DEMAND FOR HEALTH INSURANCE: A THEORETICAL AND EMPIRICAL INVESTIGATION

PREPARED FOR THE OFFICE OF ECONOMIC OPPORTUNITY

CHARLES E. PHELPS

R-1054-OEO
JULY 1973

Rand
SANTA MONICA, CA. 90406

The research reported herein was performed pursuant to a grant from the Office of Economic Opportunity, Washington, D.C. Reports of The Rand Corporation do not necessarily reflect the opinions or policies of the sponsors of Rand research.

Published by The Rand Corporation

DEMAND FOR HEALTH INSURANCE: A THEORETICAL AND EMPIRICAL INVESTIGATION

PREPARED FOR THE OFFICE OF ECONOMIC OPPORTUNITY

CHARLES E. PHELPS

R-1054-OEO
JULY 1973

Rand
SANTA MONICA, CA. 90406

PREFACE

This study was taken from the author's doctoral dissertation at the University of Chicago. It is intended to contribute to the theory of demand for insurance and to provide empirical estimates of demand for insurance.

The empirical section of this report is unique in that it analyzes specific characteristics of insurance policies held by individual families. Other studies of health insurance demand have been forced to rely upon aggregate data, where either the proportion of a population insured or the fraction of medical expenses covered by insurance is studied.

A major problem with this study is that the data are from a pre-Medicare-Medicaid period, so the effect of those governmental programs on demand for insurance is not directly estimated. Studies with a data base covering 1970 experience should remedy this problem. Finally, a problem with the empirical estimates in this study is that the effects of medical care prices on insurance demand are now well estimated. Alternative methods of specifying the medical care price are being investigated currently, and will be reported as results are available.

This report is one of a series studying demand for insurance and demand for medical care at The Rand Corporation. Other studies include:

Kenneth J. Arrow, "Optimal Insurance and Generalized Deductibles," R-1108-OEO, February 1973.

C.E. Phelps, J.P. Newhouse, "Coinsurance and the Demand for Medical Services: Evidence from The Health Insurance Marketplace," R-964-OEO, April 1973.

C.E. Phelps, J.P. Newhouse, "The Effects of Coinsurance on Demand for Physician Services," R-976-OEO, June 1972.

Jan P. Acton, "Demand for Health Care Among the Urban Poor, with Special Emphasis on the Role of Time," R-1151-OEO/NYC, April 1973.

Jan P. Acton, "Demand for Health Care When Time Prices Vary More Than Money Prices," R-1189-OEO/NYC, May 1973.

For economists, this study extends current theories of demand for insurance to take into account the reimbursement nature of the insurance payment and shows how limiting the choice open to consumers alters

the nature of demand for insurance. The empirical estimates should also be of use to economists studying the health field. Particularly, they demonstrate the endogenous character of insurance demand, which researchers must take into account when estimating demand for medical care.

For policy purposes, the study has two important features. First, it describes the currently observed distribution of insurance in a consistent pattern, so that (although the data are from 1963) it is possible to infer where insurance coverage gaps may exist today. Second, it provides estimates of factors affecting demand that can be altered by policy formation, particularly the price of insurance. Since insurance is estimated to be a relatively elastic "good" (compared with demand for medical care *per se*), provision of insurance at low loading fees can be expected to extend insurance coverage broadly. Finally, the estimated insurance demand equations suggest in what patterns demand for supplemental insurance might be observed under Medicare and under potential universal health insurance plans that do not provide full coverage of all services.

SUMMARY

Theories of demand for insurance have only recently emerged in economic literature. The majority assume that (1) insurance payments are cash "indemnity" payments, with the amount of payments specifically tied to the occurrence of a given event, and (2) the insured losses are monetary. Health insurance differs from both of these considerations: Typically, insurance payments are *proportional to* the amount of *medical care* purchased. Furthermore, underlying "loss" is not necessarily financial; it may be considered as a loss from a stock of "health" that may be corrected through the purchase of medical care.

This report develops a theory of demand for reimbursement insurance accounting for both of the above-mentioned differences. The model considers a utility function in *health* (H) and in other goods (x), where the stock of health is a random variable. The consumer is said to act as if he maximized his utility $U = U(x,H)$ within the constraints of a budget I. The level of health consumed is said to be a function of the entering level of health (H_0), the loss (illness) that occurs (ℓ), and the amount of medical care (h) purchased to counteract the illness. A known production function g(h) is said to convert medical care (h) into health (H). The production function is nonstochastic.

The final level of health consumed is $H = H_0 - \ell + g(h)$. H is random since ℓ is random.

An insurance contract is assumed to be available to cover market medical care purchases at a contractual rate. The consumer is given the choice of selecting any coinsurance rate (C) between zero and one; C is the fraction of medical bills the consumer will pay during the period of the contract. The insurer agrees to pay the remaining (1-C) percent of all medical bills, up to a predetermined level h* of medical services. The consumer pays a known premium R for this contract, which covers his *expected* expenses. Arithmetically, R may be written as

$R = \int_0^{\ell^*} (1+\theta)(1-C)p_h h(\ell)\cdot f(\ell)d\ell$, where p_h is the price per unit of h, θ is a "loading fee" on the insurance contract, $f(\ell)$ is a known

distribution of the random variable ℓ, and $\ell^* = g(h^*)$.[1] This contract is written before either the consumer or the insurer knows which value of ℓ will be drawn from the distribution $f(\ell)$.

The consumer must choose some values for C and h* at the beginning of the period, a decision under uncertainty, and he must decide how to allocate his remaining budget I-R between x and h for the remainder of the period. He makes this decision *after* observing the drawing of ℓ from the distribution $f(\ell)$, and it is a decision under *certainty*. The consumption decision (with C and h* previously chosen) is analyzed through normal comparative statics of consumer choice, with a modified budget line. Explicitly, since the insurance contract is fixed for the period, the parameters C and h* (and the resultant premium R) are invariant during the period. There is an altered budget line of the form

$$I = p_x x + Cp_h h + R$$

as long as h does not exceed h*. Once h* units have been purchased, the budget line becomes

$$I = p_x x + p_h h - (1-C)p_h h^* + R.$$

While the insurance is in force, the consumer pays only a fraction of the prices for medical care--he pays only Cp_h rather than p_h per unit of medical care. However, he has given up R income, which *on average* is equivalent to (1-C) percent of his medical expenses.

The consumer's income is said to derive from labor market participation, so that $I = w \cdot T$ (w is the wage rate, and T is the *predetermined* amount of labor force participation during the year). Any medical care purchased is said to come out of labor force time, at the rate of t per unit of h. Thus there is an implied *time price* for any medical care purchased. The first-order conditions for maximizing utility, subject

[1]The loading fee is any premium on excess of expected benefits and is usually expressed as a percent of expected benefits. Hence, θ can be defined by θ = Premiums/Expected Benefits - 1.

to a budget constraint, show that[1]

$$U_H g'(h) + \lambda(Cp_h + w \cdot t) = 0,$$

which implies that the *total price* facing the consumer is the sum of the insured money price (Cp_h) and the time price ($w \cdot t$). A positive marginal productivity of medical care is also required to elicit purchases of medical care.

From the second-order conditions the comparative statics of demand for medical care are developed, and the effects of the insurance contract on this demand are explicitly derived. The elasticity with respect to total price (η_h^T) is derived, where total price = $Cp_h + w \cdot t$. Additionally, it is shown that the elasticities of demand with respect to money price (η_{hh}) and with respect to the coinsurance rate (η_{hC}) are systematically related to the "total price" elasticity. Ignoring income effects due to premium changes, it is shown that

$$\eta_{hC} \approx \frac{Cp_h}{Cp_h + w \cdot t} \eta_h^T,$$

and

$$\eta_{hh} \approx \frac{Cp_h}{Cp_h + w \cdot t} \eta_h^T .$$

In other words, the elasticity with respect to coinsurance is identical to the elasticity of demand with respect to medical care prices (except for income effects due to premium changes as C or p_h changes).

The effects of *time costs* on demand for h are similar. Where η_{hw} is the demand elasticity with respect to the wage rate, and η_{ht} is the demand elasticity with respect to travel time, it is shown that

---

[1]$(-\lambda)$ is the marginal utility of income (I), where λ is the Lagrangian multiplier in the constrained utility maximization.

$$\dot{\eta}_{hw} \approx \eta_{ht} \approx \frac{w \cdot t}{Cp_h + w \cdot t} \eta_h^T$$

when income effects due to premium changes are ignored.

These findings suggest that strongly insured (C near zero) services involving a high time cost will be observed as having a very low own-price elasticity of demand, since $\dot{\eta}_{hh}$ will approach zero as C approaches zero (with $w \cdot t$ positive). Additionally, goods and services for which the "total price elasticity" is high and for which the time costs are low may show demand highly responsive to insurance changes. Stated differently, reimbursement insurance may induce considerable demand for some services through a reduction of the apparent money price.

As in a "normal" price system, the effects of income on demand for health are not known *a priori* from the model. A positive income elasticity of demand is assumed for both other goods and health, which implies that the derived demand for medical care will be positive.

Upon suffering a loss from the level of health, a rational consumer will never fully replace the entire loss unless the demand for H is completely inelastic to income. This suggests that, over a multiperiod model, there will be gradual erosion of the level of health until death occurs. Death may be viewed as partly random, but partly self-determined, since the consumer always has the choice of increasing his level of H_0 through prior investment, but he is never certain that some critically large drawing from $f(\ell)$ may not be observed.

The conditions for selecting an optimal insurance policy are derived by maximizing the expression for expected utility with respect to the insurance parameters C and h*. It is assumed that *for each possible illness*, the consumer will act in a rational manner, taking into account the effects of his insurance choice on his demands for medical care. The simultaneous nature of demand for health insurance and demand for medical care is fully accounted for in this treatment--in buying insurance, the consumer considers what effects that insurance will have on his demands for medical care, and when actually buying

medical care, he considers the amount of insurance he holds as a determinant of his demand for h. Since the insurance must be purchased before the illness is observed, this is much like a dynamic programming problem--an optimal decision is computed for each illness in the "second period" (when medical care is demanded), and it is always assumed that at the point of the first decision (when insurance is selected) an optimal solution will be reached at the second decision.

The expected utility function is of the following form: E(U) =

$$s_0\, U\{x,H_0\} + \int_0^{\ell^*} U\{x,H_0-\ell+g[h(\ell)]\}f(\ell)\,d\ell + \int_{\ell^*}^{\infty} U\{x,H_0-\ell+g[h(\ell)]\}f(\ell)\,d\ell$$

where ℓ^* is the exact loss that will cause the consumer to purchase h* units of medical care (the maximum covered), and s_0 is the probability that no loss will be observed during the insurance period. This expression is maximized subject to the multiple budget constraints of the following form:

$$-I + p_x x + Cp_h h(\ell) + R = 0 \qquad \text{for } \ell \leq \ell^*$$
$$-I + p_x x + p_h h(\ell) - (1-C)p_h h^* + R = 0 \qquad \text{for } \ell > \ell^* .$$

The first-order conditions for maximizing E(U) with respect to C and h* are given as follows (dependence of h on ℓ is hereafter omitted to simplify notation):

$$s_0(-\lambda)\ (-R_c - p_h h) + \int_0^{\ell^*} (-\lambda)\ (-R_c - p_h h)\ f(\ell)\,d\ell$$
$$+ \int_{\ell^*}^{\infty} (-\lambda)\ (-R_c - p_h h^*)\ f(\ell)\,d\ell = 0 \qquad \text{and}$$
$$s_0(-\lambda)\ (-R_{h^*}) + \int_0^{\ell^*} (-\lambda)\ (-R_{h^*}) f(\ell)\,d\ell$$
$$+ \int_{\ell^*}^{\infty} (-\lambda)\ (-R_{h^*}) + (1-C)p_h) f(\ell)\,d\ell = 0 ,$$

where $(-\lambda)$ is the marginal utility of income, and is always dependent upon ℓ, and where R_c and R_{h*} are the partial derivatives of the premium R with respect to C and h*. The derivatives R_c and R_{h*} are functions

and may be viewed as an offering price schedule from which the consumer may select any value of C and h* desired. These "prices" reflect the rates of change of expenditure on insurance that will be brought about by decreasing C or increasing h*.

Manipulation of the first-order conditions shows that the ratio of expected "benefits" from each insurance parameter, divided by the appropriate change in the premium as a parameter changes, must be equalized across all possible choice parameters in the insurance policy. That is, the benefits from h* divided by the additional cost of another h of coverage must equal the benefits from lowering C by another "unit" divided by the cost of obtaining that additional coverage.

The effects of income on demand for reimbursement insurance are dependent not only upon the behavior of the risk aversion measure over different income levels (as is the case with indemnity insurance) but also on the behavior of the income elasticity of demand for medical care over income levels (or over illness levels). This result demonstrates that one cannot infer whether risk aversion is increasing or decreasing with income merely by observing the income elasticity of demand for health insurance--the effects are too complicated.

The effects of an increase in the loading fee on demand for reimbursement insurance are similar to the effects of other goods in a "standard" consumption problem. As long as absolute risk aversion increases with income, an increased loading fee always diminishes demand for insurance. If risk aversion decreases with income, then the possibility of insurance being a "Giffen good" arises. This result differs from the results found in a more general formulation of the insurance package because of the limited selection of "choice variables' open to the consumer. In a model where the final income in any state of the world may be determined by the consumer, an increased price of insurance always diminishes demand.

An increase in the price of medical care (p_h) has three separate effects on demand for C: The uncompensated price rise introduces an income effect, dependent upon whether risk aversion is increasing or decreasing. If demand for medical care is own-price inelastic, then total expenditures on medical care increase, causing an increase in

demand for insurance. Finally, there is a substitution effect because of the premium increase. The net of these three effects is ambiguous. It is impossible to determine from the model whether a price increase in the medical care sector will increase or decrease demand for insurance. Interactions among income, the budget share on medical care, and the effects of p_h on demand for C are suggested.

A functional differencing technique is used to demonstrate the effects of increased mean or variance in the distribution of illness on demand for coverage. If insurers can completely identify those persons with higher average illness levels, then the effects of those increases on demand for coverage levels are uncertain. However, if insurers cannot identify "sick" persons, then there is a three-fold adverse selection problem. First, those with higher than average expenses will by definition incur more medical expenses. Second, because insurance premiums will not change in the face of the increased incidence, more coverage will be purchased by these persons, unambiguously. Third, additional coverage will induce higher expenses for any given illness, other things equal, because of the further rotation of the price line facing these consumers. Moreover, the insurer has no particular way of guarding against such contingencies; the insurance covers such a broad range of illnesses, any of which might be observed in a given period, that the actual observation of a "high" loss does not necessarily identify a person as having a high mean illness distribution. An increase in the variance of illnesses, holding the mean constant, will increase demand for coverage, provided the loading fee is not increased to compensate for the additional risk facing the insurer.

Without explicitly solving for demand curves from any specific utility function, it is shown that the demand for insurance, and the demand for medical care given the insurance policy, can be expressed in demand curves of the following form:

$$C = C(I, p_h, p_x, w, t, \theta, f(\ell), g(h), H_o, Y, h^*),$$

$$h^* = h^*(I, p_h, p_x, w, t, \theta, f(\ell), g(h), H_o, Y, C),$$

$$h = h(I, p_h, p_x, w, t, \ell, g(h), H_o, R, C, h^*, F),$$

where Ψ is a vector of parameters from the utility function. Note that insurance demand is a function of the distribution of losses $f(\ell)$, whereas demand for medical care itself is a function of the *observed* loss ℓ, and of C and h*; all equations in the system are identified.

Two sources of data are used to analyze demand for insurance--a cross-section national health care survey, taken in 1963, and a time series of total insurance premiums and benefits.

The cross-sectional study uses a micro-level set of survey data, which give exact information on the health insurance policies held by 2367 families in the United States in 1963, to study the demand for hospital, surgical, and major medical insurance, by type and source of insurance. It is shown both theoretically and empirically that group purchasing arrangements "mask" the amount of variation in consumer demand for insurance.

The cross-sectional study uses OLS family level estimates, OLS estimates on grouped data, and "Tobit" regressions, which allow the dependent variable to take on limited values, rather than the usual assumption that the dependent variable may take on unlimited values. The grouped data estimates are used to remove transitory components in the measured income of the families.

The explanatory variables in the study are permanent income (an instrumental variable estimate is used); education (to represent time costs); the expected illness level (an instrumental variable estimate is used); the price of medical care (drawn from outside sources); the loading fee on insurance (represented by the size of the work group from which insurance is purchased); and standardizing variables of age, sex of family head, race, and urban/rural. The dependent variables are total premiums, the maximum payment per hospital day, the maximum number of hospital days covered, the maximum surgical payment, and the maximum major medical insurance payment.

The results show a significant positive income elasticity of demand for insurance, in the neighborhood of .2 to .4 for hospital insurance and .8 for major medical insurance. Surgical insurance is less responsive to income changes. There is a strong positive effect of time costs (represented by education) on demand for insurance; the

elasticity of total premiums is about .4, similar to the effect on hospital insurance and surgical insurance per se.

A major finding of the study is the large own-price elasticity of demand for insurance; total premiums rise as the loading fee on insurance falls, indicating that demand for insurance may be price-elastic. The separate components of demand for insurance are in some cases inelastic, but the aggregate (as shown by total premiums) is certainly elastic. Major medical insurance is extremely sensitive to price changes; the elasticity of the maximum payment under major medical insurance is larger than one.

The effects of medical care prices on demand for insurance are not estimated with high precision, and there are contradictory results when different estimation techniques are used. The grouped data estimates show that coverage levels increase when the price of medical care increases; total premiums rise, and the explicit coverage parameters (surgical and hospital maximums) increase faster than the prices (the elasticity is greater than one).

When the data are estimated at a family level (no aggregation) a different picture emerges: Total premiums fall in response to price increases for medical care, and the coverage parameters increase less rapidly than the price increases (the elasticity of hospital and surgical maximum payments is less than one). The different results found by these two methods of estimation have not been resolved. Moreover, the time-series data show similarly erratic patterns when the functional form of the equation is changed, so that no helpful evidence can be drawn from that source.

Total premiums increase significantly with increases in expected illness levels, but *coverage* parameters appear to be invariant to that measure. If insurance is priced to reflect anticipated expenditure, then premiums *should* rise with increased average losses from health; that coverage parameters do not increase is somewhat surprising. This result may reflect errors of measurement in the "true" level of illnesses anticipated by the family. The lack of precision of this measure is shown partly by the influence of age variables on demand; theoretically, there should be little effect from age on demand for coverage if all

other variables (including expected illness) are held constant. But age probably represents increasing illness levels in these regressions, and it usually enters strongly and positively in the demand estimates.

Holding insurance coverage constant, there is a systematic reduction in demand for insurance by blacks and by rural residents. These reductions result in an *implicit* increase in the loading fee on insurance, brought about by lower average expenditure on medical care by those groups, other things equal. Also noteworthy is that community rating systems (where all insurees pay the same premium, even though their expected expenditures are different) appear to discriminate against these groups, causing an income transfer from black insurees to white insurees, from rural to urban, from poor to nonpoor, and so on.

An overview of the demand estimates, comparing methods of estimation, shows that the OLS and Tobit estimates produce quite similar expected-value predictions when the frequency of purchases approaches unity. However, when insurance purchases are rarer (as in major medical insurance, for example), the Tobit estimation technique is superior to OLS or grouped data OLS. The predicted values from the Tobit technique are more precise, especially when the explanatory variables deviate significantly from the average values.

One of the subsamples in the study uses only verified data to arrive at estimates that are not dependent upon data points observed with error. A primary problem with using the entire data sample is that the verification was systematically not complete for families with high incomes and families in large work groups, both of which should imply higher levels of insurance. Use of the entire sample includes some observations with false zero values for the dependent variable and high values for income and group size. The estimated income and group-size elasticities should be biased if these data are included, and the amount of explained variation in the dependent variable should be lowered. Both of these phenomena were observed in the actual estimates; elimination of the unverified data increased the R^2 of the equations and both the income and group-size elasticities. The verified equations are estimated with high precision for micro-data demand studies; this is to be anticipated, since insurance itself abstracts from much of the random variation that surrounds demand for medical care.

A second subsample was taken to demonstrate how demand estimates changed because of group purchase of health insurance. Put simply, an imposed level of work-group insurance would place some families in a position of being overinsured. Typically, these families would not be able to convert this excess insurance into cash, so that they will keep the insurance plan, even if it is larger than they would select privately. The net result would be to bias the estimated income elasticity toward zero if the entire sample were used. Elimination of families who purchased insurance *only* through work groups should resolve this estimation problem--those remaining can be said to be on their "private" demand curves, and the estimates should provide a clear indication of the underlying demand patterns.

The time series data source runs from 1929 to 1968, drawing information on total insurance levels from insurance industry publications (Argus Charts and Best's). Explanatory variables (income, medical care prices, age distribution of the population, state of medical technology, and a dummy variable for Medicare years) are taken from annual issues of *Statistical Abstracts*. The loading fee is calculated from observed premiums and benefits. For the entire data sample, only ordinary least-squares estimation is possible, since crucial supply curve variables are not available until 1939. Ordinary least squares (OLS) estimates are performed, and a Cochrane-Orcutt iterative technique is used to remove autocorrelation. The results show demand for insurance to be positively associated with income and the medical care price level, negatively associated with the price of insurance, and positively associated with higher illness distribution means (represented by the age variables). Investigation of the residuals shows that autocorrelation is present and that the residuals do not appear to be distributed normally in all cases.

To correct these problems, a simultaneous equation system is estimated, allowing the loading fee to be endogenous. Added to the explanatory variable list are data on the interest rate, building costs, salary costs to insurance companies, and the proportion of all enrollees in work groups. The two-stage least squares (TSLS) estimates arrive at similar conclusions to the OLS with regard to the effects of the

explanatory variables on demand for insurance, but the effects of the loading fee are less strong. This suggests (as expected) that the OLS estimates overstated the effect of the loading fee on demand for insurance. A logit transformation of the coverage ratio is shown to best satisfy the normality assumption embodied in the regression specification.

The comparative statics showing the effects of insurance on demand for medical care should give more precise estimates of the effects of insurance on total demand. These estimates can be used to improve cost estimates under various insurance plans (either private or public). The interaction of insurance with time prices suggests how non-price rationing schemes might constrain demand even under fully insured medical care plans. Evidence is presented showing that the observed money price elasticity of demand for medical services is near -.10 to -.20, smaller for hospital visits, and approximately twice as large for home visits.

The estimates of demand for insurance suggest several other policy-relevant factors. First, the variation in demand for insurance is significant and follows a consistent pattern. If little or none of the variation in demand could be accounted for by a systematic economic model, then one might infer that demand for insurance was irrational, random, or imposed from outside sources. Some people feel that group provision of insurance takes the choice of insurance out of the consumer's hands, but the demand estimates from this study show otherwise; a significant amount of the variation in insurance coverage (roughly 40 percent) is accounted for by the economic variables. Any imposed level of insurance that does not allow for such variability in demand must place some consumers in a nonoptimal situation. If employers (in group insurance) or governments (in social insurance) wish to maximize consumer welfare, automatic provision of "full" insurance is not sufficient. The consumer must be allowed to share in the "savings" from not purchasing complete insurance.

A primary empirical finding of this study is the high own-price elasticity of demand for insurance. This suggests that provision of group insurance to those purchasing insurance through individual

contracts will strongly increase the amount of insurance demanded. Group insurance has certainly performed that function in the United States to date. A major policy question is how to pass along similar "low prices" of insurance to those who do not have access to normal work-group insurance. Some insurance plans use "open enrollment" periods, during which consumers may purchase individual contracts at group rates. (The limited period of open enrollment eliminated many of the problems of adverse self-selection that the insurers face in individual contracts.) Government intervention in this area has some precedent, Medicare for persons over 65 being the most obvious. Similarly, some states specify that all drivers must carry some form of liability insurance. The extremely high price sensitivity of insurance suggests that such policies would rapidly expand the amount of insurance held, even in the absent of compulsion.

The issue of insurance pricing is also of interest. The predominant "rating" techniques in the health field are either "experience" rating, or "community" rating. The former describes a plan where the premiums are based on expected costs (in the actuarial sense). A community rating scheme is where all persons enrolled in the plan pay the same premium, even if their anticipated expenses differ. Most group insurance plans, while experience rated as a group, imply community rating for those in the group. Community rating is shown to be equivalent to price discrimination, since the benefit/cost ratios differ for different persons in the plan. An odd feature of many community rating schemes is that they produce income transfers in a direction different from the (stated) intentions of their advocates. For example, since the income elasticity of demand for health care is positive, community rating produces an income transfer from the poor to the nonpoor. Recognition of the *actual* transfer mechanisms involved in community rating should improve the ability of decisionmakers to act in the health field.

The final area of discussion involves present insurance markets and potential future insurance markets. Although I have not examined this alternative in detail, on theoretical grounds, it is possible that prepayment schemes, where the insurance against risks and the

provisions of medical services are combined, may be superior to present reimbursement insurance arrangements, given imperfections in the information market and uncertainty in the production of health through inputs of medical care.

ACKNOWLEDGMENTS

The dissertation on which this report was based was initiated at the Center for Health Administration Studies (CHAS) of the Graduate School of Business, University of Chicago, and completed at The Rand Corporation. At The Rand Corporation, the author received support from an Office of Economic Opportunity grant, "Improving Financing of Health Services for the Poor and the Near Poor." The author's studies were partly financed by traineeship awards from the U.S. Public Health Service and by a grant from the National Center for Health Services Research and Development (HS00080), administered by CHAS. The author also received a Fellowship in Medical Economics from the Pharmaceutical Manufacturers Association for one and one-half years.

Many persons contributed directly and indirectly to this study: B. Peter Pashigian and Michael Grossman of the University of Chicago and John McCall and Joseph P. Newhouse of The Rand Corporation have been of particular assistance, although none are responsible for residual errors. A substantial portion of this report is devoted to analysis of the 1963 Health Survey data from a CHAS study directed by Odin W. Anderson and Ronald Andersen. Their generosity with these data and assistance with the study have been invaluable. David J. Weinschrott of The Rand Corporation and Harold Paschner of the University of Chicago provided valuable research assistance throughout this study.

CONTENTS

I. INTRODUCTION

THE PROBLEM FOR ANALYSIS

Economic analysis of consumers' purchases of insurance is a relatively recent phenomenon. It is generally recognized that two elements must be present for a rational consumer to purchase insurance: The world must present him with some randomness, and the consumer must dislike that randomness. But specific theories for insurance have only recently been developed.

This study will focus on the demand for health insurance. Health insurance is sufficiently different from other forms of insurance that the analytic results may be of interest. Furthermore, there are several bodies of data with which a theory of insurance demand may be tested. Finally, the topic is of current interest for policy analysis, with rapidly increasing medical care prices and accompanying proposals for national health insurance.

Two sets of phenomena relating to medical insurance will be analyzed—how consumers select their insurance policies and in what manner, if any, their insurance policies affect their purchases of medical care.

PREVIOUS ANALYSIS OF THE DEMAND FOR INSURANCE[1]

Recent analysis has extended our understanding of insurance demand considerably. A general theory of demand for insurance is that by

[1]Early work on this subject may be found in the following: Daniel Bernouilli, "Exposition of a New Theory on the Measurement of Risk," trans. in *Econometrica*, 22, January 1954; John Von Neumann and Oscar Morgenstern, *The Theory of Games and Economic Behavior*, Princeton University Press, 1944; John W. Pratt, "Risk Aversion in the Large and in the Small," *Econometrica*, 32, No. 1-2, January-April 1964, 122-136; Kenneth J. Arrow, *Theory of Risk Bearing*, Markham Press, Chicago, 1970, 212-219; B. Peter Pashigian, Laurence L. Schkade, and George H. Menefee, "The Selection of an Optimal Deductible for a Given Insurance Policy," *Journal of Business*, 39, No. 1, January 1966, 35-44; John P. Gould, "The Expected Utility Hypothesis and the Selection of Optimal Deductibles for a Given Insurance Policy," *Journal of Business*, 42, No. 2, April 1969, 145-151; Vernon L. Smith, "Optimal Insurance Coverage," *Journal of Political Economy*, 76, January-February 1968, 68-77; J. Mossin, "Aspects of Rational Insurance Purchasing," *Journal of Political Economy*, 76, August 1968, 553-568.

Ehrlich and Becker, not only considering how market insurance is demanded but analyzing demand for other means of reducing risk ("self-protection" and "self-insurance"), and showing how these latter mechanisms substitute for market insurance.[1] The exposition of this theory is helpful in considering the theory developed in Section II.

Consider a two-state world, with probabilities s_0 and s_1 for states 0 and 1. Endowed income in state 0 (the state with no loss) is given as I_0^e, and endowed income in the state *with* loss is given as I_1^e. The size of the loss is defined implicitly as $I_0^e - I_1^e = L$. The consumer may purchase any amount of insurance, defined as a net transfer of income from state 0 to state 1, at a fixed price $p = dI_0/dI_1$. The amount of income transferred (the amount of insurance) is given as $I_1 - I_1^e = S$. The premium paid for the insurance (in the good state) is defined as $I_0^e - I_0 = R$. By definition, $R = S \cdot p$, where p is the price per unit of insurance. The final incomes are given as I_0 and I_1.

The appropriate amount of insurance is chosen by maximizing expected utility over all states of the world. This is given as:

(1.1) $$U^* = s_0U(I_0) + s_1U(I_1),$$

which is maximized subject to the constraint that $R = S \cdot p$. The first order condition is found to be:

(1.2) $$p = \frac{s_1U_1'}{s_0U_0'} = -dI_0/dI_1 \ .$$

The effects of a change of endowed wealth are shown by Ehrlich and Becker to be dependent upon how risks change and upon the relative risk aversion implied by the utility function.

---

[1] Isaac Ehrlich and Gary Becker, "Market Insurance, Self-Insurance, and Self-Protection," Journal of Political Economy, 80, No. 4, July-August 1972, 628. Self insurance is any activity that reduces the size of a loss; self protection is any activity that reduces the probability of a loss.

It is particularly helpful to consider how a change in the price of insurance alters the optimal amount of insurance.

$$\frac{dS}{dp} = \frac{dI_1}{dp} - \frac{dI_1^e}{dp} = \frac{dI_1}{dp} \propto -s_0U_0' + S \cdot p \cdot U_0''(s_0), \tag{1.3}$$

which is always negative for utility functions with decreasing marginal utility for income. That is, an increase in the terms of trade of income between state 0 (no loss) and state 1 (with loss) always reduces the amount transferred to state 1.

In a multiple state problem, the Ehrlich-Becker model is nearly identical. The income prospect is given as

$$s_0U(I_0) + s_1U(I_1) + \ldots + s_nU(I_n), \tag{1.4}$$

and there are (n-1) prices of insurance defined as $p_i = -dI_i/dI_0$. Expected utility is maximized by selecting the (n-1) amount of insurance that transfers income from state 0 (no loss) to states i, subject to the constraint that

$$R = \sum_{i=1}^{n} S_i \cdot p_i \,. \tag{1.5}$$

Note what happens when the ith amount of insurance changes. Since insurance is defined as *net* income transfer, it may be written as $S_i = B_i - R$, where B_i is the "gross benefit" in the ith state of the world. An increase in S_j, holding all other S_i constant, implicitly *increases* all other B_i, since the premium R increases by $p_j dS_j$.

Note also how the Ehrlich-Becker model structures the potential insurance purchased by individuals--each state of the world with loss involves a separate decision about amounts of insurance, so that in an n-state world, there are actually (n-1) insurance decisions made jointly.

An alternative structuring of insurance policies might alter insuring decisions. Health insurance policies give the consumer relatively few decision variables. For example, major medical insurance policies have only a deductible, a copayment rate, and a maximum payment covered by the policy. Many hospitalization policies have only two or one variables of choice. Further, nearly all health insurance policies in the United States are of reimbursement rather than of indemnity form. The Ehrlich-Becker model specifies insurance payments as direct income transfers (indemnities) as do other theoretical pieces in the economics of insurance. Reimbursement insurance, on the other hand, makes payment of insurance benefits conditional on, and proportional to, purchases of medical care. In effect, reimbursement insurance makes insurance payments in kind, and alters relative prices (at time of purchase) between medical care and other goods. I have chosen to analyze the demand for insurance under this restricted set of market opportunities to the consumer, and I will severely restrict the number of available decision variables to the consumer.[1] The results from such a model may be considered as a boundary case, compared with the Ehrlich-Becker paper. I allow the consumer to select *one* coverage rate, which will apply across all states of the world, and a maximum payment level for the insurance policy. At the other extreme, the Ehrlich-Becker model allows the consumer much more flexibility in his choice matrix--he may make (n-1) choices in an n-state world. Obviously, the two models

---

[1]It may be that limited choice policies have arisen, with no competition from a theoretically more desirable multiple-choice policy, because in health care, the range of states of the world is sufficiently large that transactions costs forbid making (n-1) decisions. Alternatively, it is extremely difficult to determine or measure which state of the world in fact exists when a person becomes sick. The "loss" is an illness, and expenditures on medical care are a *response* to that loss. Indemnity policies could not be written transferring a given amount of income when a given state occurs, if in fact one cannot identify which state of the world has occurred. I believe reimbursement insurance has arisen in response to this problem. Another way of looking at this is to say that transaction costs are sufficiently high in indemnity insurance (for health losses) that the loading charges become prohibitive. Reimbursement insurance, with a fixed payment rate across all states of the world, may have arisen as a "second best" alternative to the possibility of no market insurance for health losses.

coincide when n=2 (the case they initially consider in their paper). To show the differences that emerge, I consider a multiple-state world. A formal model of demand for reimbursement insurance will be developed in Section II, with particular emphasis not only on demand for insurance but on how reimbursement insurance alters demand for medical care. Whenever possible, the results from this (restricted decision-space) model will be compared with the more complete decision space model allowed by Ehrlich and Becker.

