06)	.002 (.06)	.009 (.29)	.004 (.13)
HCORCLVS	FEMALE	.057 (1.94)*	.087 (1.46)	.053 (.89)	.078 (1.31)
	FAMSIZE	-.062 (2.58)**	-.045 (1.89)*	-.060 (2.51)**	-.046 (1.95)**
	NONWHITE	.242 (1.97)**	.259 (2.15)**	.235 (1.94)**	.251 (2.10)**
	MPART	-.155 (1.94)**	-.156 (1.98)**	-.177 (2.24)**	-.170 (2.17)**
	ATTIT1	.111 (2.04)**	.083 (1.54)	.079 (1.46)	.065 (1.22)
	ATTIT2	-.086 (1.84)*	-.083 (1.80)*	-.088 (1.90)*	-.085 (1.85)*
	ATTIT3	-.082 (1.98)**	0.92 (2.27)**	-.071 (1.75)*	-.084 (2.07)**
OFFHMVS	FEMALE	.044 (.39)	.079 (.72)	.033 (.30)	.050 (.47)
	FAMSIZE	-.194 (4.33)**	-.132 (3.08)**	-.187 (4.31)**	-.135 (3.20)**
	NONWHITE	-.588 (2.62)**	-.557 (2.62)**	-.621 (2.85)**	-.593 (2.82)**
	MPART	.452 (3.02)**	.444 (3.14)**	.393 (2.71)**	.403 (2.89)**
	ATTIT1	-.068 (.66)	-.156 (1.61)	-.162 (1.62)	-.210 (2.18)**
	ATTIT2	.060 (.68)	.086 (1.02)	.056 (.64)	.082 (.99)
	ATTIT3	.286 (3.75)**	.231 (3.19)**	.317 (4.28)**	.257 (3.59)**

*Significant at 10% level.

**Significant at 5% level.

In the previous sections we reduced a large number of health variables to a more manageable set of independent health factors. In the next section we will go one step further, i.e., we will use these factors as health indicators in a structural model for health care demand in which HEALTH is treated as a one-dimensional latent variable.

5. A STRUCTURAL MODEL OF DEMAND FOR HEALTH CARE

In Sections 1-4 we showed the following:

1. By the application of principal component analysis, one can successfully reduce the dimensions of a set of data on health status. For children we were able to reduce a set of 26 variables to four independent factors. These four factors all had a very clear interpretation and explained about one-third of the total variance.
2. Various socioeconomic variables affect health. But the sign and the magnitude of the impact depends on the health measure employed.
3. Because of 2, the choice of proxy measure for health status in the analysis of the demand for medical care does influence the results of the analysis. These results should therefore be interpreted as conditional on the health measures included.

Our analysis did not result in unambiguous statements about the positive or negative effect of family characteristics on health. Nor are we able to say which health measures to include in the analysis of the demand for medical care. Both problems stem from the simple fact that no unidimensional measure of health status exists.

Ideally one would like to estimate a demand equation of the following form:

Demand = D (individual characteristics, family characteristics, availability of medical care, health status)

The individual and family characteristics are already specified in the previous section, as is the availability of medical care. But, instead of using a number of proxy measures, we would like to represent health status by one comprehensive measure.

Likewise, instead of estimating many equations of the form

Health = H (individual characteristics, family characteristics)

where Health is represented by a number of proxy measures (Section 3), we would like to represent Health by the same comprehensive unidimensional measure as used in the demand equations.

This leads to the following model specification:

(1) $H^* = \alpha' x + \varepsilon_1$

(2) $D_i = \beta_{1i}' z + \beta_{2i}' H^* + \varepsilon_{2i}$ $\qquad i = 1,4$

(3) $HP_j = \gamma_j' H^* + \varepsilon_{3j}$ $\qquad j = 1,K$

The first equation resembles the equations specified in Section 3: health is assumed to be a function of a number of socioeconomic variables, x. The dependent variable health, H^*, is unobservable.[10] The equation can be interpreted as either a production function or a demand function of health. In both cases H^* is desired health status. The second set of equations resembles the utilization equations estimated in

Section 4. The demand for medical care is a function of exogenous variables, z, and health, H^*. Thus, the proxy measures of health employed in Section 4 are replaced by one variable: H^*. The vectors x and z may partially overlap.

