Volume Title: Economic Aspects of Health

Volume Author/Editor: Victor R. Fuchs, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-26785-7

Volume URL: http://www.nber.org/books/fuch82-1

Publication Date: 1982

Chapter Title: The Status of Health in Demand Estimation; or, Beyond Excellent, Good, Fair, Poor

Chapter Author: Willard G. Manning, Jr., Joseph P. Newhouse, John E. Ware, Jr.

Chapter URL: http://www.nber.org/chapters/c6548

Chapter pages in book: (p. 141 - 184)

PART

2 Consequences of Ill Health

This Page Intentionally Left Blank

5 The Status of Health in Demand Estimation; or, Beyond Excellent, Good, Fair, and Poor

Willard G. Manning, Jr., Joseph P. Newhouse, and John E. Ware, Jr.

Several years ago we were faced with the task of determining how to measure health status for the purpose of estimating demand in the Health Insurance Study (HIS).[1] Indeed, our challenge went further, for we were to ascertain the degree to which variation in the consumption of medical care services affected health status. Although in principle we could collect data on any measurable variable, we found virtually no guidance in the economics literature on efficient ways to measure health status. Some work had been done by health services researchers, but even that literature did not offer an entirely satisfactory way to measure health status.[2]

In this paper we examine some issues pertinent to the measurement of health status for estimating demand equations. Our comments should be of particular interest to those faced with the problem of what data to collect to estimate demand. They are also relevant to the literature on demand estimation. Specifically, we address two questions.

First, what is to be gained by using better measures of health status? We define measures that we believe are better in two respects than those used in the economics literature. Our measures tend to be more reliable: they have less random measurement error. Random measurement error will typically affect not only the coefficients of the variables measured with error, but also the coefficients and t-statistics of all variables not orthogonal to those variable(s). The direction of the inconsistency in other variables and their t-statistics cannot be signed a priori (Cooper and Newhouse 1971). Also, our measures are more comprehensive. Many of

The research reported herein was performed pursuant to a grant from the U.S. Department of Health and Human Services, Washington, D.C. The opinions and conclusions expressed herein are solely those of the authors and should not be construed as representing the opinions or policy of any agency of the United States Government.

the measures in the literature are unidimensional, but health is multidimensional (Ware, Davies-Avery, Brook 1980). The gains derived from measuring the various dimensions include reduction in residual variance and potential reduction of omitted variable bias.

Second, what are the consequences of using measures of current health to explain past utilization behavior? Most empirical estimates of the demand curve based on microlevel data have been cross-sectional. Thus, they have asked individuals about their current health and their past utilization of health services, usually during the previous year. The health status measures at the time of the interview are then used to explain past utilization. We will show that the estimated coefficients of the health status variables using this procedure are inconsistent. The inconsistency will, of course, affect the coefficients of other variables that are not orthogonal to health status. The problems occur because the observed health status variables do not really predict utilization; they "postdict" it. We develop a simple model of postdiction in which the estimated coefficient of the health status variable is inconsistent, whether or not medical care affects future health status. If medical care at the margin does not affect health status, the direction of the inconsistency is toward zero; but, in general, the direction of the inconsistency cannot be signed a priori. Because of the ambiguous sign and unknown magnitude of the inconsistency, we investigate empirically its sign and magnitude. We use a data set that has similar measures of health status spaced a year apart, as well as measures of utilization during the intervening year.

Prior Work

The most frequent method used to measure health status in economic literature is the simple question: "Would you say your health, in general, is excellent, good, fair, or poor?"[3] We refer to this as EGFP. This variable has been used to estimate demand functions (sometimes with no other health status variable present) by Acton (1975), Acton (1976), Andersen and Benham (1970), Colle and Grossman (1978), Goldman and Grossman (1978), Grossman (1972), Grossman and Benham (1974), Manning and Phelps (1979), Newhouse and Marquis (1978), Newhouse and Phelps (1974, 1976), and Phelps (1975). In part, the popularity of this variable stems from its inclusion in the 1963 and 1970 Center for Health Administration Studies national probability sample surveys, which have been among the richest sources of survey data for estimating demand functions during the past decade. Indeed, nine of the twelve studies just cited used one or the other of those two surveys. The same question was also included in three other household surveys that economists have analyzed. Coincidentally, all three had New York City residents as subjects: Acton (1975) analyzed a 1965 survey of users of hospital outpatient depart-

ments; Acton (1976) analyzed a 1968 survey of two poverty neighborhoods in Brooklyn; and Goldman and Grossman (1978) analyzed a 1965–66 survey of medical care utilization by children in the Bronx.

No other measure of health status has been quite so popular, although a number of others have been used. Colle and Grossman (1978), Davis and Reynolds (1976), and Newhouse and Phelps (1976) included restricted activity days as a covariate, and Acton (1976) used bed disability days. Andersen and Benham (1970), Grossman and Benham (1974), Hershey, Luft, and Gianaris (1975), and Manning and Phelps (1979) included the number of symptoms that occurred in some previous time period; Acton (1976), Davis and Reynolds (1976), and Hershey, Luft, and Gianaris (1975) included the number of chronic conditions that the respondent had. Phelps (1975) controlled for the amount of pain felt by the respondent over the past year.

Invariably these studies find health status to be an important variable in the demand function. Health status often explains more variance than any other variable, and controlling for it can make a considerable difference in the magnitude of other estimated coefficients.

The studies just cited have all used the family or individual as the unit of observation; we call them microlevel studies. A number of demand studies in the economics literature, however, use aggregate data; i.e., the dependent variable is some measure of the demand for services in an area such as a state or region. Typically these studies do not include explicit health status measures as covariates (Davis and Russell 1972, Feldstein 1971, 1977), but assume implicitly that health status shows negligible variation across geographic entities. Fuchs and Kramer (1972) did include infant and crude mortality rates as measures of health status in analyzing demand by region, but found these measures unimportant and dropped them from their final specification.

In addition to measures of health status, two studies in the economics literature have used measures of attitude toward medical services to explain demand. Colle and Grossman (1978) used a measure they called taste for medical care consumption, and we call below attitude toward the efficacy of medical care services. They did not find this variable important. Hershey, Luft, and Gianaris (1975) used a measure they called self-reliance, and we call attitude toward the efficacy of self-care. In their study, this variable was significantly related to demand.

Two studies in the literature have some similarities with the present paper, and deserve further discussion. These are Andersen and Benham (1970) and Hershey, Luft, and Gianaris (1975). Both investigate the effect on estimated income elasticities if health status variables are included. In their paper, Andersen and Benham examined the magnitude of the bias in estimated income elasticities if health status, demographic characteristics, health insurance status, the nature of the family's usual

source of care, and whether a family member had had a physical examination in the past five years were excluded as explanatory variables. Their dependent variables were physician and dental expenditures. (We focus on their results for physician expenditures because that is approximately what we analyze in this paper.) The authors found that the simple permanent income elasticity (controlling for no other variable) was 0.63 but fell to 0.17 if all the above mentioned characteristics except health status were controlled for. When health status (measured as excellent, good, fair, and poor, and using a symptom count) was also controlled for, the income elasticity was 0.30. If measured income rather than permanent income was used, estimates were one-quarter to one-third lower.

Hershey, Luft, and Gianaris examined similar questions. They analyzed data from the sample of families from a semirural California town (population 12,000) rather than the national probability sample that Andersen and Benham had used. Their study also differs from the earlier report in that they examine the number of physician visits rather than physician expenditure. But estimates we have made suggest that expenditures and visits are highly correlated, so this difference in dependent variables should not seriously affect comparisons between the two studies.

The 1975 paper estimated measured income elasticities as being well under 0.1 (at the mean) and statistically insignificant, in marked contrast to Andersen and Benham's work. Also, the inclusion of health status variables did not cause a notable change in the elasticity estimates. Hershey, Luft, and Gianaris use the individual as a unit of observation and include income per person and family size as explanatory variables, while Andersen and Benham use the family as a unit of observation and control for family size by entering it as an explanatory variable. But the authors of the later paper say that their results are similar when the family is used as a unit of observation and total family income is the explanatory variable. Other than the difference in samples, it is difficult to reconcile these two studies.

Because our focus in this paper is on the measurement of health status, we are interested in determining how more complete and reliable measures of health status affect the estimated coefficients. Our focus differs from that of the papers discussed, which are primarily concerned with how the inclusion of a whole host of covariates affects income elasticities, not with how more complete measures of health affect income elasticities and other variables. The papers just discussed do distinguish, however, between the influence of attitudes and that of health status variables per se.[4]

All the studies in the literature that use the EGFP measure postdict utilization. In the next section we will demonstrate that postdicting can cause inconsistent estimates. In addition, health status covariates that

nominally refer to the same period as utilization (e.g., disability days in the past year) leave causality ambiguous: Did one suffer from restricted activity and therefore seek care, or did the physician advise taking it easy? In other words, apparently contemporaneous health status variables may also be endogenous.

A Simple Model of Postdiction

Suppose the true model relating utilization and health status is as follows:

$$U_t = \alpha HS_{t-1} + \gamma P_t + \mu_f + \varepsilon'_t, \tag{1}$$

$$HS_t = \beta U_t + \delta HS_{t-1} + u_t \ , \tag{2}$$

where U_t = utilization in period t,
HS_t = health status in period t, with larger values indicating better health,
P_t = price in period t,
μ_f = a person or family-specific effect that is time invariant,

and ε'_t and u_t are error terms with standard properties. The intercepts and the subscript indexing individuals have been omitted for convenience.