Many of the alternatives to market insurance (for example, self-insurance and self-protection activities produced at "home") considered by Ehrlich and Becker will not be taken up in this model. Their results may be said to be definitive with respect to that set of alternatives to market insurance. I concentrate here instead on the effects of restricting the decision space of the consumer and of introducing reimbursement insurance rather than indemnity insurance. I also concentrate on the effects of certain insurance price-setting mechanisms observed in the market, particularly group purchase of insurance (wherein all members of the group pay the same price for a given policy) and of *community* rating procedures, wherein all consumers in a given geographic area are charged the same amount for a given insurance policy. (The two methods of pricing differ only in scale; community rating policies involve larger numbers of people than work group policies.)

II. A THEORY OF REIMBURSEMENT INSURANCE

SPECIFICATION OF THE MODEL

Consider a consumer with a utility function $U = U(x,H)$ where x is a "market basket" of consumption goods, and H is a flow of "healthy days," which may vary both in quantity and quality.

This flow of services H is subjected to random losses. In other words, the rational consumer occasionally becomes ill. Since that consumer has utility for healthy days, he may wish to counteract some of his illness and increase his consumption of H. A repair and replacement service, h, is assumed to be available for this purpose. This market good, medical care, is available at a price p_h per unit of h. A technological production function g(h) will specify the relationship between the amount of medical care (h) purchased and the amount of health (H) obtained. This function g(h) will have normal production function characteristics:

$$g'(h) > 0 , \qquad g''(h) \leq 0 .$$

For any period, the consumer's final consumption of health can be expressed in terms of his beginning level of health, less any random losses that occur, plus any medical care inputs he purchases. That is,

$$H = H_0 - \ell + g(h) , \tag{2.1}$$

where ℓ is a random variable, and g(h) is the production function transforming medical care into healthy days. The random variable ℓ is distributed by:

$$d(\ell) = \begin{cases} s_0 & \text{for } \ell = 0 \\ f(\ell) & \text{for } \ell > 0 , \end{cases} \tag{2.2a}$$

and $f(\ell)$ is normalized so that

$$s_o + \int_0^\infty f(\ell)d\ell = 1 . \qquad (2.2b)$$

In principle, there is no reason why the level of health cannot be studied as if it were any other flow of goods subject to random losses. Consumption of health may be different because it cannot be directly purchased in the market, or sold or transferred once it is acquired. But this is true of many goods, and a branch of economic theory has been developed to analyze such goods.[1] Of more importance is the difficulty in measuring the actual level of H without incurring large transactions costs. The "level of health" is probably non-insurable except at prohibitive cost. The marketplace has instead produced an insurance that offsets expenditures on medical care, the level of insurance benefits being dependent upon the amount of "h" purchased. This model shows that the quantity of "h" purchased is a function not only of the random loss ℓ but of consumer income and prices in the marketplace. This formulation is a departure from previous "random wealth loss" models, which have assumed that total *monetary* loss in any state (here equivalent to the expenditure $p_h h_i$) is fixed for any consumer and is independent of his income, the prices he faces, or his insurance policy. I have introduced consumer choice as an element determining the final monetary expenditure (or "loss").[2]

---

[1] See, for example, Gary S. Becker, "A Theory of the Allocation of Time," *The Economic Journal*, 75, September 1965, 493-517; and Gary S. Becker, *Economic Theory*, Alfred A. Knopf, Inc., New York, 1971.

[2] In the Ehrlich-Becker formulation, the *size* of the loss can be reduced by the consumer through expenditures on self-insurance. This notion is quite comparable to my formulation of the "loss" as a deliberate, rational expenditure; I delineate predictable changes in this "loss" as certain exogenous variables change, using normal demand theory. Similar results could be inferred from the Ehrlich-Becker analysis of "self-insurance." The major difference between their formulation and mine, a difference to be noted in several places, is that I restrict the decision-space of the consumer by forcing him to choose one coverage level to apply to all losses.

Assume that the consumer's budget for the period is previously set at amount I. Then his budget line in the market is:

$$I = p_x x + p_h h ,$$

which may be pictured on an indifference map in x and h as shown in Fig. 1. In this map, the slope of the budget line is given by $-p_h/p_x$.

Since the consumer's utility is a function of H, the above-budget line may be transformed into an "iso-budget" line in x and H by inverting equation (2.1),

$$h = g^{-1}(H - H_0 + \ell) , \tag{2.3}$$

which establishes the derived demand for h, given H_0 (the entering level of health) and ℓ (the loss that has occurred). Demand for h is always conditioned by a specific level of loss ℓ in this model. The consumer's "iso-budget" line in x and H then becomes:

$$I = p_x x + p_h g^{-1}(H - H_0 + \ell) . \tag{2.4}$$

The slope of this "iso-budget" line may be determined through total differentiation of (2.4), keeping I constant:

$$p_x \partial x + p_h (g^{-1})' \partial H = 0 , \tag{2.5a}$$

so that

$$\frac{\partial x}{\partial H} = \frac{-p_h (g^{-1})'}{p_x} = \frac{-p_h}{p_x} \cdot \frac{1}{g'} \qquad \text{if } g' \neq 0 . \tag{2.5b}$$

As long as g' is positive, this "iso-budget" line has the usual

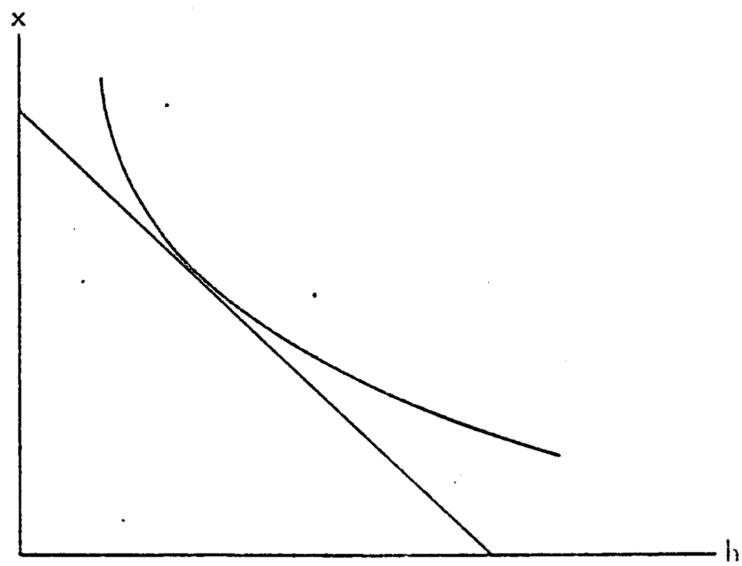

Fig. 1 — Indifference map in other goods and medical care

tradeoff characteristics: more x can be acquired only by giving up some H, and vice versa.

Differentiating again, these "iso-budget" lines are shown to be nonconvex, thus assuring a unique solution between those budget lines and a convex set of indifference curves. Differentiating equation (2.5b) with respect to H:

$$(2.6) \qquad \frac{\partial^2 x}{\partial H^2} = \frac{\partial}{\partial H}\left[\frac{-p_h}{p_x} \cdot \frac{1}{g'}\right] = \frac{-p_h}{p_x}\,(-1)(g'(h)^{-2}g''(h)) \quad \frac{dh}{dH}$$

$$= \frac{p_h}{p_x}\frac{g''(h)}{g'(h)^2} \cdot (g^{-1})' = \frac{p_h}{p_x}\frac{g''}{g'^3}\ ,$$

which is negative or zero as g" is negative or zero, and is in appropriate units of x/H^2. In the case when g" = 0, the budget lines in (x,H) differ merely by a scale factor from those in (x,h). If g" is negative, an indifference map in x,H would appear as in Fig. 2.

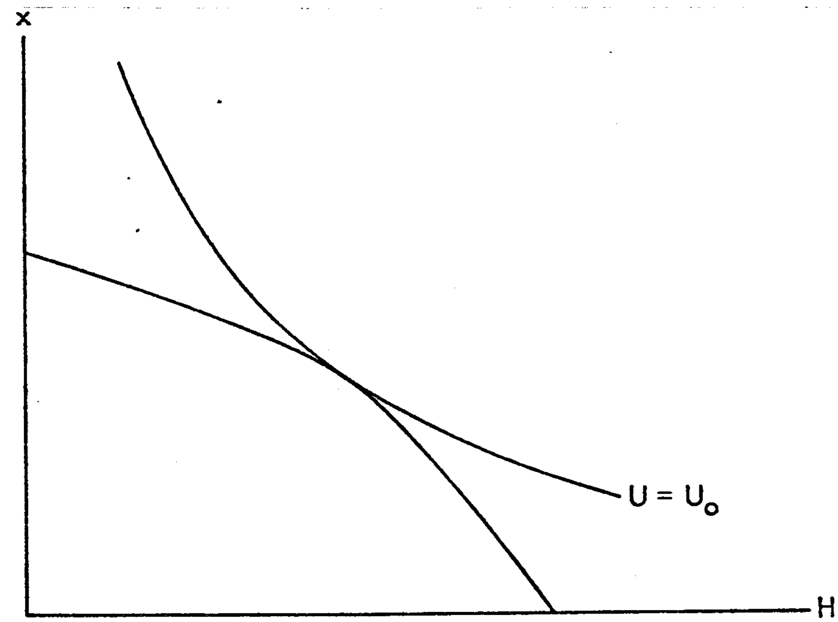

Fig. 2 — Indifference map in other goods and health

THE INSURANCE CONTRACT

Consider now the effect of an insurance policy previously purchased by the consumer. The insurance policy has two "decision variables," which the consumer can select when he buys the policy. The consumer pays Cp_h per unit of h, and the insurer will pay $(1 - C)p_h$ per unit. Once the consumer's purchases exceed h* units of medical care, his insurance policy is no longer effective, and he pays p_h per unit for all $h > h^*$. The insurance premium (R) is in part determined by the particular C chosen by the consumer. The premium may be written as:

$$R = R(C, h^*, \theta, \delta) \tag{2.7}$$

where θ is a loading charge on the insurance, and δ is a vector of parameters that systematically influence the premium (age, family size, occupation, and so on). An example of such a policy would be a hospitalization insurance contract that paid 80 percent of all expenditures up to 120 hospital days, and nothing thereafter. Such a policy has C = .20 and h* = 120 hospital days. The premium for this insurance policy may

be written as:

$$(2.8) \qquad R = (1 + \theta) \int_0^{h^*} (1 - C)p_h h \, \gamma(h)dh \; ,$$

where $\gamma(h)$ is the distribution of expenditures observed by the insurance company. If insurance reflects expected expenditure of a given individual, this is equivalent to:

$$(2.9) \qquad R = (1 + \theta) \int_0^{\ell^*} (1 - C) \, p_h h(\ell) \, f(\ell)d\ell \; ,$$

where ℓ^* is the exact loss such that $g(h^*) = \ell^*$. It is assumed throughout the remainder of this report that equation (2.9) holds as the appropriate definition of the insurance premium, unless specific indication is given to the contrary.[1]

The rate of change of the premium (R) with respect to the two insurance parameters C and h* may be said to be the marginal price of a unit of insurance. That is, the change in total expenditure (R) on insurance with respect to C is:

$$(2.10) \qquad R_c = \frac{\partial R}{\partial C} = (1 + \theta) \int_0^{\ell^*} \left[p_h \frac{\partial h}{\partial c} (1 - C) - p_h h(\ell)\right] f(\ell)d\ell \; ,$$

where $\partial h/\partial C$ is not yet defined but will be derived from economic theory. The distribution $f(\ell)$ is assumed not to be affected by the insurance parameter C. The function $-R_c$ may be thought of as a price schedule

---

[1]There may be occasions when insurers are unable to identify the exact illness distribution facing a given consumer, or insurers may deliberately choose to compute average expenditures over some group of individuals and charge each person in the group according to average experience of the group. This rating system is prevalent in group insurance policies and is also the basis of many nonprofit health insurers' rating systems.

See Appendix B for a complete discussion of R and all associated derivatives.

offered by the insurer to the individual in question, describing how R changes as the consumer varies C. This is an "honest" schedule in the sense that any change in expenditure by the consumer is fully accounted for by changes in the premium R. No dishonesty or differential information (between insurer and insured) is permitted; the premium R accounts completely for the *expected* experience of the consumer over all losses between 0 and ℓ^*.[1]

Similarly, the partial derivative of R with respect to h* gives the rate at which R changes with h*:

$$R_{h^*} = \frac{\partial R}{\partial h^*} = (1 + \theta)(1 - C)p_h h^* f(\ell^*) \quad (2.11)$$

This derivative may be considered to be the "price" of the (h*)th unit of coverage. It reflects the expenditure at the loss ℓ^* (that is, $[1 - C]p_h h^*$) times the probability density that the loss ℓ^* will occur.[2]

THE CONSUMPTION PROBLEM—A DECISION UNDER CERTAINTY

The decision to purchase medical care (h) and other goods (x) is now formulated in a manner amenable to normal economic analysis. The consumer is said to maximize utility from goods, subject to a constraining income. He has the choice, with reimbursement insurance, to alter his budget line (and spendable income) in advance of knowing which illness ℓ occurs. Once the drawing from $d(\ell)$ is known, however, he may

---

[1] Although the difference will be stressed later, this definition of "insurance prices" is very different from that of Ehrlich and Becker. In their model, the quantity of insurance is determined for each state, and the price per unit of insurance is assumed to be constant. Here, the *average* effect of a change in C is what determines the price. Since the consumer must select one C to apply across all states of the world where the loss is below ℓ^*, he cannot (as in the Ehrlich-Becker model) choose his final income in each state of the world, and there is not in fact a "price" for insurance in each state of the world.

[2] Not uncommonly, health insurance policies also have available for choice a *deductible* amount (above which expenditures must rise before C becomes effective). I omit this variable from the model primarily on the basis of expositional simplicity. Both for medical expenses less than the deductible D or above the maximum payment h*, the consumer care and demand can be analyzed similarly in either circumstance.

proceed to allocate his budget in normal fashion, taking into account the size of the loss from his health level (H) that has actually occurred. In this decision, the insurance policy is viewed as having been chosen prior to the actual drawing from $d(\ell)$, so that C and h* may be regarded as *fixed* for the consumption decision. After the insurance policy has been purchased, the budget is (I-R) rather than I, and the relevant budget lines are:

$$I - R = p_x x + C p_h h \qquad \text{for } h \leq h^* , \tag{2.12a}$$

$$I - R = p_x x + p_h h - (1 - C) p_h h^* \qquad \text{for } h > h^* . \tag{2.12b}$$

The consumer's income is said to derive from market labor participation, with the wage rate w being given and total market participation time T being determined outside of the system considered.[1] Where t is the amount of time necessary to purchase a unit of h, the effects on I of purchasing a unit of h are given as:

$$\frac{dI}{dh} = w \cdot \frac{dT}{dh} = - t \cdot w , \tag{2.13}$$

since $T = T_0 - t \cdot h$ by definition.[2]

To pick the optimal x and h, given the loss ℓ, income (before purchase of h), the production function g(h), and market prices p_x and p_h,

[1]The entire problem of time allocation in and out of the market and of nonmarket production of health and other goods is deliberately ignored here. I assume that T_0 has been predetermined by the consumer and will not change in the face of illnesses. These considerations are ignored for simplicity--time allocation in general is an extensive subject, treated in considerable detail by others. I focus here on the effects of reimbursement insurance and hold time allocation problems in abeyance.

[2]It is also conceivable that income (I) may depend directly on the level of health (H). If that is true, then a further component in dI/dh appears, and that derivative would be written as $dI/dh = -tw + (\partial I/\partial H)\, g'(h)$, where g'(h) is the marginal productivity of medical care in producing H. To the extent that H increases market time, these effects will offset the "time costs" of purchasing medical care. $\partial I/\partial H$ is set equal to zero in this model.

the consumer must maximize the Lagrangian expression Z with respect to x, h, and λ.

(2.14a) $$\max_{x,h,\lambda} Z = U(x,H) + \lambda(-I + p_x x + Cp_h h + R) \qquad \text{for } h \le h^* ,$$

or, if the computed value of h is above h*, then the consumer uses:

$$\max_{x,h,\lambda} Z = U(x,H) + \lambda(-I + p_x x + p_h h + R - (1 - C)p_h h^*) \qquad \text{for } h > h^* .$$

(2.14b)

The first order partials of (2.14) are set to zero to maximize utility:

$$(2.15) \qquad \left.\begin{aligned} &(a) \quad \frac{\partial Z}{\partial x} = U_x + \lambda p_x = 0 \\ &(b) \quad \frac{\partial Z}{\partial h} = U_H g'(h) + \lambda(Cp_h + w \cdot t = 0 \\ &(c) \quad \frac{\partial Z}{\partial \lambda} = -I + p_x x + Cp_h h + R = 0 \end{aligned}\right\} \quad \text{for } h \le h^* .$$

If he exceeds h*, the constraining budget changes, and the first-order maximizing conditions are:

$$(2.16) \qquad \left.\begin{aligned} &(a) \quad \frac{\partial Z}{\partial x} = U_x + \lambda p_x = 0 \\ &(b) \quad \frac{\partial Z}{\partial h} = U_H g'(h) + \lambda(p_h + w \cdot t) = 0 \\ &(c) \quad \frac{\partial Z}{\partial \lambda} = -I + R - (1 - C)p_h h^* + p_x x + p_h h = 0 \end{aligned}\right\} \quad \text{for } h > h^* .$$

For any loss ℓ, there is an implied pair of demand functions for x and h, of the following general form.

(2.17a) $$x = x(p_x,\ p_h,\ w,\ C,\ h^*,\ I,\ R,\ H_0,\ \ell,\ \psi,\ \Gamma) ,$$

(2.17b) $$h = h(p_x,\ p_h,\ w,\ C,\ h^*,\ I,\ R,\ H_0,\ \ell,\ \psi,\ \Gamma) ,$$

where ψ is a vector of parameters of the utility function, and Γ is a vector of parameters including age, occupational characteristics, and so on, that influence the function $g(h)$. It should be stressed that x and h are dependent, not only upon income and market prices, but also upon the observed loss ℓ, the entering level of health H_0, *and* the insurance parameters previously chosen. The dependence of demand for x and h upon the parameters C and h* is of primary importance, for it means that, in selecting optimal insurance, the consumer must take into account the effects of C and h* on demand for x and h (and consequent levels of x and H).

The comparative statics of the system (2.14) may be derived just as in any other utility maximizing framework. By differentiation of (2.14) with respect to any exogenous variable, one may ascertain the effects of that variable on the demand for x and h and upon the marginal utility of income $(-\lambda)$. Variables that may be considered as exogenous in this framework are initial income (I), the wage rate (w), the prices p_x and p_h, the level of health upon entering the period (H_0), the production function $g(h)$, and the illness ℓ.[1]

A major difference between this system and one with no insurance is that the budget line has been altered by the insurance contract. Since C and h* are fixed before the drawing of ℓ, one may treat them as being held constant for purposes of solving this problem, and in fact I may analyze what effects would occur if a different C or a different h* had been chosen by the consumer, for whatever reason. Effects of reimbursement insurance have long been discussed by economists interested in the health field,[2] but to my knowledge, no one has

---

[1]Note that x, h, and $(-\lambda)$ are all explicitly functions of the loss ℓ, and could be written $x(\ell)$, $h(\ell)$, and $-\lambda(\ell)$. For notational simplicity, I suppress the functional dependence of those variables on ℓ. Although ℓ (or $E(\ell)$) may alter w, I make no allowance for that here.

[2]See, for example, Mark Pauly, "The Economics of Moral Hazard," *American Economic Review*, 58, June 1968, 531-537; or Michael Crew, "Coinsurance and the Welfare Economics of Medical Care," *American Economic Review*, 59, No. 5, December 1969, 906-908.

systematically attempted to work out the implications of this form of insurance in a formal utility maximizing framework.[1]

DISCUSSION OF THE COMPARATIVE STATICS

The comparative statics of the medical care purchasing decision may be divided into two groups, those reflecting changes in x and h as true exogenous variables change (such as income, prices, and the loss ℓ), and those reflecting changes in x and h if different insurance variables C and h* were chosen. In these derivatives, the changes are *partial* effects on demand for x and h, induced by changes, say, in C. But demand for C is also affected by such things as income, p_h, and H_0. Response of demand to changes, say, in p_h would normally be expressed in complete terms as

$$\frac{dh}{dp_h} = \frac{\partial h}{\partial p_h} + \frac{\partial h}{\partial C} \cdot \frac{\partial C}{\partial p_h} . \tag{2.18}$$

The comparative statics in Appendix A show the *partials* appearing on the right-hand side of the full derivative dh/dp_h. The full behavior of demand in face of price changes is the sum of the two, combined with effects of changes in C as p_h changes.[2] The "normal" comparative statics effects of income and price changes are quite similar in form to those same derivatives in a model with no reimbursement insurance.

---

[1]The comparative statics from this formal utility maximization procedure are derived in C. Phelps and J. Newhouse, *Coinsurance and the Demand for Medical Services,* R-964-OEO/NC, The Rand Corporation, April 1973. The results are summarized in Appendix A and are discussed below.

[2]Note that if $\partial C/\partial p_h$ is sufficiently negative (implying that more coverage is purchased in response to price increases), then the net effect of an increase in p_h could conceivably be to increase demand. This situation is somewhat unlikely, but it does point out how demand estimation of medical care could be led astray if effects of reimbursement insurance are not fully considered.

Effects of Income on Demand for Goods

As in the case of the no-insurance model, the derivatives $\partial x/\partial I$ and $\partial h/\partial I$ cannot be signed in theory. Both are assumed to be positive. The rationale for this assumption is that in a world with two goods, at least one must be a normal good. It would seem inappropriate to specify x as inferior, since it represents a consumption bundle. It also seems inappropriate to specify h as inferior, since we do not wish to study the demand for insurance covering purchases of an inferior good.

It is also assumed that $\partial\lambda/\partial I$ is positive. Since $(-\lambda)$ is the marginal utility of income, that is equivalent to specifying that the second derivative of utility with respect to income is negative. This assumption provides us with a risk-averse consumer, a necessary condition for observing insurance purchases in a rational world.

Effects of Total Price on Demand for h

This model establishes the total net price for medical care as $NPH = Cp_h + w \cdot t$; total price consists of both a money price and a time price, with the money price being subject to insurance. The response of demand for medical care to total price is given as

$$\frac{\partial h}{\partial NPH} = \left(-h - \frac{\partial R}{\partial NPH}\right) \frac{\partial h}{\partial I} + \frac{\lambda p_x^2}{|M|} , \tag{2.19}$$

where

$$\frac{\partial R}{\partial NPH} = \frac{\partial R}{\partial C} \cdot dc + \frac{\partial R}{\partial p_h} \cdot dp_h + \frac{\partial R}{\partial w} \cdot dw .$$

As will be demonstrated below, the responsiveness to changes in any of the components of total price will be systematically related to this total price elasticity and to the relative proportions of the total price accounted for by the components. As a frame of reference, I refer to a "pure" price response to total price as $\partial h/\partial TP$ and a

corresponding *elasticity* as η_h^T (or total price elasticity). This expression will be *defined* as

$$\frac{\partial h}{\partial TP} = -h\,\frac{\partial h}{\partial I} + \frac{\lambda p_x^2}{|M|}, \tag{2.20a}$$

and the corresponding total price elasticity η_h^T (ignoring premium changes) as

$$\eta_h^T = \frac{\partial h}{\partial TP} \cdot \frac{(Cp_h + w \cdot t)}{h} = \frac{\partial h}{\partial TP} \cdot \frac{TP}{h}. \tag{2.20b}$$

This definition will be used to establish relationships among money price, time price, and coinsurance elasticities.

Effects of Changes in Money Price on Demand for h

It is of obvious interest to ascertain how demand for h changes in response to medical care price changes, when reimbursement insurance is in effect. The *partial* derivative $\partial h/\partial p_h$ is given in Appendix A as

$$\frac{\partial h}{\partial p_h} = C\left(-h - \frac{\partial R}{\partial p_h}\right)\frac{\partial h}{\partial I} + \frac{C\lambda p_x^2}{|M|}, \tag{2.21a}$$

which will be negative as long as h is a superior good, and $\partial R/\partial p_h$ is not both negative and "large."[1]

If we ignore the effects of price changes on insurance premiums, the demand slope $\partial h/\partial p_h$ is simply the total price slope $\partial h/\partial TP$ multiplied by the coverage ratio C. Then the elasticity of demand with

---

[1] See Appendix B for a derivation of $\partial R/\partial p_h$ and a discussion of its sign.

respect to changes in the money price (p_h) can be written as:

$$\text{(2.21b)} \qquad \eta_{hh} = \frac{\partial h}{\partial p_h}\frac{p_h}{h} = \frac{Cp_h}{TP}\left(\frac{\partial h}{\partial TP}\cdot\frac{TP}{h}\right) = \frac{Cp_h}{Cp_h + w\cdot t}\cdot \eta_h^T ,$$

when we ignore changes in premiums due to changes in money prices. That is, the elasticity with respect to money price is approximately given by the *total-price* elasticity times the percent of total price accounted for by money price.

If, as an approximation, we believe that the total-price elasticity is constant, this formulation suggests that the elasticity of demand with respect to money prices will fall as insurance coverage increases (as C approaches zero). If this formulation is correct, then one would expect measured money-price elasticities of demand for insured services to be quite low (if insurance coverage is high), especially if the time prices for the service are a significant proportion of the total price. Hospitalization might be a good example of such a service. Reimbursement insurance is quite common (and generally quite complete) in this country, and the time-price for receiving a hospital day is quite large. The observed money price elasticities of demand for hospitalization are expected to be quite low.

Several additional comments are in order regarding goods apparently made "free" through full reimbursement insurance. First, even if there were no time price associated with a good or service covered by reimbursement insurance, the full derivative $\partial h/\partial p_h$ would include income effects of premium changes as p_h changed, and in general these changes could be expected to provide a downward sloping demand curve, even at full insurance. In such a case, the *average* effects of price changes (averaged over all losses ℓ) are reflected in total premiums, and this change is transmitted to each individual h through an income effect.

Second, in services typically covered by reimbursement insurance, the time prices will act as a resource-allocation mechanism even with full reimbursement insurance. The experience of nations with "free" medical care demonstrates that demand is effectively constrained by time prices, even in the absence of any money price.

Effects of Wage Rates and Service Time on Demand for h

The comparative statics of this system show that changes in the opportunity costs of time (as reflected by wage rates) have similar effects on demand as do changes in money prices. Since the net price of medical care (NPH) is written as NPH = $Cp_h + w \cdot t$, it is natural that the results are similar. The equations in Appendix A indicate that a change in the wage rate has the following effects on demand:[1]

$$\text{(2.22a)} \qquad \frac{\partial h}{\partial w} = \left[\left(-h - \frac{\partial R}{\partial w} \right) \left[\frac{p_x U_{xH} - NPH \cdot U_{xx}}{|M|} \right] + \frac{\lambda p_x^2}{|M|} \right] t \ .$$

Ignoring the effect of changes in w on the premium R, it is easily seen that the demand slope with respect to w is simply the total-price slope $\partial h / \partial TP$ times the service time t. It follows immediately that the wage-rate elasticity is approximately given as:

$$\text{(2.22b)} \qquad \eta_{hw} = \frac{\partial h}{\partial w} \cdot \frac{w}{h} = \frac{w \cdot t}{TP} \left(\frac{\partial h}{\partial TP} \frac{TP}{h} \right) = \left(\frac{w \cdot t}{Cp_h + w \cdot t} \right) \cdot \eta_h^T \ .$$

This expression is simply the total-price elasticity times the proportion of total price accounted for by time price.

This formulation gives the total-price elasticity as the simple sum of wage-price and money-price elasticities. That is, by adding the expressions for η_{hw} and η_{hh},

---

[1]In a full analysis, changes in w would be expected to alter the labor force participation and hours worked, and therefore to alter total income. I ignore such effects in this analysis by holding total hours worked constant, and assuming that *all* medical care is received at the expense of working hours. Although this is a somewhat constrained assumption, I feel that it will effectively isolate the crucial problem here: how opportunity costs affect demand for services. It almost goes without saying that provisions of sick-leave time or use of non-market time (which may have a lower opportunity cost than market time) to purchase medical care would reduce the effects of the time price. I also ignore income changes induced through $\partial w \cdot T_0$; in effect, I study a "semi-compensated" wage effect.

$$(2.23) \qquad \eta_{hw} + \eta_{hh} \approx \frac{w \cdot t}{Cp_h + w \cdot t} \eta_h^T + \left(\frac{Cp_h}{Cp_h + w \cdot t} \right) \eta_h^T \approx \eta_h^T .$$

The time price rises with the wage rate, so that time prices discriminate against higher level wage earners. In general, time prices offer a reliable method of price discrimination on an earned income basis. This fact has been observed in countries with national health care systems (such as Sweden and the United Kingdom) where private clinics arise in competition with "free" public medical facilities. The private clinics offer higher monetary prices (they charge a positive price) and reduce queueing time. Persons with high opportunity costs for their time rationally seek medical care at the private facilities. For example, about 10 percent of the outpatient medical care in Sweden is delivered outside the "free" system. This is almost a textbook case of how profit-seeking behavior will induce entry when a discriminating monopolist attempts to select out those with highest demand for services; other suppliers will enter and offer price reductions to those paying the largest amount. In this case, the "price reduction" is brought about by increasing the monetary price, the only unusual feature of this example.[1]

Time cost changes due to increases in t (the time required per unit of service) have identical effects on behavior as do changes in w. If travel or service time is higher for a given person, his total costs will be higher as well, independent of the wage rate or the money price. The elasticity of demand with respect to t is identical to the "price" portion of the wage elasticity. This suggests that, holding constant such factors as income and education, people living

[1]Suppose a person were faced with the choice of two sources of otherwise identical medical care: One has no money price but a waiting time of two hours to receive the service; the other has a money price of $5.00 but a waiting time of one-half hour. For a person with an hourly opportunity wage of $1.00, the total price would be $3.00 for the "free" care, and $5.50 for the other source of care. If the opportunity wage were $3.00, the total price of the "free" care would be $9.00 and the other source of care would cost $6.50. The low-wage person would rationally choose the "free" care, and the high-wage person would rationally choose the other source of care.

in rural areas with longer travel distance will have fewer medical visits than people living in urban areas with immediate access to medical care. It also suggests that there may be serious pitfalls to using aggregated data (such as state averages) in estimating demand for medical care. In a given geographic region (encompassing an entire "market" for medical care), there will tend to be interactions between money prices and queues. If money prices rise, queues may fall, so that total prices will change less than money prices. Alternatively, in areas where there are large demand shifts (such as rapid population shifts in a community), one may observe both rising money prices and increasing queue lengths. In such a situation, the total price will have increased even more than the money price increases. In either case, use of state aggregated data may give inappropriate results.

Effects of ℓ on Demand for Goods

The loss ℓ from the health level H is assumed to be random and not subject to consumer control.[1] Demand for medical care is always conditional upon a given loss ℓ in this model, and we may investigate how changes in the size of that loss will change the demand for medical care.

The derivative $\partial x/\partial \ell$ is shown to be a negative function of $\partial x/\partial I$, and $\partial h/\partial \ell$ is shown to be a positive function of $\partial x/\partial I$! These results show that there is a positive tradeoff between dollars spent on x and dollars spent on h, given the constraining budget function. Similarly, the derivative $\partial \lambda/\partial I$ shows that a loss ℓ is equivalent to an income loss, multiplied by the shadow price of health NPH/g'(h). The marginal utility of income increases with ℓ, which is equivalent to placing the consumer at a lower level of income.

[1]In a broader context, the probabilities of losses may be controlled through such activities as vaccinations, exercise, and good housing. In addition, the demand for market medical care (h) may be changed if household production of health may be substituted for market production. I have chosen not to investigate these phenomena but to concentrate on effects of reimbursement insurance on medical care purchases. For an extensive treatment of self-protection and self-insurance as alternatives to market insurance, see Ehrlich and Becker, "Market Insurance, Self-Insurance, and Self-Protection."

This result may be used to determine if the consumer would ever return his health level H to its initial level H_0, given that he has suffered some positive loss ℓ. Let $D = H - H_0 = g(h) - \ell$. Then the derivative $dD/d\ell$ will establish whether or not H is ever returned to its original level. This derivative is given as:[1]

$$(2.24a) \qquad \frac{\partial D}{\partial \ell} = g'(h)\,\frac{\partial h}{\partial \ell} - 1 = p_x\,\frac{\partial x}{\partial I}\,\frac{g'(h)}{g'(h)} - 1 = p_x\,\frac{\partial x}{\partial I} - 1 \ .$$

Converting to elasticities, and setting $p_x x/I = \omega_x$ = the budget share of x,

$$(2.24b) \qquad \frac{\partial D}{\partial \ell} = \frac{p_x x}{I}\,\eta_{xI} - 1 = \omega_x \eta_{xI} - 1 \ ,$$

which must be negative by the Engel aggregation. The income elasticities of all goods in the budget constraint, weighted by the average budget shares for each good, must sum to unity. Therefore, unless the demand for h is completely income inelastic, the consumer will never fully replace any losses from his level of health H.

Intuitively, we expect this to occur. A loss ℓ has been shown to be equivalent to an income loss. In general, when a consumer's income is reduced, he cuts back on the consumption of all goods rather than leaving one at its original level and reducing all others to absorb the income loss. This model has confirmed that when a consumer suffers a loss ℓ, he reduces the levels of both x and H consumed. He is worse off because of the loss ℓ, but he is able to improve his position to some extent by purchasing some h to offset some of the loss.

---

[1] I assume here that $g''(h) = 0$ to simplify the expressions for $\partial x/\partial \ell$, $\partial h/\partial \ell$.

Effects of H_0 on Demand for Goods

We may similarly investigate the effects on medical care purchases if the consumer had entered the period with a higher level of health (H_0). Presumably, since $H = H_0 - \ell + g(h)$, changes in H_0 will have identical income effects to changes in ℓ, except that the sign will be reversed. However, there is an additional factor operating in health level changes. If insurers recognize that persons with more H_0 will purchase less medical care, then the insurance premiums may reflect that fact also. If that is true, and the persons with high health levels can be identified, then there will be an offsetting income effect through changes in the premium R. This reduced demand for h due to a higher H_0 has nothing to do with the illness distribution, which has been held constant. Changes here are simply a reflection of diminishing marginal utility for H; if a person enters with a high level, he will be less inclined to purchase additional medical care, given any loss ℓ.