The model includes K additional equations. These equations state that the proxy measures of health, HP_j, $j=1,K$, are proportional to the overall measure H^*. Thus the probability of an illness increases as health, H^*, decreases. The number of days in bed will decrease if H^* increases, etc. That is, provided the estimation results show the correct signs for the coefficients γ_j. This model, which has the form of a MIMIC model (see Jöreskog and Sörbom, 1978, for more detail), is estimated for adults and children separately. For children we use the four health factors of Section 2 as indicators (HP_j, $j=1,4$). For adults we use the six original health proxies (HP_j, $j=1,6$).

The model has been estimated under the assumption that all disturbance terms are normally distributed.[11] Furthermore, we assume $E(\varepsilon_1, \varepsilon_{2i}) = E(\varepsilon_1, \varepsilon_{3j}) = E(\varepsilon_{2i}, \varepsilon_{3j}) = 0$, $i = 1,4$, $j=1,K$. And $E(\varepsilon_{3j}, \varepsilon_{3l}) = 0$, $j \neq l$. The disturbance terms added to the utilization equations, ε_{2i}, are allowed to be freely correlated with each other.

In order to identify all parameters in the model, we fix the constant term of the HOSPERVS equation to be equal to its value obtained from the regression analyses in the previous section. HEALTH is dimensioned by setting its impact on HCORCLVS equal to -1.0. Thus a one unit increase in health results in one less visit to a health center.

At the bottom of Table 15, we see that a one unit increase of $HEALTH^*$ decreases HSTAT by -2.9, the number of days ill by 1.6, and the

Table 15

Estimation Results of the Structural Model of Demand for Medical Care (Adults)

	HEALTH*	HOSPERVS	HOSPOPVS	HCORCLVS	OFFHMVS
HEALTH*		-.339 (2.90)**	-.946 (2.93)**	-1.00 (-)	-6.16 (3.69)**
HOSP		-.001 (.54)	.003 (.69)		
ALL		-.001 (.19)			-.007 (.35)
HMO				.001 (.29)	
XHMO				-.005 (1.11)	
MCAID		.008 (.82)	.486 (2.49)**	1.667 (6.79)**	-.627 (1.19)
PRIVINS		-.290 (4.37)**	-.077 (.53)	-.106 (.58)	.538 (1.37)
HMOINS		-.016 (.32)	-.068 (.69)	.803 (6.56)**	-.104 (.40)
AT1		-.003 (.28)	-.043 (1.86)*	-.014 (.49)	-.095 (1.53)
AT2		-.011 (.89)	-.002 (.09)	-.052 (1.72)*	.046 (.72)
AT3		.010 (.77)	.057 (2.33)**	-.044 (.14)	.020 (1.30)
NONWHITE	-.057 (1.52)	-.128 (2.54)**	.090 (.91)	.237 (1.88)*	-.790 (2.80)**
55+	-.033 (.47)				
FEMALE	-.066 (2.34)**				
EDUC	.005 (1.31)	.002 (.32)	.004 (.38)	-.014 (1.11)	.004 (.11)
FULL	-.012 (.50)	.025 (.96)	-.117 (2.15)*	.004 (.06)	-.603 (3.87)*
PART	.046 (1.46)	-.048 (1.16)	.030 (.37)	-.151 (1.47)	.132 (.56)
FAMSIZE	-.005 (.62)	-.018 (1.89)*	-.015 (.75)	.032 (1.25)	-.187 (3.23)*
MARRIED	.025 (.63)				
MEDINC	.003 (.10)				
TFAMINC	.005 (2.35)**	-.000 (.06)	-.001 (.21)	.002 (.32)	.053 (4.28)*
Constant	-.653 (3.47)	.274 (-)	-.293 (.16)	.053 (2.36)	-.189 (3.47)

HSTAT	HCAP	LIM	ILL	DAYS ILL	DAYS BED
-2.889 (3.80)**	-.089 (3.35)**	-.144 (3.32)**	-1.029 (3.79)**	-1.605 (3.62)**	-8.376 (3.56)**

- = value fixed.
*Significant at 10% level.
**Significant at 5% level.

number of days in bed by 8.4. Furthermore, it reduces the probability that HCAP = 1 or LIM = 1 by .09 and .14 respectively.