Suppose time is measured in years, and one has observations on health status at the end of each year (health status is assumed constant during the year) and on utilization in year two. In our data, P_2 is (approximately) orthogonal to all other variables. If (1) is estimated as specified (e.g., U_2 is estimated as a function of HS_1 and P_2), "prediction" occurs.

If, however, HS_2 is used in lieu of HS_1, "postdiction" occurs. We derive the plim of the estimated coefficients in Appendix A. In the general postdiction case the estimated coefficients of price and health status are inconsistent and the direction of the inconsistency cannot be signed. If, however, β equals zero (medical care in the observed range does not affect current health status), the inconsistency in the estimated coefficient of health status, α, is toward zero, and the estimated coefficient of price, γ, is consistent.

Because the sign of the inconsistency is in general not known, nor is the magnitude of the problem, we turn to empirical methods to ascertain the possible magnitude of the problem.

Measures of Health Status, Attitudes, Behavioral Propensities, and Knowledge

The World Health Organization defines health as "a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity" (WHO 1948). Two aspects of this definition have

implications for the comprehensive measurement of health status. First, health is multidimensional with distinguishable physical, mental, and social components. (A fourth dimension, physiological health, is outside the scope of this paper; see Brook, Goldberg, Applegate, et al., [in press] for further discussion and definition). Second, the continuum of health for each of its components should extend beyond the negative states defined by different levels of illness and beyond the absence of illness to include the degree to which levels of positive states are enjoyed.

Construction of Health Status, Attitudinal, and Other Measures

Details of the health status and other measures used below can be found in series of monographs, which are referenced and summarized in a supplement of *Medical Care* (see Brook, Ware, Davies-Avery et al. 1979). In summary, these measures were constructed by using standard psychometric scaling techniques intended to achieve several desirable measurement properties: (a) variability and roughly symmetrical score distributions (as opposed to extremely skewed score distributions), (b) reliability (the proportion of measured variance that is true score, as opposed to random error), and (c) validity (an understanding of what is being measured and how differences in scores should be interpreted).

Variability and Symmetry

Health status and attitudinal measures focus on concepts that vary substantially in general populations. For example, with regard to mental health, the measures emphasize prevalent symptoms of emotional instability (for example, anxiety and depression) rather than psychotic disorders that are relatively rare in general populations.

Reliability

To enhance measurement reliability, the HIS fielded unambiguous questionnaire responses, such as "My health is now excellent" as opposed to "Health is good"; the latter response could refer to either the value placed on health or to the goodness of one's health, and its time frame is ambiguous. Multiple-response, as opposed to dichotomous, choices further improved reliability. For example, respondents were asked: "During the past month, how much of the time have you felt depressed?" Six choices, ranging from "All of the time" to "None of the time," were offered, as opposed to asking "During the past month, have you been depressed?" with responses of "Yes" and "No." Finally, rather than depending on less reliable single-item scales (e.g., the EGFP variable), we constructed our measures to combine items whenever appropriate to form multi-item scales.[5]

Validity

John E. Ware, Jr. has studied the validity of each measure, using both a variety of empirical methods and also an analysis of the content of the items contained in each measure. The primary empirical method employed was construct validation (Ware, Davies-Avery, and Brook 1980).[6] Construct validity studies are based on a theoretical model of how a valid measure should relate to other measured variables. For example, a valid measure of physical health should correlate (a) negatively with age, (b) positively with personal ratings of health status, (c) substantially with other physical health indicators that employ different measurement methods (such as another data source or a different scaling technique), and (d) negatively with medical care consumption.

In this paper we have used four categories of health-related measures to explain medical care consumption: (1) health status, (2) attitudes toward self-care and medical care, (3) behavioral propensities, and (4) sophistication or knowledgeability regarding the medical care delivery system. Tables 5.1 and 5.2 contain summaries of information about each variable, including (a) variable labels used in presenting regression results, (b) the kind of concept defined by each variable, (c) number of questionnaire items used in scoring, and (d) a brief operational definition.

Each variable is scored in such a way that a higher score indicates more of what the variable measures; thus, a higher score on the CURRENT health scale indicates more (better) current health, and a higher score on the WORRY scale indicates more worry. Additional background information on the development and rationale for these measures is briefly summarized below. Appendix B lists items by major category.

Physical Health

We measure the physical component of health in terms of limitations on functional status, a count of chronic diseases, and a count of acute symptoms. Functional status refers to the performance of or capacity to perform a variety of activities that are normal for an individual in good health (Stewart, Ware, and Brook 1981). Our functional status variable (PHYSLIM) measures limits on functional status and combines physical performance and capacity items in three areas: self-care (e.g., feeding, bathing), mobility (e.g., confinement indoors), and physical activities (e.g., walking, running). Because some differences in physical health may increase medical care consumption without affecting functional performance or capacity, we included two other measures of physical health, a count of chronic diseases (DISEASE) and a count of acute symptoms of illness whether treated or untreated (ACSIL). We do not at

Table 5.1 **Operational Definitions and Labels Assigned to Health Status Measures**

Type of Scale Variable Label	Health Concept[a]	Number of Items	Definition
Single-item scales			
POOR	G	1[b]	1 if health rated poor, zero otherwise
FAIR	G	1[b]	1 if health rated fair, zero otherwise
GOOD	G	1[b]	1 if health rated good, zero otherwise
Multi-item scales			
PHYSLIM[c]	P	9	Acute limitations in self-care, mobility, and physical activities
DISEASE[c]	P	32	Simple count of the number of physical disease conditions (26 possible)
ASCIL[c]	P	27	Simple count of the number of acute physical symptoms, past month
EMOINS	M	19	Emotional instability (anxious, depressed downhearted, tense, worried), past month
PWB	M	8	Positive well-being (in good spirits, cheerful, feeling loved, cared for), past month
SOCACT	S	9	Frequency and nature of social contacts and group memberships
CURRENT	G	9	Rating of health in general, at present
PRIOR	G	3	Rating of past health status
OUTLOOK	G	4	Rating of expected future health status
RESIST	G	4	Perceived bodily resistance to health threats
CONCERN	G	2	Amount of concern over personal health
WORRY	G	2	Amount of worry over personal health
HPQTOT	G	22	Health perceptions summary score; unweighted sum of CURRENT, PRIOR, OUTLOOK, RESIST, and negative of WORRY
LCU[c]	G	20	Life change units indicating stressful events

[a]Physical Health (P), Mental Health (M), Social Health (S), and General Health (G).

[b]Dummy variables based on single-item rating of health in terms of excellent, good, fair, or poor.

[c]In the regressions, the variable was replaced by log (variable + 1) to diminish the effect of skewness as well as to produce homoskedastic error plots.

Table 5.2 Operational Definitions and Labels Assigned to Measures of Attitude, Knowledgeability, and Behavioral Propensity

Variable Label	Health Concept[a]	Number of Items	Description
TRSIL[b]	B	27	Simple count of acute physical symptoms that care was sought for, past 30 days
ATGD	A	2	Favorable attitude toward going to the doctor
CONSOPH	K	8	Sophistication or knowledgeability about the medical care delivery system
EFFDOC	A	4	Favorable attitude toward the efficacy of doctors and medical care services
EFFSLF	A	4	Favorable attitude toward self care and home remedies
REJECT	B	4	Conscious avoidance of the sick/patient role

[a]Attitude (A), Behavioral propensity (B), Knowledgeability/Sophistication (K).

[b]In the regressions, the variable was replaced by log (variable + 1) to diminish the effect of skewness as well as to produce homoskedastic error plots.

this time have data on disability days. Thus, even our most comprehensive set of measures is not as comprehensive as possible.

Mental Health

We define the mental component of health in terms of selected phenomena of psychological disorders about which there is considerable conceptual agreement in the mental health literature. In addition we include positive states of well-being that are often ignored in general population studies (Ware, Johnston, Davies-Avery et al. 1979). Psychological disorders include two negative constructs: anxiety and depression. Because these constructs are highly collinear, they were combined into a single indicator of emotional instability (EMOINS). To distinguish between persons not experiencing a psychological disorder, highly correlated measures of positive emotional states and ratings of quality of life were combined to define positive well-being (PWB). All mental health items focus on psychological states, rather than on physiological and somatic ones (such as those measured by ACSIL), because inclusion of the latter in mental health would confound definitions of physical and mental health components.

Social Health

We measure the social component of health status by the frequency of social activities and by ratings of social resources in several distinct categories: (a) visits with friends and relatives, (b) memberships and participation in group functions, and (c) quality of social supports (e.g., having close friends that can help solve personal problems). To minimize the overlap between mental and social health measures found in previous studies (Donald, Ware, and Brook 1978), the social health measure excluded ratings of subjective feeling states related to social well-being (as, for example, feeling cared for and loved). Instead, these feelings are included in the positive well-being measure within the mental health concept. Psychometric studies of social activities have identified a common component, which can be summarized by a single health scale (SOCACT). Moreover, prediction of medical care consumption does not improve if we disaggregate social dimensions.

General Health Perceptions

In addition to the physical, mental, and social components suggested by the WHO definition of health status, ratings of general health perceptions were also tested (Ware, Davies-Avery, and Donald 1978). The health perception measures are distinct from the health status measures discussed above in that they do not focus on any specific health component. Instead, they ask for a personal assessment of health in general. In theory, general health ratings allow people to consider not only the objective information they have about their health but also their evaluation of that information. Measures were defined with respect to time (PRIOR, CURRENT, and OUTLOOK) and three other constructs: resistance-susceptibility to illness (RESIST, WORRY, and CONCERN). A unidimensional health component underlying these dimensions was defined by a summary indicator (HPQTOT), which combined the scales just mentioned excluding CONCERN. (The score for WORRY is multiplied by -1 prior to summing; CONCERN did not load on the same factor as the other five.) We also included the number of life change units (LCU), a method of weighting the stressful life events that befell the individual (Ware, Davies-Avery, and Brook 1980).