THE COMPARATIVE STATICS WITH RESPECT TO INSURANCE PARAMETERS

Although the insurance parameters C and h* are in fact endogenous, and will be selected by the consumer differently for different I, p_h, and so on, we may investigate how any change in C (or h*) affects demand for x and h for a given loss. These partials may be used, combined with the other derivatives in the comparative statics, to determine the full effects of changes in such variables as income, p_h, the insurance loading fee (θ), and the wage rate. As noted before, the full derivative with respect to income would be:

$$\frac{dh}{dI} = \frac{\partial h}{\partial I} + \frac{\partial h}{\partial C} \cdot \frac{\partial C}{\partial I} + \frac{\partial h}{\partial h^*} \cdot \frac{\partial h^*}{\partial I} \,. \tag{2.25}$$

Effects of C on Demand for Goods

The partial effect of C on demand for h is given in Appendix A as:

$$\frac{\partial h}{\partial C} = -R_c \frac{\partial h}{\partial I} + p_h\left(-h \frac{\partial h}{\partial I} + \frac{\lambda p_x^2}{|M|}\right). \tag{2.26}$$

The expression in (2.26) is easily seen to be equivalent to p_h times the total-price derivatives $\partial h/\partial TP$ (plus an income effect due to changes in R as C changes). That is,

$$\frac{\partial h}{\partial C} = p_h \frac{\partial h}{\partial TP} + (-R_c) \frac{\partial h}{\partial I}, \tag{2.27a}$$

and by deleting the "income" effects due to changes in R, it is apparent that the elasticity of demand for h with respect to coinsurance is similar to the medical price elasticity η_{hh}:

$$\eta_{hC} = \frac{\partial h}{\partial C} \cdot \frac{C}{h} \approx \left(\frac{Cp_h}{Cp_h + w \cdot t}\right) \frac{\partial h}{\partial TP} \frac{TP}{h} \tag{2.27b}$$

$$= \left(\frac{Cp_h}{Cp_h + w \cdot t}\right) \eta_h^T \approx \eta_{hh}.$$

The complete expression for η_{hC} is given (including premium changes) as

$$\eta_{hC} = \left(\frac{Cp_h}{Cp_h + w \cdot t}\right) \eta_h^T + \frac{(-R_c)}{I} \eta_{hI}, \tag{2.28}$$

where η_{hI} is the income elasticity of demand for h.

Since $(-R_c)$ is in part determined by η_{hC}, it is possible to collect terms from the right hand side of (2.28) involving $\partial h/\partial C$. One way to accomplish this is to assume $\partial h/\partial C$ is constant over all losses, so that

$$(2.29) \qquad \frac{\partial h}{\partial C} = p_h \frac{\partial h}{\partial TP} + \left(\frac{\partial h}{\partial I}\right)(-R_c)$$

$$= p_h \frac{\partial h}{\partial TP} + \frac{\partial h}{\partial I}\left((1 + \theta)\int_0^{\ell^*} p_h h\left(1 - \frac{(1 - C)}{h}\frac{\partial h}{\partial C}\right)f(\ell)d\ell\right) .$$

Collecting terms in $\partial h/\partial C$ gives:

$$(2.30a) \qquad \frac{\partial h}{\partial C} = \frac{p_h \frac{\partial h}{\partial TP} + \frac{\partial h}{\partial I}(1 - C)(1 + \theta)p_h\bar{h}}{1 + (1 - C)(1 + \theta)p_h \frac{\partial h}{\partial I}} ,$$

which (in elasticity form) is:

$$(2.30b) \qquad \eta_{hC} = \frac{\partial h}{\partial C} \cdot \frac{C}{h} = \frac{\frac{Cp_h}{Cp_h + w \cdot t}\eta^T + \bar{\omega}_h \eta_{hI}(1 - C)C}{1 + (1 - C)(1 + \theta)\omega_h \eta_{hI}} ,$$

which is empirically close to η_{hh} if ω_h, η_{hI}, and $\bar{\omega}_h$ are "small."

Several studies have estimated the effects of changes in C on demand for medical care, which have been shown here to be approximately equivalent to the own-price elasticity for medical care. One study used quotations from actual group health insurance premiums to estimate the responsiveness of demand to changes in C, showing that (between coinsurance rates of .10 and .25) the coinsurance elasticity was -.07.[1] Another study observed the results of a natural experiment where the coinsurance rate in a prepaid group practice was changed from 0 to .25; the estimated coinsurance elasticity over this range was -.14 for physician services and -.07 for "ancillary" services

---

[1]Charles E. Phelps and Joseph P. Newhouse, *Coinsurance and the Demand for Medical Services*, R-964-OEO/NC, The Rand Corporation, Santa Monica, April 1973.

(laboratory tests, X-ray examinations, and so on).[1] In another study using hospital claims data, a "coverage ratio" elasticity was estimated to be between zero and -.08 for various diagnoses, where the dependent variable was hospital length of stay. That author did not correctly identify the coinsurance elasticity as being equivalent to a medical care price elasticity.[2]

The *partial* effects of altering the coverage level C can be shown with the use of indifference curves. Assume, as was done when the comparative statics were analyzed, that the size of the loss ℓ is known by the consumer. Then Fig. 3 shows how a change in C alters the equilibrium values of x and h: Budget line AD is where no insurance is in effect. On budget line AB, the price of medical care has become Cp_h, but no adjustment has been made to reflect the insurance premium. Budget line EC is parallel to AB but is lower by the amount of the insurance premium R. The total effect of dh/dC (for this loss ℓ) consists of a movement from h_1 to h_3 due to substitution and income effects as price is lowered, and from h_3 to h_2 due to income effects from premium increases. The net adjustment, and the one actually observed, would be $h_2 - h_1$. It is clear that income effects might outweigh substitution effects if the premium R were sufficiently large.

The partial effects of a change in h* on demand for medical care are given as:

$$\frac{\partial h}{\partial h^*} = -R_{h^*} \frac{\partial h}{\partial I} . \tag{2.31}$$

In other words, there is strictly an income effect on demand for h generated through shifts in premiums as h* changes.

---

[1] Charles E. Phelps and Joseph P. Newhouse, *The Effects of Coinsurance on the Demand for Physician Services*, R-976-OEO, The Rand Corporation, Santa Monica, June 1972. An abridged version is published in the *Social Security Bulletin*, June 1972, as "Effects of Coinsurance: A Multivariate Analysis."

[2] Gerald Rosenthal, "Price Elasticity of Demand for Short-Term General Hospital Services," in Herbert E. Klarman, ed., *Empirical Studies in Health Economics*, The Johns Hopkins Press, Baltimore, 1970.

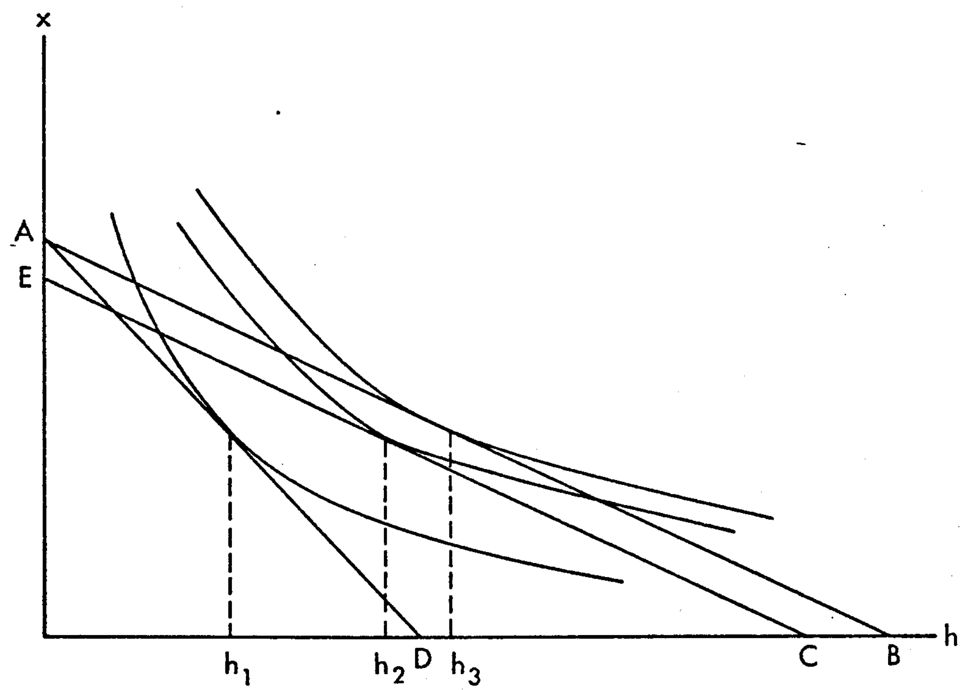

Fig. 3 — Indifference map showing effects of coinsurance on demand for medical care

Indifference curves may be used to demonstrate that a change in h* leads to an income effect and not a substitution effect on demand for h. A decrease in h* reduces the range over which C is operative and decreases the amount of h purchased in the range between the old and the new values of h*. The inclusion of the parameter h* can be visualized as a kink in the budget line at the point h = h*. On the indifference map in Fig. 4, h* is reduced from a very large value ("infinity") to some finite level h^*_I.

Budget line AK is with no insurance. In budget line AB an insurance policy with coinsurance C and h* = ∞ is in effect. Budget line GC adjusts the consumer's income by a premium R_2, to pay for the insurance contract specified in AB. In budget line FEC the consumer has selected a finite maximum benefit h*, and the premium has been reduced by a commensurate amount (to R_0). The budget line FEC has a kink in it at $h = h^*_I$ (at point E) to reflect the shift in price from Cp_h to p_h at that point. For this loss, when h* = ∞, dh/dC is given by $(h_2 - h_1) - (h_2 - h_3) = (h_3 - h_1)$. If the insurance policy is altered

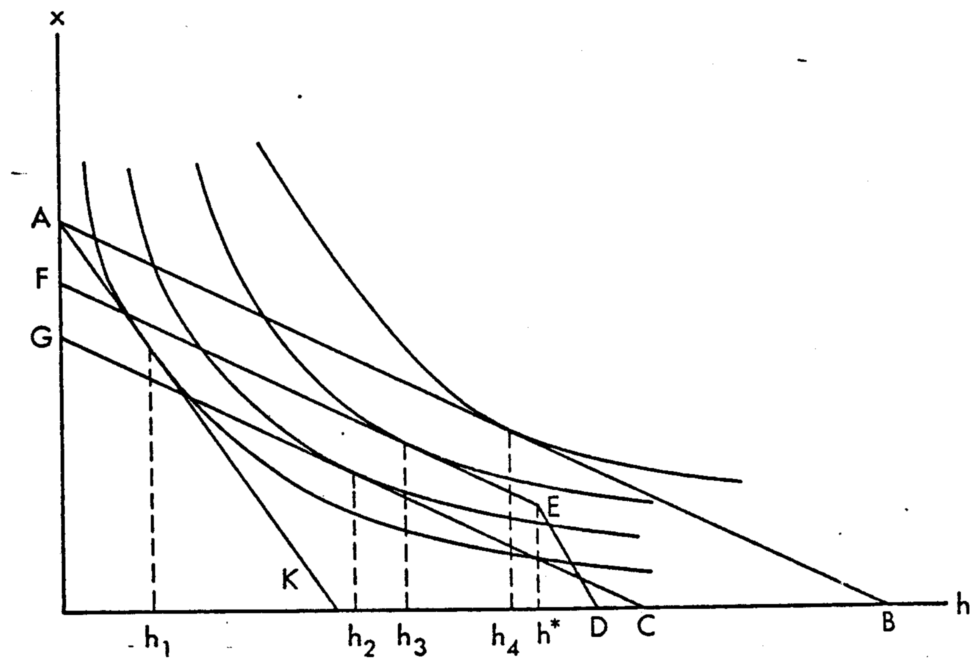

Fig. 4 — Indifference map showing effects of reducing upper coverage limit on demand for medical care

so that $h^* = h^*_1$ (budget line FEC), the premium is reduced by $(R_2 - R_1)$ and the income effect of dh/dI moves the final purchase of h to h_4. This shows that, if the solution level of h is less than h*, the income effect of changing h* is negative (lower h* causes higher level of h).

An example of the behavior of dh/dh* when the solution value of h is above h* is shown in Fig. 5. When h > h*, the sign $\partial h/\partial h^*$ is positive. A reduction in h* reduces the amount of h purchased. This result may be confirmed in the comparative statics of Appendix A, where it is shown that for $h \leq h^*$, $\partial h/\partial h^* = -R_{h^*}(\partial h/\partial I)$, and for h*, $\partial h/\partial h^* = (-R_{h^*} + (1 - C)p_h)(\partial h/\partial I)$, which is necessarily positive when $\partial h/\partial I$ is positive.

The above results have been derived for the situation where the loss ℓ is known and may be considered a problem in certainty. *On average,* does insurance induce demand for medical care? That is, does the average person with reimbursement insurance buy more medical care than the average uninsured person, other things equal?

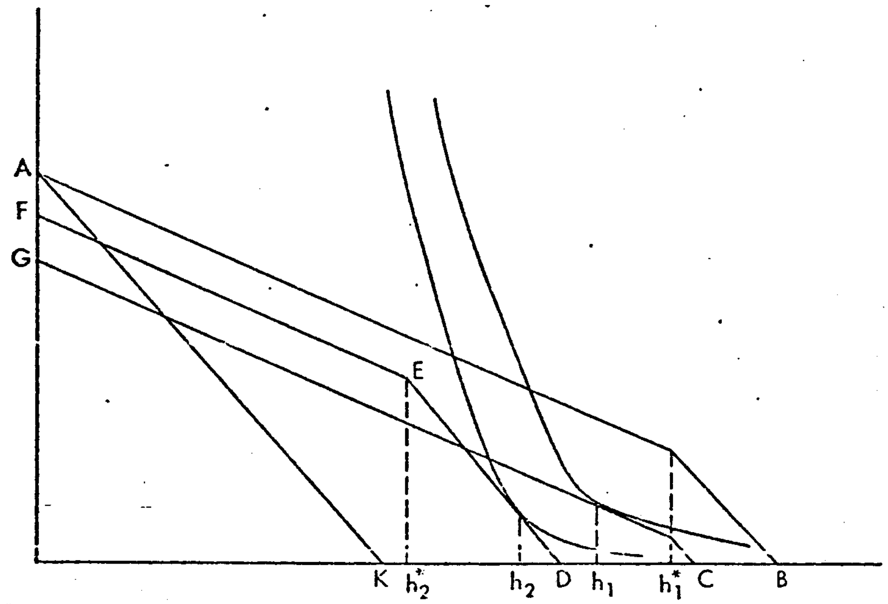

Fig. 5 — Indifference map showing effects of reducing upper coverage limit on demand for medical care when upper limit is exceeded

Initially, any loading in the insurance premium is ignored, so that expected benefits equal the premium R. Also, for any given consumer, the insurance policy does not result in an income transfer (on average).[1] As shown in Fig. 6, the *average* equilibrium solution must have the characteristics that (1) the amounts of x and h consumed are such that the consumer is still on his original budget line, showing no shift in total purchasing power; (2) since suppliers of x and h still face the relative price on p_x/p_h, they must be in equilibrium at the solution value; (3) the consumer must be making marginal decisions based on the altered price line that the insurance contract imposes.

In Fig. 6, let AE be the uninsured budget line. Then AB is the relevant "rotated" price line that reflects the copayment rate C chosen. The problem is to adjust the budget line by some premium R so that, on

---

[1]Actually, some insurers (notably Blue Cross and Blue Shield) deliberately attempt to effect income transfers among subscribers. But for the community as a whole, no income transfers can occur. The insurance policy will not create any new resources to be used in the production of x and h.

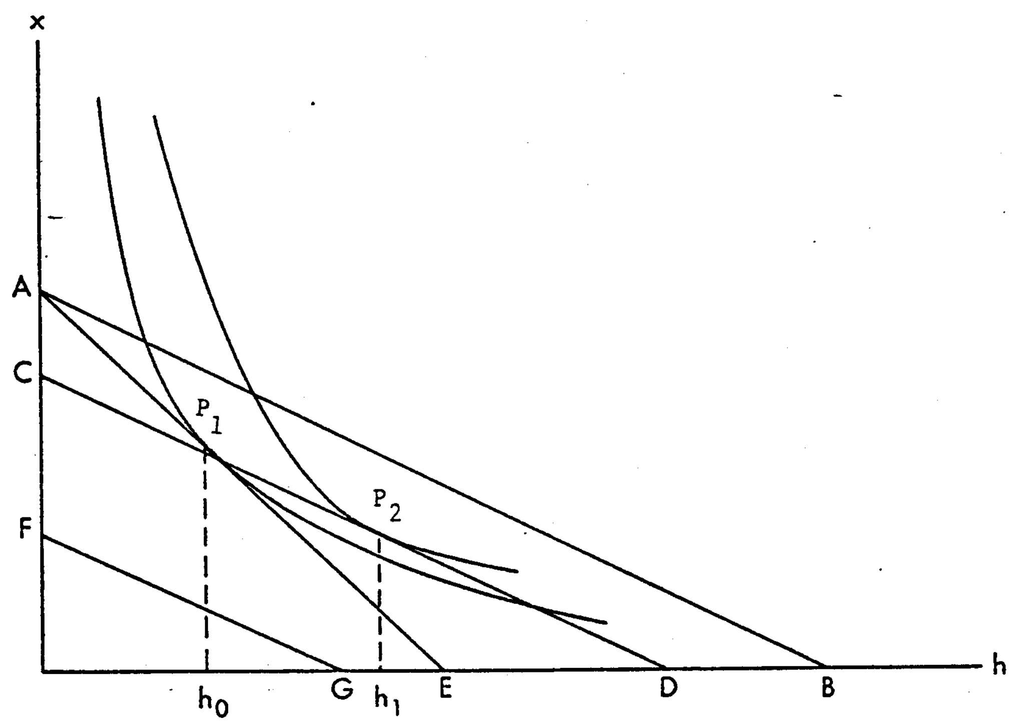

Fig. 6 — Average effects of coinsurance on demand for medical care

average, the consumer is still constrained by the same real income. The line CD provides such a solution, and the point P_2 indicates the equilibrium values of x and h that will be purchased. In this situation (without considering the income effects of a loading fee on the premium R), the final level of h consumed, on average, must exceed the quantity consumed when no insurance is present.

When the loading fee is taken into account, no *a priori* determination of the effects of C on demand for h can be made. There is some level of loading fee that will drive the final budget line sufficiently low that the average value for $\partial h/\partial C$ is positive. The line FG in Fig. 6 is such a budget line.

THE PURCHASE OF HEALTH INSURANCE

In purchasing a health insurance contract, the consumer selects values of C and h* to maximize expected utility in x and H, over all possible values of the loss ℓ. These decisions are made under

uncertainty, since the consumer does not know at the time of the insurance purchase what loss will be experienced.

Assume, as set forth in Equations (2.1) and (2.2) that the consumer's level of health H is random, with losses from H distributed by a density function $d(\ell)$. For any specified loss, the consumer acts in the manner formulated above, and takes into account his predictable behavior in purchases of x and h, when he is choosing the optimal values of C and h*. The problem is to pick C and h* so as to maximize expected utility in x and H, where expected utility is described by:[1]

$$(2.32a)\qquad E(U) = s_0 U[x,H_0] + \int_0^{\ell^*} U[x,H_0 - \ell + g(h)]f(\ell)d\ell$$

$$+ \int_{\ell^*}^{\infty} U[x,H_0 - \ell + g(h)]f(\ell)d\ell \ ,$$

where ℓ^* is the loss that causes the consumer to purchase exactly h* units of medical care. The constraining budgets are:

$$(2.32b)\qquad I = p_x x + Cp_h h + R \qquad \text{for } h \le h^* \ ,$$

$$(2.32c)\qquad I = p_x x + p_h h - (1 - C)p_h h^* + R \qquad \text{for } h > h^* \ .$$

The maximizing conditions with respect to C and h* are given as:

$$(2.33)\qquad \frac{\partial EU}{\partial C} = s_0\left[U_x \frac{\partial x}{\partial C} + U_H g'(h) \frac{\partial h}{\partial C}\right] + \int_0^{\ell^*} \left[U_x \frac{\partial x}{\partial C} + U_H g' \frac{\partial h}{\partial C}\right] f(\ell)d\ell$$

$$+ \int_{\ell^*}^{\infty} \left[U_x \frac{\partial x}{\partial C} + U_H g'(h) \frac{\partial h}{\partial C}\right] f(\ell)d\ell = 0 \ ,$$

---

[1] Recall that x, h, and $(-\lambda)$ are all explicit function of ℓ. For notational simplicity, they are written without specifying the functional dependence.

$$(2.34)\quad \frac{\partial EU}{\partial h^*} = s_0\left[U_x \frac{\partial x}{\partial h^*} + U_H g' \frac{\partial h}{\partial h^*}\right] + \int_0^{\ell^*} \left[U_x \frac{\partial x}{\partial h^*} + U_H g' \frac{\partial h}{\partial h^*}\right] f(\ell)d\ell$$

$$+ \int_{\ell^*}^{\infty} \left[U_x \frac{\partial x}{\partial h^*} + U_H g'(h) \frac{\partial h}{\partial h^*}\right] f(\ell)d\ell = 0 \ .$$

Using explicit expressions for $\partial x/\partial C$, $\partial h/\partial C$; $\partial x/\partial h^*$, and $\partial h/\partial h^*$ (for $h \leq h^*$ and for $h > h^*$), equations (2.27) simplify to the following:[1]

$$(2.35a)\quad \frac{\partial EU}{\partial C} = s_0(-R_c - p_h h)(-\lambda_0) + \int_0^{\ell^*} (-R_c - p_h h)(-\lambda) f(\ell)d\ell$$

$$+ \int_{\ell^*}^{\infty} (-R_c - p_h h^*) f(\ell)d\ell = 0 \ ,$$

$$(2.35b)\quad \frac{\partial EU}{\partial h^*} = s_0(-R_{h^*})(-\lambda_0) + \int_0^{\ell^*} (-R_{h^*})(-\lambda) f(\ell)d\ell$$

$$+ \int_{\ell^*}^{\infty} (-R_{h^*} + (1 - C)p_h)(-\lambda) f(\ell)d\ell = 0 \ ,$$

where $(-\lambda_0)$ is the marginal utility of income when $\ell = 0$, and $(-\lambda)$ is the marginal utility of income in general.

If we think of C and h* as being two different "goods," then rearrangement of these first-order conditions suggests:

$$(2.36a)\quad E(-\lambda) = \frac{s_0(-\lambda_0)p_h h + \int_0^{\ell^*} (-\lambda)(p_h h) f(\ell)d\ell + \int_{\ell^*}^{\infty} p_h h^*(-\lambda) f(\ell)d\ell}{(-R_c)} \ ,$$

and

[1] In performing these differentiations of integrals, I used the Leibnitz Rule. The limits of integration enter the general formula given by Leibnitz, but they always cancel in these first-order equations.

$$(2.36b) \qquad E(-\lambda) = \frac{(1 - C)p_h \int_{\ell^*}^{\infty} (-\lambda) f(\ell) d\ell}{R_{h^*}} = \frac{\int_{\ell^*}^{\infty} (-\lambda) f(\ell) d\ell}{\gamma(h^*)(1 + \theta)h^*} .$$

The numerators of (2.36a) and (2.36b) may be interpreted as the expected marginal gains from each of the components of the insurance policy. If utility is measured in "utils," the marginal gains from C are in "dollar-utils," and the marginal gains from h* are in "price-utils." These marginal gains derive directly from the marginal utility of income $(-\lambda)$, and from the nature of demand for the insured good. Gains in C involve a weighting of all possible expenses on h, weighted not only by $d(\ell)$, but by the marginal utility of income $[-\lambda(\ell)]$ for each loss ℓ. Gains involving h* come from extending the price differential--$(1-C)p_h$--over a range not previously covered. The denominators of (2.36a) and (2.36b) are simply the partial derivatives of R with respect to the variables C and h*, and may be thought of as the marginal prices for additional coverage. By equating (2.36a) and (2.36b) we see that the ratio of expected gains from adding coverage (through either C or h*), divided by the appropriate "marginal price," must be identical for all elements of the insurance policy.[1]

[1]Note the difference in the results between this model and the Ehrlich-Becker formulation, where there are (n-1) choice variables, corresponding to each of the states of the world with hazard. In this model, there are only two choice variables, in a world with an infinitely large set of possible states. Insurance, in such a world, must *average* benefits across states of the world. Obviously, the constrained expected utility cannot be as high under this type of insurance as under (n-1) choice variable insurance *unless* loading fees on the latter are prohibitive. But since we do not observe insurance against every potential type of health loss, it seems plausible that the load on that type of insurance is sufficiently high to make single-coverage insurance more attractive to consumers (at current price levels for insurance). A major reason for this is the high transactions cost involved in (n-1) state insurance--it is extremely expensive just to identify which state of the world has occurred in health insurance, making payment of indemnity (income transfer) insurance infeasible.

How reimbursement insurance "averages" across states of the world is best shown by comparing the first-order equilibrium conditions of

It may be seen from equation (2.35b) that an infinitely large maximum payment is not optimal under this type of insurance. Since $-R_{h*}$ must be negative under rational insurance pricing, it follows that $[-R_{h*} + (1 - C)p_h]$ must be positive for (2.35b) to equal zero. But if that equation is zero, the final integral (from h* to infinity) must be positive; in other words, h* cannot be set equal to infinity. Since θ is constant over all losses, the question immediately arises as to why full coverage above a deductible is not optimal.[1] The answer is found in the Ehrlich-Becker formulation of the insurance problem. Since we are dealing with gross rather than net insurance, the relevant "terms of trade" (in the Ehrlich-Becker sense) are *not* given as $-R_{h*}$ but by the rates at which income may be exchanged between the 0th state and any other state. If the "real price" (in the Ehrlich-Becker sense) is defined as $\bar{p}$, then

$$(2.37) \qquad \bar{p} = \frac{-\partial I_0}{\partial I_1} \cdot \frac{s_0}{f(\ell^*)} = \frac{-R_{h*}}{(-R_{h*} + (1 - C)p_h)} \frac{s_0}{f(\ell^*)}$$

$$= \frac{(1 + \theta)(1 - C)p_h h^* f(\ell^*)}{(1 - C)p_h - (1 - C)(1 + \theta)p_h h^* f(\ell^*)} \cdot \frac{s_0}{f(\ell^*)}$$

$$= \frac{(1 + \theta)h^* \cdot s_0}{1 - (1 + \theta)h^* f(\ell^*)} ,$$

---

this model with those in Ehrlich and Becker. In their model, it must hold for (n-1) first-order conditions that $(-\lambda)_i/(-\lambda)_0 = P_i$ (where p_i is the price of insurance in the ith state). In this model, as equations (2.36) show, the "prices" $-R_j$ and R_{h*} must be equal to the ratio of expected benefits divided by average utility of income over all states of the world. "Benefits" in these equations are weighted combinations of utility and expenditure, as is appropriate to match the benefit/utility ratios to the appropriate "prices."

[1]Full coverage above a deductible is shown to be optimal by both Ehrlich and Becker, "Market Insurance, Self-Insurance and Self-Protection," and Kenneth J. Arrow, "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, 53, December 1963, whenever the "loading fee" is constant across all losses. Here a constant θ is not equivalent in fact to constant "real price" (in the Ehrlich-Becker sense) as h* increases.

which clearly increases as h^* increases. (The denominator of (2.37) must be positive, as is shown in Appendix B, so that $\bar{p}$ always remains positive.) It is easily seen that the "real price" of adding a unit of coverage (h^*) in reimbursement insurance becomes increasingly expensive as h^* becomes large, so that coverage is eventually terminated.

When comparing the two models, note that the terms of trade for any state of the world are *not* given in this model by the "prices" of the insurance policy, $-R_c$ and R_{h*}, but by ratios of $(R_{h*})/[(1 - C)p_h - R_{h*}]$ and of $(R_c)/(-R_c - p_h h)$. This feature will assist understanding of how reimbursement insurance demand changes in the face of changes in the world, as will be shown below.[1]

It is not possible to determine from equation (2.35a) whether full coverage ($C = 0$) is specifically optimal or nonoptimal. In fact, no level of coverage is theoretically eliminated from the set of optimal solutions by this equation.[2]

Equations (2.35a) and (2.35b) are two equations in two unknowns and imply a solution for C and h^*, given income, prices, the loading fee θ, the function g, and the distribution $d(\ell)$. The optimal values of C and h^* are specific numbers and do not involve a random component. In other words, C and h^* are not a function of a particular loss ℓ, whereas demand for medical care (h) will always be a function of ℓ. In terms of estimation, this is sufficient to allow identification of demand curves for medical care, and for insurance in joint estimation. The loading fee should be an excluded variable in estimating demand for

---

[1]Since p_h enters into the "real" price of changing C, it follows that patterns of demand for C will depend on how h is demanded over various illness states ℓ. In other words, the "terms of trade" price is determined internally by the consumer's preferences.

[2]I have not allowed C to vary except between zero and unity in this model. To be completely correct mathematically, I should establish slack variables in the expected utility maximization equations and use Kuhn-Tucker conditions to determine the true optimum. Such a formulation would allow "corner solutions" at $C = 0$ or $C = 1$, in a more formal manner. For simplicity, and because this methodology does not add any information to the knowledge of consumer behavior, I omit the Kuhn-Tucker formulation and assume that $0 \leq C \leq 1$. An alternative assumption is that loading fees are "very high" for other values, so they are excluded from consumer selection.

medical care itself, and the loss ℓ will be excluded from the insurance equation.

The implied demand equations are of the form:

$$C = C(I, p_h, p_x, g(h), d(\ell), R_c, R_{h*}, \theta, w, H_0, \psi) , \tag{2.38a}$$

$$h^* = h^*(I, p_x, p_h, g(h), d(\ell), R_c, R_{h*}, \theta, w, H_0, \psi) , \tag{2.38b}$$

$$h = h(I, p_x, p_h, C, h^*, g(h), w, \ell, H_0, \psi) , \tag{2.38c}$$

where ψ is a vector of parameters of the utility function, and other variables are as defined previously. Without solving for explicit demand functions,[1] note that the comparative statics of the insurance purchase decision may be computed, just as in any other demand system, by differentiation of the first-order conditions with respect to exogenous variables and use of Cramer's rule to solve the resultant equations for the values of interest.

For any exogenous variable V, the total differentials of (2.35a) and (2.35b) with V allowed to change will provide two equations in the two unknowns $\partial C/\partial V$ and $\partial h^*/\partial V$. The solutions will be of the form

$$\frac{\partial C}{\partial V} = \frac{\begin{vmatrix} \dfrac{\partial^2 EU}{\partial C \partial V} & \dfrac{\partial^2 EU}{\partial C \partial h^*} \\ \dfrac{\partial^2 EU}{\partial h^* \partial V} & \dfrac{\partial^2 EU}{\partial h^{*2}} \end{vmatrix}}{|D|} = \frac{\begin{vmatrix} -b_1 & a_{12} \\ -b_2 & a_{22} \end{vmatrix}}{|D|} \tag{2.39a}$$

$$\frac{\partial h^*}{\partial V} = \frac{\begin{vmatrix} \dfrac{\partial^2 EU}{\partial C^2} & \dfrac{\partial^2 EU}{\partial C \partial V} \\ \dfrac{\partial^2 EU}{\partial C \partial h^*} & \dfrac{\partial^2 EU}{\partial h^* \partial V} \end{vmatrix}}{|D|} = \frac{\begin{vmatrix} a_{11} & -b_1 \\ a_{12} & -b_2 \end{vmatrix}}{|D|} , \tag{2.39b}$$

[1] Attempts have been made to solve for demand equations for C and h*, with explicit utility functions. Without going into the details, these attempts were unsuccessful in general, if the distribution of $f(\ell)$ was allowed to be continuous. When the problem was reduced to having only two illness states, with discrete probabilities, there was indeed a solution for C, but the equation was highly complex and noninformative in general.

where

$$(2.39c) \qquad |D| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} \frac{\partial^2 EU}{\partial C^2} & \frac{\partial^2 EU}{\partial C \partial h^*} \\ \frac{\partial^2 EU}{\partial h^* \partial C} & \frac{\partial^2 EU}{\partial h^{*2}} \end{vmatrix} .$$

The first-order conditions for optimal reimbursement insurance are given in (2.33) and (2.34). The second-order partial derivatives are:

$$(2.40a) \quad a_{11} = \frac{\partial^2 EU}{\partial C^2} = \frac{\partial}{\partial C}\left[s_0(-R_c - p_h h_0)(-\lambda_0) + \int_0^{\ell^*} (-R_c - p_h h)(-\lambda) f(\ell) d\ell + \int_{\ell^*}^{\infty} (-R_c - p_h h^*) f(\ell) d\ell\right]$$

$$= \left[-s_0(-R_c - p_h h_0)\frac{\partial\lambda}{\partial C} - \int_0^{\ell^*} (-R_c - p_h h)\frac{\partial\lambda}{\partial C} - \int_{\ell^*}^{\infty} (-R_c - p_h h^*)\frac{\partial\lambda}{\partial C}\right]$$

$$+ s_0(-\lambda_0)\left(-p_h \frac{\partial h_0}{\partial C} - \frac{\partial^2 R}{\partial C^2}\right) + \int_0^{\ell^*} (-\lambda)\left(-p_h \frac{\partial h}{\partial C} - \frac{\partial^2 R}{\partial C^2}\right) + \int_{\ell^*}^{\infty} (-\lambda)\left(-\frac{\partial^2 R}{\partial C^2}\right) f(\ell) d\ell .$$

Using information from Appendix A, this becomes

$$(2.40b) \qquad a_{11} = -\left[s_0(-R_c - p_h h_0)^2 \frac{\partial\lambda}{\partial I} + \int_0^{\ell^*} (-R_c - p_h h)^2 \frac{\partial\lambda}{\partial I} f(\ell) d\ell + \int_{\ell^*}^{\infty} (-R_c - p_h h^*)^2 f(\ell) d\ell\right]$$

$$-\left[s_0(-R_c - p_h h_0)(-\lambda)p_h \frac{\partial h}{\partial I} + \int_0^{\ell^*} (-R_c - p_h h)(-\lambda)p_h \frac{\partial h}{\partial I}\right]$$

$$+ s_0(-\lambda_0)\left(-p_h \frac{\partial h_o}{\partial C} - \frac{\partial^2 R}{\partial C^2}\right) + \int_0^{\ell^*} \left(-p_h \frac{\partial h}{\partial C} - \frac{\partial^2 R}{\partial C^2}\right)(-\lambda) f(\ell) d\ell$$

$$+ \int_{\ell^*}^{\infty} (-\lambda) - \frac{\partial^2 R}{\partial C^2} f(\ell) d\ell \ ,$$

where

$$-\frac{\partial^2 R}{\partial C^2} = \frac{\partial}{\partial C}(-R_c) = (1 + \theta)\int_0^{\ell^*} 2\ p_h \frac{\partial h}{\partial C} - (1 - C)p_h \frac{\partial^2 h}{\partial C^2} f(\ell) d\ell$$

$$\doteq (1 + \theta) 2\ p_h\left(\frac{\partial \bar{h}}{\partial C}\right) .$$

For a risk averter, this expression is obviously negative for all values of C *if* $\partial h/\partial I$ and $\partial h/\partial C$ are zero for all losses. In other words, if demand for h is not affected by either income or the insurance contract itself, a maximum in expected utility is always established. However, if $\partial h/\partial I$ and $\partial h/\partial C$ are nonzero, it is necessary to postulate that a_{11} is negative. Intuitively, what equations (2.40) show are that the "pure" components of reimbursement insurance are such that optimal purchases are insured if the first-order conditions are met, but because demand for health care responds to the insurance contract, confounding effects are introduced. A *sufficient* (but not necessary) condition for a_{11} to be negative for all values of C is that both $\partial h/\partial I$ and the absolute value of $\partial h/\partial C$ diminish as losses become larger. That is, if demand for h becomes less responsive to changes in income and coinsurance as losses become larger, then a_{11} will of necessity be negative, insuring that any C chosen will be optimal. Alternatively, if those derivatives increase as losses increase, it is possible that a_{11} is positive, thus making a solution value of C a local minimizing value of expected utility, rather than a local maximizing value.