Each column in the top part of Table 15 represents an equation in our model. With respect to HEALTH* we find a significant positive effect of family income. This, of course, is a "summary" of the findings presented in Table 8. HEALTH* in fact can be viewed as a weighted sum of the health indicators used. The HEALTH* equation indicates a negative effect of FEMALE and a slight negative effect of NONWHITE on health. The utilization equations show that, "holding health constant," nonwhites have fewer visits to private offices and hospital emergency rooms, and more visits to health centers.

The other variables included in the health equation show no impact on health. Given the analyses in the previous section, this result is not surprising. As we have seen, various socioeconomic variables have positive, negative, or no effect at all on health, depending on which health measure we use. Consequently, when we obtain a unidimensional health measure, based on the various health measures previously employed, the effect of the socioeconomic variables can be expected to be small at best.

A similar result is obtained for children (see Table 16): FAMINC is the only significant variable in the health equation, apart from the familiar age and sex effects.

HEALTH*, in the model for children, correlates highly with HSTAT, but does not show much relationship with the four health factors. In fact, the coefficient for the second factor (RESPIRATORY) has the "wrong" sign (bottom of Table 16).

Table 16

Estimation Results of the Structural Model of
Demand for Medical Care (children)

	HEALTH*	HOSPERVS	HOSPOPVS	HCORCLVS	OFFHMVS
HEALTH*		-.009 (3.90)**	-.527 (3.50)**	-1.00 (-)	-3.07 (4.82)**
HOSP		-.003 (1.39)	-.003 (.71)		
GPPED		-.018 (1.06)			-.049 (.70)
HMO				-.007 (1.57)	
XHMO				-.002 (.40)	
MCAID		.195 (2.53)**	.119 (.90)	.397 (.22)	-.588 (1.75)*
PRIVINS		.112 (1.68)*	-.005 (.46)	-.382 (2.46)**	.043 (.15)
HMOINS		-.047 (.95)	-.006 (.75)	1.104 (9.05)**	-.618 (2.94)**
AT1		.002 (.16)	.115 (.52)	.144 (1.49)	-.109 (1.98)**
AT2		-.010 (.73)	.000 (.02)	-.073 (2.38)**	.119 (2.08)**
AT3		.001 (.10)	-.004 (.16)	.021 (.63)	-.231 (3.90)**
NWHITE	-.290 (1.42)	.029 (.58)	.218 (1.64)*	.035 (.16)	-1.635 (2.69)**
LT6	-.210 (4.47)**				
12-17	.001 (.04)				
FEMALE	-.047 (1.65)*				
MEDUC	.034 (1.36)	-.004 (.68)	.012 (.74)	.041 (1.57)	.171 (2.33)**
MFULL	-.214 (1.37)	.020 (.54)	.154 (1.55)	-.269 (1.61)	-.642 (1.38)
MPART	-.167 (1.18)	-.017 (.50)	.018 (.20)	-.334 (2.20)**	-.066 (.16)
FAMSIZE	.025 (.66)	-.002 (.16)	.021 (.84)	-.031 (.74)	-.102 (.90)
LMAGE	.080 (1.04)				
MARRIED	-.083 (1.34)				
MEDINC	-.004 (.63)				
FAMINC	.031 (1.90)*	-.002 (.92)	.016 (1.68)*	.027 (1.59)	.091 (1.87)*
Constant	-4.391 (2.08)	.111 (-)	-2.245 (1.88)	-3.711 (1.75)	-12.28 (1.97)

HSTAT	FACT1	FACT2	FACT3	FACT4
-3.81 (2.17)	-.003 (.31)	.017 (1.58)	-.020 (1.58)	-.011 (1.08)

- = value fixed.
*Significant at 10% level.
**Significant at 5% level.

When interpreting both Tables 15 and 16 it is useful to remember the scale we used for the children's and adults' health variables. For both children and adults, we scaled the health measure so that a one unit increase on health results in one less visit to a health care center. The corresponding reduction for hospital outpatient visits is also nearly one less visit for adults, approximately one-half for children, and for hospital emergency room visits is slightly greater than one-third for adults but only approximately one-hundredth for children. These results seem plausible and give some basis for our claim that the unobserved Health* factor can serve as a comprehensive unidimensional health index. However, the results for private office visits (OFFHMVS) imply a very high response to a one unit increase in Health*---3.07 and 6.16 for children and adults respectively.[12] Given the average values of OFFHMVS (approximately 1.6 for both children and adults), these results seem implausible. On the other hand, many of the results look quite reasonable and are consistent with those based on the regression analysis in the previous sections.