For comparison, we also measured health perceptions by the single-item rating of current health in terms of EGFP.[7] EGFP should perform less well than CURRENT; CURRENT, in principle, measures the same construct but should be more reliable because it is a nine-item scale with multiple responses for each item (Ware, Davies-Avery, and Donald 1978).

Other Health-related Variables

In addition to the health status measures discussed above, we examine six measures of attitude (tastes/sentiments), behavioral propensity, and

knowledgeability regarding medical care services (see Table 5.2 for a summary of operational definitions). In contrast to health status, these measures focus directly on medical care seeking behaviors, such as whether one does or does not like to go to the doctor (ATGD; Ware 1976) or seeks care conditional upon having symptoms (TRSIL). Self-care attitudes, such as whether one believes in the efficacy of home remedies (EFFSLF) or the efficacy of medical care (EFFDOC), assess sentiments regarding treatment options (Lau and Ware, forthcoming). Consumer sophistication (CONSOPH) assesses knowledge of the medical care delivery system (Newhouse, Ware, and Donald, 1981). It should be a conceptually more appropriate measure of knowledge than education because it measures specific rather than general knowledge. Education was, however, tried and had insignificant effects in every specification.

Data and Nonhealth Status Variables

The data come from the first year of experience in three of the six sites of the Health Insurance Study (Newhouse 1974): Seattle, Washington; Fitchburg, Massachusetts; and Franklin County, Massachusetts. The first site—Dayton, Ohio—is omitted because the measures of health status changed somewhat between those taken at enrollment and those taken at the end of the first year of participation. The fifth and sixth sites, which are in South Carolina, are omitted because complete data from those sites are not yet available.

For this analysis we used a random sample of the population of the three sites, with the following exceptions:

1. Those over 61 at the time of enrollment were not eligible.
2. Those with incomes in excess of $25,000 (1973 dollars) were not eligible. This restriction excludes approximately the upper 5% of the income distribution.
3. The military and their dependents, veterans with service connected disabilities, and the institutionalized population (e.g., those in jails or state mental hospitals) were not eligible.
4. In Seattle, those who belonged to the Group Health Cooperative of Puget Sound (approximately 15% of the Seattle population) are not included in the sample analyzed here.
5. In the two Massachusetts sites, the low-income population was oversampled. Specifically, those families with incomes that were within 150% of the poverty line had a 33% greater chance of being included. We have not reweighted the sample of reflect this oversampling because we are not interested in predicting site means.

The sample we analyze is a subset of the adults (aged 18 and older) enrolled in the three sites. Adults who failed to fill out the health questionnaire at enrollment or after the first year of participation were excluded to avoid problems of missing data with the health status measures.

Individuals who died, attrited, or were suspended for part of the year, typically because they were in military service, were also excluded. The sample consists of 1,557 adults; there were 165 exclusions. Approximately half the exclusions were individuals who died, were suspended, or attrited; the remainder did not return the questionnaire.

Health Status Data-gathering Methods

All data on health status were gathered by means of a standardized questionnaire that was self-administered in each respondent's home. Heads of households each received $20 for completing questionnaires containing up to 531 relevant items; dependents received $5 each. The rate of returned questionnaires approximated 95%. While still in the field, we checked returned questionnaires for missing items; callbacks (in person or by telephone) were initiated whenever more than six items were left blank. This procedure produced very few (less than 1%) missing responses for returned questionnaires.

The Nonhealth Status Variables

In this paper we have confined our analysis to covered annual outpatient expenditures (OUTP) for health care,[8] except those for mental health care, dental care, and drugs and supplies. Outpatient care includes services provided by medical doctors, osteopaths, and some nonphysicians, such as chiropractors, podiatrists, speech therapists, physical therapists, and optometrists. Roughly 90% of the outpatient expenditure, however, is for physician services. Claims filed by the participants, including those for unreimbursed expenses, provide data on the amount and type of expenses.

We have used two variables to specify the coinsurance coverage provided to the participant by the HIS.[9] The first is a logarithmic coinsurance function LC = ln (coinsurance percent + 1) for those individuals facing a family coinsurance rate, i.e., where all members of a family face the same coinsurance rate at the same time.[10] The second is an indicator variable for the individual deductible plan (IDP). That plan approximates an outpatient care deductible of $150 per person or $450 per family (actually 95% coinsurance to a maximum of $150 out-of-pocket per person or $450 out-of-pocket per family), with all inpatient care free and outpatient care free beyond the deductible.

The explanatory variables also include three other indicator variables for experimental treatment: whether a household was given a preenrollment screening examination (EXAM),[11] whether the family was exempted from having to file biweekly diaries reporting sick days (NOHR);[12] and whether the family was enrolled for three rather than five years(YR3).[13]

The remaining nonhealth status explanatory variables control for variation in socioconomic factors. They include income for the two years

prior to enrollment, family size, age, sex, and race. We selected functional forms for the continuous variables that would yield homoskedastic residual plots. Table 5.3 contains the formal variable definitions and Table 5.4 describes the sample characteristics. Data on all of these socioeconomic variables were derived from pre-enrollment interviews.

The Expenditure Model

The distribution of outpatient expenses has three characteristics that require special attention if one is to obtain reliable estimates of the demand for care. First, part of the distribution is clustered at zero; second, the distribution of positive expenditure is highly skewed; and third, the error terms for different family members are positively correlated. In other work (Manning, Morris, Newhouse et al. 1981), we have suggested that an appropriate model for expenditures is one with two parts (or equations) with variance components in the error term. The first equation models the decision to seek care and the second estimates the logarithm of nonzero expenses, conditional upon positive expenditure. The first equation appropriately handles the zero mass. The logarithmic transformation of positive expenses reduces the estimation problems caused by the skewness of positive expenses. In this case the variance components specification closely approximates the pattern of intrafamily correlation.

Here we will use a simpler, more tractable model. Instead of the two-part model, we will use a single dependent variable, the logarithm of expenses plus $5. Five dollars was chosen as the constant that left the data

Table 5.3 Socioeconomic Variables

Indicator Variables (0,1)		
BLACK	= 1	if race of the head of family is black
FEMALE	= 1	if female
AFDC	= 1	if someone in the family received Aid to Families with Dependent Children
INCMIS	= 1	if information was missing
FITC	= 1	if in the Fitchburg site
FRAN	= 1	if in the Franklin County site
Continuous Variables		
LNAGE	= ln	(age)
LINC	= ln	(average of 1973, 1974 family income in constant 1972 dollars)[a]
LFAM	= ln	(family size)

[a]Income was set equal to $1,000 if reported to be less. Individuals with INCMIS = 1 received the site mean LINC.

Table 5.4 **Sample Characteristics**

Variable	Average	Standard Deviation	Minimum	Maximum
LC	1.58	1.95	0.00	4.56
IDP	0.25	0.43	0.00	1.00
NOHR	0.13	0.33	0.00	1.00
EXAM	0.65	0.48	0.00	1.00
YR3	0.75	0.43	0.00	1.00
LINC	8.88	0.67	6.91	10.01
INCMIS	0.02	0.15	0.00	1.00
LFAM	1.00	0.59	0.00	2.40
BLACK	0.02	0.13	0.00	1.00
AFDC*	0.02	0.14	0.00	1.00
FEMALE	0.53	0.50	0.00	1.00
AGE	35.50	11.80	18.00	62.00
LNAGE	3.52	0.33	2.89	4.13
FITC	0.25	0.43	0.00	1.00
FRAN	0.31	0.46	0.00	1.00
OUTP	129.91	191.13	0.00	2454.16
ln (OUTP + 5)	4.08	1.42	1.61	7.81

*Data on AFDC status collected only in Seattle.

most nearly normal. If the goal of the analysis were to predict expenses, then this simpler model would provide biased estimates of raw means because the size of the zero mass is covariate-related in a manner different from the response in the nonzero expenses. But the inferences that can be drawn from the simpler model appear to be robust in spite of this misspecification in the case of these data. Because our goal is inference rather than prediction, we are willing to make a sacrifice for computational simplicity.

An alternative to the logarithmic transform plus $5 would be to use raw dollar expenses as the dependent variable. Unfortunately, with our sample size, the nature of health expenditure data causes least squares to yield imprecise results. Least squares would be appropriate if expenditure had a normally distributed error term or if there were enough data to rely on the Central Limit Theorem. In either case, the *t*- and *F*-statistics would be well behaved and the coefficients would be robust. Unfortunately, neither condition is met. Even sample sizes in excess of 1,000 yield imprecise results given the error distribution of our data. A few very large expenses have an undue influence on the results.

The logarithmic transformation of the expenses provides more robust and more efficient estimates. Nonzero outpatient expenses are very close to lognormally distributed, and so the use of a logarithmic transform reduces the departure from normality. The smaller the departure, the more robust the estimates. Thus, the logarithmic transform lessens the

likelihood that a few large expenses will have an undue influence on the coefficients. The use of a logarithmic transformation also reduces the coefficient of variation in the dependent variable. The coefficients will therefore be more precise with the logarithmic transform than with the raw dollar sale if lognormality holds. For the HIS data, the increase in precision with a logarithmic transform is roughly equivalent to a three-to-fivefold increase in sample size.

We have used a random-effects variance-components estimator. With this estimator we can obtain efficient estimates of the regression coefficients and consistent estimates of the standard errors. The data exhibit a nearly constant intrafamily correlation across family sizes. Hence, the residual correlations are similar to those of a variance-components model with a family-specific error term. The expenditure equation is estimated by maximum likelihood, iterating over the coefficients and the proportion of the error variance accounted for by the family component.