I postulate here that a_{11} is negative for all observed purchases of reimbursement insurance. That is, consumers will in fact be observed at expected utility maximizing positions in their demands for reimbursement insurance.[1]

The second-order partial of expected utility with respect to the maximum payment h* is given in equation (2.41). This derivative is everywhere negative for risk averters, the common result found when insurance does not alter demand for a particular good.

$$(2.41)\qquad a_{22} = \frac{\partial^2 EU}{\partial h^{*2}} = \frac{\partial}{\partial h^*}\left[s_0(-\lambda_0)(-R_{h^*}) + \int_0^{\ell^*} (-\lambda)(-R_{h^*})f(\ell)d\ell \right.$$

$$\left. + \int_{\ell^*}^{\infty} (-\lambda)(1 - C)p_h(-R_{h^*})f(\ell)d\ell\right]$$

$$= -\left\{s_0(-R_{h^*})^2 \frac{\partial\lambda}{\partial I} + \int_0^{\ell^*} (-R_{h^*})^2 \frac{\partial\lambda}{\partial I} f(\ell)d\ell\right.$$

$$\left. + \int_{\ell^*}^{\infty} (-R_{h^*} + (1 - C)p_h)^2 \frac{\partial\lambda}{\partial I} f(\ell)d\ell\right\}$$

$$+ E(-\lambda)\left(\frac{-\partial^2 R}{\partial h^{*2}}\right) - (1 - C)p_h(-\lambda_{\ell^*})g'(h^*)f(\ell^*) \ ,$$

where the last expression comes from the Leibnitz formula for differentiation of integrals (arising from the evaluation of the limits of integration), and $\partial^2 R/\partial h^{*2} = -(1 + \theta)(1 - C)p_h g'(h^*)f(\ell^*) \equiv -(1 + \theta)(1 - C)p_h\gamma(h^*) < 0$. Thus any choice of h* that meets the first-order condition will also meet the second-order condition for maximizing expected utility.

[1] This assumption is not novel to the choice of reimbursement insurance. In the "standard" theory of consumer demand, it is necessary to hypothesize that all second-order partials and the determinant of the Hessian matrix $|M|$ are of correct sign. Economic theory is not strong enough to make such predictions unambiguously.

To evaluate the selection of both C and h* simultaneously, it is necessary to evaluate the determinant $|D|$, as given in equation (2.39c). For that, it is necessary to evaluate the cross-partial derivative $a_{12} = \partial^2 EU/\partial C \partial h^*$ and then the total determinant. The sign of a_{12} cannot be determined *a priori*, which in itself is sufficient to make the sign of $|D|$ unknown. I must hypothesize that $|D| > 0$ for the joint insurance choice of C and h* to be a maximizing value, just as I had to hypothesize that a_{11} is negative. Finally, since a_{12} (the second-order cross-partial of expected utility with respect to C and h*) cannot be signed, the derivatives of C and h* with respect to any exogenous variable cannot be determined in the full model (see equation 2.39).

Since the full model cannot be used to make any predictions about demand for insurance, the model must be treated more simply by evaluating changes in demand for C and x for h* separately, as if each were a single-equation model. (This has the same net effect as assuming that a_{12} is "approximately" equal to zero, so that it does not enter into the expression in (2.39). In this more simplified model, the comparative statics can be analyzed in the following fashion:

(2.42a)
$$\frac{\partial C}{\partial V} = \frac{b_1}{-a_{11}} = \frac{\dfrac{\partial^2 EU}{\partial C \partial V}}{\dfrac{-\partial^2 EU}{\partial C^2}},$$

(2.42b)
$$\frac{\partial h^*}{\partial V} = \frac{b_2}{-a_{22}} = \frac{\dfrac{\partial^2 EU}{\partial h^* \partial V}}{\dfrac{-\partial^2 EU}{\partial h^{*2}}},$$

for any exogenous variable V. Since a_{11} is assumed to be negative, and a_{22} is always negative for risk averters, the derivative $\partial C/\partial V$ will be the same sign as the partial derivative b_1, and likewise for the derivative $\partial h^*/\partial V$ and the partial derivative b_2. The derivatives $\partial C/\partial V$ and $h^*/\partial V$ will simply be signed by the appropriate second-order cross-partial derivatives of the expected utility equation.

THE COMPARATIVE STATICS OF DEMAND FOR INSURANCE

It may be useful to reinterpret the reimbursement insurance model I have developed, casting it in a framework similar to that developed by Ehrlich and Becker. Since "income" in my model is defined by $I = p_x x + Cp_h h + R$, it is easy to see that a change in the coinsurance rate produces a net change in I of $R_c + p_h h$. The substitution effects cancel one another. Put differently, if C falls by one unit, then for a given illness, the net income increase (for $h \leq h^*$) is $p_h h$ less the change in the premium R. Thus, the expression $(-R_c - p_h h)$ is equal to $-dI/dC$. The same analogy can be drawn for losses above h*, and for changes in income as h* changes itself.[1] The first-order conditions for selecting optimal insurance can then be written as

$$(2.43) \qquad s_0(-\lambda_0)\,\frac{dI_0}{dC} + \int_0^{\ell^*} (-\lambda)\,\frac{dI}{dC}\,f(\ell)d\ell + \int_{\ell^*}^{\infty} (-\lambda)\,\frac{dI}{dC}\,f(\ell)\;\partial\ell = 0,$$

which is identical in substance to the Ehrlich and Becker model, since they define the price of insurance in a given state as $p = -\partial I/\partial I_0$. As I will show, changes in parameters of the model (such as income, the loading fee on insurance, or the price of medical care) will systematically change these "real prices" of insurance in each state of the world, producing a pure price effect in the Ehrlich and Becker sense. In addition, changes in these parameters have systematic "income effects," which depend in sign upon the behavior of the risk aversion measure r(I) as losses change.

Effects of a Change in Income on Demand for Insurance

I will derive the income effect on demand for C to demonstrate how both income and price effects are observed when income changes in my model. The derivative $\partial C/\partial I$ can be signed in this model (see 2.42) by signing the expression $\partial^2 EU/\partial C \partial I$ (which is derived in Appendix C).

---

[1]For losses above h*, $dI/dC = -R_C - p_h h^*$; when h* changes, the analogous changes are $-R_{h^*}$ and $-R_{h^*} + (1-C)\,p_h$.

That expression is:

$$(2.44) \qquad b_1 = \frac{\partial}{\partial I}\left[s_0(-R_c - p_h h_0)(-\lambda_0) + \int_0^{\ell^*} (-R_c - p_h h)(-\lambda) f(\ell) d\ell + \int_{\ell^*}^{\infty} (-R_c - p_h h^*)(-\lambda) f(\ell) d\ell\right]$$

$$= -\left\{ s_0(-R_c - p_h h_0)\left(\frac{\partial\lambda_0}{\partial I}\right) + \int_0^{\ell^*} (-R_c - p_h h)\left(\frac{\partial\lambda}{\partial I}\right) f(\ell) d\ell + \int_{\ell^*}^{\infty} (-R_c - p_h h^*) \frac{\partial\lambda}{\partial I} f(\ell) d\ell \right\}$$

$$+ s_0(-\lambda_0)\left(\frac{-\partial^2 R}{\partial C \partial I} - p_h \frac{\partial h_0}{\partial I}\right) + \int_0^{\ell^*} (-\lambda)\left(\frac{-\partial^2 R}{\partial C \partial I} - p_h \frac{\partial h}{\partial I}\right) f(\ell) d\ell + \int_{\ell^*}^{\infty} (-\lambda)\left(\frac{-\partial^2 R}{\partial C \partial I}\right) f(\ell) d\ell$$

The expression in braces is the "pure" income effect on demand for insurance. Note that $r(I) \cdot (-\lambda) = -\partial\lambda/\partial I$, so the expression in brackets is simply the first-order condition (2.33) weighted by Pratt's risk aversion measure $r(I)$. It may be shown that if either $r(I)$ or $r^*(I)$ increases with losses, then the bracketed expression becomes negative; if either $r(I)$ or $r^*(I)$ decreases as ℓ increases, the opposite is true. In other words, the pure income effect is dependent upon the risk aversion sign and cannot be known *a priori*.[1]

[1]This result is also found by Mossin, "Rational Insurance Purchasing," p. 563. Relative risk aversion is defined as $r^*(I) = I \cdot r(I)$. The *change* in $r^*(I)$ as income changes is given as $r(I) + I \cdot r'(I) = r(I)(1 + \eta_{r(I),I})$. In other words, $r^*(I)$ may increase with income while $r(I)$ decreases with income, as long as the elasticity of $r(I)$ with respect to income is less than one in absolute value. For these purposes, declining *absolute* risk aversion is a stronger condition than declining relative risk aversion. Arrow has argued that $r'(I) < 0$, while $(r^*(I))'$ is positive. See his *Essays in the Theory of Risk Bearing*, Markham Publishing Company, Chicago, 1971, p. 96.

Income also affects the change of income in each state $(-R_c - p_h h)$, thus altering the implicit "real" prices of insuring each state of the world. Since the average income effect is incorporated into $\partial(-R_c)/\partial I$, how the "real price" of any state of the world changes depends upon how $\partial h/\partial I$ (the income effect) changes as losses increase or decrease.[1] Thus, *it is impossible to infer risk aversion behavior by observing the income elasticity of demand for reimbursement insurance* unless it is known with certainty that $\partial h/\partial I = 0$ for all states of the world. An additional prediction from this model is that the income elasticity of demand for various types of reimbursement insurance will vary, depending on how the demand response to income changes over states of the world, for the commodity or service being insured. I cannot, however, establish a predicted sign for $\partial C/\partial I$.

A second general proposition also follows. Given increasing risk aversion, if insurance premiums are not adjusted for income shifts (that is, if $\partial^2 R/\partial C \partial I = 0$), then demand for reimbursement insurance will increase with income. That may sound like a trivial statement, but it is of considerable policy interest. A substantial number of insurance policies sold in this country do not, in fact, adjust either total premiums or the rate $(-R_c)$ as income changes.[2] If increasing

---

[1]Where $-R_c = (1 + \theta) \int_0^{\ell^*} \left[p_h h - (1 - C)p_h \frac{\partial h}{\partial C}\right] f(\ell)d\ell$, the derivative $\frac{-\partial^2 R}{\partial C \partial I} = (1 + \theta) \int_0^{\ell^*} \left[p_h \frac{dh}{\partial I} - (1 - C)p_h \frac{\partial^2 h}{\partial C \partial I}\right] f(\ell)d\ell \approx (1 + \theta)p_h \overline{\frac{\partial h}{\partial I}} \approx (1 + \theta)\overline{w_h \eta_{hI}}$ if the income-coinsurance interaction term $\frac{\partial^2 h}{\partial C \partial I}$ is "small."

[2]The reasons for such policies being offered, and the reasons why they can remain in existence, even in a competitive market, are discussed in Section V. Briefly, the plans were historically offered because it was thought that the average rating policy (no income adjustments) would assist some low-income families in obtaining insurance. Even now, many group health insurance policies perform the same function, by charging every person in a given employment group the same amount for a given insurance policy, even though use differs systematically by income. It may be advantageous for a person (even if a low user on average) to remain in such a group if the group insurance is sold at substantially lower rates than individually purchased insurance policies. When these arrangements are found in equilibrium, there

risk aversion holds, then there will be a systematic demand for these insurance policies by persons with high income, since the "real price of insurance" declines to them as they purchase more medical care.

If $\partial^2 R/\partial C \partial I = 0$, the "pure" income effect on demand for C remains unchanged, as in (2.43). However, the substitution effect acts to reduce the costs of insurance to persons with higher incomes, since they systematically purchase more medical care for any given loss but have no adjustment made in their rate schedule $(-R_c)$. If the "price" of insurance for a given loss is $\bar{p} = (-R_c)/(p_h h + R_c)$—analogous to the Ehrlich and Becker formulation—then $\bar{p}$ will decrease for larger losses as income increases, if $\partial h/\partial I$ is positive and if no changes in $(-R_c)$ are made as income changes.

Effects of income on demand for h* involve only pure income effects. If risk aversion is increasing with income, then $\partial h^*/\partial I$ will be positive; the opposite is true if there is decreasing risk aversion.[1] Note that because C is the proportion of bills paid for by the consumer, a positive income elasticity of demand for insurance is indicated by having $\partial C/\partial I$ negative.

The Effects of Medical Care Prices on Demand for Insurance

Changes in the market price of medical care induce complicated changes in demand for reimbursement insurance; I will not derive the mathematics of the changes here, since the results are both difficult to describe and ambiguous in prediction.[2] There are three major

---

is a systematic subsidy from the poor to the rich through insurance premiums.

[1]It may be somewhat tempting at this point to use changes in demand for upper limits of coverage as a proxy for measuring whether risk aversion is increasing or decreasing. It *is* true that if only the upper payment bound is available for choice, the behavior of risk aversion could be inferred from the income slope of demand for insurance. But if more than one insurance parameter may be chosen by the consumer, the full model (as given in (2.39)) should be evaluated, and we cannot use the simple result for $\partial h^*/\partial I$ to evaluate risk aversion.

[2]The interested reader may obtain a more detailed description of the comparative statics of demand for C as p_h changes from C. Phelps, "Demand for Health Insurance: A Theoretical and Empirical Investigation," Ph.D. dissertation, University of Chicago, 1973.

effects of price changes on demand for C that are either unsigned or act in opposing directions. First, there is a term dependent upon how "relative risk aversion in medical care" changes as income changes. This is a concept unique to this model and is the difference between the usual risk aversion measure ($r^*(I)$) and the income elasticity of demand for medical care (η_{hI}). Terms involving this measure of risk aversion in health arise because the insurance is of reimbursement form, where payment is conditional on actual consumption of medical care. A second set of effects emerges as a basic income effect; as the price of medical care changes, the total insurance premium R also changes, inducing an income effect on demand for insurance proportional to the changes in R as p_h changes. Since R increases as p_h increases (given inelastic demand for h), increases in p_h in effect reduce available income, and the income effect will act in accordance with that. Finally, there is a set of terms depending on the *rate* that *expenditures* on medical care change as illnesses increase, terms that are dependent upon how $(1 + \eta_{hh})$ changes as the illness level changes. This last set of terms can be shown to alter the real price of insurance in each state of the world in a systematic fashion and can be interpreted as price effects in the Ehrlich and Becker sense. The net effect of these three components is ambiguous, so no formal prediction emerges from my model on the effects of changes in p_h on demand for reimbursement insurance. Analogous effects are observed in the comparative statics of demand for h*.

Effects of Changes in Time Price on Demand for Insurance

The time price of medical care can change either because the time component (t) changes or because the value of a unit of time (w) changes. The analysis of these changes is similar to that involving changes in p_h, except that the assumption can be made that the insurer cannot measure the values of w or t for specific individuals, so the premium is invariant to those derivatives. It can then be shown that the dominant factor in assessing how demand for C is altered as the time price changes is how the measure of risk aversion in health ($r^*(I) - \eta_{HI}$) changes over losses. If this measure increases with

losses to health (decreases as income increases), then demand for insurance is positively associated with increases in time costs.

The Effects of an Increase in the Loading Fee on Demand for Insurance

The loading fee (θ) is defined in my model as the excess of premiums over expected benefits and is proportional to expected benefits. It is important to realize that θ is not a price of insurance in the Ehrlich and Becker sense, although it has effects on demand that are exactly analogous to price effects in a standard demand model. There is a basic substitution effect that reduces demand for insurance as θ increases, just as in usual demand theory, and an income effect that is dependent upon the effects of income on demand for insurance. If insurance is a "normal good" ($\partial C/\partial I$ is negative, implying that more insurance is purchased as income rises), then demand for reimbursement insurance is always negatively related to θ. However, if $\partial C/\partial I$ is positive (that is, if insurance is an "inferior good") and if the budget share of insurance premiums is sufficiently large, then the possibility arises that demand for insurance could increase if θ rose. Further, it can be shown that as θ increases, the "real price of insurance" (in the Ehrlich and Becker sense) in each state of the world increases as losses increase, thus reducing demand for insurance as θ increases. Since demand for insurance is empirically found to be a normal good (see Sections III and IV), my model predicts that demand for insurance will be negatively related to the loading fee.[1]

Effects of Changes in the Illness Distribution $f(\ell)$ on Insurance Demand

An important question is how demand for insurance changes as shifts occur in the distribution of illnesses facing the consumer. Since the consumer may vary only the coinsurance rate C (which applies to all covered states of the world) or the maximum payment h^*, it is not intuitively clear that shifts in the illness distribution will increase demand for insurance. Without presenting the proofs,[2] I will

[1]Further details of this analysis are available from the author upon request.

[2]The proof is available from the author upon request.

describe the effects of shifts in the distribution $f(\ell)$ on demand for insurance.

The first effect observed arises because the premium changes as $f(\ell)$ shifts (recall that the premium includes all possible illnesses faced by the consumer). If the illness distribution shifts to the right (more serious illnesses on average), then premiums will rise, producing a proportional income effect on demand for insurance. Second, there is a substitution effect as the "real" price of insurance changes, depending on how the income effect systematically alters demand for medical care over various states of the world. Both of these effects are almost exactly proportional to the total income effect described earlier. Finally, there is an additional assumption of risk by the consumer when $f(\ell)$ shifts to the right, which will unambiguously increase demand for insurance. Thus, unless the income effect dominates the decision, I expect that persons who are on average sicker will choose more complete coverage. This is important to insurance pricing (because of the self-selection involved) and important for persons estimating demand for medical care, since it suggests that the amount of insurance held by individuals will be a function of their *expected* illness (but not actual illnesses observed at the time of medical care purchase).

The effect of self-selection is strengthened to the extent that insurance premiums do not reflect individual illness distributions, but rather some average illness distribution. This conclusion is hardly startling, but it is of strong consequence in the American health insurance market. Since it is expensive for insurance companies to detect the true illness distribution of each customer, many insurance policies (those group rated or community rated) reflect only average, not individual, illness distributions, and there are strong incentives for "sickly" persons to choose more complete coverage.

It can also be shown that increases in the variance of the illness distribution (holding constant the expected value) will increase demand for coverage, unless premiums rise strongly to compensate for the extra risk taken on by the insurer. If insurers act in a risk-neutral fashion, then my model unambiguously predicts that persons with higher variance in their illnesses will choose more complete insurance.

SUMMARY OF RESULTS

I have investigated demand for reimbursement insurance, attempting to sign the comparative statics of a general expected utility maximizing model. The results have been unsatisfactory in the sense that the model generates few clear, unambiguous predictions. In general, income effects associated with "pure" income changes, as well as "income effect" components of changes in market prices, time prices, and health levels, all depend upon the sign of the risk aversion measure for signature. If the assumption of increasing risk aversion (either absolute or relative) is made, the income effects are positive (insurance is a superior good), and effects of changes in market and time prices *for medical care* are shown to have a complex and unsignable effect on demand for reimbursement insurance. The model *does* reveal that factors affecting the average demand for medical care will enter into the demand for health insurance (including income, prices in the market, time and travel costs, and the *moments* of the illness distribution facing the consumer). The existence of actual demand curves for coinsurance (C) and for maximum coverage (h^*) have not been demonstrated, but it is shown that the mathematical conditions for such demand curves to exist have been fulfilled (there are a sufficient number of equations to theoretically solve for the unknown values of C and h^*, given all other parameters in the system).

The major benefit of this model is viewed in terms of model specification. Although the effects of the exogenous variables have not been established, a model has been rigorously derived that suggests the appropriate variables to include in a demand curve for insurance, and the model gives some insights into expected signs. Additionally, the model shows how demand for insurance and demand for medical care interact, establishing that a simultaneous equation system would be necessary to estimate demand for medical care and demand for health insurance. Furthermore, empirical estimates of the income elasticity of demand for insurance *cannot* be used to infer how risk aversion changes with income, since other factors enter the partial derivative $\partial C/\partial I$ beyond the risk aversion measure. The notion that "anything increasing demand for medical care will increase demand for coverage" is shown to be too

simplistic. Complex interactions between insurance and such factors as income, medical care prices, and entering health levels are more predominant than any simple effects.

Finally, the model developed a rigorous interpretation of the effects of insurance on demand for medical care. The majority of the testable hypotheses were derived from this work. Explicitly, the model predicts that the elasticities of demand for medical care with respect to money prices and with respect to coinsurance changes are identical; a relationship between those elasticities and the elasticity with respect to time price for medical care is also obtained. These relationships are not tested here, since the purpose of this study is to investigate demand for insurance.[1]

The next two sections describe two estimates of the demand for reimbursement insurance--one using a time-series study where average insurance coverage rates for medical care are used as the dependent variable, and the other where actual insurance contracts of individual families are investigated in a micro-cross section study. Both of these estimates rely upon the theory developed in this section for model specification, and both rely upon the existence of the demand curves for insurance that are implied by the model presented here.

---

[1]Empirical evidence on the coinsurance elasticities, time-price elasticities, and money-price elasticities is presented by Phelps and Newhouse, *Coinsurance and the Demand for Medical Services*.

III. ANALYSIS OF HOUSEHOLD INTERVIEW DATA ON DEMAND FOR INSURANCE

The theory derived in Section II suggests components of demand to be used in any estimation. Although no explicit demand curves are obtained, the mathematics permit indirect inferences to be made about the nature of demand. Most of the comparative statics of demand yielded ambiguous predictions, although non-zero effects of the various independent variables are to be anticipated in general. In this section, I attempt empirical estimation of the demand for reimbursement insurance with a body of cross-section data from 1963. In Section IV, I use highly aggregated annual time series observations to estimate similar demand curves. Wherever appropriate, comparisons between the two studies will be drawn.

THE CROSS-SECTION DATA

Data are available from a national health care survey, conducted by the Center for Health Administration Studies of the University of Chicago, in 1963.[1] The study investigates use of the health services system by 2367 randomly selected families representative of the national population. Among other questions, the families are asked what types of insurance they hold (and with which companies) and how much total health insurance premium is paid by each family. That is one source of data about the level of insurance of families in this sample.

Additional data are provided by the insurer listed for each family. An *insurance verification form* was sent to each listed insurer, asking for explicit details of the insurance contract. These forms verify the premium amounts paid by the consumer and give the amount of premium contributed by the employer (if any). The insurer also provided details of the insurance contract--the maximum payment per hospital day, the maximum number of hospital days covered, the maximum payment to be made under surgical insurance, the maximum payment under major

[1] Ronald Andersen and Odin W. Anderson, *A Decade of Health Services*, University of Chicago Press, 1967.

medical insurance (if any), and so on. These data can be used as dependent variables in a regression study; they represent quantities of insurance demanded and total expenditures on insurance.

Unfortunately, the return rate on the insurance verification forms is considerably less than 100 percent, and it is particularly low for those with group insurance, as seen in Table 1.

For policies for which no verification was received, the only information available is the total family premium, as reported in the family interview, and some information relating to the presence or absence of various types of insurance. In the family interview data there is no comparable information to that obtained in the verification process about specific details of the policy. For families where verification is incomplete, that information is lost.

Non-reporting by insurers raises several estimation problems. The families for whom insurance information is missing (who are employed and have group insurance) probably are more highly insured than the "average" family. Yet if we use the "reported" values for their insurance contracts, we obtain an artificial "zero" value rather than the true value of insurance carried by the family. Non-reporting by group insurers should bias the estimated coefficients, since we know that larger values of insurance are systematically missing. Even if the

Table 1

VERIFICATION RATES OF HEALTH INSURANCE POLICIES

	Total Reported by Family	Total Verified by Insurers	Percent Verified	Total Rejected[a] by Insurers	Number Rejected as Percent of Total Number of Policies
Individual policies	685	558	81.5	55	8
Group policies	2014	1065	52.5	46	2.3
All policies	2699	1623	60.3	101	3.72

[a]The insurer responded to the verification but indicated no records of enrollment by the family.

coefficient estimates are not biased, we should observe a reduction in the goodness of fit of the estimation, because of the additional errors of measurement of the dependent variable. This problem may be circumvented to some extent by eliminating from the sample those families for whom verification is incomplete, but there is then no guarantee that the remaining sub-sample is representative of the population as a whole.

Verified Data

To establish the effects of incomplete verification, I have eliminated from one set of regressions all families for whom insurance verification was incomplete. The remaining data can be characterized as being relatively free of measurement errors in the dependent variables, but that sub-sample cannot necessarily be viewed as representative of the population as a whole. Comparison of data means in Appendix E shows that families remaining after this procedure have markedly lower incomes than the entire population ($5797 for the sub-sample, with verified data; $6370 for the whole sample). It is difficult to establish whether the remaining sample purchased significantly less insurance than the sample as a whole, because that is precisely the information missing for the unverified families. Comparison of expenditures for these two groups show that the entire sample had an average purchase of $116, whereas the "verified" sample had an average purchase of $109. This comparison may be misleading, since the unverified data will not include any estimates of employer-paid premiums for those families with missing verification.[1] Families without verified insurance will have downward biased reports on their insurance level. Use of the entire sample should bias toward zero the income and own-price

[1]Estimates of the total insurance premium can be derived from the data, even with incomplete verification. In those families for which verification was complete, the average verified premium was $102, $71 of which was paid by the family, and $31 of which was paid by employers. Assuming that ratio holds for the remainder of the sample, and that MAXPRM measures only family-paid premiums in the families for which verification was incomplete, the estimated average premium for the entire population can be computed, using the information about rates of verification. I computed that figure to be approximately $120.

coefficient estimates, since high income and large group-size families will have systematically low reported values for premiums.

Group Insurance Purchases and Limited Choice of Insurance

A second sub-sample is used to demonstrate the problems associated with estimating demand for group-purchased insurance. Group insurance can be sold at substantially lower loading fees than can individual insurance, so that availability of a group plan reduces the price of insurance to the consumer, but typically there is very limited choice among types of insurance in these group plans. The consumer frequently must choose between an "optimal" plan at a high price and a "sub-optimal" plan at a lower price. Often his expected utility is higher under the group insurance plan, and he will rationally make that choice.[1] In such cases, the theory of demand for insurance I have developed may not be well supported by observed purchases, because individual consumers will have little choice among insurance policies (holding price constant at the group insurance level). Put differently, the revealed preferences of the consumer will demonstrate only an overwhelming "own price" effect, but the predicted effects of income, the price of medical care, and other factors may not be observed in group purchase data.

To understand how this phenomenon affects estimation of demand for insurance, a sub-sample of families was taken where the family purchased either *no* insurance or purchased *at least one* individual policy; these families may also have some group insurance, with the individual plan supplementing their group coverage. The demand characteristics of the population should be better estimated from this sub-sample, since the families in this sub-sample will be at an equilibrium demand position that reflects their private demand rather than the average demand of all members of the group through which they are insured. I hypothesize that demand will be estimated with better precision for this sub-sample than for the entire population, and that the estimated elasticities of

---

[1]The terms "optimal" and "sub-optimal" are used advisedly. If the two policies were available at identical prices, the consumer would choose the plan I term "optimal" in preference to that I term "sub-optimal." In context of the actual prices observed, the consumer will always rationally choose an "optimal" plan.

demand will be further from zero for every independent variable except the price of insurance.[1]

For example, suppose that insurance demanded is a positive function of income. In any work-group insurance plan, there will be a variety of persons with different income levels, each of whom demands insurance according to his income level if he were purchasing insurance individually. If some "floor" is placed on the amount of insurance available (at reduced price), then some persons will purchase more insurance than they would purchase individually (if prices were constant over all policies). Those with low incomes will (rationally) purchase the group policy, placing them at the "floor" level of insurance offered through the group. Any least squares regression estimate of the "income effect" on insurance demand will be biased toward zero, as shown in Fig. 7. By using a sub-sample of only those families who purchase at least one individual policy (or those who purchase no insurance), we should estimate demand curves for persons on their "private demand curves," persons on demand line AA in Fig. 7. In that sub-sample, there is a poor estimator of the loading fee on insurance, however, since my estimate for that variable depends on work-group size.[2] For this set of regressions, the own-price variable (the loading fee, represented by work-group size, may have a biased coefficient, since for individually purchased policies, the group-size measures with error the actual loading fee faced by those persons without group insurance possibilities.

Those families observed with high levels of work-group insurance could be characterized as being overinsured, especially when the employer pays part of the insurance premium. In such cases, the family cannot normally obtain the equivalent amount of wages, in lieu of insurance, and cannot effectively use their *reimbursement* insurance policy as a means of securing loans in capital markets.

---

[1]Because of self-selection, this sub-sample may be atypical of the entire population; the results must be considered with that in mind.

[2]See p. 53 for a discussion of how work-group size may be used as an instrument for the price of insurance.

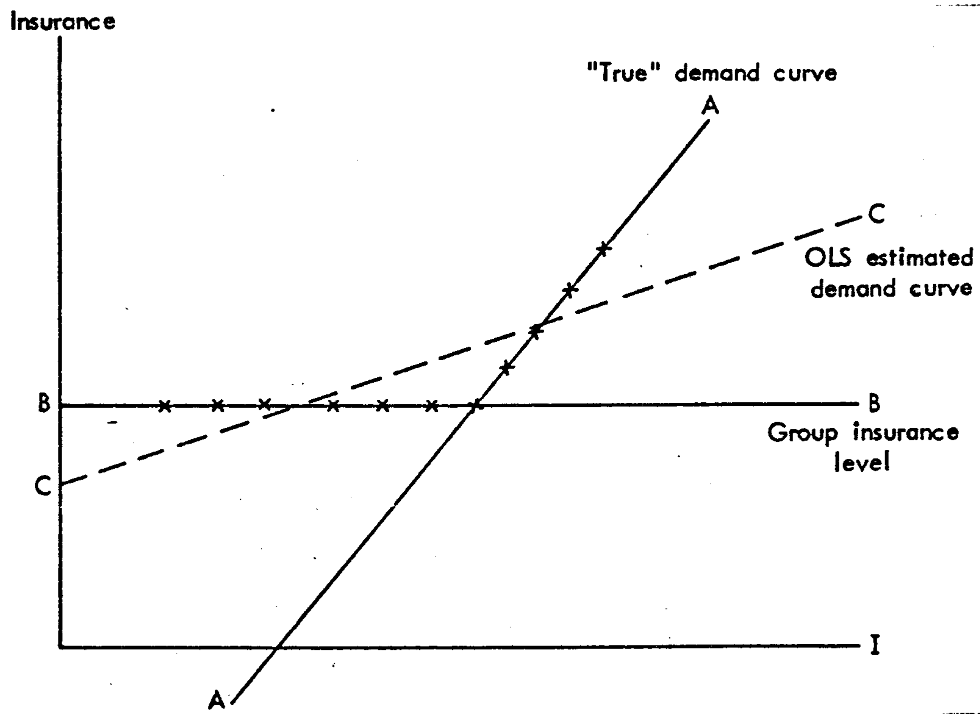

Fig. 7 — Distortion of demand estimation with imposed group insurance

DEPENDENT VARIABLES

Expenditure on Insurance

In the estimates of demand for insurance, I used several different dependent variables. The means and standard deviations of these variables are described in Appendix E.

MAXPRM: *Max*imum estimate of the total *prem*iums of the family. This variable is used to obtain a "best" estimate of the total insurance premium of the family. For each family unit, MAXPRM is the larger of the family-reported premium (from the household survey) or the verified premium (from the insurance verification form received from the insurer). If the policy is verified, then MAXPRM normally includes premium payments by both the family and the employer. If no verification was received, then MAXPRM will refer only to family-paid premiums.