For adults, racial differences in utilization patterns cannot be attributed solely to differences in health status. Family size has a significant negative impact on health care utilization, "holding health constant." To the extent that employment status influences health care utilization, the effect is direct, not through health status. Finally, total family income shows a positive effect on health status and on the number of private office visits, again "holding health constant."

For children, we find no significant racial differences in health status but, "holding health constant," we find that nonwhites go more often to hospital outpatient clinics and less often to private physician practices. The effect of mother's education on children's health is positive, but not significant. The direct effect of mother's education on private physician visits is positive and significant. The employment status of the mother shows only direct effects on utilization (i.e., not through health). Family income shows positive effects both on health and utilization.

6. CONCLUSION

It has been common practice to add one or more proxy variables for health in demand equations for medical care "to control for variation in health status." The choice of these proxy variables is almost always guided by the availability of the data.

In this paper we show that this habit is not as innocent as it seems. Health measures should obviously be included in demand equations for medical care. But the choice of the variables representing health will have an impact on the estimation results regarding various socioeconomic variables.

As shown in Section 3, health should be treated as endogenous. But doing this does not solve the problem presented by the fact that not one of the available health measures is by itself a sufficient proxy for health. A variety of proxy measures must therefore be used.

In Section 5 we showed how these proxy measures can be used as indicators for a unidimensional health measure. This measure, which is unobservable, is introduced in a structural model for health care demand. Thus HEALTH becomes a latent variable in a Multiple Causes-Multiple Indicators (MIMIC) model.

As indicators we use the health proxy measures and utilization of health care. The latter can be used as an indicator for health once we adequately control for income, insurance, availability, and taste differences. As causal factors in a health production function, the socioeconomic variables that were correlated with one or more of the health measures analyzed in Section 3 were included.

The results are encouraging, especially for children. The model yields reasonable estimates, as compared to the unrestricted OLS regressions on utilization, and the latent variable HEALTH does have the impact one would expect if it represents a true measure of health status.

Some caveats, however, should be mentioned. We did not solve the question of how to choose among various health-proxy measures. We merely pushed the problem one step back by including one latent variable, HEALTH, in the demand equations and by stating that the proxy measures were proportional to this overall measure. Thus the ex-post interpretation of HEALTH is conditional upon the choice of the health indicators used and the weight they get in the estimation process.

The estimation assumes normality of the disturbances. For adults we use various 0-1 dummy variables as health indicators, which makes the normality assumption less plausible. For children we transformed many discrete health-proxy variables into a small set of continuous factor scores, which is an important improvement over our earlier work. But the problem remains with respect to the health-care utilization data which are bounded from below by zero, and usually take only a few discrete values. This problem seems particularly severe with respect to private office visits. This variable has a large concentration of zeros and (other than HOSPOPVS) correlates strongly with a number of other variables. This might explain our implausible results with respect to OFFHVS, both for adults and children.

Finally, we should mention that part of the model is contructed in an ad hoc manner, with little a priori knowledge and without a firm theoretical base. The utilization module can easily be shown to be derived from a general demand framework. But the "production function of health" should be viewed as a first attempt to show the impact of various socioeconomic variables on a comprehensive measure of health status. The formulation fits within Grossman's theory of the demand for health. But the analyses lack the input of other disciplines, e.g., epidemiology. A further understanding of the causal relationships between, say, income or family size or education and health is needed to improve the specification of the health production function in the model.

NOTES

[1]There are, of course, major exceptions: well-baby care, immunizations, some gynecological care, some screening tests, care during pregnancy.

[2]Manning et al. (1981) take a similar view.

[3]The HIS numbers are for adults aged 17-44 years, both sexes. See U.S. Department of Health and Human Services (1981, p. 24).

[4]In order to assess what information is contained in the variable HSTAT, we performed the principal component analyses with and without this variable. In both cases the same health factors were obtained. HSTAT correlates with two of these factors, HANDICAP and ACUTE. However, as we see in Section 4, HSTAT also contains independent information relevant to the prediction of health care utilization. This information did not show up as an independent factor in the principal component analyses with HSTAT included. In the remainder of this paper we will delete HSTAT from the principle component analyses. But we will treat it as an additional variable to explain utilization in Section 4. This also permits the comparison of our results with other work using HSTAT, such as Colle and Grossman (1978), and Goldman and Grossman (1978).