Empirical Results

In examining the empirical results obtained from our use of alternative sets of health status variables, we focus on the two questions posed in our introduction. What have we gained in explanatory power by using more comprehensive health status variables than the simple excellent, good, fair, poor (EGFP) question? What are the consequences of using postdictive variables, i.e., of using measures of current health to explain past utilization behavior?

Gains from Comprehensiveness

There are two potential gains from employing more comprehensive health status variables than EGFP. The EGFP measure provides only four responses to a unidimensional concept of health. Alternative measures could reduce the coarseness of the measurement by (a) providing a finer and more reliable division in a unidimensional measure, and (b) providing measures of several dimensions of health.

In comparing EGFP with CURRENT, we can observe the effect of greatly increasing the number of responses. In comparing CURRENT with HPQTOT, we can observe the gains from a broader definition of health still restricted to a scalar measure. Comparing HPQTOT with its components (CURRENT, PRIOR, OUTLOOK, WORRY, and RESIST) provides a test of whether health perceptions are unidimensional or multidimensional. Finally, adding measures of physical, mental, and social health, and attitudes toward the efficacy of medical care extends the number of dimensions of health that may affect overall expenditures.

Our results show that there are indeed gains from using more comprehensive measures. Table 5.5 contains test statistics that compare one

Table 5.5 Tests for Alternative Specifications of the Health Variables

H_0	versus	H_1	Pesaran's N_0*	F** Value	F** df
EGFP		CURRENT	−2.56	—	—
CURRENT		HPQTOT	−0.48	—	—
HPQTOT		HPQ components	—	4.79	4,1537
HPQ Components		Full set	—	4.74	14,1523

*Distributed $N(0, 1)$. A negative sign implies a rejection of the null hypothesis in favor of the alternative model.

**The one-percent critical value for $F_{4,1000}$ is 3.34, and for $F_{14,1000}$ it is 2.09.

alternative health specification with another. Two sets of tests are presented. First, for testing nonnested specifications, we provide a statistic N_0 that is a standardized normal variate when the null hypothesis (the first health specification) is true.[14] Second, for nested specifications, the standard F-statistic is calculated. For both types of tests, the data have been transformed to remove the effect of intrafamily correlation. As Table 5.5 indicates, we can reject EGFP in favor of CURRENT at an $\alpha < .01$. However, we cannot reject CURRENT in favor of the broader HPQTOT (scalar) measure.

Further gains come when HPQTOT is disaggregated. At an $\alpha < .001$ one can reject the hypothesis that HPQTOT contains all the behavioral information contained in its components. If a single health perceptions scale HPQTOT were appropriate, then the (unstandardized) coefficients on its components would be equal to each other and to the (unstandardized) coefficient on the scale itself. As the test statistic in Table 5.5 indicates, this equality does not hold. Thus, we can reject the hypothesis that behavior reflects a single health perception.

Further, one can reject the hypothesis that only health perceptions matter. A fuller specification, including physical limitations (PHYSLIM), counts of recent symptoms (ACSIL), attitudes about going to the doctor (ATGD), and the efficacy of self-treatment (EFFSLF), exhibits significantly increased explanatory power. The health perception components as a group are still significant, but the size of their coefficients is now reduced because these variables no longer act as proxies for other omitted health status variables.

Explained variation increases from 0.1292 in a specification with only EGFP to 0.1844 with the full model. Another way to describe this increase is that a more comprehensive measure of health status, if available, will yield a gain in precision at least equivalent to a 7% increase in sample size [$1.07 = \exp\ (1. - .1292)/(1 - .1844)$]. Seven percent is a

lower bound because the more comprehensive measure not only causes the residual variance to fall but also decreases the proportion of error variance that comes from intrafamily correlation.[15] We have calculated the decrease in the confidence interval for a mean individual on the free plan; the decrease equals 12%.

The introduction of more comprehensive health status and attitudinal variables could affect the coefficients on other nonhealth covariates. Any variable that is correlated with the omitted variable will have some bias in its coefficient. Table 5.6 presents the coefficients for selected variables, under alternative specifications. The results are rather mixed. The coinsurance function and individual deductible plan coefficients are, of course, unaffected, because the assignment of the sample to coinsurance plans was designed to leave the plan, for practical purposes, orthogonal to other covariates. The coefficients on income (LINC) and family size (LFAM) are significant and also change little as the specification is enriched.

The coefficient on BLACK moves toward zero (becomes less negative) by about one-third of its standard deviation as the health specification becomes more comprehensive. Thus, blacks appear to have relatively poorer health in the dimensions omitted from the EGFP specification. We caution against overinterpreting this result, however, because there are few blacks in our sample (Table 5.4). The age (LNAGE) coefficient increases as the specification is increased, moving about one-half of its estimated standard deviation. Finally, in the most comprehensive specification, the FEMALE coefficient falls by 1 to 1.5 standard deviations. Note, however, that the FEMALE coefficient is always strongly positive, in contrast to the belief of Hershey, Luft, and Gianaris (1975), who suggest that it will be negative if health status is controlled for.

Predictive Results

The coefficients of the health status and attitudinal variables from the various specifications are given in Table 5.7; to facilitate interpretation, we generally present standardized coefficients. The addition of health, attitudinal, and other covariates generally reduced the coefficients of health perceptions. Consumer sophistication (CONSOPH) is insignificant, but its impact is indeterminant because a sophisticated consumer may consume more efficacious services and fewer inappropriate ones. Physical limitations, chronic disease, and symptoms all have a positive impact on utilization. Favorable attitudes about going to the doctor are mirrored in higher utilization of medical services, whereas individuals who believe in the efficacy of self-treatment use fewer services. This latter result is also found by Hershey, Luft, and Gianaris (1975). We replicate Colle and Grossman's (1978) result on the unimportance of attitudes concerning the efficacy of medical care. There is strong support for the

Table 5.6 **Selected Variation in Coefficient Estimates as Specification Changes ($|t|$)**

	Coefficient Estimates of Selected Variables						
Specification	LC	IDP	LINC	LFAM	BLACK	LNAGE	FEMALE
EGFP	−.15	−.45	.04	−.04	−.83	.49	.55
	(7.20)	(4.80)	(.62)	(.56)	(2.94)	(4.30)	(8.58)
CURRENT	−.15	−.44	.04	−.03	−.87	.52	.57
	(7.07)	(4.71)	(.62)	(.45)	(3.10)	(4.69)	(8.92)
HPQTOT	−.15	−.44	.04	−.03	−.86	.54	.56
	(7.08)	(4.68)	(.63)	(.42)	(3.08)	(4.84)	(8.83)
HPQ components	−.15	−.44	.05	−.03	−.82	.57	.58
	(7.15)	(4.78)	(.84)	(.39)	(2.95)	(4.90)	(8.98)
Full list	−.15	−.43	.04	−.03	−.77	.54	.46
	(7.32)	(4.72)	(.72)	(.46)	(2.85)	(4.38)	(6.59)

Table 5.7 Coefficients for Alternative Health Variables

Specification	Variable[a]	$\hat{\beta}$	$\lvert t \rvert$
EGFP[b]	GOOD	+.17	2.26
	FAIR	+.67	4.89
	POOR	+.64	1.08
CURRENT	CURRENT	−.20	5.73
HPQTOT	HPQTOT	−.21	6.23
HPQ components	CURRENT	−.13	2.76
	OUTLOOK	+.01	1.99
	PRIOR	−.13	3.55
	RESIST	−.02	0.47
	WORRY	+.10	2.50
Full list	ACSIL	+.08	1.83
	DISEASE	+.08	1.89
	PHYSLIM	+.18	2.07
	EMOINS	+.14	2.82
	PWB	+.13	2.69
	SOCACT	+.05	1.29
	CONCERN	+.06	1.51
	CURRENT	−.07	1.37
	LCU	+.04	1.11
	OUTLOOK	+.06	1.49
	PRIOR	−.10	2.66
	RESIST	−.02	0.44
	WORRY	+.03	0.76
	ATGD	+.11	3.06
	EFFDOC	−.02	0.48
	EFFSLF	−.08	2.36
	REJECT	−.01	0.27
	TRSIL	+.08	2.18
	CONSOPH	−.04	1.11

[a]Except for the indicators for EGFP, all health status variables are in standardized form: $(x_{ij} - \bar{x}_i)/\sigma_{x_i}$.

[b]The standardized coefficients are .08, .18, .04.

notion that individuals with anxiety and depression do use the medical system more. But surprisingly, positive well-being has the wrong sign and is significant.[16] The variable social contacts and support (SOCACT) is positive but insignificant. A medical (contagion) model predicts a positive sign, whereas the psychological support model (the more social support, the less reliance on the medical care system for support) predicts a negative sign.

Postdiction vs. Prediction

As noted earlier, most microlevel studies in the literature have used current health measures to explain or postdict past medical utilization. With the panel data we have, we can test whether the inconsistency that a postdictive model causes is important.

We have used Wu's second and preferred test (1973, 1974) to detect any dependence between the postdictive health variables and the error in the expenditure equation.