VERPRM: *Ver*ified *prem*iums for the family. VERPRM gives the total of all verified premiums for this family, including both family and employer contributions. In the

sub-sample of families for which all insurance policies were verified, MAXPRM and VERPRM should be identical. For the entire population with complete verification, MAXPRM has a mean value of $108 and VERPRM a mean value of $102; the two variables have a simple correlation of .98. Presumably, differences between the two are caused by families overstating their insurance payments during the family interview.

These premium variables will be used as one measure of the demand for insurance, just as any other expenditure data might be used for such an estimate. Relationships between demand for premiums and demand for coverage itself are derived in Appendix B, and are discussed more extensively later.

Coverage Levels

Other insurance variables relate more directly to the level of "coverage" provided by insurance. The potential variety of health insurance policies precludes direct comparison of all such policies; I shall estimate demand for the most common types of insurance parameters. The three most common types of health insurance relate to hospitalization, medical-surgical (designed to pay physician's fees), and "major medical" insurance, which covers a broad range of medical expenses. In hospitalization insurance, two parameters are of importance in determining the amount of "coverage" purchased--the total number of hospital days covered, and the maximum payment allowed per day of hospitalization.

The explicit variables used to estimate demand for hospitalization insurance are:

HOSMAX: The *max*imum payment allowed per *hos*pital day. This variable gives the explicit maximum payment specified in the contract for any hospitalization. If the policy specifies that no maximum limit is set, or if the policy pays for the "full semi-private room rate," a value of $30 per day was coded for that policy. In the NORC sample, 32 percent of the families have a policy that specifies "full semi-private room rate," and they received a coded value of $25.

HOSDAY: The maximum number of *hos*pital *day*s covered by the policy. Most hospitalization policies place some upper bound on the number of hospital days covered. If no limit is specified in the policy, a value of 365 days was coded in for HOSDAY. Only 1.5 percent of the policies in this sample set no upper bound on coverage.

Another type of health insurance is medical-surgical insurance, which pays surgeons' and physicians' fees, primarily for work performed in hospitals. The payment mechanisms vary considerably from policy to policy. Some pay according to a specific fee schedule for any procedure performed by the physician. Others pay "usual and customary" fees of the physician. Generally, the relative levels of payment for any given medical procedure are fairly constant within a given policy, so their specifications for the "largest amount to be paid for any procedure" can be taken as a good approximation of the coverage level of the policy.

In estimation, I refer to this variable as

SURMAX: The *max*imum payment for *sur*gical procedure. This variable gives the maximum payment for any surgical or medical procedure, as specified in the insurance contract; if "no limit" is specified, a value of $900 has been coded, representing (approximately) the 95th percentile of all surgical maximum payments observed in this sample.

Finally, I consider the purchase of major medical insurance. These policies, rather than restricting coverage to a specific type of medical treatment, cover a broad range of treatments. They specify some dollar deductible (ranging from $50 to $500 or $1000, $100 being the most common value), an explicit copayment rate for all expenditures above the deductible (almost uniformly 20 percent payment by the consumer, 80 percent by the insurer), and a maximum dollar payment under the plan, typically large relative to surgical maximum payments. The smallest "maximum payment" observed of those actually purchasing major medical insurance is $5000, with many policies specifying amounts between $10,000 and $20,000. The largest maximum observed in this sample is $40,000.

The explicit variable used in the study is:

MM-MAX: The *m*ajor *m*edical *max*imum payment. The largest total expenditure covered under the major medical policy. If the family has no major medical policy (as is true for 74 percent of the families), a value of "zero" is coded for MM-MAX.

In every case, if the family has no insurance of the type being considered, a value of "zero" is coded for that family. In families with more than one insurance policy, the insurance policies have been aggregated within the family. In all "premium" variables, the insurance premiums have simply been added together. In the "maximum payment" variables, the largest coded value for any of the policies was taken as the level of coverage held by the family. For example, if a family held two hospitalization policies, one of which paid \$15 per day for hospital room and board and another \$20 per day, a value of \$20 would be coded in HOSMAX for that family.

INDEPENDENT VARIABLES

The theory of demand for reimbursement insurance developed in this study suggests which exogeneous variables should influence demand for insurance. These include income, the price of insurance, the price of medical care, and the distribution of illnesses facing the consumer. Since direct measurements are not available for any of the actual economic variables suggested by the theory, instrumental variables are used in each case.

Income and Education

I have used an estimate of permanent income for the family data regressions in this study. An instrumental variable for income was constructed by regressing observed family income on a variety of factors that should be free of transitory behavior. This instrument was used as a measure of permanent income. The exact estimation is given in Appendix E. The measure of permanent income had a multiple correlation of .65 with observed income, and the variance of the "permanent" income measure is approximately 25 percent lower than the variance of observed income.[1]

---

[1]In M. Friedman's studies of the consumption function, he found a higher proportion of variation on observed income accounted for by permanent income; the range in his budget studies was .63 to .87. His method of estimation is entirely different from mine, however, and involves observations of total consumption by family units. The lower R^2 in my instrument may be attributed to actual measurement error of

In the grouped data regressions, average income for each cell was used as a measure of permanent income. In the actual grouping for this study, all observations were sorted into one of 192 groups, according to (1) education of the head of the family, (2) geographic region (North or South), (3) race of head of family, (4) area within a city (central city, suburban, fringe, or open country), and (5) type of dwelling unit (home or multiple unit dwelling). Each family fell into one of the 192 groups when all five of these characteristics of the family are taken into account. The data means for these 192 groups then become the "new" observations in a regression of the form $Y_i = BX_i + U_i$, where the capital lettered values for X, Y, and U indicate aggregated data. Since the variance of the error term from such an equation will be systematically related to the number of observations used to calculate Y_i and X_i, the regressions must be weighted to eliminate the inherent heteroscedasticity. In particular, since $VAR(U_i)$ is proportional to $1/n_i$ (n_i = number of observations in the ith cell), the transformed regression $\sqrt{n_i}Y_i = \sqrt{n_i}BX_i + \sqrt{n_i}U_i$ is used, which will have the homoscedastic properties assumed for ordinary least squares.

The effects of education on demand are complicated. It is well known that education will affect permanent income of a consumer unit--indeed, the family head's education was one of the predictive variables used in the instrumental variable estimate of income for the families. Thus, inclusion of head's education as a separate independent variable may obscure actual income effects. Second, education may affect the family's ability to produce non-market alternatives to market health insurance (home production of medical care may be one such alternative). Finally, holding income constant, education should indicate the opportunity wage rate of the household head. I initially included head's education to reflect the latter phenomenon. That is, I interpret "education" as a measure of opportunity costs of time, since "income" is held constant in the regression.

---

observed income, and by omitted variables in the regression predicting "permanent income from observed income." See Milton Friedman, *A Theory of the Consumption Function*, Princeton University Press, 1957, Table 4, p. 67.

<u>The Price of Insurance</u>

In a cross-section, I do not have good estimates of the effective loading fee for any given insurance policy. I use an instrumental variable that should be well-correlated with actual loading charges, at least for a large portion of the insurance observed in the sample. That variable is the *work-group size* of the head of the household (GRPSIZ). In the health insurance market there are significant economies of scale in selling to large groups of consumers. Most insurers offer explicit group discounts on loading fees; the discounts increase as the contract size increases. As noted in one text on health insurance, "The discount scale is a part of the filed rate structure of each group health company."[1]

MacIntyre[2] reports that the average payout ratio for all individual insurance policies in 1960 was .51, implying a gross loading fee of .96. He also notes that the average group insurance payout rate was .91, implying a gross loading fee of .1. Blue Cross and Blue Shield plans (non-profit insurers, not subject to taxation, whose business is almost exclusively from group insurance) reported an average payout rate of .94 and .90 respectively, implying gross loading rates of .06 and .11. The gross differentials between group and individual insurance are obvious. Of more interest are figures provided for reductions in premiums as total expenditure increases. Table 2 shows representative group insurance volume discounts. The loading fee reduction can be derived from any premium reduction. For example, if the gross monthly premium includes a gross loading fee of 20 percent, even a 1 percent total reduction in premium implies a reduction in the loading fee of about 5 percent. A 10 percent reduction in the gross premium is equivalent to a 50 percent reduction in the gross load.

In the 1963 NORC sample, the average monthly premium per family is approximately $10, so that the 3 percent discount would be available to a group of approximately 60 persons. An average group of 500

---

[1] O. D. Dickerson, *Health Insurance*, Richard D. Irwin & Co., Homewood, Ill., 1959.

[2] Duncan V. MacIntyre, *Voluntary Health Insurance and Rate Making*, Cornell University Press, 1962.

Table 2

SCALE ECONOMIES IN INSURANCE PURCHASING

Gross Monthly Premium	Percent Reduction in Premium
$ 500-750	3
750-2,500	5
2,500-3,000	6
3,000-3,500	7
3,500-4,000	8
4,000-5,000	9
5,000-7,500	10
7,500-10,000	11
10,000-15,000	12
15,000-20,000	13
20,000-25,000	14
25,000 +	15

Source: O. D. Dickerson, *Health Insurance*, Richard D. Irwin Co., Homewood, Ill., 1959, p. 376.

persons would have a monthly premium of $5000, qualifying them for a 10 percent reduction on premiums.

The measure of work-group size available in the NORC survey is a scaled variable, ranging from 1-8 (where 1 = 1-2 persons, and 8 = 500+ persons). From these data I have estimated a conversion factor that can be used to translate demand elasticities with respect to "group size" into demand elasticities with respect to the gross leading fee. At average values for the data sample, the appropriate conversion rate is -2/3. That is, if the estimated "group size" elasticity of demand were +1.0, the comparable estimate of elasticity of demand with respect to gross loading fee would be approximately -.67. This rate is not intended to be exact, but only to provide an approximate conversion between the estimated elasticities with respect to GRPSIZ and the corresponding loading fee elasticities.

When individual insurance is purchased, the indicator of the effective loading fee is less reliable. The explicit variable used in the regressions is:

GRPSIZ: Work *Group Size* of the head of the household. GRPSIZ is a variable reported in the household survey, on a scale from 1-8. The scale values, the corresponding work group size, and the estimated loading fee for a group of that size are given in Table 3.

If a family has no group health insurance, the value given for GRPSIZ is 1, indicating that the full individual policy loading fee would be faced by the family.

The Price of Medical Care

In cross-sectional data, it is difficult to obtain meaningful estimates of the "true" price of medical services faced by a given consumer. There are well-known price differentials between urban and rural areas and differences between different geographic areas of the country; regional dummy variables have been used in many studies of demand for medical care and then interpreted as "price" coefficients. I have attempted to provide more precise data by using BLS expenditure data.[1] "Standard" family expenditures for medical care are given for 39 cities throughout the country and four regional "non-urban" areas. Expenditures are estimated for three "standards of living" within each of the areas sampled. The expenditure data given are said to be estimated from samples where the quantity of the medical services is held constant; therefore, the expenditure differentials should accurately reflect medical care price differentials among the 43 different areas of the country sampled by the BLS. The problem of matching the appropriate "prices" from this study with each family in the NORC survey was solved by locating each family in a "primary sampling unit," providing information on actual city of residence. These data, combined with information about type of housing of a specific family, are sufficient to give a reasonable matching algorithm between the BLS price data and the 1963 NORC survey data. In the ten largest cities in the country, a direct match between the two is possible. In other cities, the closest comparably sized city in the same geographic region was

[1] *Standards of Living for an Urban Family of Four Persons,* Bureau of Labor Statistics, 1969.

Table 3

ESTIMATED LOADING FEE FOR VARIOUS WORK-GROUP CATEGORIES

GRPSIZ	Number of Employees	Estimated "Loading Fee"
1	1-2	.95
2	3-4	.75
3	5-9	.45
4	10-20	.25
5	21-50	.18
6	51-100	.15
7	100-499	.10
8	500 +	.08

matched. The BLS "standard of living" was assigned to each family on the basis of type of housing and residence in either central city or suburbs (for example, a homeowner in a suburb was assigned a "high" standard of living, whereas an apartment dweller in a central city area was assigned a "low" standard of living). All persons living in non-urban areas were assigned to the appropriate "rural" category, based on geographic region. The use of home ownership as a criterion is approximately equivalent to that of the BLS study--in "moderate" standard of living brackets, 75 percent of families were homeowners; and in "higher" standard of living brackets, 85 percent were homeowners. The living arrangement in the lower standard was limited to rental housing only, "the prevalent pattern at the lower end of the consumption scale."

These matched prices obviously measure with some error the actual medical care prices faced by the families in the NORC study. The errors may be systematic rather than random, so that the usual errors in variables (EV) result may not hold. Since the medical price variable used in this study is, in effect, an average price for a region, those families in a given region who actually face high prices will have a measured price that is too low, and vice versa. Hence, a negative covariance between the price variable and its measurement error has been established, and the usual EV result (which assumes zero

covariance) will not hold. It is easily shown that in a simple regression model, where p is the true price and u is the measurement error in price, the probability limit of the coefficient estimate is given by:[1]

$$(3.1) \qquad \text{plim } b = B\left(\frac{\text{var}(p) + \text{cov}(pu)}{\text{var}(p + u)}\right)$$

$$= B\left(1 - \frac{\text{cov}(pu) + \text{var}(u)}{\text{var}(p + u)}\right).$$

The bias in the coefficient estimate can be in any direction. If the cov(pu) is negative and larger in absolute value than var(p), then the estimated coefficient can actually change sign. Alternatively, if the cov(pu) is negative but larger in absolute value than var(u), then the estimate is biased away from (not toward) zero. Hence, with systematic errors of measurement, the coefficient estimate may be upward or downward biased, and may even have the wrong sign. It is also possible, if var(u) + cov(pu) exactly equals zero, that there is *no* bias in the coefficient estimate. The possibility of systematic error in medical price measurement should be kept in mind when considering the medical price coefficient estimates with the resulting possible biases in that coefficient estimate and other (non-orthogonal) coefficient estimates in the equation.

The Distribution of Losses

My theory of demand for insurance shows that demand for insurance coverage may increase as larger losses become more probable. Insurance premiums should also increase, since the expected insurance payouts for medical expenses will increase. Many problems arise when one actually attempts to estimate this "illness distribution" for a given family. In my theory, I discuss changes in the probabilities of a

---

[1] Joseph P. Newhouse and Charles E. Phelps, *On Having Your Cake and Eating it Too: A Review of Estimated Effects of Insurance on Demand for Health Services,* R-1149-NC, The Rand Corporation, Santa Monica, forthcoming.

loss from the "stock of health." In the real world we cannot actually observe those losses, which makes estimation of their distribution somewhat difficult. I used a variety of measures to estimate the mean of this distribution, none of which gave ultimately satisfactory results in the demand equations. The technique finally used was to regress reported family "disability days" on a set of demographic variables, using the estimated value as a predictor of the mean of the illness distribution. Other estimates of the distribution of illness were considered and rejected, simply because they did not add significant information to the empirical estimates of demand for insurance.

ESTIMATION TECHNIQUES

Three basic estimating techniques have been used with these data, each with apparent strengths and weaknesses in providing an accurate analysis of the data. The basic techniques involved are individual (micro-unit) regression, grouped data regression, and "Tobit" regression. The first of these, normal multiple regression, is used as the basic estimation technique. Estimates were also made on grouped data, primarily as a tool for eliminating random errors in independent variables. The methods of grouping have been discussed previously. The third estimating technique, "limited dependent variable estimation," was proposed by J. Tobin.[1]

Multiple Regression

The virtues and vices of multiple regression are well known. I am concerned primarily whether or not the underlying assumptions of multiple regression are well met in the data. The basic assumptions of the linear regression model are (1) the dependent variable is linearly related to the coefficients and to a random error term; (2) the error term has zero mean and constant variance over all observations, and the errors of different observations are uncorrelated. Finally, it is usually assumed that the independent variables are nonstochastic

---

[1] James R. Tobin, "Estimation of Relationships for Limited Dependent Variables," *Econometrica,* 26, No. 1, January 1958, 26-36.

and therefore uncorrelated with the error term of the equation.[1] Modifications of this theory provide proper estimating techniques when the homoscedasticity assumption does not hold and provide insight into estimation problems when the explanatory variables are random, with their stochastic component uncorrelated with the error term of the equation. Of more importance in my sample is the assumption that the dependent variable is linear in a random error term. That is,

$$y_j = X_j B + u_j, \text{ where } E(u) = 0 \tag{3.2}$$

implies that, for sufficiently small values of $X_j B$, a negative value of y is permitted. That assumption is violated by my sample data. Tobin, in his article on the "Tobit" technique, discusses this issue.[2]

> There are indeed conceivable values of the independent variables for which multiple regression would give negative estimates for the expected value of [the observed dependent variable]. It is true that the absence of negative observations in the sample tends to keep the regression above the axis until extreme values of the independent variables are reached. *But this protection is purchased at the cost of making the regression line so flat that expenditure is underestimated at the opposite end.* These discrepancies could be important in predicting aggregates which include extreme cases. [Emphasis added.]

Grouped Data Regressions

One technique that may alleviate the problems of estimating a limited dependent variable is to use grouped data means as data points. Grouping data observations will substantially reduce the frequency of cases where the dependent variable has a value of zero. The coefficient estimates from grouped data regressions are unbiased and are efficient if the regressions are weighted by the number of observations in each grouping.[3] As in any aggregated data, the average value of insurance

---

[1] Arthur S. Goldberger, *Econometric Theory*, John Wiley & Sons, New York, 1969, p. 162 ff.

[2] Tobin, "Limited Dependent Variables," p. 36.

[3] Goldberger, *Econometric Theory*, pp. 235-236.

purchases for a group "cell" will be

$$(3.3) \qquad y_j = \frac{\sum_{i=1}^{n_j} y_{i_j}}{n_j} = \frac{\sum_{i=1}^{n_j} y_{i_j}}{k_j}\left(\frac{k_j}{n_j}\right),$$

where y_j is the mean of the dependent variable in the jth cell, n_j is the mean of the dependent variable in the jth cell, and k_j is the number of actual purchases of insurance in the jth cell. Grouped data means may be thought of as a joint estimate of the average value of positive purchases in the group and the probability of a positive purchase. The resulting average value for the group may not represent the actual demand of any member of the group.[1]

Limited Dependent Variable ("Tobit" Analysis)

Tobin has proposed an alternative method for estimating data when the dependent variable has only limited range. This one-step procedure jointly estimates the probability of a purchase and its value if positive. The technique is an extension of *probit* analysis, hence the name "Tobit." The basic model considered by Tobin specifies

$$(3.4) \qquad y_j = \begin{cases} 0 & \text{if} \quad I_j < u_j \\ I_j - u_j & \text{if} \quad I_j > u_j \end{cases}$$

where y_j is the observed purchase and u_j is a $(N(0,\sigma)$ variable.[2] The

---

[1]For example, for major medical insurance maximum payments, the "grouped data" means range from \$0 to \$4000, whereas the minimum *actual* purchased major medical insurance policy had a "maximum payment" of \$5000. The modal payment under major medical insurance in 1963 was \$10,000. The discrepancy between the "grouped data" averages and individual data simply reflect that only one quarter of the families in the sample made any positive purchases of major medical insurance.

[2]The basic model considers general "limit" for the dependent variable; here the limit is always zero.

index I_j is written as

$$I_j = X_j B \,, \tag{3.5}$$

where X_j is a vector of independent variables and B is a coefficient vector. For any value of I, there will be a concentration at $y = 0$, and there will be a distribution of positive y's. Obviously, the higher the value of the index I_j, the higher the probability of a positive purchase being observed and the higher the expected value of the purchase. A maximum likelihood estimate of the coefficient vector B is formed, in which estimating equations are non-linear. Figure 8 shows this graphically.

In Tobit regressions, a log-likelihood (chi-square) test provides a test of the hypothesis that the entire coefficient vector is identically zero, and asymptotically normal ratios of coefficients to their standard errors are computed from the estimated covariance matrix.

Estimation of Total Premium Demand Curves

In a classic demand estimation, one may estimate either quantity demand curves or expenditure curves and theoretically obtain the same coefficient estimates. The exception is the coefficient on the own-price for the good, which is an estimate of $(1 + \eta_{xp_x})$ rather than of η_{xp_x} itself. In insurance demand, the situation is more complicated. In Appendix B I derive expressions relating premium elasticities to elasticities of demand for coverage. These relationships show approximately what elasticities for C are implied by an observed elasticity for premiums (with respect to any explanatory variable). For all variables, it is necessary to know the price and income elasticities of demand for health care itself to be able to make the appropriate inferences, as the expressions from Appendix B show.

RESULTS OF ESTIMATION

Demand for health insurance has been estimated for the entire set of dependent variables discussed above, and with several estimating techniques. The results of those estimations are set forth in Tables

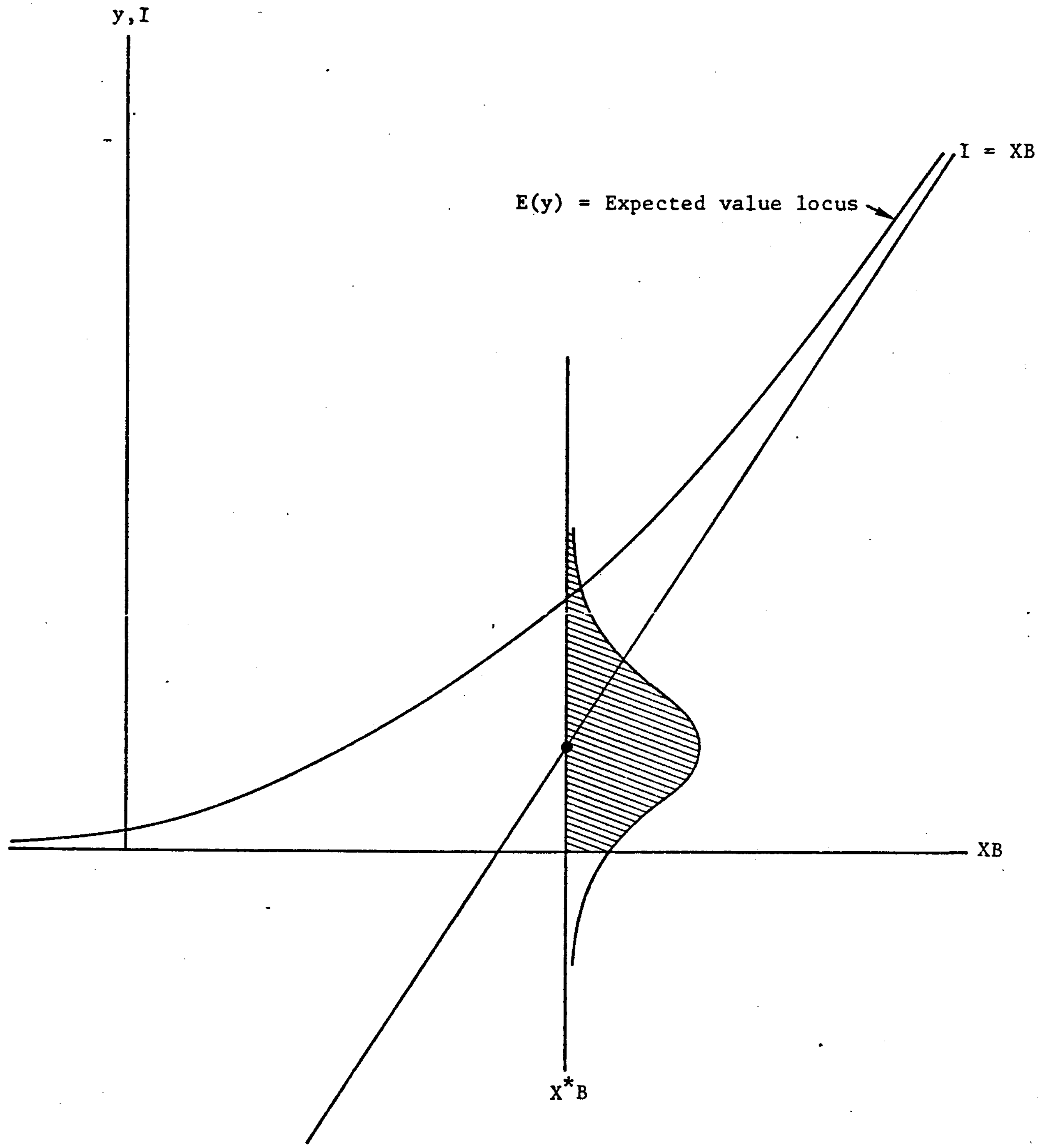

Fig. 8 — The total expected value locus (shaded area shows $P(y>0|X^*,B)$)

4 through 9. Each of these tables gives the estimated elasticities at mean values of the variables, reported with the associated t-statistic computed for the coefficient. Dummy independent variables do not have continuous ranges, so elasticities at mean values are somewhat meaningless. I have reported the estimated *coefficients* for these variables, along with the associated t-statistic. In every case, these coefficient estimates are preceded by a "$" since the units of measurement imply dollar amounts. Those values without the dollar sign are always elasticity estimates.

Total Premiums: MAXPRM and VERPRM

Total insurance premiums are estimated by MAXPRM (the maximum value reported either by the family at point of interview, or by the insurer through the verification process), and by VERPRM (total verified premiums). The regressions estimating total health insurance expenditures are straightforward. The estimated income elasticity of demand (using permanent income) is in the neighborhood of .33 to .47, depending on the sample used. The expected bias in the income elasticity using the total sample was observed, suggesting that the higher of these estimates is probably closer to the "true" value. In the verified data, the estimated total health insurance premium had an income elasticity of .42 (MAXPRM) and .47 (VERPRM). The t-ratios on these estimates (4.25 and 4.3) place a very low probability on the hypothesis that the income elasticity of premiums is nonpositive--the .95 confidence limit is .22 to .62 for MAXPRM. Similarly, the effects of education on total premiums are estimated to be positive, but the precision is lower than for income elasticity estimates.

The education effects have been estimated in both quadratic and logarithmic form, with somewhat differing results. With the logarithmic form of education, the estimated elasticity is .26, with a t-ratio of 3.25; but the estimated elasticity in quadratic form at mean values is .15 to .20 with t-ratios of 2.2 in the complete and verified samples. There appears to be significant multicollinearity between income and the education variables in those equations using the quadratic form for education.

Table 4

DEPENDENT VARIABLE: MAXPRM[a]

Data Base and Estimation Technique	Permanent Income (η)	Education (η)	Age (η)	Expected Illness Level (η)	Work-group Size (η)	Medical Care Prices(η)	Welfare Medical Care ($)	Black ($)	Rural ($)	R^2	F
1. Individual family data n = 2367	.33 (3.63)	.21[b] (2.25)	.27[b] (4.71)	.55 (3.52)	.34 (12.15)	-.15 (0.96)	-46.6 (3.74)	-10.0 (1.39)	-5.9 (1.16)	.203	42.87
2. Individual family data n = 2367	.22 (3.06)	.54[c] (3.56)	.60[c] (5.91)	.50 (3.31)	.29 (12.15)	-.08 (0.53)	-46.0 (3.69)	-16.5 (2.40)	-7.7 (1.53)	.200	49.04
3. Verified individual family data n = 1579	.42 (4.24)	.15[b] (1.61)	.37[b] (5.81)	.51 (2.85)	.45 (17.91)	-.39 (2.17)	-38.3 (2.82)	-19.6 (2.52)	-5.7 (1.01)	.359	62.53
4. Grouped data--base n = 2367	.20 (2.20)	.20[b] (2.14)	.49[b] (3.56)	.13 (1.37)	.35 (6.47)	.26 (1.18)	-104.0 (3.30)	--	--	.666	32.69
5. Grouped data--base n = 1579	.30 (3.16)	.47[b] (4.72)	.87[b] (5.78)	.21 (3.80)	.34 (6.73)	.12 (.47)	-11.5 (.36)	--	--	.710	40.14
6. Tobit regressions[d] n = 2367	.34 (3.30)	.35 (3.09)	.45 (7.00)	.42 (2.55)	.33 (13.83)	-.22 (1.30)	-69.9 (4.82)	-21.0 (2.65)	-6.2 (1.14)	--	613.6[e]
7. Families with no insurance or at least one individual plan n = 1208	.41 (3.18)	.41[b] (2.77)	.65[b] (5.87)	.62 (2.50)	.40 (14.33)	-.29 (1.09)	-41.0 (2.62)	-7.9 (.86)	-4.4 (.65)	.31	38.50

[a]Elasticities reported unless column is headed by a "$." t-levels in parentheses.

[b]Variable entered in quadratic form. Elasticity reported is at mean values of x and y, and t-test is on joint vector $b_1 + 2b_2x$.

[c]Entered in logarithmic form.

[d]Elasticities of E(y). t-statistic is not exact.

[e]Chi-square test with 14 d.f.

Table 5

DEPENDENT VARIABLE: VERPRM[a]

Data Base and Estimation Technique	Permanent Income (η)	Education (η)	Age (η)	Expected Illness Level (η)	Work-group Size (η)	Medical Care Prices (η)	Welfare Medical Care ($)	Black ($)	Rural ($)	R^2	F
1. Individual family data n = 2367	.43 (3.42)	.10[b] (.71)	.36[b] (4.89)	.36 (1.69)	.37 (11.01)	.04 (.21)	-40.4 (3.30)	-19.1 (2.69)	-4.7 (.94)	.1498	29.59
2. Individual family data n = 2367	.27 (2.78)	.19[c] (1.86)	.25[c] (4.42)	.33 (1.60)	.37 (11.08)	.13 (.62)	-40.0 (3.26)	-24.7 (3.65)	-6.3 (1.26)	.148	33.96
3. Verified individual family data n = 1579	.47 (4.30)	.12[b] (.96)	.37[b] (5.59)	.44 (2.23)	.45 (16.31)	-.42 (2.13)	-40.2 (2.90)	-17.3 (2.16)	-5.4 (.96)	.318	52.09
4. Grouped data--base n = 2367	.29 (2.39)	.29[b] (2.41)	.83[b] (4.58)	.20 (1.64)	.37 (5.11)	.86 (2.93)	-15.3 (.51)	--	--	.608	25.36
5. Grouped data--base n = 1579	.30 (2.85)	.48[b] (4.34)	.89[b] (5.46)	.31 (3.19)	.34 (6.33)	.19 (.67)	-28.8 (.87)	--	--	.672	33.53
6. Tobit regressions[d] n = 2367	.38 (2.61)	.34 (2.13)	.47 (5.18)	.11 (.49)	.45 (11.57)	-.09 (.36)	-57.1 (3.93)	-35.3 (4.49)	-5.1 (.99)	--	396.36[e]
7. Families with no insurance or at least one individual plan n = 1208	.46 (2.75)	.43[b] (2.16)	.66[b] (4.84)	.49 (1.54)	.48 (13.10)	-.76 (2.18)	-24.3 (1.78)	-21.8 (2.90)	-5.4 (.92)	.263	30.42

[a] Elasticities reported unless column is headed by a "$." t-levels in parentheses.

[b] Variable entered in quadratic form. Elasticity reported is at mean values of x and y, and t-test is on joint vector $b_1 + 2b_2x$.

[c] Entered in logarithmic form.

[d] Elasticities of E(y). t-statistic is not exact.

[e] Chi-square test with 14 d.f.

The effects of age on demand for health insurance are similarly affected by the functional form. The estimated elasticities with respect to age are .37 using verified data samples in the quadratic form, and .23 in the log form, with a high degree of precision surrounding both estimates (the t-ratios are 5.84 and 4.97). A positive coefficient on age is somewhat unexpected, if the expected illness of the household has indeed been held constant. This suggests that the measure of expected illness (ESTLOS) does not completely account for expected variations in illness levels, and that the "age" variables are reflecting increases in illness levels with age rather than an independent effect of age on demand for insurance.

The total premiums estimates show the anticipated positive relationship between the premiums and the expected loss facing the family. The elasticity of premiums with respect to ESTLOS is .51 using the verified data, with a t-statistic of 2.8. This elasticity obviously reflects both changes in coverage as expected losses change and marginal cost pricing on the part of the insurer.

The elasticity of total premiums with respect to work-group size (GRPSIZ) is positive and is estimated to be .45 using verified data, with an associated t-statistic of 17.0. This estimate has very high precision (two standard errors below the estimate would be .40); since group-size increases imply reductions in the load on insurance, these estimates suggest that total premiums rise as price falls. Although expenditure rises as price falls, it is not necessarily true that insurance is a price-elastic good. With a conversion factor between group size and loading fee of -2/3, an estimate is implied of $\eta_{R\theta} = -.30$. Assuming $\bar{\theta} = .20$ for the sample, and that the average value for C is .40, then the implied elasticity of C with respect to θ is .54, and the comparable "coverage" elasticity is -.80.[1] I should point out that conversion of a value for $\eta_{R\theta}$ to a value of $\eta_{C\theta}$ is very dependent

[1]This conversion uses equation (B.16b). The "best estimates from the time series show a corresponding elasticity of -.84 (90 percent confidence interval from -.27 to -1.11) in the logarithmic logit equation; the estimate from the linear logit equation was -.63 (90 percent confidence interval from -.16 to -1.14). See Section IV.

on values assumed for θ, C, and η_{hc}. An example at the end of Appendix B shows this clearly.

The effects of medical care price levels on total insurance premiums are estimated to be negative in the individual data regressions and positive in grouped data regressions. These estimates are somewhat displeasing, not only because of the sign-reversal associated with the different methods of estimation, but also because of the low degree of precision of the estimates. The net impression from these estimates is that total price does not have a significant effect on total insurance premiums, and that we should consider the appropriate coefficient to be zero. This finding is somewhat surprising, in view of the time-series estimates (in Section IV) showing a *coverage ratio* elasticity of approximately .80 as medical care prices change. It is difficult to rationalize coverage *ratios* that increase with medical care prices, while premiums remain constant with prices, unless demand for medical care is own-price elastic. Yet most estimates of the own-price elasticity of demand for medical care show considerable inelasticity. A further analysis of this phenomenon follows below, in a discussion of effects of medical prices on demand.