[5]Health Maintenance Organization. In this type of arrangement consumers pay a fixed amount--a capitation fee--for all services for a specified period of time.

[6]We reestimated the equations of columns 2 and 4, replacing the two health factors with the five original health measures on which they were based. The results were almost the same, showing that the two constructed health factors adequately represent the variation in the five original health measures.

[7]There are three included insurance variables: Medicaid (MCAID), private insurance (PRIVINS) and HMO insurance (HMOINS). The omitted category is "no insurance." The insignificance of PRIVINS may be due to the high correlation between Medicaid and PRIVINS variables (-.726).

[8]The results for HOSPERVS and TOTAL appeared not to be sensitive to alternative health specifications. They are therefore not included in Table 14.

[9]As for adults, we also ran the regressions including all health variables. The results confirmed our results using the four health factors plus HSTAT and are therefore not presented.

[10]In the past couple of years a number of studies have been published using this approach. Work based on microdata includes Van de Ven and van der Gaag (1979) and Lee (1979). The work of Wolfe and van der Gaag (1981) indicates a preliminary version of the model presented here, using only part of the health information. Hooymans and Van de Ven (forthcoming) present a useful discussion on the identification of such a model and the subsequent dimensions (and interpretation of the resulting health index).

[11]This assumption is likely to be violated, given the limited character of some of the endogenous variables. Lee (1979) deals with this problem when deriving the likelihood function of his model. For children, we reduce the problem by replacing the health indicators (usually binary during variables) by the continuous health factor scores. We do not provide a solution, however, for the limited character of the health care utilization data. Comparison of our results with the ones obtained in the previous sections does not suggest that the possible bias

due to the violation of the normality assumption is of major importance. Some oddities in our results, however, do call for caution (see text).

[12]Turning to Table 13, the results for HSTAT also suggest a much larger response among children from a one unit decrease in health on OFFHMVS than HCORCVS, HOSPOPVS or HOSPERVS. The ratio of coefficients (for example, OFFHMVS to HCORCLVS) is similar for HSTAT in DLS and H* in MIMIC. Among adults (see Table 11), the OLS results for HSTAT also follow a similar pattern to the adult MIMIC results.

REFERENCES

Colle, Ann, and Grossman, Michael. 1978. "Determinants of Pediatric Care Utilization." Journal of Human Resources, 13 (Supplement):115-58.

Davis, Karen, and Reynolds, Roger. 1976. "The Impact of Medicare and Medicaid on Access to Medical Care." In R. Rosett, ed., The Role of Health Insurance in the Health Services Sector. New York: National Bureau of Economic Research.

Edwards, Linda, and Grossman, Michael. 1980. "Children's Health and the Family." In R. Scheffler, ed., Annual Series of Research in Health Economics, Vol. 2. Greenwich, Conn.: JAI Press.

Goldman, Fred, and Grossman, Michael. 1978. "The Demand for Pediatric Care: A Hedonic Approach." Journal of Political Economy, 86:259-80.

Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." Journal of Political Economy, 80/2: 223-55.

Hooymans, E.M., and van de Ven, W.P.M.M. Forthcoming. "Implementing a Health Status Index in a Structural Health Care Model." In van der Gaag, J., Neenan, B., and Tsukuhara, T., eds., Economics of Health Care. New York: Praeger.

Hyman, J. 1971. "Empirical Research on the Demand for Health Care." Inquiry, 8;61-71.

Jöreskog, K.G., and Sörbom, D. 1978. Estimation of Linear Structural Equation Systems by Maximum Likelihood Methods. Chicago: International Educator Services.

Lee, L.F. 1979. "Health and Wages: A Simultaneous Equation Model with Multiple Discrete Indicators." Working Paper No. 79-127, Department of Economics, University of Minnesota.

Manning, Willard G., Newhouse, Joseph P., and Ware, John E., Jr. 1981. "The Status of Health in Demand Estimation: Beyond Excellent, Good, Fair and Poor." NBER Conference Paper 86. January.