Consider the following model of medical behavior:

$$U_i = Y_i'\beta + Z_i'\delta + \varepsilon, \tag{3}$$

$$Y_i = \mu_i + \nu_i, \quad i = 1, \ldots, N, \tag{4}$$

where for person i, U_i is the logarithm of outpatient expenditure plus \$5, Y_i is the $G \times 1$ vector of postdictive health status variables, Z_i is the $K_1 \times 1$ vector of known nonstochastic regressors, and μ_i is a $G \times 1$ vector of unknown constants. The errors ε and ν are assumed to be multivariate normal with covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_{\varepsilon\varepsilon} & \delta \\ \delta' & \sigma_{\nu\nu} \end{bmatrix}$$

Wu proposed that if K_2 ($> G$) nonstochastic variables Z_2 were available, then one test for δ equal to zero [i.e., zero covariance between equations (3) and (4)] is

$$T_2 = \frac{Q^*/G}{(Q - Q^*)/(N - K_1 - 2G)},$$

where

$$Q^* = (b_1 - b_2)'[(Y'A_2Y)^{-1} - (Y'A_1Y)^{-1}]^{-1}(b_1 - b_2),$$

b_1 = OLS estimate of β,

b_2 = instrumental variable estimate of β,

$$A_1 = I - Z_1(Z_1Z_1)^{-1}Z_1' = M_{Z_1},$$

$$A_2 = Z(Z'Z)^{-1}Z' - Z_1(Z_1'Z_1)^{-1}Z_1',$$

$$Z = (Z_1, Z_2),$$

$$Q = (U - Yb_1)'A_1(U - Yb_1) = N \cdot \text{OLS estimate of } \sigma_{\varepsilon\varepsilon}.$$

Under the null hypothesis ($\delta = 0$), T_2 is distributed as an F-statistic with $(G, N - K_1 - 2G)$ degrees of freedom.

The instruments that we used for this analysis were the (predictive) enrollment values for the health variables and the physician visits in the year prior to the one being observed. The data have been transformed by

using the postdictive estimate of intrafamily correlation to remove the effect of intrafamily correlation.

The values for Wu's test in Table 5.8 confirm that we can detect in our sample the dependence between the error in the equation and the postdictive variables. At each level of complexity, ranging from EGFP to the full list of health measures, we can reject the hypothesis that the postdictive measures are independent of the error at α levels of 1% or less. Postdicting raises the R^2; depending on the specification, postdicting increases the R^2 from 5% (CURRENT) to 26% (full list) (Table 5.9).

As discussed above, the direction of the inconsistency is, in general, theoretically indeterminant. Empirically, postdicting tends to move the coefficients away from zero. Table 5.10 provides a side-by-side comparison of predictive and postdictive coefficients. (For the richest specification, collinearity makes side-by-side comparisons more difficult.) For the EGFP, CURRENT, HPQTOT, and HPQ components specifications, the coefficients of the postdictive variables are larger in absolute value in all but two cases: the indicator for good health and the insignificant RESIST variable. In some cases the differences are very striking. The postdictive coefficient on WORRY in the HPQ components specification is double and the indicator for poor health in the EGFP specification is more than double their respective predictive coefficients. The result on WORRY is even more striking in the full specification, where it increases by a factor of 5. The coefficient of PHYSLIM also doubles in that specification.

Table 5.8 **Tests for Independence of the Error and the Postdictive Variables**

Specification	Wu's T_2	df	Significant at $\alpha \leq$
EGFP	10.26	3,1534	.001
CURRENT	16.46	1,1538	.001
HPQTOT	9.43	1,1538	.010
HPQ components	4.08	5,1530	.010
Full list	2.86	17,1506	.001

Table 5.9 **R^2 For the Predict and Postdict Models**

Specifications	Predict R^2	Postdict R^2
EGFP	.1292	.1415
CURRENT	.1338	.1403
HPQTOT	.1371	.1462
HPQ components	.1480	.1741
Full list	.1844	.2316

The results suggest that frequent use of outpatient services during the past year is associated with worry, physicial limitation, and preoccupation with poor health. Whether this reflects exogenous variation in health status that caused utilization (e.g., an auto accident that resulted in physical limitation and medical treatment) or whether it represents an

Table 5.10 **Coefficients for Predictive and Postdictive Health Variables**

		Predictive		Postdictive	
Specification	Variable[a]	β	|t|	β	|t|
EGFP[b]	GOOD	+.17	2.26	+.16	2.15
	FAIR	+.67	4.89	+.73	5.78
	POOR	+.64	1.08	+1.63	4.22
CURRENT	CURRENT	−.20	5.73	−.23	6.70
HPQTOT	HPQTOT	−.21	6.23	−.26	7.47
HPQ Components	CURRENT	−.13	2.76	−.13	2.66
	OUTLOOK	+.09	1.99	+.13	2.97
	PRIOR	−.13	3.55	−.17	4.67
	RESIST	−.02	0.47	−.005	0.12
	WORRY	+.10	2.50	+.20	4.94
Full list	ACSIL	+.08	1.83	−.02	0.37
	DISEASE[c]	+.08	1.89	+.07	1.78
	PHYSLIM	+.08	2.07	+.16	4.32
	EMOINS	+.14	2.82	+.02	0.37
	PWB	+.13	2.69	+.03	0.69
	SOCACT	+.05	1.23	−.01	0.16
	CONCERN	+.06	1.51	+.01	0.34
	CURRENT	−.07	1.37	−.03	0.58
	LCU	+.04	1.11	−.01	0.34
	OUTLOOK	+.06	1.49	+.09	2.01
	PRIOR	−.10	2.66	−.14	3.79
	RESIST	+.02	0.44	+.02	0.39
	WORRY	+.03	0.76	+.16	3.71
	ATGD	+.11	3.06	+.17	4.79
	EFFDOC	−.02	0.48	−.02	0.63
	EFFSLF	−.08	2.36	−.07	2.02
	REJECT	−.01	0.27	+.07	1.95
	TRSIL	+.08	2.18	+.22	6.11
	CONSOPH[c]	−.04	1.11	−.03	0.94

[a]Except for the indicators for EGFP, all health variables are in standardized form: $(x_{ij} - \bar{x}_i)/\sigma_{x_i}$.

[b]The standardized predict coefficients are .08, .18, .03; the standardized postdict coefficients are .08, .22, .14.

[c]Measured only at enrollment.

effect of the use of medical care (e.g., detection of previously undiagnosed disease, greater appreciation of the consequence of poor health habits such as smoking) is an intriguing question that remains for future research.

The use of the postdictive measures can also cause the coefficients of other variables that are not orthogonal to the health variables to be inconsistent. Postdiction generally causes the coefficients of LNAGE to decrease in absolute value and BLACK to increase in absolute value (Table 5.11). The largest changes occur in the Full List specification, where the coefficients for LNAGE and BLACK move by 1.4 and 0.6 standard deviations, respectively. In contrast, postdiction has almost no effect on the insurance coverage variables.

Conclusion

In answer to the first question we posed in this paper, we can say that the gain in explanatory power from using more comprehensive measures of health than those traditionally used is substantial. Compared with the simple EGFP question, the use of our most comprehensive definition of health was equivalent to an increase of around 10% in sample size. Much of this gain could potentially be achieved by sampling from those dimensions that predict utilization instead of employing the entire battery used in the HIS; e.g., one might include only two to four items from the current health scale rather than nine, and drop items related to the efficacy of medical care altogether. In general, the cost of a 10% increase in sample size will exceed the cost of collecting data on the more comprehensive measures of health (it would have greatly exceeded it in the

Table 5.11 Predictive and Postdictive Coefficients For LNAGE and BLACK

		Predictive		Postdictive	
Specification	Variable	β	$\|t\|$	β	$\|t\|$
EGFP	LNAGE	+.49	4.30	+.45	3.99
	BLACK	−.83	2.94	−.91	3.26
CURRENT	LNAGE	+.52	4.69	+.48	4.30
	BLACK	−.87	3.10	−.88	3.15
HPQTOT	LNAGE	+.54	4.84	+.49	4.45
	BLACK	−.86	3.08	−.84	3.01
HPQ components	LNAGE	+.57	4.90	+.56	4.89
	BLACK	−.82	2.95	−.83	3.06
Full list	LNAGE	+.54	4.38	+.37	3.14
	BLACK	−.77	2.85	−.92	3.46

Health Insurance Study), and so more comprehensive definitions seem cost effective.

In principle, more comprehensive definitions should also reduce omitted variable bias in other coefficients, but in our data these reductions were moderate to small. The estimated effect of coinsurance was essentially unchanged, but coinsurance was constructed to be orthogonal to other variables in our sample at the time of enrollment and so our result would not necessarily replicate in nonexperimental data. Likewise there was little effect on the estimated income elasticity, but in our sample the estimated income elasticity itself was near zero. We caution against generalizing this result; e.g., in the Dayton site, our results (not reported here) indicate an income elasticity of 0.2 to 0.4 in each of the first two years (Newhouse, Manning, Morris, et al., 1981). We did not include Dayton data in this paper because the health status measures prior to enrollment are not the same as those at the end of one year. Had we been able to analyze the Dayton data in a fashion similar to those of the other sites, we might have found that the inclusion of a more comprehensive health status measure did have an effect on the estimated income elasticity.

Including more health status variables has a moderate effect on the coefficient measuring race, but we must point out that there are few blacks in either Seattle or the two Massachusetts sites, and so this result may not replicate. We also found a moderate effect on the variable measuring sex, and we can think of no caveats about this result.

One can show in a simple theoretical model that the use of a health status measure from a later period to predict utilization in an earlier period (postdiction) leads to inconsistent estimates. We can detect this inconsistency in our data. Empirically, postdiction raised both the measured R^2 and the absolute values of the coefficients of the health status variables. Postdiction also increased the absolute value of the estimated coefficient for blacks and decreased it for age.