Demand for Hospitalization Insurance

Two dependent variables reflect consumer demand for hospitalization insurance—the maximum amount per hospital day to be paid by the insurance policy and the maximum number of days covered. If the consumer knows the price of hospital care in advance, then selection of a HOSMAX parameter implies a coinsurance (coverage) rate on hospital expenses. The selection of the maximum number of days covered (HOSDAY) defines the other major choice variable open to the consumer. I have separately estimated the demands for these variables, with the results reported in Tables 6 and 7.

The estimated income elasticity of demand for hospitalization coverage is lower than that of total insurance premiums, although the differences cannot be substantiated at usual levels of probability. Nevertheless, the point estimates suggest that the income elasticity of demand for hospital coverage is one-half to one-third the size of

Table 6

DEPENDENT VARIABLE: HOSMAX[a]

	Data Base and Estimation Technique	Permanent Income (η)	Education (η)	Age (η)	Expected Illness Level (η)	Work-group Size (η)	Medical Care Prices ($)	Welfare Medical Care ($)	Black ($)	Rural ($)	R^2	F
1.	Individual family data n = 2367	.15 (1.34)	.28[b] (2.35)	.24[b] (3.58)	.06 (.31)	.45 (15.21)	.58 (4.19)	-2.4 (1.69)	-2.8 (3.45)	.2 (.35)	.174	35.42
2.	Individual family data n = 2367	.12 (1.34)	.25[c] (2.93)	.16[c] (3.31)	.02 (.14)	.45 (15.24)	.59 (4.34)	-2.4 (1.67)	-3.0 (3.91)	.1 (.21)	.173	41.13
3.	Verified individual family data n = 1579	.18 (2.00)	.25[b] (2.17)	19[b] (3.33)	.03 (.21)	.54 (23.36)	.37 (3.07)	2.4 (1.64)	-2.9 (3.38)	.3 (.47)	.390	71.38
4.	Grouped data--base n = 2367	.09 (.66)	.41[b] (3.31)	.72[b] (3.54)	-.02 (.14)	.40 (4.94)	.98 (5.11)	-.2 (.04)	--	--	.554	20.32
5.	Grouped data--base n = 1579	.03 (.29)	.52[b] (2.25)	.58[b] (3.20)	.07 (.66)	.40 (6.25)	.81 (4.00)	-12.5 (1.97)	--	--	.597	24.24
6.	Tobit regressions[d] n = 2367	.15 (.95)	.45[b] (3.02)	.34[b] (4.04)	-.01 (.033)	.52 (14.4)	.47 (2.76)	-4.1 (2.36)	-4.5 (4.71)	-.2 (.32)	--	434.23[e]
7.	Families with no insurance or at least one individual plan n = 1208	.23 (1.42)	.64[b] (3.37)	.51[b] (4.33)	.26 (.85)	.42 (11.85)	.27 (1.04)	-2.1 (1.38)	-3.1 (3.63)	-.03 (.06)	.230	25.42

[a] Elasticities reported unless column is headed by a "$." t-levels in parentheses.

[b] Variable entered in quadratic form. Elasticity reported is at mean values of x and y, and t-test is on joint vector $b_1 + 2b_2x$.

[c] Entered in logarithmic form.

[d] Elasticities of E(y). t-statistic is not exact.

[e] Chi-square test with 14 d.f.

Table 7

DEPENDENT VARIABLE: HOSDAY[a]

Data Base and Estimation Technique	Permanent Income (η)	Education (η)	Age (η)	Expected Illness Level (η)	Work-group Size (η)	Medical Care Prices (η)	Welfare Medical Care (Days)	Black (Days)	Rural (Days)	R^2	F
1. Individual family data n = 2367	.49 (2.80)	-.15[b] (.80)	.30[b] (2.92)	.48 (1.76)	.38 (8.62)	.21 (.98)	-.20 (1.55)	-17.7 (2.42)	.79 (.15)	.076	13.64
2. Individual family data n = 2367	.32 (2.44)	.01[c] (.08)	.24[c] (3.19)	-.41 (1.60)	.39 (8.79)	.26 (1.26)	-19.9 (1.54)	-20.5 (2.92)	0.0 (0.0)	.075	15.70
3. Verified individual family data n = 1491	.42 (2.75)	-.02 (1.22)	.21 (2.43)	-.62 (2.39)	.53 (13.73)	.03 (.15)	-18.8 (.22)	-22.0 (2.49)	-.8 (.13)	.190	24.75
4. Grouped data--base n = 2367	.11 (.61)	-.14[b] (.78)	.49[b] (1.89)	.03 (.19)	.16 (1.60)	.86 (3.47)	7.0 (.18)	--	--	.338	8.37
5. Grouped data--base n = 1579	.06 (.42)	.43[b] (2.21)	.48[b] (2.07)	.03 (.24)	.28 (3.70)	.55 (2.25)	-40.8 (1.00)	--	--	.372	9.68
6. Tobit regressions[d] n = 2367	-.06 (.34)	.60[b] (3.12)	.38[b] (3.59)	-.02 (.06)	.40 (8.53)	-.14 (.62)	-58.7 (1.90)	-69.3 (4.09)	3.2 (.29)	--	167.0[e]
7. Families with no insurance or at least one individual plan n = 1208	.38 (1.72)	.30[b] (1.16)	.46[b] (2.52)	-.19 (.47)	.39 (8.15)	-.38 (1.08)	-12.6 (.78)	-28.2 (3.21)	.97 (.14)	.128	12.18

[a] Elasticities reported unless column is headed by "days." t-levels in parentheses.

[b] Variable entered in quadratic form. Elasticity reported is at mean values of x and y, and t-test is on joint vector $b_1 + 2b_2x$.

[c] Entered in logarithmic form.

[d] Elasticities of E(y). t-statistic is not exact.

[e] Chi-square test with 14 d.f.

total premium income elasticities. The elasticity with respect to GRPSIZ is about .4 to .5, indicating that the own-price elasticity of demand is about -.25 to -.30. This estimate appears to conflict with that involving total premiums, which shows that total premiums rise as loading fees fall (that is, as group size increases). These two results are not necessarily inconsistent—total premiums include not only hospitalization insurance but surgical, major medical, and other types of insurance. Their combined effects lead to the estimated elasticity on total premiums, in contrast to the partial effects of demand for hospitalization insurance. Further, even with hospitalization insurance alone, the consumer may increase his coverage either by increasing the amount per day or the number of days covered or by reducing any deductible that might be present.

The estimated elasticity of HOSMAX and of HOSDAY with respect to *expected* illness levels is effectively zero; the coefficients are very small, and standard errors around those coefficients are quite large in comparison with the coefficients. (The same phenomenon was found to be true of SURMAX, the maximum payment to be made under surgical insurance.) This suggests that the coverage level chosen by the consumer is effectively invariant to the *mean* of expected illnesses he may encounter and that changes in premiums as expected illnesses change are due simply to expected marginal cost pricing.[1]

The theory of demand for reimbursement insurance shows that substitution effects will increase demand for coverage as expected illnesses increase, but income effects (premium increases) decrease demand. The net result is ambiguous; empirically, the two effects appear to cancel one another. There are some empirical "hints" that, if demand elasticities were not obscured by group purchase arrangements, we would observe a positive relationship between expected illness levels and coverage levels. In the regressions using the "private demand curve" subsample of families--those who purchased no insurance (by choice), or at least one non-group policy--the demands for HOSMAX and for HOSDAY show positive "expected loss" elasticity estimates, but

---

[1]No measure of the variance in the loss distribution $f(\ell)$ is available in a cross-section study.

the t-ratios are smaller than unity. (The SURMAX and MM-MAX variables both show estimated elasticities near one with respect to ESTLOS in that sample, but the precision standard errors are approximately the same size as the estimate.) Whether a better estimator of the theoretical concept of "expected loss from health stock" would show more effects of expected losses on demand for coverage is a moot point--such estimators do not exist.

Demand for Surgical-Medical Insurance

Surgical-medical insurance generally pays physicians' fees for in-hospital procedures, occasionally covering out-of-hospital care also. Some policies pay billed charges by the physician. Others pay according to complex fee schedules, predetermined by the insurer. As described previously, I have used a single variable to represent the coverage provided by any surgical insurance held by the family--the maximum payment that can be made under the plan.

As shown in Table 8, the income elasticity of SURMAX was estimated to be .30 in the grouped data estimates but near zero in the individual family estimates. However, the "education" elasticity of demand for SURMAX is quite high--.28 in the family regressions (t = 2.71) and .39 in the verified data estimates (t = 2.86). These estimates indicate that the "time cost" effects on demand for surgical insurance are quite strong. Because of collinearity problems, the partial income effects may not be well estimated, and the t-statistics may be biased downward.

Medical care prices have a strong effect on demand for surgical insurance. In the grouped data estimates, using a physician price variable, the estimated elasticity was 1.71 for the entire sample (t = 7.20), and 1.18 in the verified data grouped regressions (t = 4.70). In the individual family estimates, these elasticities were approximately .50 (t = 3.41). The grouped data estimates suggest that there is an increase in implicit "coverage levels" for physician's care. The individual data estimates suggest some decrease in coverage (since the elasticity of SURMAX with respect to medical prices is less than unity). I suggest that the grouped data estimates are more reliable here, primarily because of errors in the medical care price series that

Table 8

DEPENDENT VARIABLE: SURMAX[a]

Data Base and Estimation Technique	Permanent Income (η)	Education (η)	Age (η)	Expected Illness Level (η)	Work-group Size (η)	Medical Care Prices (η)	Welfare Medical Care ($)	Black ($)	Rural ($)	R^2	F
1. Individual family data n = 2367	-.01 (.07)	.45[b] (3.09)	.18[b] (2.31)	.03 (.13)	.51 (14.29)	.50 (3.41)	-28.9 (1.08)	-63.7 (4.16)	-7.1 (.63)	.144	28.16
2. Individual family data n = 2367	.06 (.55)	.28[c] (2.71)	.10[c] (1.71)	-.08 (.39)	.50 (14.13)	.49 (3.44)	-28.3 (1.06)	-64.5 (4.45)	-7.6 (.69)	.142	32.45
3. Verified individual family data n = 1493	.05 (.44)	.39[b] (2.86)	.18[b] (2.28)	.00 (.01)	.62 (20.41)	.41 (2.85)	-43.8 (1.42)	-59.3 (3.41)	-7.0 (.53)	.320	50.10
4. Grouped data--base n = 2367	.30 (2.05)	.29[b] (1.96)	.77[b] (3.51)	.04 (.27)	.16 (1.79)	1.71 (7.18)	42.00 (.58)	--	--	.541	19.26
5. Grouped data--base n = 1579	.08 (.65)	.41[b] (2.89)	.54[b] (2.47)	.09 (.65)	.30 (4.06)	1.18 (4.73)	-92.00 (1.08)	--	--	.520	18.07
6. Tobit regressions[d] n = 2367	.03 (.23)	.55[b] (3.31)	.27[b] (2.88)	.04 (.18)	.58 (14.43)	.36 (2.17)	61.4 (2.08)	-95.3 (5.70)	-6.8 (.62)		412.96[e]
7. Families with no insurance or at least one individual plan n = 1208	.14 (.70)	.53[b] (2.34)	.56[b] (3.47)	.29 (.81)	.50 (11.63)	-.11 (.42)	-31.4 (1.15)	-57.2 (3.87)	-11.7 (.97)	.190	19.25

[a] Elasticities reported unless column is headed by a "$." t-levels in parentheses.

[b] Variable entered in quadratic form. Elasticity reported is at mean values of x and y, and t-test is on joint vector $b_1 + 2b_2x$.

[c] Entered in logarithmic form.

[d] Elasticities of E(y). t-statistic is not exact.

[e] Chi-square test with 14 d.f.

face a given family, errors that should be averaged out by the data grouping process.

The own-price elasticity of demand for SURMAX is estimated to be less than unity, converting the GRPSIZ estimate to an own-price estimate. Again, I do not find this result contradictory to the possibly elastic demand for all insurance suggested by the total premium estimates, since some of the various components of insurance may individually be price-inelastic, but the total package of insurance may be price-elastic.

As previously noted, the SURMAX variable does not seem to respond to expected illness levels, suggesting that the movement of premiums to E(L) is primarily due to marginal cost pricing. That SURMAX has a positive and statistically significant elasticity with respect to age of the head of the family suggests that the ESTLOS variable may not completely control for expected illnesses and that age of family head may reflect some additional effects of expected losses.

Demand for Major Medical Insurance

In 1963, major medical insurance was a relatively rare phenomenon. Of the families in the sample, only 23 percent held any type of major medical insurance, and in that group the amount of that insurance varied considerably. The estimate of major medical insurance involved estimating the maximum payment specified in the major medical policy, with a "maximum" of zero coded if the family did not hold any major medical policy. These equations were not estimated with the same precision as those for other insurance variables, probably because of the nature of the data. With a large number of families making no purchase, the dependent variable was frequently "zero" (but never negative), although the theory of multiple regression specified that the dependent variable be allowed to take on unlimited values. The estimation problems with this set of data should be more serious than, say, for total premiums, where non-zero values are observed in about 75 percent of the sample.

As shown in Table 9 the best predictor of major medical purchases was simply work-group size. These estimates suggest that only persons

Table 9

DEPENDENT VARIABLE: MM-MAX[a]

Data Base and Estimation Technique	Permanent Income (η)	Education (η)	Age (η)	Expected Illness Level (η)	Work-group Size (η)	Medical Care Prices (η)	Welfare Medical Care ($)	Black ($)	Rural ($)	R^2	F
1. Individual family data n = 2367	.12 (.38)	.80[b] (2.30)	.06[b] (.34)	.85 (1.62)	.93 (11.34)	.97 (1.88)	105.0 (.16)	143.0 (.38)	127.0 (.47)	.100	18.63
2. Individual family data n = 2367	.61 (2.58)	.20[c] (.85)	.08[c] (.58)	.60 (1.17)	.91 (11.16)	.82 (1.61)	-106.0 .16	361.0 (1.0)	180.0 (.67)	.097	21.00
3. Verified individual family data n = 1579	.10 (.34)	.74[b] (2.20)	-.22[b] (.56)	.54 (.98)	1.03 (13.44)	1.05 (1.92)	702.0 (.09)	88.0 (.18)	306.0 (.90)	.174	23.51
4. Grouped data--base n = 2367	.18 (.43)	-.04[b] (.10)	-.72[b] (.28)	-.09 (.22)	1.35 (5.28)	.16 (.14)	705.0 (.31)	--	--	.348	8.75
5. Grouped data--base n = 1579	-.06 (.18)	.43[b] (1.19)	.02[b] (.02)	.04 (.11)	1.08 (5.83)	-.07 (.06)	-633.0 (.28)	--	--	.36	10.84
6. Tobit regressions[d] n = 2367	.79 (2.16)	.57[b] (1.45)	-.45[b] (1.76)	1.01 (1.72)	1.11 (11.61)	1.02 (1.78)	-1011.2 (1.80)	-91.9 (.34)	-29.3 (.17)	--	324.66[e]
7. Families with no insurance or at least one individual plan n = 1208	.18 (.32)	1.216[b] (1.88)	-.333[b] (.72)	1.04 (.98)	1.49 (12.38)	-.72 (.62)	-41.0 (.07)	-119.0 (.39)	-60.0 (.25)	.161	16.36

[a]Elasticities reported unless column is headed by a "$." t-levels in parentheses.

[b]Variable entered in quadratic form. Elasticity reported is at mean values of x and y, and t-test is on joint vector $b_1 + 2b_2x$.

[c]Entered in logarithmic form.

[d]Elasticities of E(y). t-statistic is not exact.

[e]Chi-square test with 14 d.f.

in very large work groups (where the loading fee is at a minimum) purchase major medical insurance; the amount of such insurance may be almost exogenous once the purchase decision has been made, since few work groups offer more than one level of coverage under major medical plans. In the individual family estimates, the elasticity with respect to GRPSIZ was estimated at .93 (all families) and 1.03 in the verified data sample. The t-ratios on both of these estimates exceeded 11.0, indicating the high precision of the estimates. These estimates imply an own-price elasticity of approximately -.7, using the conversion factor I have constructed. This factor may still be a low estimate of the true own-price elasticity of demand for major medical insurance, since the GRPSIZ variable will not be a completely accurate measure of the loading fee on insurance.

These individual family estimates also show an elasticity with respect to medical care prices of approximately 1.0 (t = 1.90), suggesting that persons who hold major medical policies adjust the maximum payment to approximately compensate for medical care price changes. The grouped data estimates, however, show no significant effects of prices on MM-MAX, although the grouping of the dependent variable may obscure true effects in this case, since many non-purchasers are present in each cell.

Because of the nature of the dependent variable, I expect the Tobit regressions to be most accurate when estimating demand for MM-MAX. Equation 6 of Table 9 gives the results of that estimation. The estimated income elasticity is .8 (t = 2.16), the education elasticity is .57 (t = 1.45), and the age effect is strongly negative and significant. The elasticity of work-group size is 1.11 (t = 11.61), from which I infer a true price elasticity of approximately -.8. Also of note (and exceptional in this study) is that expected illness is strongly and significantly associated with demand for MM-MAX (η = 1.01, t = 1.72), and medical care prices have roughly the same effect. The major difference between these results and the OLS estimate on the same sample (Equation 1, Table 9) is the higher estimated income elasticity using Tobit methods and the negative effect of age on demand for MM-MAX.

SUMMARY OF EFFECTS OF INDEPENDENT VARIABLES

Effects of Income and Education

I attempted to ascertain the effects of nonreporting on income elasticity estimates by estimating the identical equations with a subsample of families for whom all insurance reporting was complete. That sample includes 1579 of the original 2367 families (.667 percent of the data base). The remaining one-third of the families were excluded if they had at least one insurance policy for which verification was not received. The results confirmed my hypothesis that the estimated income elasticity would be low in the sample using unverified data; the elasticity of premiums with respect to income was .42 ($t = 4.25$) in this sample using individual families as data points, and .30 ($t = 3.16$) in the grouped data regressions. These estimates are higher than the corresponding estimates of income elasticity in the entire sample, and the precision of the estimates is much better.

The issue of which income measure to use can also be raised. Although the equations reported previously use an instrumental variable estimate of permanent income, there is some theoretical reason to believe that transitory (or measured) income might affect demand for health insurance in its own right. The argument is that health care, like any capital stock, might be purchased with transitory increases in income. To the extent that the level of health (H) can be increased through medical procedures, even when no illness is observed, transitory income might be positively associated with demand for *medical care*. That demand might be translated into a derived demand for additional health insurance, and thus transitory income might be a predictor of that demand as well. Since the argument is that transitory increases in income lead to desired increases in health, one cannot assume that the risk reduction (the implicit function of health insurance) is the goal of additional purchases of health insurance, but that the differential knowledge between insurer and insuree makes health insurance particularly attractive (the insuree knows that he plans on purchasing some extra medical care). To the extent that this differential information exists, transitory purchases of health insurance

would be a way of *extending* or increasing any transitory income gains and thus purchasing an extra amount of medical care (and therefore health). I anticipated *a priori* that any such purchases of health insurance associated with transitory income would be for specific services that can be postponed, scheduled, and "stored." Surgical procedures of many types fit that description (hernia repairs, plastic surgery, a whole variety of corrective orthopedic procedures). Regressions estimated from the verified sample (n=1579) showed that this hypothesis was verified. The results from all the estimated equations are given in Table 10.[1]

The only types of insurance that were more sensitive to observed income than to permanent income were surgical insurance and major medical insurance, as suggested by the above reasoning. Apparently, demand for those types of insurance is influenced not only by risk aversion in general but by differential knowledge between insurer and insuree. Those insurees who intend to invest windfall gains in surgical procedures appear to take advantage of that differential information through purchase of surgical or major medical insurance as an "efficient" method of paying for the (planned) investment.

Another factor influencing the income estimate is education. I included an education variable in the regressions on the grounds that, holding income constant, the wage rate (opportunity costs of time) of the family would be well correlated with the education of the head of the household. Inclusion of education in the regressions reduced the precision of estimates of the income elasticity of demand for insurance, as shown in Table 11. The education of the head of household, of course, is not only highly correlated with wate rates (or dependence of total income on earning capacity of the head) but with total income. I give the regressions results with education included, with the

---

[1]The equations were also estimated with pure "transitory" income--the residual difference between predicted and observed income. Since that transitory component has a mean very near zero, the computed elasticities at mean values of transitory income have no interpretable meaning. However, the slope coefficients in the regression have been compared with those obtained in the "observed income" regressions and were found to be similar in each case.

Table 10

INCOME ELASTICITIES WITH PERMANENT AND OBSERVED INCOME

Dependent Variable	Permanent Income Elasticity	Observed Income Elasticity
Total premiums (MAXPRM)	.43 (4.27)	.18 (6.64)
Maximum hospital payment per day	.15 (2.21)	.01 (.80)
Maximum number of hospital days	.35 (2.17)	.02 (.40)
Maximum surgical payment	.05 (.44)	.16 (4.86)
Maximum major medical insurance payment	.08 (.28)	.18 (2.02)

caveat that these coefficients may be biased downward, and the t-statistics are biased downward in most cases.[1]

If one is attempting to estimate the income elasticity, then the regressions with education omitted are preferable, at least to estimate the correct precision of the coefficient. However, for predictive purposes, including the education variables is preferable. These results also verify the general suggestion that education variables should be interpreted as wage-rate effects rather than as indicators of production efficiency for non-market alternatives to insurance.

The Effects of Medical Care Prices

The effects of medical care prices on total insurance expenditures are estimated with somewhat less precision than other variables, possibly because of the systematic measurement error in the medical price variable, which can cause a bias of indeterminant direction and size

[1] Richard V. L. Cooper and Joseph P. Newhouse, "Further Results on the Errors in the Variables Problem," P-4715, The Rand Corporation, Santa Monica, October 1971.

Table 11

COMPARISON OF INCOME ELASTICITIES WITH EDUCATION EXCLUDED AND INCLUDED

	n = 2367		n = 1579	
Dependent Variable	Income Elasticity No Education Variables	Income Elasticity with Education Variables	Estimated Income Elasticity Education Omitted	Estimated Income Elasticity Education Included
MAXPRM	.28 (t = 4.26)	.20 (t = 2.19)	.38 (t = 4.92)	.30 (t = 3.16)
VERPRM	.33 (t = 3.69)	.29 (t = 2.39)	.35 (t = 4.03)	.30 (t = 2.85)
HOSMAX	.22 (t = 2.17)	.09 (t = .66)	.123(t = 1.23)	.04 (t = .27)
HOSDAY	.29 (t = 2.25)	.11 (t = .61)	.18 (t = 1.54)	.06 (t = .42)
SURMAX	.34 (t = 3.21)	.30 (t = 2.05)	.20 (t = 1.85)	.08 (t = .65)
MM-MAX	.41 (t = 1.36)	.18 (t = .43)	.465(t = 1.72)	-.06 (t = .18)

in the coefficient estimate.[1] In the individual family regressions, the estimated elasticity of premiums with respect to medical prices was -.15 (t = .96) in the entire sample and -.39 (t = 2.17) in the verified sample; and in the grouped data regressions, a positive (but not significant) effect was estimated in both the verified and unverified samples. As previously discussed, even with a positive "true" coefficient it is possible to have an estimated negative coefficient, because of systematic measurement error. The time series study (Section IV) and the grouped data results suggest that this may have occurred; both of those estimates show that increases in medical prices increase demand for coverage.

Further information about the total effects of price on medical insurance can be found by analyzing components of insurance. My tentative conclusion is that total expenditure on health insurance is invariant with the price of medical care. This result is somewhat difficult to interpret, since various insurance policy parameters increase with medical care prices. Assuming that demand for medical care is own-price inelastic (so that total expenditures increase as medical prices increase), one consistent interpretation of premium's being invariant to medical prices is that net coverage decreases as prices increase. If maximum payments for various medical procedures increase with medical care prices, but with an elasticity of less than one, the implicit coverage has decreased; and premiums could remain constant or decline under those circumstances, even with medical prices increasing.

We also see that the medical-care price elasticity of the maximum payment per hospital day (HOSMAX) is estimated to be considerably less than unity in the individual data, which implies that coverage ratios of hospital care are declining as medical care prices increase. This result would be consistent with constant (or slightly declining) health insurance premiums. Similar results are estimated with respect to surgical maximum payments, suggesting that surgical coverage ratios also decline in the face of medical care price increases.

---

[1] See the discussion on p. 56, especially (3.1).

One anomaly produced by these estimates (and the only major difference encountered in estimation with grouped data, as compared with individual data) is that in the grouped data regressions a positive "medical care price elasticity" of premiums was estimated, and elasticity of "SURMAX" was estimated to be well above unity (1.71 for the entire sample), with the elasticity of HOSMAX being estimated near unity. These estimates imply that coverage ratios increase (or at least are constant) with medical care price increases and that premiums rise accordingly. The estimates *within* grouped data are consistent, but the overall result regarding demand patterns is inconsistent between grouped data estimation and individual family estimation. The disparity between these results is disturbing. The results may merely reflect random "noise" in the coefficient estimates, and the "true" coefficient may lie somewhere near zero for premiums or unity for hospital and surgical maximum payments. This explanation is somewhat unsatisfactory, especially in light of the high degree of precision of the 1.71 elasticity estimate for surgical payments. Alternatively, the "medical care price" variable may contain systematic measurement errors of substantial magnitude that are at least partly eliminated through grouping of the original data.

If the EV (with negative covariance) explanation is correct, then the grouped data estimates may better reflect the effects of medical care prices on demand for health insurance. Those estimates suggest that medical care prices indeed increase coverage levels and that premiums increase accordingly.

Effects of Loading Fees

I have used the size of the work group of the household head as a proxy for the loading fee on insurance; the pattern of reduction in the load as work-group size increases is shown in the initial description of that variable. The estimated coefficients of work-group size on total premiums are estimated with extremely high precision, and the results are consistent across data subsamples and estimating techniques. The estimate of the "group size" elasticity of premiums is about .35 in the total sample, which translates roughly into an estimator of -.2 for the change in premiums with respect to the loading fee.

As with the estimated income elasticities, there is *a priori* reason to believe that the total sample of 2367 families will provide a downward biased estimate of the effects of work-group size on premiums, since it is the largest groups that systematically did not report on the insurance verification forms. The subsample of families with complete verification should not only be estimated with better precision overall, but the coefficient for GRPSIZ should be biased downward. This indeed occurred in the estimations. For the total sample, estimated in individual families, the elasticity with respect to group size was .35 (t = 12.18). In the verified subsample (with no false values for the dependent variable), the elasticity estimate was .45 (t = 17.91). These estimates converge (at approximately two standard deviations from mean values) at approximately .40, although there is no reason to believe *a priori* that they should be identical; indeed, I would expect a difference to occur in the direction observed. The magnitude of difference of the implied own-price elasticity is not as large as the estimated differences in these coefficients; the *implied* difference is between elasticities of -1.2 and -1.3 between the two samples.

Expected Illness Levels

Insurance premiums are almost certain to vary directly with expected illnesses, either through changes in coverage or through expected marginal cost pricing, or both. I would have been extremely surprised to find an estimated elasticity of premiums with respect to $E(\ell)$ that was negative. The estimates from individual family data strongly suggest a positive effect; in the entire sample, the estimated coefficient is .55 (t = 3.52) and in the "verified" sample, the estimate was .51 (t = 2.8).

In the grouped data regressions, the estimates are lower and less precise. I believe that this difference is due to the specification of the independent variables. As in the case with "income," an instrumental variable estimate has been used in the individual family regressions, and simple group means have been used in the grouped data regressions. If illness is highly variable, even within these groups there may be a substantial transitory element in the value for

"estimated illness," which would bias the estimated coefficient toward zero[1] and bias the computed t-statistic downward.[2] Both of these phenomena appear to have occurred in these estimates. Further, systematic patterns of illness expectations as family age and size change have not been incorporated into the grouped data variable for expected illness. I expect age variables to enter positively and significantly in these regressions, with the age variables simply reflecting illness patterns. If that is indeed true, those age effects should vanish in the individual regressions, since age and family composition have entered the instrumental variable estimate of expected illness used in individual family regressions. This has occurred in estimates of insurance premiums, as can be verified by comparing values in the tables.

The Effect of Race

The estimates show a systematic reduction in the amount of insurance purchased by nonwhite families, holding constant the income of the families and the education of the family head. It could be argued that this difference merely reflects income differentials, but the "race" variable was one of the predictor variables in the permanent income instrumental equations. I suggest here that the tendency for blacks to purchase less medical care than whites is related systematically to the *effective* loading fee on health insurance. This phenomenon may be because of more difficult access to medical care, fewer physicians in areas where blacks live, lower incomes, different education or perceptions of the medical system, or different tastes. For whatever reasons, it has the effect of increasing the loading fee to blacks on either community rated or group insurance, since blacks pay the same premium but have lower expected expenses.

---

[1]Goldberger, *Econometric Theory*, pp. 282-284. To the extent that the expected illness variable is correlated with other included variables, the other coefficient estimates in the equation will be biased as well.

[2]Cooper and Newhouse, "Further Results on the Errors in the Variables Problem."

Some simple comparisons show this difference. In 1963-1964, whites had 42 percent more physician visits, on average, than blacks (4.7 visits for whites, 3.3 visits for blacks); in 1968-1970, the difference was still 25 percent (4.4 visits for whites, 3.5 visits for blacks).[1] To establish the amount of differential utilization appropriate for this problem, one would want utilization for whites and blacks under the same insurance plan, but holding nothing else constant. The best examples of such data that I am aware of are Medicare expense data for persons over 65. As shown in Andersen *et al.*,[2] expenditures for white people under Medicare are greater than for black people. For hospital care, the ratio is 1.54; for physician services, the ratio is 1.33.[3] If an average black person and an average white person each paid the same amount for Medicare coverage, the effective loading fee would be much higher for the black person. Suppose that the loading fee for the white person were .1, so (if expected benefits were $154 for hospital care) the premium would be $170. For the same $170 policy, the black person would have expected benefits of $100. The loading fee for a black person is .7; for a white person with the same policy the loading fee is .1. If the same loading fee were applied to a physician services policy (with a differential utilization factor of 1.33), the premium for whites would be $146 (for expected benefits of $133), and the loading fee for blacks would be .46. Obviously, the larger the relative demand of the whites in an insurance plan, the larger the effective loading fee facing blacks.

---

[1] These data are from National Center for Health Statistics, Series 10, Nos. 18 and 70.

[2] Ronald Andersen *et al.*, *National Trends and Variations in Expenditures for Personal Health Services*, Center for Health Administration Studies, University of Chicago, January 1973.

[3] The ratios are similar in other studies. National Center for Health Statistics data from 1968-1969 give a ratio of 1.43 for hospital care and 1.33 for physician visits. See *Current Medicare Survey Report, Hospital Insurance Sample, Inpatient Hospital Utilization 1966-69*, CMS-2S, Office of Research and Statistics, Social Security Administration, March 1973, for the hospital data; and *Age Patterns in Medical Care, Illness and Disability, United States, 1968-69*, Series 10, No. 70, National Center for Health Statistics, U.S. Department of Health, Education and Welfare, March 1972.

Blacks normally are unable to obtain insurance policies that reflect this fact, since group insurance and community rated insurance pool people no matter what their expected expenses. Because of this, as I have shown, blacks face systematically higher *effective* loading fees on health insurance, which will reduce their demand for insurance. If my measure of permanent income in fact controls for the lower income of blacks, then this is a reasonable explanation of the lower demand by blacks for health insurance.[1] A similar argument can be made for rural families, female heads of household, or any other sub-group of consumers with expected expenses lower than average. Note that it may still be optimal for these persons (blacks, persons living in rural areas, and so forth) to join group insurance plans, if the effective loading fee is lower than for a comparable policy purchased through non-group sources. The substantially higher loading fees for non-group policies allow this unusual feature of community rated insurance to continue to exist in a competitive market.

Comparison of Different Subsample Estimates

Two subsamples were used in these studies to isolate particular problems. First, I used only those observations where all insurance data were verified (to establish biases in the coefficients and the R^2 measure due to incomplete reporting by the insurance carriers). Those regressions are reported in equations (3) and (5) of Tables 4-9. As hypothesized previously, the precision of fit in these regressions is considerably higher than when the entire sample was used. The R^2 measures are roughly doubled in the verified data equations, compared with the total-sample individual family regressions. Moreover, the income and work-group size coefficients are generally larger in the verified subsample, which is consistent with the systematic nonreporting discussed earlier.

[1]This is one example of how "community rating," which is nominally designed to "bring poor people into the insurance market" through risk pooling, actually performs an opposite function; that is, I am suggesting that community rating of health insurance "robs from the poor and gives to the rich." For a further discussion of this phenomenon, see Section V.

These equations may be used for predictive purposes by applying means of the explanatory variables from the total sample, as long as the coefficients remain stable over all persons, and as long as there are no omitted interaction terms in our specification. Tables 12-14 show the demand elasticities found by using the verified subsample coefficient estimates, with national average means used for all explanatory variables (the means from the entire sample of 2367). These equations are viewed as the "best" predictors of demand, since they involve the highest degree of precision of all estimates and apply appropriate levels of explanatory variables (income, work-group size, medical care prices, age, education, and so on). The results from OLS are shown in Tables 12 and 13 (individual and grouped data), and the expected value elasticities from Tobit regression are shown in Table 14.

The R^2 measures are increased significantly by omitting the unverified families, as one might expect; the unverified policies represent measurement errors in the dependent variable. In the MAXPRM regressions, the R^2 increased from .20 to .36; in estimations of individual components of the insurance, the increase was even more dramatic. The HOSMAX regression R^2 improved from .174 to .390; the HOSDAY regressions showed an increase of .076 to .191; the SURMAX regression R^2 increased from .144 to .320, and the MM-MAX R^2 increased from .100 to .174. These increases show the amount of variation in the observed dependent variables that is due to measurement error, caused by incomplete verification. The results suggest that, if complete verification were the rule, the percent of variation in insurance demand explained by the regressions would be well above .35 for total premiums and for HOSMAX, the two "best" regressions in these data. This is a higher percentage of explained veriation than is typically found in cross-section demand regressions using survey data.[1] This is to be anticipated, since

[1]For example, Anderson and Benham report R^2 measures of .098 and .124 for medical and dental expenditures, using the same data source as this study uses. See Ronald Andersen and Lee Benham, "Factors Affecting the Relationship between Family Income and Medical Care Consumption," in Herbert Klarman, ed., *Empirical Studies in Health Economics,* The Johns Hopkins Press, Baltimore, 1970, p. 82.