Newhouse, Joseph P. 1981. "The Demand for Medical Care Services: A Retrospect and Prospect." In J. van der Gaag and M. Perlman, eds., Health, Economics, and Health Economics. Amsterdam: North-Holland.

Newhouse, Joseph P., and Phelps, Charles E. 1974. "Price and Income Elasticities for Medical Care Services." In Mark Perlman, ed., The Economics of Health and Medical Care. New York: John Wiley and Sons.

Robinson, P.M., and Ferrara, M.C. 1977. "The Estimation of a Model for an Unobservable Variable with Endogenous Causes." In Aigner, D.J., and Goldberger, A.S., eds., Latent Variables in Socioeconomic Models. Amsterdam: North-Holland.

Rosett, R., and Huang, L.P. 1973. "The Effect of Health Insurance on the Demand for Health Care." Journal of Political Economy, 81:281-305.

Shakotko, R.A. 1980. "Dynamic Aspects of Children's Health, Intellectual Development, and Family Economic Status." New York: NBER Working Paper No. 451.

U.S. Department of Health and Human Services. 1981. Current Estimates From the National Health Interview Survey: United States, 1980. DHHS Publication No. (PHS) 82-1567. December.

van de Ven, W.P.M.M., and van der Gaag, J. 1979. "Health as an Unobservable: A MIMIC-Model for Health Care Demand." Discussion Paper No. 544-79, Institute for Research on Poverty, University of Wisconsin-Madison.

Wolfe, Barbara L. 1980. "Children's Utilization of Medical Care." Medical Care 18 (December): 1196-1207. (Institute for Research on Poverty Reprint No. 419).

Wolfe, B., and van der Gaag, J. 1981. "A New Health Status Index for Children." In van der Gaag, J., and Perlman, M., eds., Health, Economics, and Health Economics. Amsterdam: North-Holland.

APPENDIX A

Means and Standard Deviation of
Non-Health, Non-Attitude Variables

	Means	Standard Deviations
AGE	8.62(C), 37.07(A)	4.73(C), 8.01(A)
< 6	.277(C)	.448(C)
12-17	.307(C)	.462(C)
> 55	.012(A)	.109(A)
FEMALE	.491(C), .482(A)	.500(C), .500(A)
EDUCATION	13.33(A)	2.66(A)
FAMSIZE	4.96(C), 4.47(A), 4.38(F)	1.36(C), 1.17(A), 1.25(F)
NONWHITE	.117(C), .075(A), .099(F)	.322(C), .264(A), .299(F)
MARRIED	.881(C), .946(A), .889(F)	.324(C), .227(A), .314(F)
WORKFULL	.585(A)	.493(A)
WORKPART	.115(A)	.320(A)
OCCUPATION	41.15(A)	25.93(A)
MEDINC[1]	1.339(C), 1.379(A), 1.347(F)	.307(C), .288(A), .301(F)
FAMINC[1]	1.517(C), 1.636(A), 1.549(F)	.679(C), .657(A), .692(F)
BIRTHORD	1.96(C)	1.10(C)
LMAGE	.037(C)	.189(C)
FFULL	.823(C), .831(F)	.382(C), .375(F)
FPART	.008(C), .008(F)	.089(C), .088(F)
FOCC	49.67(C), 50.39(F)	21.13(C), 20.83(F)
MFULL	.189(C), .202(F)	.392(C), .402(F)
MPART	.197(C), .202(F)	.398(C), .402(F)
MOCC	21.33(C), 23.06(F)	25.66(C), 26.29(F)

	Means	Standard Deviations
MEDUC	12.57(C), 12.75(F)	2.46(C), 2.53(F)
AGE HEAD	37.43(F)	8.60(F)
OWN HOME	.837(F)	.370(F)
MCAID	.071(C), .038(A)	.257(C), .192(A)
PRIVINS	.908(C), .955(A)	.290(C), .208(A)
HMOINS	.081(C), .062(A)	.273(C), .242(A)
HOSP	12.00(C), 12.20(A)	6.06(C), 6.06(A)
ALL	.0006(A)	.003(A)
GPPED	.0002(C)	.0008(C)
HMO	12.51(C), 13.05(A)	8.93(C), 8.86(A)
XHMO	15.51(C), 16.18(A)	8.02(C), 7.63(A)

[1]MEDINC and FAMINC measured in units of $10,000.

Code: (C) = Children
(A) = Adults
(F) = Families