What practical advice can we offer? Many researchers will not be able to avoid the problem of analyzing health status variables from a period subsequent to their measures of utilization. If possible, one would like to treat such health status variables as endogenous. But it is difficult to think of good instrumental variables; a natural choice might be age, but age appears to have an independent effect on demand. At a minimum one should be aware of the problem.

Furthermore, one should use all the health status variables at hand, but the natural tendency would be to do this anyway, so such advice is not particularly helpful. Perhaps our findings about the gains from more comprehensive measures are most helpful to those who have a chance to influence the data that will be collected. At least for demand estimation—and probably for most other purposes as well—the resources necessary to

obtain more comprehensive health status measures than EGFP appear worth the sacrifice in sample size.

What future steps seem indicated in research on health status measures as covariates? We see two. First, the remaining measurement error in the covariates should be accounted for in coefficient estimates. Second, one might consider specifying health status as an unobserved or latent variable, thereby exploiting the covariance of demand and health status when deriving scales.

Appendix A

The Inconsistency of Postdiction

Let the true model be

$$U_t = \alpha HS_{t-1} + \gamma P_t + \mu_f + \varepsilon_t', \tag{A.1}$$

$$HS_t = \beta U_t + \delta HS_{t-1} + u_t, \tag{A.2}$$

where U_t is utilization in time t, HS_t is health status in time t, P_t is the price of medical services in time t, μ_f is a time-invariant family-specific error term and ε_t' and u_t are error terms that are independent of the explanatory variables and are not autocorrelated. Suppose t is measured in units of years and P_2 is orthogonal to HS_1. (This latter supposition is not necessarily true in nonexperimental data.)

Let period 1 be the pre-experimental period; we observe HS at the end of that period and denote it as HS_1. We have another observation one year later, which we will denote as HS_2. The utilization we seek to explain is U_2, i.e., utilization during the year bounded by the two observations on health status. Let all variables be measured as deviations from their means.

Postdicting U_2 means using HS_2 in place of HS_1 when estimating α and γ. Let a and g be estimates of α and γ in a postdictive model.

$$\operatorname{plim}\begin{pmatrix} a \\ g \end{pmatrix} = \begin{bmatrix} \operatorname{var} HS_2 & \operatorname{cov}(HS_2, P_2) \\ \operatorname{cov}(HS_2, P_2) & \operatorname{var} P_2 \end{bmatrix}^{-1} \begin{bmatrix} \operatorname{cov}(HS_2, U_2) \\ \operatorname{cov}(P_2, U_2) \end{bmatrix}$$

$$\operatorname{plim} g = \gamma + \frac{\gamma(\beta\gamma \operatorname{var} P_2)^2 - \gamma(\beta^2 \operatorname{var} P_2 \operatorname{var} U_2) - \gamma(\alpha\beta\delta \operatorname{var} P_2 \operatorname{var} HS_1)}{\operatorname{var} HS_2 \operatorname{var} P_2 - (\beta\gamma \operatorname{var} P_2)^2}$$

If $\beta = 0$ (medical care at the margin does not affect health status), the expression on the right is zero, and plim $g = \gamma$; g is consistent. If $\beta < 0$, the direction of the inconsistency cannot be signed with certainty. If $\beta > 0$, the direction of the inconsistency is toward zero. This can be shown as follows. Consider the fraction on the right side. Because the determinant of a variance-covariance matrix is positive, the denominator is positive. Moreover, the sum of the first two terms in the numerator is positive, since var $U_2 > \gamma^2$ var P_2 (by the definition of U_2) and γ is negative. Because γ is negative and δ is positive, the third term in the numerator has the sign of β; if β is positive (medical care improves health status), plim g will be biased toward zero.

$$\text{plim } a = \frac{\beta \operatorname{var} P_2 \operatorname{var} U_2 + \alpha\delta \operatorname{var} P_2 \operatorname{var} HS_1 - \beta(\gamma \operatorname{var} P_2)^2}{\operatorname{var} HS_2 \operatorname{var} P_2 - (\beta\gamma \operatorname{var} P_2)^2}$$

If $\beta = 0$, this expression becomes

$$\frac{\alpha\delta \operatorname{var} HS_1}{\delta^2 \operatorname{var} HS_1 + \operatorname{var} u} = \frac{\alpha}{\delta} - \frac{\alpha}{\delta}\left(\frac{\operatorname{var} u}{\delta^2 \operatorname{var} HS_1 + \operatorname{var} u}\right)$$

$$= \frac{\alpha}{\delta}\left(1 - \frac{\operatorname{var} u}{\delta^2 \operatorname{var} HS_1 + \operatorname{var} u}\right)$$

Because $0 < \delta < 1$, plim a could be greater or less than α. If $\beta \neq 0$, a term is added to this expression that is positive if β is positive and conversely.

Appendix B

Questionnaire Items Used to Construct Measures of Health Status, Attitudes, Behavioral Propensity, and Knowledgeability

Category	Item
Physical Limitations (PHYSLIM)	1) Are you able to drive a car?
	2) When you travel around your community, does someone have to assist you because of your health?
	3) Do you have to stay indoors most or all of the day, because of your health?
	4) Are you in bed or a chair for most or all of the day because of your health?
	5) Does your health limit the kind of vigorous activities you can do, such as running, lifting heavy objects, or participating in strenuous sports?
	6) Do you have trouble either walking several blocks or climbing a few flights of stairs because of your health?
	7) Do you have trouble bending, lifting, or stooping because of your health?
	8) Do you have any trouble either walking one block or climbing one flight of stairs because of your health?
	9) Are you unable to walk unless you are assisted by another person or by a cane, crutches, artificial limbs, or braces?
Chronic Diseases (DISEASE)	1) Has a doctor ever said you have glaucoma (increased pressure in the eye)?
	2) Can you usually hear and understand what a person says, without seeing his face and without a hearing aid, if he whispers to you from across a quiet room?

Appendix B (continued)

Category	Item
	3) Have you ever had hay fever or other allergies to plants and grasses?
	4)* Do you have any natural teeth at all? (Your own teeth, not artificial?)
	5)* Has a dentist ever told you that you have gum problems?
	6) In the past 12 months, have you had trouble with pimples on your face?
	7) Has a doctor ever said you had goiter (GOY-ter) or thyroid trouble?
	8) During the past 12 months, have you had any pain, aching, swelling, or stiffness in your joints—for example, your fingers, hip, or knee? (Do not count problems caused by an injury.)
	9) During the past 12 months, have you ever felt short of breath?
	10)* Has a doctor ever told you that you had an enlarged heart or heart failure?
	11) Has a doctor ever said that you had high blood pressure?
	12)* Has a doctor ever said that you had a heart attack?
	13)* Has a doctor ever said that you have angina? (An-JI-na or AN-ji-na)
	14) Has a doctor ever told you that you had chronic bronchitis (bron-KY-tis) or emphysema (em-feh-SEE-ma)?
	15) Has a doctor ever said that you had tuberculosis (T.B.)? (Tuberculosis pronounced "too-burr-cue-LO-sis")
	16) Has a doctor ever said that you had a peptic ulcer, stomach ulcer, or duodenal ulcer (ulcer of the small bowel)?
	17) Did a doctor ever say you had kidney disease?

18) Has a doctor ever said you have high blood cholesterol?
19) Has a doctor ever said that you had anemia (a-NEE-mee-a) or low blood?
20) Has a doctor ever said to you that you had diabetes or pre-diabetes?
21) Has a doctor ever told you that you had cancer?
22) Have you had hemorrhoids (piles) within the past 12 months? (Hemorrhoids pronounced "HEM-or-royds")
23) Have you had a hernia within the past 12 months?
24) During the past 12 months, have you noticed varicose veins in your legs?
25) Do you have any missing limbs—that is, arms, legs, or fingers that are missing or have been amputated?
26) As far as you know, during the past 12 months, have you had bursitis?
27) As far as you know, during the past 12 months, have you had arteriosclerosis or hardening of the arteries?
28) As far as you know, during the past 12 months, have you had chronic hepatitis or yellow jaundice?
29) As far as you know, during the past 12 months, have you had chronic gall bladder trouble or gallstones?
30) As far as you know, during the past 12 months, have you had phlebitis (thrombophlebitis)?
31) As far as you know, during the past 12 months have (women only) you had any disease of the uterus or ovary?
32) As far as you know, during the past 12 months have (women only) you had any lumps in your breasts?

Appendix B (continued)

Category		Item
Acute Physical Symptoms (ACSIL)	1)	During the past 30 days, did you have a cough, without fever, which lasted at least 3 weeks?
	2)	During the past 30 days, did you have a sore throat or cold, with fever, lasting more than 3 days?
	3)	During the past 30 days, did you have a weight loss of more than 10 pounds (unless you were dieting)?
	4)	During the past 30 days, did you have an upset stomach, for less than 24 hours?
	5)	During the past 30 days, did you have stiffness, pain or swelling of joints, lasting more than 2 weeks?
	6)	During the past 30 days, did you have backaches or sciatica?
	7)	During the past 30 days, did you have trouble falling asleep at night?
	8)	During the past 30 days, did you get up exhausted in the mornings, even with the usual amount of sleep?
	9)	During the past 30 days, did you have a skin rash, or breaking out on any part of the body?
	10)	During the past 30 days, did you have shortness of breath with light exercise or light work?
	11)	During the past 30 days, did you have chest pain when exercising?
	12)	During the past 30 days, did you have a stopped up nose, or sneezing or allergies for 2 weeks or more?
	13)	During the past 30 days, did you have swollen ankles when you woke up?
	14)	During the past 30 days, did you have headaches almost every day?