Table 12

SUMMARY OF "BEST" ESTIMATES OF DEMAND USING COEFFICIENTS FROM VERIFIED DATA ESTIMATES AND AVERAGE VALUES OF ALL VARIABLES FROM TOTAL SAMPLE ORDINARY LEAST SQUARES--INDIVIDUAL FAMILY REGRESSIONS

Dependent Variables	Explanatory Variables										
	Income (η)	Education[a] (η)	Age[a] (η)	Expected Illness (η)	Group Size (η)	Medical Prices (η)	Welfare Care ($)	Black ($)	Rural ($)	R^2	F
Total Premiums	.45 (4.24)	.15 (1.61)	.35 (5.81)	.52 (2.85)	.52 (17.91)	−.39 (2.17)	-38.3 (2.82)	-19.7 (2.52)	-5.7 (1.01)	.359	62.53
Maximum Payments Per Hospital Day	.19 (2.00)	.24 (2.17)	.18 (3.33)	.03 (.21)	.63 (23.36)	.37 (3.07)	-2.4 (1.64)	-2.9 (3.38)	.3 (.47)	.390	71.38
Maximum Number of Hospital Days	.42 (2.75)	−.02 (1.22)	.21 (2.43)	−.62 (2.59)	.53 (13.73)	.03 (.15)	-18.8 (days) (1.21)	-22.0 (days) (2.49)	−.8 (days) (.13)	.190	24.75
Maximum Surgical Payments	.07 (.43)	.47 (2.77)	.21 (2.30)	.00 (.01)	.89 (20.41)	.50 (2.85)	-44.0 (1.42)	-59.3 (3.41)	-7.0 (.54)	.322	50.10
Maximum Major Medical Payment	.13 (.34)	.98 (2.35)	−.23 (1.02)	.66 (.98)	1.45 (13.44)	1.26 (1.92)	79.4 (.09)	88.1 (.18)	305.9 (.90)	.173	23.51

[a]Entered in quadratic form t-test on $\frac{\partial y}{\partial x} = a_o + a_1\bar{x}_1$.

Table 13

SUMMARY OF "BEST" ESTIMATES OF DEMAND USING COEFFICIENTS FROM VERIFIED DATA ESTIMATES AND AVERAGE VALUES OF ALL VARIABLES FROM TOTAL NATIONAL SAMPLE GROUPED DATA REGRESSION RESULTS

Dependent Variables	Explanatory Variables								
	Income (η)	Education[a] (η)	Age[a] (η)	Expected Illness (η)	Group Size (η)	Medical Prices (η)	Welfare Care ($)	R^2	F
Total Premiums	.33 (3.16)	.45 (4.72)	.85 (5.78)	.34 (3.80)	.42 (6.73)	.12 (.47)	-11.5 (.36)	.693	40.14
Maximum Payments Per Hospital Day	.05 (.29)	.62 (2.25)	.76 (3.20)	.10 (.66)	.64 (6.25)	1.04 (4.00)	-12.5 (1.97)	.57	24.24
Maximum Number of Hospital Days	.09 (.42)	.40 (2.21)	.73 (2.07)	.04 (.24)	.42 (3.66)	.67 (2.25)	-40.0 (days) (1.01)	.37	9.68
Maximum Surgical Payments	.09 (.65)	.39 (2.89)	.55 (2.47)	.10 (.78)	.36 (4.06)	1.15 (4.73)	-92.1 (1.08)	.496	18.07
Maximum Major Medical Payment	-.07 (.18)	.58 (1.19)	.02 (.02)	-.04 (.11)	1.61 (12.38)	-.07 (.06)	-633.0 (.28)	.362	10.84

[a] Entered in quadratic form t-test on $\frac{\partial y}{\partial x} = a_o + a_1\bar{x}_1$.

Table 14

SUMMARY OF "BEST" ESTIMATES OF DEMAND USING COEFFICIENTS FROM VERIFIED DATA ESTIMATES AND AVERAGE VALUES OF ALL VARIABLES FROM TOTAL NATIONAL SAMPLE LIMITED DEPENDENT VARIABLE (TOBIT) REGRESSIONS [a]

Dependent Variables	Explanatory Variables										
	Income (η)	Education [b] (η)	Age [b] (η)	Expected Illness (η)	Group Size (η)	Medical Prices (η)	Welfare Care ($)	Black ($)	Rural ($)	P(y)[c]	$X^2_{14}/14$
Total Premiums	.48 (3.76)	.31 (2.59)	.43 (5.65)	.39 (1.85)	.62 (17.52)	-.50 (2.32)	-70.5 (3.87)	-44.0 (4.47)	-7.7 (1.19)	.75	53.49
Maximum Payments Per Hospital Day	.23 (1.91)	.37 (3.02)	.27 (3.48)	.06 (.20)	.75 (20.37)	.23 (1.45)	-4.4 (2.10)	-5.5 (4.75)	-.2 (.24)	.76	52.56
Maximum Number of Hospital Days	.43 (2.70)	.06 (.24)	.32 (2.20)	-.49 (1.89)	.69 (16.07)	-.08 (.42)	-42.6 (days) (2.27)	-41.3 (days) (4.02)	.7 (days) (.11)	.65	35.01
Maximum Surgical Payments	.12 (.77)	.50 (3.03)	.25 (2.66)	.01 (.05)	.74 (19.26)	.35 (1.94)	-86.3 (2.16)	-107.7 (4.87)	-8.6 (.57)	.67	45.79
Maximum Major Medical Payment	.79 (1.96)	.72 (1.59)	-.57 (2.03)	1.01 (1.48)	1.34 (12.79)	.60 (.91)	-1460.0 (1.71)	-176.0 (.44)	-9.1 (.04)	.14	25.85

[a] t-tests are not exact.

[b] Entered in quadratic form: t-test on $\frac{\partial y}{\partial x} = a_o + a_1\bar{x}_i$.

[c] Predicted probability that $y > 0$ given the average values of all x_i.

insurance abstracts from much of the randomness shown in the demand curves for (say) physician services *per se*.

An alternative sample was taken of families who either purchased *no* insurance or purchased at least one non-group insurance policy during 1963. These families were thought to be exhibiting their personal demands more effectively than families buying insurance solely through group insurance plans. The hypothesis being tested was that *all* elasticities would be further from zero using this sample than using the total sample, as long as work-group insurance did not allow sufficient variation in types of insurance to satisfy individual demands.[1] I also anticipated that the R^2 in these equations would be higher than for the total sample, if the behavioral equations truly describe systematic demand for insurance. That hypothesis was sustained in general, as shown by comparison of equations (1) with equations (7) of Tables 4-9. The population in this subsample may not be characteristic of the entire population; there may be self-selection involved in persons who either bought *no* insurance or who bought *some* non-group insurance, which makes this sample atypical. The results should be viewed with that in mind. For every dependent variable, the income, age, education, expected illness, and work-group size elasticities were higher in the "private demand curve" sample than in the total sample; there were minor aberrations with the medical care price variables, and the dummy variables for race, free medical care, and rural care did not generally follow this pattern.

POSSIBLE PROBLEMS OF THE ANALYSIS

Several possible problems arising from this cross-section analysis should be kept in mind. First, as is true with any regression, the results are conditional on the functional form chosen. There may be interaction effects between some of the variables (such as income and work-group size) that my formulation will not show. Obviously, a large

---

[1]As an extreme case, if all work-group insurance in the country were identical, then the estimated elasticities would go to zero as the proportion of persons buying insurance through groups rose.

variety of alternative formulations could be used in these regressions; budget and time constraints limit the analysis.

Second, the results of the different types of regressions differ somewhat, so the factors affecting demand for insurance cannot be said to be stable across either estimating techniques or functional forms. In particular, the effects of medical care prices are poorly estimated in this study; refinement of the "price" variables would be desirable, but that sort of information is typically missing in cross-sectional regressions, and it was necessary to use outside data sources even for the prices I included in these regressions. The effects of medical care prices on demand for insurance are not well established and would be a fruitful subject for further investigation.

A major problem with any cross-sectional study of this sort is that errors in the variables are to be expected. All of my coefficients may contain some error--the permanent income measure may not completely abstract from random components; a "year" of education may not be identical for all persons in the sample; work-group size may not be a perfect estimator of the loading fee; the medical care price index almost certainly contains errors; and the expected illness measure probably contains a considerable error component as well. The net effect of all of these on the coefficients is simply unknown. Unfortunately, there are no other data sources available with which to make comparisons, and no published studies examine actual insurance purchases of individuals in a manner similar to this study.

The errors in variables problem aso alters the standard errors of the coefficient estimates. If only one variable contains measurement errors, it has been shown that the t-statistic on the corresponding coefficient is biased toward zero; the bias of the t-statistics for all other coefficients cannot be established.[1]

On net, the benefits gained from examining micro-data are (a) more intricate effects can be analyzed than with macro-data, (b) there is a general information loss with aggregation that is avoided by using

[1] Cooper and Newhouse, "Further Results in the Errors in Variables Problem," 1971.

the micro-data, and (c) possible instability involved in use of aggregated time-series data is avoided. The tradeoffs between micro- and macro-data face every researcher. My view is that the micro-data represents a more helpful data source to analyze demand for health insurance than the time-series data, especially since "health insurance" changed so markedly over the period involved in the time series study.

VALIDITY OF THE UNDERLYING MODEL

How do these regressions strengthen (or weaken) belief in the underlying model proposed in Section II? I set forth a model of demand for reimbursement insurance that did not in fact establish a great variety of refutable predictions. Nevertheless, the model proved helpful in formulating these estimates, and the "success" of the estimations in accounting for variation in demand for insurance may be said to strengthen belief in the underlying model. Several propositions suggested by the model were borne out by the data: The income elasticities of demand for various types of insurance appear to differ significantly from one another. The own-price elasticity is negative and quite large in absolute value; moreover, the own-price of insurance seems to dominate the decision to purchase most types of insurance. The effects of time prices (concerning demand for medical care) indeed influence the demand for reimbursement health insurance, suggesting that higher time prices increase demand for insurance. The effects of prices of medical care on demand for insurance are still unresolved from this study, although the time series and the aggregated micro-data suggest a strong positive association between medical care prices and demand for insurance coverage. Finally, the effects of expected illness levels on demand for insurance appear to be negligible; although premiums increase with expected illness levels, coverage parameters are invariant with those changes. I could not establish how increased variance in illness levels altered demand for insurance.

One aspect of the model worth noting is the caution the model introduces in interpretation of empirical results. Previous models of demand for insurance suggest that behavior of the risk aversion measure might be inferred from the income elasticity of demand for

insurance. The results show that no unambiguous interpretation can be made of the income elasticity, and no inferences can be made concerning the risk aversion measure.[1] Similarly, a common belief is that increasing medical care prices will increase demand for health insurance. The theory presented shows how that might or might not occur; although the empirical results are not uniformly in agreement, there is some suggestion that (at least in the range of data considered in these studies) increasing medical prices do indeed increase demand for health insurance. Similar complexities exist in the interpretation of effects of illness anticipations, although it can be shown that under current insurance mechanisms observed in this country, those at higher risk have strong inducements to purchase more insurance (see Section V).

Further support for the model can be derived from the predicted effects of insurance on demand for medical care *per se*. Although this study does not consider those effects in full detail, other evidence and studies show that the basic demand model hypothesized is supported in many respects.[2]

---

[1]Ehrlich and Becker reach the same conclusion, for different reasons.

[2]Phelps and Newhouse, *The Effects of Coinsurance on Demand for Physician Services;* Phelps and Newhouse, *Coinsurance and the Demand for Medical Services.*

IV. TIME SERIES ESTIMATES

Alternative estimates of demand for insurance are based on annual time series data extending from 1929 to 1968. This estimate of demand involves aggregation over all persons in a given time period, with the average values of the entire population taken as data points. Such aggregation involves several assumptions that should be made explicit. First, it assumes that the micro-units being aggregated all have identical demand curves (which is equivalent to stating that every person has the same utility function). Given this assumption, there is no instability in the regression estimates over time; otherwise, the regressions would have no predictive value. Zellner shows that when micro-unit coefficients differ randomly about some mean value, there is no aggregation bias.[1] This result holds even if there are errors in the variables, so long as the error vectors for the coefficients and those of the variables themselves are uncorrelated. It is also shown that the normal two-stage least squares technique for estimating simultaneous equations provides a consistent estimate of the mean of the coefficient vectors of the micro-units, and that the usual measure of the standard errors will be biased in a direction that overstates the precision of the coefficient estimates.

Use of time series data implies that the "quality" of the dependent variable has not changed in any manner during the period of estimation. I implicitly assume that the nature of health insurance has remained constant over the period, an assumption that may well not be borne out by fact. I have chosen dependent variables for the time series estimation that should be relatively constant over time--the level of health insurance premiums and the average *coverage ratio* that results from the community's health insurance benefits and its expenditures on medical care.

---

[1]Arnold Zellner, "On the Aggregation Problem: A New Approach to a Troublesome Problem," in K. A. Fox, ed., *Economic Models, Estimation and Risk Programming: Essays in Honor of Gerhard Tintner*, Springer-Verlag, Inc., New York, 1969, pp. 365-374.

DATA SOURCES

The primary series on health insurance premiums and benefits was taken from annual issues of Argus Charts of Insurance[1] and from Best's Reports.[2] These reports summarize health insurance premiums and benefits over all reporting companies, including profit-maximizing insurance companies and nonprofit health insurers. The premium series (PREM) and benefit series (BEN) represent aggregations of all hospital, surgical, major medical, and accident insurance policies. During the early years of this study, almost all "health insurance" was issued in accident and injury policies, which paid specific amounts upon the occurrence of specific events. Those policies represent true "indemnity" insurance. Beginning about 1935, there was considerable growth of hospitalization insurance plans, covering expenses incurred by the insured while in approved hospitals. Concurrent growth of surgical insurance (paying doctors for in-hospital surgery) was observed. The most recent phenomenon in the health insurance field has been growth of "major medical" insurance policies, which offer much higher maximum payments than other types of health insurance, coupled with large deductibles and non-zero coinsurance rates. These different types of insurance should be evaluated separately. Data limitations necessitate aggregation over all types of insurance and deal only with the resultant benefits and premiums for all insurance. The remaining data have been obtained from annual issues of *Statistical Abstracts of the United States*,[3] and from *Historical Statistics of the United States from Colonial Times to 1957*,[4] as indicated in data definitions given below.

---

[1] *Argus Chart of Health Insurance*, National Underwriter Co., Cincinnati, annual for 1954 to 1968 data. *Argus Chart of Casualty Insurance*, National Underwriter Co., Cincinnati, annual for 1933 to 1953 data.

[2] *Best's Casualty Reports*, Best's Inc., New York, annual for 1929 to 1933 data.

[3] U.S. Bureau of the Census, *Statistical Abstract of the United States*, Washington, annual.

[4] U.S. Bureau of the Census, *Historical Statistics of the United States from Colonial Times to 1957*, Washington, 1960.

DEPENDENT VARIABLES

The theory of demand for insurance developed in Section II does not give an explicit functional form for estimating demand curves for health insurance. I have chosen to estimate two basic forms of the dependent variable--an average coverage ratio and total insurance premiums.

Coverage Levels

The basic coverage level variable used in the regressions is the variable COVR, defined as insurance benefits divided by medical expenditures. This variable suffers from the deficiency that it aggregates many types of medical expenditures, some of which the population may not have intended or wished to insure (drug expenses, for example) and some that may be highly insured (such as hospitalization). The average coverage level also includes expenses of many people who have no insurance as well as of people with positive amounts of insurance. Data limitations prevent further disaggregation with time series figures, necessitating this pooling of data.

A transformation of the COVR variable has also been used. Standard regression theory specifies that the dependent variable not be restricted to a limited range of values. The COVR variable will take on values between zero and one for any person, and the average values will fall between zero and about one-half. To correct for this restriction the logit transformation has been used. A new variable is defined by specifying a functional form for COVR as:

$$\text{COVR} = \frac{1}{1 + e^{-f(x)}} \ . \tag{4.1}$$

The variable LOGIT is then defined as

$$\text{LOGIT} = \ln\left(\frac{\text{COVR}}{1 - \text{COVR}}\right) = f(x) = b_0 + b_1x_1 + \ldots + b_kx_k \ . \tag{4.2}$$

The logit transformation suggests a relationship between COVR and the explanatory variables ($f(x)$) as indicated in Fig. 9. A linear estimate of LOGIT implies a curvilinear relationship between the explanatory variables and COVR. In the actual time series, the values of COVR ranged from .03 to near .5, suggesting that the lower half of the curve in Fig. 9 describes the relationship between $f(x)$ and COVR. That is, a positive change in $f(x)$ will lead to an increasingly large increase in COVR as the value of $f(x)$ increases.

Insurance Premiums

Time series data on total insurance premiums may provide some information about the demand for insurance itself. In this simple time series framework, assume that the premium for a given person is written as

$$R = (1 + \theta)k(p_h \bar{h}) \ , \tag{4.3}$$

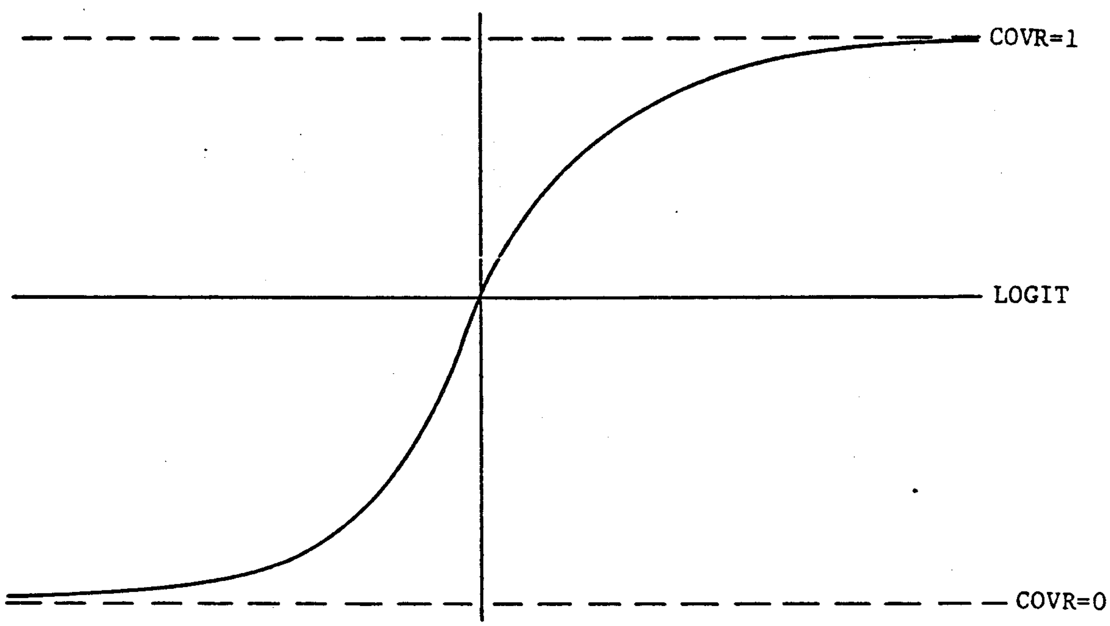

Fig. 9 — The logit transformation

where $k = (1 - C) = $ COVR, the average coverage ratio, and θ, p_h, and $\bar{h}$ retain their usual definitions. Then effects of changes in I, p_h, and θ are given in elasticity form as

$$(4.4) \qquad \eta_{RI} = \frac{\partial R}{\partial I} \cdot \frac{I}{R} = \frac{I}{R}(1 + \theta) \cdot \left(kp_h \frac{\partial \bar{h}}{\partial I} + p_h\bar{h}\frac{dk}{dI} + kp_h \frac{d\bar{h}}{dk} \cdot \frac{dk}{dI}\right)$$

$$= \frac{(1 + \theta)\left[kp_h\bar{h}\eta_{\bar{h}I} + kp_h\bar{h}\eta_{kI} + kp_h\bar{h}\eta_{\bar{h}k}\eta_{kI}\right]}{(1 + \theta)kp_h\bar{h}}$$

$$= \eta_{hI} + \eta_{kI}(1 + \eta_{\bar{h}k}) ,$$

where $\eta_{\bar{h}I}$ is the income elasticity of $\bar{h}$, $\eta_{\bar{h}k}$ is the "coverage elasticity" of $\bar{h}$, and η_{kI} is the elasticity of k with respect to I. Similarly,

$$(4.5) \qquad \eta_{Rp_h} = \frac{\partial R}{\partial p_h} \cdot \frac{p_h}{R} = p_h \frac{(1 + \theta)\left(k\bar{h}(1 + \eta_{\bar{h}\bar{h}}) + \overline{kh}(\eta_{kp_h}(1 + \eta_{\bar{h}k})\right)}{(1 + \theta)k(p_h\bar{h})}$$

$$= (1 + \eta_{\bar{h}\bar{h}}) + \eta_{kp_h}(1 + \eta_{\bar{h}k}) ,$$

and

$$(4.6) \qquad \eta_{R\theta} = \frac{\partial R}{\partial \theta} \cdot \frac{\theta}{R} = \frac{\theta kp_h\bar{h} + (1 + \theta)kp_h\bar{h}\eta_{k\theta}(1 + \eta_{\bar{h}k})}{(1 + \theta)kp_h\bar{h}}$$

$$= \frac{\theta}{1 + \theta} + \eta_{k\theta}(1 + \eta_{\bar{h}k}) .$$

From these expressions, we may solve for η_{iK}, η_{kp_h}, and $\eta_{k,\theta}$ as:

(4.7a) $$\eta_{kI} = \frac{\eta_{RI} - \eta_{\overline{h}I}}{1 + \eta_{\overline{h}k}} ,$$

(4.7b) $$\eta_{kp_h} = \frac{\eta_{Rp_h} - (1 + \eta_{\overline{hh}})}{1 + \eta_{\overline{h}k}} ,$$

(4.7c) $$\eta_{k\theta} = \frac{\eta_{R\theta} - \frac{\theta}{1 + \theta}}{1 + \eta_{\overline{h}k}} .$$

Obviously, it is necessary to know both the response of $\overline{h}$ to income and p_h and the response of $\overline{h}$ to changes in k if one is to infer elasticities of demand for k with respect to I, p_h, and θ. I make such inferences conditional on specific values of those demand elasticities for $\overline{h}$, while noting that the inferences are subject to fluctuation, depending on what values are assigned to those elasticities. In the time series estimates, I use real health insurance premiums per capita as the measure of R (adjusted for the price index for all goods and services), denoted as RPRMPC.

EXPLANATORY VARIABLES

Income

The measure of income used in this study is *real per capita national income*. Real income followed a well-known pattern during the period of this study, beginning with a one-third decline during the Depression, followed by rather steady growth from 1939 to the present.[1]

The Price of Medical Care

My price variable (PRATIO) describes the behavior of the price index of medical care relative to movement of the general price index.

[1] *Statistical Abstracts*, relevant years.

That ratio moved from .8 in 1929 to over 1.1 in 1933, then declined erratically until after World War II, from which time it has shown steady upward movement. It is of some interest that the real price of medical care was as high in 1933 as it was in 1963.[1] The price level of medical care is taken as exogenous in this study, an assumption that may be incorrect. If, in fact, the amount of reimbursement insurance in the community affects total demand for medical care, and if the supply of medical care is relatively inelastic, then increases in insurance coverage may indeed cause increases in the equilibrium price of medical care. I have not attempted to incorporate such a system into these estimates, primarily because no information was available on many variables that could be considered crucial to estimating a supply function in the medical care market. Without such variables as the real wage rate paid to physicians, it seemed inadvisable to attempt to build a model that allowed the price level of medical services to be endogenous.

The Price of Insurance

The price of insurance is defined as the gross loading fee on premiums. That is, the variable THETA is defined as (PREM/BEN) - 1. That variable is computed from the data on insurance for each year in the study. There is considerable possibility that this variable, as defined, is measured with random error, which would bias the estimated price coefficient toward zero. There is an added problem of simultaneous equation bias; strictly speaking, the price of insurance is endogenous, dependent on a supply function for insurance as well as demand for insurance. I report demand estimates based on the computed values of THETA from a time series extending from 1929 to 1968. For 1939 to 1968 I have estimated a simultaneous model, using indexes of building costs, labor wage rates for insurance firms, and the interest rate paid on insurance investment portfolios (used as a measure of the cost of capital) as additional instruments. Enrollment is considered

[1] *Statistical Abstracts,* 1958 to 1969; and Elizabeth Langford, "Medical Care in the Consumer Price Index, 1936-1956," *Monthly Labor Review,* 86, No. 9, September 1963, 1053-1058.

endogenous. That estimate is limited to the period after 1939 because no data are available on enrollment figures for previous years. The estimated demand curves for the period from 1939 to 1968 consider THETA endogenous; as is shown in the regression results, there is some evidence that simultaneous equation bias indeed hampered the ordinary least square estimates.

Illness Distributions in the Population

Precise estimates of the level of potential illnesses facing the population in a given year cannot be provided. As an indicator of those levels, I include two variables describing the age characteristics of the population--the percent of the total population over age 65 (which should have a positive coefficient), and the proportion of the population between ages 14 and 25 (a group of people who face little illness and hence a variable that should have a negative coefficient).

The State of Medical Technology

The theory suggests that the marginal productivity of medical care in producing *health* (the underlying component of the utility function), should influence the demand for insurance. If productivity of medical care increases, a given level of health can be produced with a smaller input of medical care. Under normal circumstances, that would reduce demand for insurance, since total expenditures would be lower. I include as an estimate of the level of technology the proportion of all deaths in the community due to heart disease. There is reason to consider that variable a reliable indicator of the state of medical technology in a community. As medical skills increase, more people will be kept alive who would ordinarily die of various diseases. Those people most commonly eventually die of heart disease, so an increase in the incidence of death from heart disease can be taken as evidence that other diseases in the community are being controlled.

Also included in the regressions is a dummy variable (MEDICR), set to one for years when Medicare is in force (1965-68) to allow for possible effects of the introduction of this social insurance scheme on the private insurance system. All equations are estimated with a

constant term. The explanatory variables used in the insurance demand estimates are summarized in Table 15.

ORDINARY LEAST SQUARES REGRESSIONS

To use the entire data sample on insurance premiums and benefits (ranging from 1929 to 1968), I have estimated demand for insurance using ordinary least squares methods for that 40-year series. Data necessary to allow the loading fee (THETA) to be endogenous are not available for the first ten years of this sample, so the observed THETA cannot be estimated. The ordinary least squares equations should provide some information about demand for insurance, but the known problems of OLS regression will be evident. In particular, the effects of insurance loading fee on demand for insurance will probably be overstated, and the remaining coefficients may also be subject to bias because of the simultaneity involved. The OLS regressions are shown in Table 16.

These estimates show an income elasticity for coverage ranging between .39 and .71 (for various dependent variables, with moderately high precision around those coefficients). The effects of medical prices (PRATIO) on demand for C are estimated as elasticities between .63 and 1.87, again depending on the functional form chosen. The own-price elasticity for insurance is estimated to be between -.97 and -.41, with high degrees of precision for each of those estimates. I anticipate that the effects of THETA are overstated in the OLS regressions. The age-percentage variables have the predicted effects (OVR65 > 0 and UNDR25 < 0), and the technology variable has the anticipated negative sign.

The OLS estimates have several problems. First, the estimates of various elasticities diverge, depending on functional form. Second, there is strong evidence of autocorrelation of the residuals, especially in the COVR regression. Finally, the residuals do not appear to be normally distributed in the COVR regression, with the largest discrepancy at the low portion of the cumulative probability distribution.[1]

[1]The residuals were plotted on normal probability paper. This technique gives one method of choosing between the COVR and the LOGIT

Table 15

EXPLANATORY VARIABLES USED IN TIME SERIES STUDY

Variable Name	Meaning of Variable Name	Definition
REALY	Real income (Y) per capita	$\frac{\text{National Income}}{\text{Population}}$
PRATIO	Price ratio of medical prices to all consumer prices	$\frac{\text{Medical Prices (1953=100)}}{\text{Consumer Prices (1953=100)}}$
THETA	The loading fee on insurance	$\frac{\text{Premiums}}{\text{Benefits}} - 1$
OVR65	Percent of the population over 65 years of age	$\frac{\text{Persons Over 65}}{\text{Total Population}}$
UNDR25	Percent of the population between ages 18 and 25	$\frac{\text{Persons 18-25}}{\text{Total Population}}$
MEDICR	A dummy variable for years when Medicare is in effect	=1 for 1965-68 =0 otherwise
TECH	Level of medical technology	Percent of deaths due to heart attack
C	Constant term	=1 for all observations

The LOGIT regression residuals also appear to diverge from a normal distribution, but not as severely as with the COVR regression.

The problem of autocorrelation may be solved by using a differencing estimation technique. All observations are transformed, so that $y_t - \rho y_{t-1} = Y_t$ (and all corresponding explanatory variables are similarly differenced). An iterative procedure picks the value of rho that minimizes the residual sum of squared errors. The results are shown in Table 17. The autocorrelation appears to be effectively corrected, as indicated by the Durbin-Watson statistics (particularly for the LOGIT regression). The coefficient estimates change moderately; the most noticeable difference between these regressions and standard OLS

regressions. The residual pattern more closely fits the normality assumption in the LOGIT equation.

Table 16

ELASTICITIES OF DEMAND FOR COVERAGE

OLS 1929-1968[a]

Explanatory Variable	Dependent Variable: COVR	LOGIT	RPRMPC[b]	(Inferred Coverage Elasticity)[c]
REALY	.71 (2.99)	.56 (3.49)	.66 (2.05)	.39
PRATIO	1.87 (3.86)	.63 (1.95)	1.37 (2.10)	.54
THETA	-.41 (4.98)	-.59 (10.67)	-.68 (6.10)	-.97
OVR65	2.16 (1.85)	1.49 (1.91)	7.61 (4.88)	--
UNDR25	-2.20 (2.93)	-1.43 (2.85)	1.86 (1.85)	--
MEDICR	-.01 (1.10)	-.004 (.15)	-.003 (.33)	--
TECH	-2.43 (2.49)	.15 (.023)	-6.00 (4.60)	--
C	.1892 (1.46)	-2.884 (4.13)	-82.74 (3.24)	--
R^2	.9915	.9944	.9826	--
F(7,32)	266.473	816.205	257.488	--
D.W	0.85	1.42	.97	--

[a]Elasticities at mean values reported, "t-ratios" in parentheses.

[b]Elasticity of premiums.

[c]Inferred coverage elasticities using (3.7); $\eta_{\bar{h}I} = .25$ and $\eta_{\bar{h}k} = .05$ are assumed for these calculations.

Table 17

ELASTICITIES OF DEMAND FOR COVERAGE
AUTOCORRELATION DIFFERENCING ESTIMATES
1930-1968[a]

Explanatory Variable	Dependent Variable: COVR	LOGIT	RPRMPC[b]	(Inferred Coverage Elasticity)[c]
REALY	.30 (1.35)	.35 (1.61)	.24 (1.34)	-0.01
PRATIO	.96 (1.98)	.82 (1.85)	.96 (2.33)	.04
THETA	-.16 (1.71)	-.35 (4.18)	0.00 (.03)	-.33
OVR65	3.00 (3.50)	2.58 (2.73)	11.66 (7.83)	. .
UNDR25	-.69 (.83)	-1.50 (2.11)	3.00 (3.81)	. .
MEDICR	-.02 (1.62)	-.06 (.80)	.01 (1.05)	. .
TECH	-.46 (.55)	.05 (.07)	-.70 (.96)	. .
C	-.43 (1.94)	-4.399 (3.24)	-24.3 (7.18)	. .
R^2	.9926	.9952	.906	. .
F(7,31)	592.39	920.218	17.08	. .
D.W.	1.67	2.044	1.04	. .
RHO (ρ)	.854	.724	.926	. .

[a]Elasticities at mean values reported; "t-ratios" in parentheses.

[b]Elasticity of premiums.

[c]Inferred coverage elasticities using (3.7); $\eta_{hI} = .25$, $\eta_{hK} = .05$ are assumed for these calculations.

is the lower precision indicated for the THETA coefficient, indicating that part of the apparent precision of the OLS results is due to the autocorrelation.

The results from OLS and the autocorrelation-OLS suggest that the income elasticity of demand for coverage is positive, with the "best" point estimate (derived from the autocorrelation OLS LOGIT equation) of .35 (t = 1.61), so that the 90 percent confidence region is approximately (0.0, .70). The effect of medical care prices on demand for coverage is estimated to be positive; the point estimate of the elasticity is .82 (t = 1.85), and the 90 percent confidence region is (.07, 1.56). The effects of the loading fee on demand for coverage are estimated to be negative; the point estimate is -.35 (t = 4.18), so that the 90 percent confidence region is (-.20, -.48). None of these estimates is marked by reasonably high precision, and the problem of possible simultaneous equation bias still persists, even with the autocorrelation removed. Two-stage least squares estimates are used to account for simultaneous equation effects.

SIMULTANEOUS EQUATION ESTIMATES

For the sample period from 1939 to 1968, sufficient data exist to calculate two-stage least squares estimates of demand for insurance, allowing the price of insurance (the loading fee THETA) to be endogenous. Exogenous variables included in the instrument matrix in these estimates are REALY, PRATIO, OVR65, UNDR25, MEDICR, TECH, WEEK, BUILD, INTER, and a constant term, where WEEK, BUILD, and INTER are (respectively) indexes of the weekly salary of employees of insurance companies and brokerage firms, the general construction cost index, and the rate of return on investment portfolios of insurance companies, taken to be their opportunity cost of capital. These variables enter the insurance supply equation. The endogenous variables in the system included the amount of insurance (measured by COVR, LOGIT, or RPRMPC), the loading fee (THETA), the enrollment rate for insurance (ENROLL), and the percent of enrolled persons buying group insurance (GROUP).