15) During the past 30 days, did you have a cough without fever, which lasted for less than a week?
16) During the past 30 days, did you have loss of consciousness, fainting, or passing out?
17) During the past 30 days, did you have acid indigestion or heartburn after many meals?
18) During the past 30 days, did you have a sprained ankle, but you could still walk?
19) During the past 30 days, did you have a toothache?
20) During the past 30 days, did you have stomach "flu" or virus (gastroenteritis) with vomiting or diarrhea?
21) During the past 30 days, did you have bleeding (other than nose bleed or periods) not caused by accident or injury?
22) During the past 30 days, did you have an eye infection?
23) During the past 30 days, did you feel nervous or anxious most of the time?
24) During the past 30 days, did you feel depressed or sad most of the time?
25) During the past 30 days, did you (men only) have difficulty passing urine or prostate trouble?
26) During the past 30 days, did you (women only) have difficulty controlling urine, or bladder or kidney problems?
27) During the past 30 days, did you (women only) have irregular periods, or bleeding between periods?

Emotional Instability (EMOINS)

1) How often did you become nervous or jumpy when faced with excitement or unexpected situations during the past month?
2) Did you feel depressed during the past month?
3) How much of the time, during the past month, have you been a very nervous person?

Appendix B (continued)

Category	Item
	4) During the past month, how much of the time have you felt tense or "high-strung"?
	5) During the past month, have you been in firm control of your behavior, thoughts, emotions, feelings?
	6) During the past month, how often did your hands shake when you tried to do something?
	7) How much of the time, during the past month, have you felt downhearted and blue?
	8) How often have you felt like crying, during the past month?
	9) During the past month, how often did you feel that others would be better off if you were dead?
	10) How much have you been bothered by nervousness, or your "nerves," during the past month?
	11) How often, during the past month, have you felt so down in the dumps that nothing could cheer you up?
	12) During the past month, did you ever think about taking your own life?
	13) During the past month, how much of the time have you felt restless, fidgety, or impatient?
	14) During the past month, how much of the time have you been moody or brooded about things?
	15) During the past month, how often did you get rattled, upset, or flustered?
	16) During the past month, have you been anxious or worried?
	17) How often during the past month did you find yourself having difficulty trying to calm down?

	18)	During the past month, how much of the time have you been in low or very low spirits?
	19)	During the past month, have you been under or felt you were under any strain, stress, or pressure?
Positive Well-being (PWB)	1)	How happy, satisfied, or pleased have you been with your personal life during the past month?
	2)	During the past month, how much of the time have you felt that the future looks hopeful and promising?
	3)	How much of the time, during the past month, has your daily life been full of things that were interesting to you?
	4)	During the past month, how much of the time have you generally enjoyed the things you do?
	5)	When you got up in the morning, this past month, about how often did you expect to have an interesting day?
	6)	During the past month, how much of the time has living been a wonderful adventure for you?
	7)	How much of the time, during the past month, have you felt cheerful, lighthearted?
	8)	During the past month, how much of the time were you a happy person?
Social Activities (SOCACT)	1)	About how many families in your neighborhood are you well enough acquainted with, that you visit each other in your homes?
	2)	About how many close friends do you have—people you feel at ease with and can talk with about what is on your mind? (You may include relatives.)
	3)	Over a year's time, about how often do you get together with friends or relatives, like going out together or visiting each other's homes?

Appendix B (continued)

Category	Item
	4) During the past mônth, about how often have you had friends over to your home? (Do not count relatives.)
	5) About how often have you visited with friends at their homes during the past month? (Do not count relatives.)
	6) About how often were you on the telephone with close friends or relatives during the past month?
	7) About how often did you write a letter to a friend or relative during the past month?
	8) In general, how well are you getting along with other people these days—would you say better than usual, about the same, or not as well as usual?
	9) How often have you attended a religious service during the past month?
	10) About how many voluntary groups or organizations do you belong to—like church groups, clubs or lodges, parent groups, etc. ("Voluntary" means because you want to.)
	11) How active are you in the affairs of these groups or clubs you belong to? (If you belong to a great many, just count those you feel closest to. If you don't belong to any, circle 4.)
POOR, FAIR, GOOD	1) In general, would you say your health is excellent, good, fair, or poor?
Current Health (CURRENT)	1) According to the doctors I've seen, my health is now excellent.
	2) I feel better now than I ever have before.
	3) I am somewhat ill.
	4) I'm not as healthy now as I used to be.
	5) I'm as healthy as anybody I know.

	6)	My health is excellent.
	7)	Doctors say that I am now in poor health.
	8)	I feel about as good now as I ever have.
Prior Health (PRIOR)	1)	I was so sick once I thought I might die.
	2)	I've never had an illness that lasted a long period of time.
	3)	I have never been seriously ill.
Health Outlook (OUTLOOK)	1)	I will probably be sick a lot in the future.
	2)	In the future, I expect to have better health than other people I know.
	3)	I think my health will be worse in the future than it is now.
	4)	I expect to have a very healthy life.
Resistance-susceptibility (RESIST)	1)	I seem to get sick a little easier than other people.
	2)	Most people get sick a little easier than I do.
	3)	I'm as healthy as anybody I know.
	4)	When there is something going around, I usually catch it.
Health Concern (CONCERN)	1)	Others seem more concerned about their health than I am about mine.
	2)	My health is a concern in my life.
Health Worry (WORRY)	1)	I never worry about my health.
	2)	I worry about my health more than other people worry about their health.
Sickness Orientation (ORIENT)	1)	Getting sick once in a while is a part of my life.
	2)	I accept that sometimes I'm just going to be sick.
Stressful Life Events (LCU)	1)	During the past 12 months, have you been fired, or laid off, from any job?
	2)	During the past 12 months, has there been any major change in your responsibilities at work?

Appendix B (continued)

Category	Item
	3) During the past 12 months, has there been any major change in your responsibilities at home?
	4) During the past 12 months, would you say that you have been arguing with each other more than usual, or less than usual?
	5) How about your in-laws—during the past 12 months, would you say you have been arguing with your in-laws more than usual, or less than usual?
	6) During the past 12 months, would you say that you and your girlfriend/boyfriend have been arguing with each other more than usual, or less than usual?
	7) At any time in the past 12 months, did you split up with your girlfriend/boyfriend?
	8) During the past 12 months, have you been arguing with your parents more than usual or less than usual?
	9) During the past 12 months, have there been any major changes in your personal habits—that is, the way you talk, dress, eat, or spend time?
	10) Within the past 12 months, did any close family member die?
	11) Within the past 12 months, did any close friend of yours die?
	12) During the past 12 months, have there been any major changes in your living conditions—like moving to a new place, or the neighborhood getting better or worse, or your house or apartment in better or worse shape?
	13) During the past 5 years, how many different homes (houses, apartments, trailers, etc.) have you lived in, including the one you live in now?

	14)	During the past 12 months, were you attacked or assaulted in any way by another person—like in a fist fight, or being beaten up or mugged?
	15)	During the past 12 months, did anyone rob or steal something from you—that is, rob you on the street or take money or property from your home or car?
	16)	During the past 12 months, were you in any kind of accident which involved property damage, but no personal injuries?
	17)	During the past 12 months, did you have any legal problems?
	18)	Did you change to a new school during the past 12 months?
	19)	Did you have to leave school during the past 12 months?
	20)	During the next 12 months, do you expect to have any problems making payments on any debts or financial obligations you have—like taxes, mortgage payments, consumer loans or installment debts?
Consumer Sophistication (CONSOPH)	1)	Some operations done by surgeons are not really necessary.
	2)	If you have doubts about your own doctor's advice, it's a good idea to get another doctor's opinion.
	3)	Stomachaches and headaches are hardly ever caused by your emotions.
	4)	A medicine prescribed by a doctor can have very different prices, depending on whether or not it has a brand name.
	5)	If you have to go into the hospital, your doctor can get you admitted to any hospital you prefer.
	6)	Doctors are checked every few years, before their licenses are renewed.
	7)	For many illnesses, doctors just don't have any cure.
	8)	Two doctors who are equally good at their job may still suggest very different ways of treating the same illness.

*Combined with one other item to score same disease.

Notes

1. The HIS is a social experiment designed, inter alia, to estimate the response of demand and health status to variation in the price of medical care services.

2. Elsewhere John E. Ware, Jr. has reviewed the literature on various measures of health status (Brook, Ware, Davies-Avery et.al. 1979). See also Aday and Eichorn (1972) and Freeburg, Lave, Lave et.al. (1979).

3. This is the wording in the 1963 Center for Health Administration Studies survey and the Health Insurance Study; the Health Interview Survey prefaces the question with: "Compared to other persons of your age, . . ."

4. Andersen and Benham are interested in the difference between measured and permanent income elasticities. We are interested in this difference as well, but have, at the moment, a relatively poor measure of permanent income. We have averaged two years of income (put in real terms), and so our measure should cut the variance of transitory income by a factor of two. Besides the issues discussed in the text, Hershey, Luft, and Granaris are interested in disaggregated measures of utilization, e.g., patient-initiated visits and check-ups.

5. We have not corrected our estimates for any remaining measurement error in the multi-item scales.

If errors in the responses to individual items are not perfectly correlated, error variance will be a smaller proportion of total variance, the greater the number of items in the scale (Nunnally 1967).

6. Construct validation can be very useful when an agreed-upon criterion variable does not exist or is not available for the measure being validated.

7. EGFP was entered as three indicator variables.

8. These expenditures reflect closely the pattern of visits to physicians.

9. In other analyses of these data, we have also attempted to control for other characteristics of the insurance policy. We dropped those variables for this analysis because they were insignificant. In particular, we had included a variable for the upper limit on out-of-pocket expenditures. The lack of significance does not necessarily mean that a cap on out-of-pocket payments has no effect on expenses; instead, it means that we did not discern such an effect in these annual data.

10. In other analyses we have used an ANOCOVA specification with indicator variables for each of the coinsurance plans (free, 25%, 50%, 95%, and individual deductible). The LC specification can explain more than 90% of the between-plan variation in expenditures.

11. Some families were given a screening exam at enrollment to improve the precision with which changes in physiological health could be measured at the end of the experiment. The results were reported to the family's physician. Because one could expect followup of abnormal results, a random (within plans) one-third were not examined.

12. The HIS mailed a diary to the families biweekly so that they could report disability days and information on medical utilization not contained on claims forms (e.g., telephone visits). Because the diary could stimulate both better reporting of utilization and more true utilization, we compared a random 25% of the Massachusetts sites who did not receive the diary with those who did.

13. The duration of the enrollment period was varied to help determine the effect of the length of enrollment on expenses.

14. Pesaran (1974) and Pesaran and Deaton (1978) proposed this variant of Cox's (1961, 1962) test for choosing among nonnested specifications.

Let $H_0: y = Xb_0 + u_0 , \quad u_0 \sim N(0, \sigma_0^2 I) ,$

$H_1: y = Zb_1 + u_1 , \quad u_1 \sim N(0, \sigma_1^2 I) .$

Then $N_0 = T_0 / [V(T_0)]^{1/2} ,$

where $T_0 = \frac{n}{2} ln[\hat{\sigma}_1^2/(\hat{\sigma}_0^2 + \frac{1}{n} b_0' X' M_Z X \hat{b}_0)]$,

$V(T_0) = \hat{\sigma}_0^2 \hat{b}' X' M_Z M_X M_Z X \hat{b}_0 / \hat{\sigma}_{10}^4$

$M_X = I - X(X'X)^{-1}X'$, $M_Z = I - Z(Z'Z)^{-1}Z'$,

$E(\sigma_{10}^2) = \sigma_0^2 + b_0' X' M_Z X b_0/(n - k_1)$

15. Comparisions of R^2 are not legitimate when the intrafamily correlation changes. If and only if such correlation is constant, the change in R^2 is a sufficient statistic for a change in the log-likelihood ratio. Note that when retransforming the logarithm of dollars to raw dollars the gain in precision is exponentiated.

16. Recall, however, that mental outpatient care is not included in these expenses. Some of this anomaly may reflect the separation of medical and mental treatment of emotionally related problems.

References

Acton, Jan P. 1975. Nonmonetary factors in the demand for medical services. *Journal of Political Economy* 83: 595–614.

———. 1976. Demand for health care among the urban poor, with special emphasis on the role of time. In Richard N. Rosett, ed., *The role of health insurance in the health services sector.* National Bureau Conferences Series, No. 27. New York: National Bureau of Economic Research.

Aday, Lu Ann, and Eichhorn, R. L. 1972. *The utilization of health services: indices and correlates.* DHEW Publ. No. (HSM) 73–3003. Washington: U.S. Government Printing Office.

Andersen, Ronald, and Benham, Lee. 1970. Factors affecting the relationship between family income and medical care consumption. In H. E. Klarman, ed., *Empirical Studies in Health Economics.* Baltimore: The Johns Hopkins Press.

Brook, R. H.; Goldberg, G. A.; Applegate, L. J.; et al. In press. *Conceptualization and measurement of physiologic health.* R-2262-HHS. Santa Monica, Cal.: The Rand Corporation.

Brook, R. H.; Ware, J. E.; Davies-Avery, A., et al. 1979. Overview of adult health status measures fielded in Rand's Health Insurance Study. *Medical Care* 17 (Supp.): 1–131.

Colle, Ann D., and Grossman, Michael. 1978. Determinants of pediatric care utilization. *Journal of Human Resources* 13 (Supp.): 115–58.

Cooper, Richard V., and Newhouse, Joseph P. 1971. *Further results on the errors-in-the-variables problem.* P-4715. Santa Monica, Cal.: The Rand Corporation.

Cox, D. R. 1961. Tests of separate families of hypotheses. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1.* Berkeley: University of California Press.

——— 1962. Further results on tests of separate families of hypotheses. *Journal of the Royal Statistical Society, Series B* 24: 406–24.

Davis, Karen, and Reynolds, Roger. 1976. The impact of medicare and medicaid on access to medical care. In Richard N. Rosett, ed., *The Role of Health Insurance in the Health Services Sector.* National Bureau Conference Series, No. 27. New York: National Bureau of Economic Research.

Davis, Karen, and Russell, Louise B. 1972. The substitution of hospital outpatient care for inpatient care. *Review of Economics and Statistics* 54: 109–20.

Donald, C. A.; Ware, J.E.; Brook, R. H., et al. 1978. *Conceptualization and measurement of health for adults in the Health Insurance Study: Vol. IV, Social health.* R–1987/4–HEW. Santa Monica, Cal.: The Rand Corporation.

Feldstein, Martin S. 1971. Hospital cost inflation: a study of nonprofit price dynamics. *American Economic Review* 61: 853–72.

——— 1977. Quality change and the demand for hospital care. *Econometrica* 45: 1681–1702.

Freeburg, Linnea C.; Lave, Judith R.; Lave, Lester B.; and Leinhardt, Samuel. 1979. *Health status, medical care utilization, and outcome: an annotated bibliography of empirical studies.* DHEW Publ. No. (PHS) 80–3263. Washington: U.S. Government Printing Office.

Fuchs, Victor R., and Kramer, Marcia. 1972. *Determinants of expenditures for physicians' services in the United States, 1948–1968.* New York: National Bureau of Economic Research.

Goldman, Fred, and Grossman, Michael. 1978. The demand for pediatric care: an hedonic approach. *Journal of Political Economy* 86: 259–80.

Grossman, Michael. 1972. *The demand for health: a theoretical and empirical investigation.* New York: Columbia University Press.

Grossman, Michael, and Benham, Lee. 1974. Health, hours and wages. In Mark Perlman, ed., *The economics of health and medical care.* New York: John Wiley.

Hershey, John C.; Luft, Harold S.; and Gianaris, Joan M. 1975. Making sense out of utilization data. *Medical Care* 13: 838–54.

Lau, R., and Ware, J. E. Forthcoming. *Refinements in the measurement of health-specific locus of control dimensions.* Santa Monica, Cal.: The Rand Corporation.

Manning, Willard G., Jr.; Morris, Carl N.; Newhouse, Joseph P., et al., 1981. A two-part model of the demand for medical care: preliminary results from the Health Insurance Study. In *Health, Economics and Health Economics,* eds. Jacques van der Gaag and Mark Perlman; Amsterdam; North Holland.

Manning, Willard G., Jr., and Phelps, Charles E. 1979. The demand for dental care. *The Bell Journal of Economics* 10: 503–25.

Newhouse, Joseph P. 1974. A design for a health insurance experiment. *Inquiry* 11: 5–27.

Newhouse, Joseph P., Manning, Willard G., Morris, Carl N. et al. 1981. Some interim results from a controlled trial of sharing in health insurance. *New England Journal of Medicine* 305 (in press).

Newhouse, Joseph P., and Marquis, M. Susan. 1978. The norms hypothesis and the demand for medical care. *Journal of Human Resources* 13: (Supp.): 159–82.

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.

——— 1976. New estimates of price and income elasticities. In Richard N. Rosett, eds., *The role of health insurance in the health services sector.* National Bureau Conference Series, No. 27. New York: National Bureau of Economic Research.

Newhouse, Joseph P.; Ware, John E.; and Donald, Cathy. 1981. How sophisticated are consumers about the medical care delivery system? *Medical Care* 19:316-328.

Nunnally, Jum C. 1967. *Psychometric theory.* New York: McGraw-Hill Book Company.

Pesaran, M. H. 1974. On the general problem of model selection. *Review of Economic Studies* 41: 120–31.

Pesaran, M. H., and Deaton, A. S. 1978. Testing nonnested nonlinear regression models. *Econometrica* 46: 677–94.

Phelps, Charles E. 1975. Effects of insurance on demand for medical care. In R. Anderson et al, eds., *Equity in health services.* Cambridge: Ballinger Publishing Company.

Stewart, A. L.; Ware, J. E.; and Brook, R. H. 1981. Advances in the measurement of functional status: construction of aggregate indexes. Medical Care 19:473–488

Ware, J. E. 1976. Scales for measuring general health perceptions. Health Services Research 11:396–415.

Ware, J. D.; Davies-Avery, A.; and Brook, R. H. 1980. *Conceptualization and measurement of health status for adults in the Health Insurance Study: Vol. VI, Analysis of relationships among health status measures.* R–1987/6–HEW. Santa Monica, Cal.: The Rand Corporation.

Ware, J. D.; Davies-Avery, A.; and Donald, C. A. 1978. *Conceptualization and measurement of health for adults in the Health Insurance Study: Vol V, General health perceptions.* R–1987/5–HEW. Santa Monica, Cal.: The Rand Corporation.

Ware, J. E.; Johnson, S. A.; Davies-Avery, A., et al. 1979. *Concep-*

tualization and measurement of health status for adults in the Health Insurance Study: Volume III, Mental health. R–1987/3–HEW. Santa Monica, Cal.: The Rand Corporation.

World Health Organization. 1948. Constitution. In *Basic Documents.* Geneva: World Health Organization.

Wu, De-Min. 1973. Alternative tests of independence between stochastic regressors and disturbances. *Econometrica* 41:733–50.

——— 1974. Alternative tests of independence between stochastic regressors and disturbances: finite sample results. *Econometrica* 42:529–46.