The results from the two-stage least squares (TSLS) estimates are given in Tables 18-19; in the first set of estimates, the COVR variable,

Table 18

DEMAND ELASTICITY ESTIMATES--TWO-STAGE LEAST SQUARES
X-MATRIX LINEAR[a]

	COVR	LOGIT	RPRMPC	Implied COVR Elasticities[b]
REALY	.40 (1.88)	.40 (1.93)	.09 (.39)	(-.15)
PRATIO	1.82 (5.54)	.81 (2.53)	1.26 (3.35)	(.44)
THETA	-.40 (1.53)	-.63 (2.45)	-.61 (2.04)	(-.89)
OVR65	.99 (1.01)	1.49 (1.53)	5.46 (4.83)	--
UNDR25	-1.34 (2.50)	-1.30 (2.44)	2.40 (3.87)	--
MEDICR	-.01 (1.15)	-.004 (.61)	-.004 (.61)	--
TECH	-.77 (.44)	-.63 (.36)	-2.50 (1.17)	--
C	.064 (.137)	-1.83 (.59)	-143.1 (1.75)	--
R^2	.9884	.9928	.9918	--
F(7,22)	182.322	247.253	209.601	--
D.W.[c]	1.41	1.80	1.61	--
Dhrymes t-Ratio Adjustment Factor	.8254	.757	.7415	--

[a]Elasticity estimates at mean values reported; asymptotic normal (z) levels are shown in parentheses.

[b]Inferred from premiums elasticities using (3.7); values of η_{hI} = .25 and η_{hK} = .05 are assumed for these calculation.

[c]D.W. = Durbin-Watson statistic.

Table 19

DEMAND ELASTICITY ESTIMATES--TWO-STAGE LEAST SQUARES LOGARITHMIC FORM[a]

	COVR	LOGIT	RPRMPC	Implied COVR Elasticities[b]
REALY	.19 (1.03)	.19 (1.14)	.08 (.38)	-.16
PRATIO	-.21 (.45)	.07 (.16)	.18 (.35)	-.59
THETA	-.84 (2.81)	-.79 (2.79)	-.61 (1.77)	-.88
OVR65	1.65 (1.62)	1.58 (1.65)	6.15 (5.23)	--
UNDR25	-1.01 (2.07)	1.00 (2.16)	.70 (1.24)	--
MEDICR	(.01) (.19)	0.00 (.01)	-.01 (.15)	--
TECH	-.47 (.26)	-.61 (.36)	-1.35 (.65)	--
C	.11 (.014)	-.74 (.14)	23.67 (2.73)	--
R^2	.9916	.9918	.9931	--
F(7,22)	345.506	325.639	286.844	--
D.W.[c]	1.56	1.52	1.59	--
Dhrymes t-Ratio Adjustment Factor	.9633	.9259	.7919	--

[a]Asymptotic normal (z) scores in parentheses.

[b]Inferred from premiums elasticities using (3.7); values of η_{hI} = .25 and η_{hK} = .05 are assumed for these calculations.

[c]D.W. = Durbin-Watson statistic.

and all explanatory variables are in linear form.[1] In the second set of estimates, all variables are in logarithmic form.

The income elasticity estimates from these equations are reasonably similar. The regressions with income in linear form show an estimated elasticity of .40 (z = 1.93) in the LOGIT equation; the 90 percent confidence interval is (.05, .75).[2] In logarithmic regressions, the estimated income elasticity is .19 (z = 1.14); the 90 percent confidence interval is (-.09, .47).

The estimated effect of medical care prices on demand for insurance is dependent upon the functional form chosen. In the linear X-matrix equations, the estimated elasticity of COVR with respect to medical care prices is 1.82 (z = 5.54); the 90 percent confidence interval is (1.26, 2.38), indicating a highly responsive demand for COVR to changes in PRATIO. Alternatively, if the LOGIT demand curve is used, the estimated elasticity of PRATIO on demand for coverage is .81 (z = 2.53); the 90 percent confidence interval is (.49, 1.13). This response of coverage to medical prices is large, but not as much as shown by the COVR equations. Finally, if the logarithmic transformations are used, the results are strikingly different. The estimated elasticity in the COVR equations is -.21 (z = .45), and in the LOGIT equation the estimated elasticity is .07 (z = .16). In these equations there is insufficient precision to make any definite statements about the value of the elasticity of coverage with respect to medical prices.

The estimates of coverage elasticity with respect to THETA are similar among functional forms. The asymptotic z-scores of these estimates are considerably lower than in the OLS estimates, indicating that the OLS equations may overstate the precision. The COVR equation (Table 18) shows an elasticity of -.40 (z = 1.53); the 90 percent confidence region is (-.84, +.04). The LOGIT equation with THETA linear gives an elasticity estimate of -.63 (z = 2.45); the 90 percent

---

[1] The LOGIT variable is, of course, in logarithmic form in all equations.

[2] All confidence intervals in the simultaneous equation estimates are *large sample* intervals. All "z-scores" are asymptotically normal statistics.

confidence region is (-1.06, -.20), suggesting (with about 90 percent confidence) that insurance coverage behaves as a normal good (negatively sloped demand curve) but that demand is inelastic.

In the logarithmic form equations, the own-price elasticity of COVR is estimated to be -.84 (z = 2.81); the 90 percent confidence region is (-1.35, -.34). In the logarithmic LOGIT equation, the estimated own-price elasticity of coverage is -.79 (z = 2.79); the 90 percent confidence retion is (-1.26, -.31). These estimates suggest that demand for coverage is more highly elastic than the linear form estimates show.

As is obvious from these results, the estimates of income, medical price, and own-price elasticities of insurance coverage are highly dependent upon the functional form of the equation. There is little to choose among the equations on the basis of explained variance--the LOGIT estimates are typically higher, but the LOGIT variable is in logarithmic form, whereas the COVR variable is (in one set of equations) in linear form. A comparison of the F-statistics for the equations shows that the LOGIT equations dominate the COVR equations no matter what the functional form of the explanatory variables, but these measures do not give accurate evidence on which equation is "better." To establish some helpful way of choosing among these equations, the pattern of the residuals was investigated.

First, the Durbin-Watson (D.W.) statistics between the regressions may be compared. The D.W. statistic should be near 2.0 if the error terms are non-autocorrelated. If there is positive autocorrelation, the D.W. statistic will become smaller, approaching 0.0 as the autocorrelation increases. In these equations there are 30 observations with eight explanatory variables. The D.W. upper and lower acceptance and rejection levels for those values are not tabled (the standard tables reach only six explanatory variables), but the tables may be extrapolated to provide some indication of what the appropriate significance levels of the D.W. test are. For 30 observations, the 5 and 1 percent rejection regions are given in Table 20.

We may reject the hypothesis of no positive serial correlation if the actual D.W. statistic falls below d_L and accept the null hypothesis

Table 20

CRITICAL VALUES FOR THE DURBIN-WATSON STATISTIC, 30 OBSERVATIONS

Probability Level	K = 6		K = 8[a]	
	d_L	d_U	d_L	d_U
.05	1.07	1.83	.93	2.01
.01	.88	1.61	.74	1.93

Source: J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression I," *Biometrica*, 37, December 1950.

[a]These values are extrapolations from the complete D.W. tables, and should be viewed only as approximations to a calculated value.

if the D.W. statistic is above d_U. For intermediate values, no result may be inferred.

Application of the Durbin-Watson statistic to my regressions is generally inconclusive. Only in the OLS regression of COVR and RPRMPC is there a sufficiently small D.W. test to accept the hypothesis of serial correlation of the residuals.

A second test of the regresions is whether there is heteroscedasticity in the residuals, in which case weighted regressions would be appropriate. Visual examination of the residuals in my equations did not reveal any significant heteroscedasticity, particular attention being paid to possible changes in the variance of the residuals with the dependent variable.

Finally, residuals were plotted on cumulative normal probability paper for the three dependent variables, with the linear and the logarithmic specification. The patterns of the residual plots clearly establish that the logarithmic form regressions have normally distributed residuals, whereas the linear formulations depart somewhat from normality. The LOGIT transformation demand curves show normality either with the X-matrix linear or in logarithmic form offering one discrimination factor to choose from among the regressions.

A final test might be to examine which of these formulations gives the best fit for COVR, through transformation of the various equations. I performed that test for each of the COVR and LOGIT equations studied--the OLS estimates and the two-stage estimates in logarithmic and linear form. The residual standard errors of COVR were calculated from the fitted values, with the result that the LOGIT transformation of the coverage ratio, estimated in linear form for the X-matrix, provided the best fit for COVR variables. (That remained true in both OLS and with TSLS, under various formulations of the exogenous variable list in the TSLS estimates.) The linear form of the X-matrix in the COVR equations, alternatively, provided the poorest fit, and the linear LOGIT equations therefore seem to be the most appropriate. Given the satisfactory pattern of residuals, and the appropriate D.W. statisics from these equations, the linear LOGIT equations are the most valuable in predicting coverage levels. Table 21 summarizes these results.

The linear form LOGIT equation from the TSLS estimates provides the following picture of demand for reimbursement insurance: The income elasticity is about .40 (95 percent confidence interval is -.006 to .806), the elasticity with respect to the price of insurance is -.63 (-.12 to -1.14), and the elasticity with respect to the price of medical care is .81 (.17 to 1.45). Demand for coverage is increased by higher risk (OVR65), decreased by lower risk (UNDR25), and is not a significant function of the level of medical technology (TECH). No measurable effect of the introduction of Medicare upon insurance coverage of other members of society was found. In the sample studied, over 99 percent of the variation in COVR was explained by the linear LOGIT formulation.

SUMMARY OF RESULTS

The regressions using time series data on total insurance premiums and on average coverage ratios for medical insurance are somewhat inconclusive. The estimated own-price elasticity for insurance (THETA) is always negative, with the best equations providing point estimates of about -.60 to -.70 for that elasticity. The income elasticity and medical price elasticity of COVR are estimated with less precision,

Table 21

STANDARD ERROR OF RESIDUALS FROM PREDICTING COVERAGE RATIO FROM VARIOUS FUNCTIONAL FORM SPECIFICATIONS

Dependent Variable (original specification)	X-Matrix Linear	X-Matrix Logarithmic	Ratio of Standard Errors (LINEAR/LOG)
COVR (OLS)	.01504	.00958	1.57
LOGIT (OLS)	.00794	.01114	.71
Ratio of standard errors (COVR/LOGIT)	1.89	.86	
COVR (TSLS)	.01194	.01081	1.10
LOGIT (TSLS)	.00948	.00998	1.05
Ratio of standard errors (COVR/LOGIT)	1.25	1.08	

and the estimated values are dependent upon the functional form for the equation. The regressions using total premium data provide no extra insight into the "true value" of those elasticities; the precision of the coefficient estimates is quite low in both the linear and logarithmic form equations of RPRMPC. The variables reflecting the expected illness distribution in the population (OVR65 and UNDR25) strongly enter the demand curves for coverage, and suggest that those persons facing higher *mean* illness levels purchase more insurance coverage. The variable reflecting the state of medical technology (TECH) entered negatively in the OLS regressions but was generally insignificant in the TSLS estimates.

Factors making estimation of individual coefficients imprecise or inefficient include high multicollinearity among the explanatory variables, omitted variable bias, and possible nonspherical disturbances. Tests for the last show that the LOGIT equations and the logarithmic COVR and RPRMPC equations conform to the usual regression assumptions. Tests of residual variance in predicting COVR show that the linear X-matrix LOGIT equations give the best fit.

A more complete model would allow the price of medical care (PRATIO) to be endogenous, which might correct the problems surrounding those estimates. In such a model, demand and supply curves would be estimated for the medical sector as well as the health insurance sector, and the effect of insurance coverage on demand for medical care would be included in the specification of the model. Unfortunately, I was unable to obtain sufficient data to estimate the supply of medical services equation. To do so correctly would necessitate data on the price of labor to hospitals (for example, nurses), available only erratically over the sample period; the amount and price of construction capital available for hospitals, much of which does not come through normal capital markets; the wage rate for physicians, which is not available in any meaningful sense for most of the years in the sample; and so on. Given the absence of these highly relevant factors, I decided to omit any attempts at a four-equation model with the time series data.

V. POLICY IMPLICATIONS

The primary issues of importance for actual insurance schemes that might be drawn from this study include (1) pricing systems for insurance, (2) compulsory enrollment, (3) the variety of insurance policies provided under mandatory schemes, and (4) the effects of having insurance payments made in kind (reimbursement insurance) rather than in money (indemnity insurance).

INSURANCE PRICING ARRANGEMENTS

The issue of pricing has troubled the health insurance field since its inception. In the early 1900s, most "health insurance" in the United States was supplied through casualty insurers and was of indemnity form.[1] The pricing of those policies was almost uniformly based on expected marginal cost, where premiums fully and accurately reflected all costs associated with the insured party.

When Blue Cross and Blue Shield plans were organized in the 1930s to provide insurance against hospital and physician expenses, a different pricing system was adopted. Almost uniformly, these plans used a *community rating* scheme,[2] wherein every person purchasing insurance from that organization paid the same premium, no matter what the person's age, occupation, family size, or expected medical expenses. This pricing system was explicitly adopted to provide health insurance to certain groups in the population. Somers and Somers characterize Blue Cross-Blue Shield's philosophy as based "on the view that it was to serve the total community, and to apply rates that would maximize community enrollment."[3] Since persons of different ages (for example)

---

[1] That is, a flat cash settlement was made upon the occurrence of a particular event (such as loss of a limb).

[2] For a discussion of the history of these plans, see Herman M. and Anne R. Somers, *Doctors, Patients, and Health Insurance*, Brookings Institution, Washington, 1961, Part Five; or Duncan M. MacIntyre, *Voluntary Health Insurance and Rate Making*, Cornell University Press, Ithaca, 1962, Chapters II and IV.

[3] Somers and Somers, *Doctors, Patients, and Health Insurance*, p. 309.

have predictably different expenditures for medical care, the community rating system may be equated with price discrimination. Since Blue Cross-Blue Shield had an effective monopoly in health insurance during this period, it may be seen that their community rating would maximize enrollment, as would price discrimination for any monopolist.

Community rating produces an income transfer from persons with low expected medical expenditures to those with high expected medical expenditures. Thus, if the income elasticity of demand for medical care is positive, community rating could be viewed as a mechanism to transfer income from those persons with low incomes to those with high incomes.

An alternative view of community rating is that it is actually able to bring poor risk persons (typically referred to as low-income persons) into the insurance market, and that the consequences are desirable for those persons. Somers and Somers suggest that "private health insurance might have been able to provide near-universal enrollment had the original Blue Cross concept of community rating prevailed, *thereby affording the financial base for underwriting the high-cost low income groups whose insurance could not be self-sustaining.*"[1]

One can infer what the actual net transfers through community rating are by investigating demand for community rated insurance. If net effects benefit the poor, then states with a heavy incidence of poor families should be characterized by strong demand for community rated Blue Cross insurance, given that it is available in the market. If net income transfers benefit the rich, then high income persons should systematically demand Blue Cross coverage. One test of this proposition would be to investigate the effects of state income levels on demand for community rated insurance. Data from four years show a positive effect of state income levels on Blue Cross's share of the insurance market, where community rating was practiced in that state. An increase of 1 percent in state income was estimated to produce a .19 percent increase in the Blue Cross market share in these community rating states. In states where experience rating (expected marginal

[1]*Ibid.*, p. 518. Emphasis added.

cost pricing) was uniformly practiced, this relationship vanished.[1] It appears that community rating systematically produces income transfers from the poor to the non-poor, so that the non-poor are likely to exhibit a preference for community rated insurance, all other things equal.

Although I know of no studies demonstrating this phenomenon, there is reason to believe that other variables that systematically increase demand for medical care will also systematically increase demand for community rated insurance, if it is available. As long as there are some low-cost consumers included in the rate base, for whatever reasons, high-cost users would systematically benefit from the community rating system.

Coexistence of community rated insurance and experience rated insurance would be theoretically impossible in a competitive market with freely available information about probabilities of events and complete free choice of insurance. Two factors may allow community rated insurance to exist even with experience rated insurance available in the marketplace. First, it may be difficult for information to be acquired about probabilities of illnesses (or expenditures), leading to some breakdowns in otherwise "perfect" markets. Second, the purchase of insurance through groups may bring some low-cost users into a community rated plan, even if they know that they are subsidizing other users. To a "healthy" person, group-purchased community rated insurance implies a higher loading fee than normal, because of the subsidy to other persons in the plan. But the group sale of insurance reduces the loading fee, compared with experience rated individual policies. On net, a rational consumer may prefer to purchase a community rated group insurance plan rather than an experience rated individual insurance plan. Apparently this occurs to some extent, since community rated insurance still holds a non-zero share of the insurance market.

[1] Charles E. Phelps, "Regulation of Non-Profit Health Insurance," unpublished, 1970.

There is some evidence that the health insurance market is moving toward a pure experience rated (competitive) equilibrium. As noted previously, an entire community purchasing community rated insurance is consistent only with monopoly in the provision of insurance. This was essentially the condition of the insurance market in the 1930s. But by 1952, of 87 Blue Cross plans throughout the country, 13 offered experience rated or "cost-plus" hospital plans, involving approximately 1.6 million persons (less than 4 percent of the total Blue Cross enrollment).[1] By 1958, only 22 of the Blue Cross plans offered only community rated contracts (with total enrollment of about 12.5 million persons).[2] One observer reports that by 1968, "Virtually all Blue Cross plans now use the technique of experience rating at least to some extent."[3] From 1950 to 1968, the share of the hospitalization insurance market held by Blue Cross (measured in percent of persons enrolled) declined from about 50 percent to less than 40 percent, indicating the increasing market position of the experience rating commercial insurance companies.

COMPULSORY INSURANCE

Setting aside the pricing aspects of health insurance, another feature of current health insurance proposals is the implied or explicit compulsory enrollment. The social arguments for and against compulsory enrollment are based on other grounds than positive economics, and I prefer not to deal with that issue here. However, three separate effects can be expected to occur if compulsory insurance is adopted.

Income Transfers

The combination of compulsory enrollment and community rating will introduce a recognizable pattern of income redistribution. Indeed, the income redistribution aspects of some plans are cited by proponents

---

[1] MacIntyre, *Voluntary Health Insurance*, p. 164.

[2] *Ibid*, p. 165.

[3] Fredric R. Hedinger, "The Social Role of Blue Cross," *Inquiry*, 5, No. 2, June 1968, p. 11.

of those plans as being their main positive feature. The net income transfers of a compulsory health insurance system may work in a direction other than is intended by proponents of such plans. Because of the positive income elasticity of demand for medical care (even given constant health insurance levels) the income flows may run from the poor to the non-poor. If the existing location of medical service providers is such that low income persons face high travel time to obtain medical care, then even with comprehensive health insurance, travel costs may reduce the consumption of medical services by low income persons. This would, in effect, magnify the income transfer from low to high income persons.

Changes in Incentives Regarding Self-Selection and Self-Insurance

Ehrlich and Becker show that non-market alternatives to insurance are available in general cases--self-insurance (here, self-treatment of medical problems that reduces demand for market medical services) and self-protection (taking steps to avoid illness and injury).[1] They also show that (given some theoretical abstractions) an increase in market insurance will reduce efforts in self-protection and self-insurance, especially if such efforts are not recognized in the market insurance premiums. Their results suggest that compulsory insurance will increase the demand for market medical care (since self-protection and self-insurance would be reduced in the face of increased insurance coverage). In other words, there will be induced demand for market medical services resulting from a diminished non-market production of health. The extent of substitution between market and non-market production of health has not been measured. Additionally, those who are at present very efficient at self-production of health will, in the face of compulsory health insurance, reduce their expenditures on "self-insurance," producing shifts in demand toward the formal medical care sector. Similar arguments may be extended with regard to "self-protection" expenditures, as long as such expenditures are not recognized in the total price of insurance. Presumably, if national health

[1]Ehrlich and Becker, "Market Insurance, Self-Insurance, and Self-Protection," p. 633.

involved insurance "community" rating, then the demand for self-protection in the community would decrease and medical care use increase.

Overinsurance

One may also consider the amount and types of insurance available under compulsory national schemes. My estimates of demand for insurance show significantly non-zero demand elasticities with respect to income, medical care prices, time costs, and illness levels facing the consumer. Presumably, any compulsory national health insurance scheme would place some people in a position of being overinsured relative to their private demands.[1] Moreover, if health insurance is of the reimbursement form, the consumer will be unable to reduce his amounts of insurance directly, since health insurance benefits would not be acceptable as collateral for monetary loans and cannot be sold to other persons. Unless variability in coverage is provided under compulsory insurance, some persons will obviously be placed in a non-optimal position. To determine the optimal amount of variability, it is necessary to consider the costs of producing that variability, and I have no estimate of those costs. However, the revealed preferences of consumers in purchasing individual insurance policies (at higher loads than for group insurance) suggest that consumers may be willing to pay a significant price for variability.

The Medicare Act of 1965 established two possible levels of insurance for persons over 65 years of age, with enrollment in one level (hospitalization insurance) being mandatory, and enrollment in the other (supplemental medical insurance) being voluntary. Despite the heavy subsidy involved for persons enrolling in the supplemental medical

[1]The same argument may be made about employer contributions to health insurance premiums, in lieu of wages to an employee. If a consumer is to collect his entire wage, he must accept whatever insurance policy is being purchased by his employer. His resulting level of insurance may be "too high" relative to his private demands. Presumably, competition in provision of wages will establish, at least on average, the correct level of employer-paid insurance premiums. I expect that amount to increase with income, since wages paid in kind are not subject to income taxes. The estimated income elasticity of demand for employer-paid premiums is .97, whereas the income elasticity for total premiums is .48.

insurance, only 84 percent of the eligible population voluntarily enrolled in this plan.[1] Payments by state welfare organizations enrolled another 11 percent of the eligible population, bringing the total enrollment to 95 percent of the eligible population. M. Feldstein has studied voluntary enrollment for "Part B" of Medicare by state and found the enrollment proportions to be positively associated with state income, percent of the population that was white, and percent living in urban areas. These results are consistent with my demand equations, which show a positive income elasticity and a positive elasticity with respect to medical care prices (here, represented by the percent of the population living in urban areas). Feldstein also included age variables and variables describing general demand for insurance in the state. These variables are somewhat difficult to interpret, especially because the age and race variables are so highly correlated with income.[2]

My empirical estimates of demand for insurance suggest that as their various characteristics change, consumers would demand significantly different amounts of insurance, even if all insurance policies were offered at the same loading fee levels. Any compulsory scheme that would not allow for variability would suppress consumer preferences to some extent.

Following is an illustration of the amount of overinsurance that might emerge under a "typical" compulsory plan. Assume that a compulsory insurance policy was instituted for every person in the United States and was set (to use an arbitrary example) at the level purchased privately by college graduates with an income of $14,000, living in an urban area, with the head of household between 45 and 54 years old. I estimated that in 1963, such a family would privately demand $280 of health insurance, if insurance were actually purchased. This family's

---

[1]The subsidy occurred through underpricing of the plan. Initially, premiums were set at $36 per person per year; those premiums increased by 1971 to $67.20 per year. In 1967, when premiums were $36 per year, benefits were $38 per year on average. In 1968, annual premiums were $44, and average annual benefits were $77.

[2]Martin S. Feldstein, "An Econometric Model of the Medicare System," *Quarterly Journal of Economics*, No. 338, February 1971, p. 1-20.

insurance is considered to be purchased at the minimum commercial loading fee, implied by the head of house being in the largest possible work group. Since the loading fees for the largest private groups approach those on governmental policies, I assume that this same loading fee will characterize the governmental plan. Table 22 shows the expected value differences between the *hypothetical* governmental insurance plan and the amount that would be purchased privately by various individuals, *at the minimum load rate* implied by a governmental policy. Obviously, this exercise shows only the amount of overinsurance that would result from this particular plan. If the insurance level were set higher, the amount of overinsurance would increase, and conversely.

Summation over all age-education-income groups in society would give the total amount of overinsurance implied by this particular insurance plan. The amount of overinsurance could be taken as an estimate of an opportunity cost of not providing variability in compulsory insurance. This is not a true measure of the welfare loss, since the compulsory insurance would have some value to the families (but not as much value as alternative uses of the equivalent amount of income).

As long as demand elasticities for insurance are non-zero, and as long as a fixed level of insurance is made compulsory, some people will probably be overinsured. The normal arguments about subsidies in kind versus subsidies in money apply as much to insurance as to any other good. If consumers privately demand variability in coverage, then some welfare loss is implied by compulsory insurance.

THE EFFECTS OF REIMBURSEMENT INSURANCE ON MEDICAL CARE DEMAND

One common characteristic of health insurance is the method of payment. Unlike home insurance, property insurance, and some forms of automobile insurance, market medical insurance typically computes the benefits to be paid by observing expenditures on medical care and repaying the consumer in some fixed proportion, agreed upon contractually. This payment mechanism alters the marginal price line facing consumers, reducing it below market prices. Arrow views this type of insurance as payment in kind.[1] Section II indicated that there are two effects

[1]Arrow, *Theory of Risk Bearing*, p. 203.

Table 22

AMOUNT OF OVERINSURANCE THAT WOULD OCCUR FOR VARIOUS INCOME-EDUCATION GROUPS IF COMPULSORY HEALTH INSURANCE OF VALUE $280 HAD BEEN INSTITUTED IN 1963[a]

Family	Education of Head							
Income	None	1-4	5-6	7-8	9-11	12	13-15	16+
2,000	216.8	190.6	166.5	147.2	134.2	128.5	130.2	139.3
4,000	203.9	175.1	149.3	129.0	115.5	109.5	111.2	120.7
6,000	189.6	158.5	131.2	110.0	96.0	89.8	91.7	101.5
8,000	174.0	140.8	112.4	90.4	76.0	69.7	71.6	81.6
10,000	157.3	122.4	92.8	70.7	55.6	49.2	51.1	61.3
12,000	139.6	103.2	70.3	49.8	35.0	28.4	30.3	40.7
14,000	121.1	83.4	52.3	29.0	14.0	7.4	9.4	19.8

[a]These figures show the difference between $280 premiums and the expected value of total premiums, predicted from the verified data Tobit regressions, using average values for ESTLOS, TOTPRI, BLACK, FEMALE, RURAL, FREE CARE, and WELFARE CARE. The value of GRPSIZ has been set at 8 (the largest possible value), indicating minimum loading fees for all insurance. The figure of $280 represents the value of insurance for a family of $14,000 income, with college education, *if insurance was purchased*.

involved in reimbursement insurance, acting sometimes in offsetting fashion, that alter demand for medical care; the marginal price of insurance facing the consumer is lowered, causing a substitution toward medical care, and there is an income effect, which is either positive or negative as the *net* insurance benefit is positive or negative. For serious illnesses, both effects act in the same direction, indicating an increase in demand for medical care. Overall, however, there is no prediction about the net effects of reimbursement insurance on demand for medical care. Unless the resources used in producing insurance are large, and η_{hp} is quite small, reimbursement insurance will increase demand for care on average.

One study of expenditures on prescription drugs showed that quantities consumed and total expenditures were twice as high in an insured group of persons as in uninsured persons in the same city. This study

compares actual expenditures of insured families with a sample of the remainder of the population of Windsor, Ontario. The results are shown in Table 23.

Of particular interest is that age-adjusted expenditures in this study are higher than gross expenditures, indicating that there was not an abnormally high age group of persons enrolled in the insurance plan. Consumers paid a $0.35 deductible for each prescription, with no coinsurance provisions. At mean prescription prices, consumers paid about .10 of the market price for the "average" prescription. Since the marginal cost of high-priced prescriptions was zero, insured persons purchased more expensive drugs, as well as increased quantities of drugs.[1] This is one piece of evidence on the effects of insurance on demand for medical goods and services.

Rosett and Huang have also estimated the effects of reimbursement insurance on demand for medical care, using data from the 1960 survey of consumer expenditures.[2] Their estimates of changes in demand for 1960 suggest that medical care is quite elastic to coinsurance levels. If a policy were instituted that had no deductible, the demand shifts they predict for various coinsurance levels would be as shown in Table 24.

Rosett and Huang overestimate the effects of insurance on demand for medical care for a variety of reasons. First, they treat insurance as exogenous, introducing simultaneous equation bias. They use an average, rather than the appropriate marginal coinsurance rate (of necessity, since they had no information on deductibles), and lump together insured and uninsured services (a similar problem--uninsured services can be treated as having an infinitely large deductible). Of

---

[1]The implied arc-elasticity of demand for drugs (d) (using the average price) is

$$\bar{\eta} = \frac{\Delta d}{\Delta p_d} \cdot \frac{\bar{P}_d}{\bar{d}} = \frac{1.89}{-3.43} \cdot \frac{2.06}{3.15} = -\ .36.$$

[2]Richard N. Rosett and Lien-Fu Huang, "The Effect of Health Insurance on the Demand for Medical Care," *Journal of Political Economy,* 81, No. 2, Part 1, March/April 1973, 281-305.

Table 23

DRUG EXPENSES WITH AND WITHOUT INSURANCE

	Total Population	95 Percent Confidence Interval of Sample	Insured Population	Age-Adjusted Insured Population
Annual number of prescriptions	2.19	1.81-2.57	4.08	4.20
Annual expenditures	$8.29	$6.89-9.69	$16.48	$16.64
Mean $/prescription	$3.78	$3.70-3.86	$ 4.03	$ 3.96

Source: M. R. Greenlick and B. J. Darsky, "A Comparison of General Drug Utilization Under a Drug Prepayment Plan," *American Journal of Public Health*, 58, No. 11, November 1963, 212.

Table 24

ESTIMATED EXPENDITURE, NO DEDUCTIBLE, 1960

Coinsurance Rate	Expenditure (in $ millions)
.1	$29.4
.2	25.3
.3	21.4
Actual expenditures--1960	10.4

Source: R. Rosett and L. Huang, "The Effects of Health Insurance on the Demand for Medical Care."

most serious consequence is that they have no measure of employer-paid premiums (which are an important fraction of total insurance), and they must assume (for lack of information) that all insurance policies have the same loading fee, an assumption that is clearly shown to be incorrect in my data. Since coverage varies inversely with the loading fee,

this causes a systematic error to be established in their measure of coverage, hence biasing their coefficient estimate away from zero.[1]

Other empirical studies show the effects of health insurance on demand for medical care. Newhouse and Phelps use the same data base (the 1963 CHAS-NORC health survey) as used in this analysis of demand for health insurance.[2] This study treats insurance as endogenous and uses the marginal coinsurance rate facing the consumer as the explanatory "price" variable, thus correcting problems faced by other researchers in this area. The estimated price elasticities for hospital care are near -.1 to -.2; for physician office care, they are between -.05 and -.15, depending on the particular specification used.

Additional evidence is available on the effect of coinsurance on the demand for medical services. Data from a study on demand for medical care in a prepayment plan have been analyzed by Phelps and Newhouse.[3] This study shows that the introduction of a 25 percent coinsurance rate on all services in a prepaid group practice setting reduced the quantities of service demanded by 28 percent (significant at .00007 probability). The study has data on use and expenditures by 2567 persons in 1966 (when no coinsurance was present in the prepayment plan) and in 1968 (when the 25 percent coinsurance was instituted). All characteristics of the families remained unchanged, so any changes in demand could be attributed only to the coinsurance payments.

Other available evidence on the effects of insurance on demand for medical care has been collected by Phelps and Newhouse,[4] showing that (if coinsurance ranged between 25 percent and zero) the aggregate

---

[1]This and other medical care demand studies are criticized more carefully in Joseph P. Newhouse and Charles E. Phelps, *On Having Your Cake and Eating it Too: A Review of Estimated Effects of Insurance on Demand for Medical Care*, R-1149-NC, The Rand Corporation, Santa Monica, forthcoming.

[2]Joseph P. Newhouse and Charles E. Phelps, *Price and Income Elasticities for Medical Care*, R-1197-NC, The Rand Corporation, Santa Monica, forthcoming.

[3]Phelps and Newhouse, *The Effects of Coinsurance on Demand for Physician Services*.

[4]Phelps and Newhouse, *Coinsurance and the Demand for Medical Services*.

demand elasticity was approximately -.1, varying by type of service. The implications of this result for prospective national health insurance plans suggest that demand under the "minimum" of the proposed plans (the Administration bill, which features a 25 percent coinsurance) would be about 20 to 25 percent lower than demand under the "maximum coverage" plan (such as the Kennedy-Griffiths-Corman proposal, which features no coinsurance or deductibles). How this demand would be met and the distributional consequences in the short run are discussed by Rockwell *et al.*[1]

These studies do not provide the information necessary to assess the costs of some prospective national health insurance plans. For example, some plans primarily feature a "catastrophic coverage" scheme, where there is little coverage for routine events but complete (C = 0) coverage for any expenses above a certain level. Implicit in such proposals is that there is (relatively) high elasticity of demand for medical care for routine services, but (relatively) inelastic demand in the case of "severe" or high-loss events. The studies cited above do not show how the elasticity changes as the quantity of medical care demanded changes; they show either average elasticities over broad ranges of expenditures or elasticities estimated in constant elasticity form. Also not yet shown empirically is how the response to coinsurance changes over different levels of coinsurance. (Put differently, no researchers have yet shown how the response to price changes with price levels.) If the coinsurance elasticity of medical care increases in absolute value as C approaches zero, then the national health insurance proposals that offer complete coverage (C = 0) may be inordinately more expensive than those with some coinsurance provision (typically proposed is 20 percent). Alternatively, if the coinsurance elasticity of demand for medical care approaches zero as C approaches zero, then there may be little induced demand by addition of that final coverage. The question is empirical, but it has not yet been demonstrated. Both effects of coinsurance should be learned before complete assessment of

---

[1]Marshall A. Rockwell *et al.*, *Public Policy and Economics of National Health Insurance*, The Rand Corporation, Santa Monica, forthcoming.

national health insurance plans is possible. One would show whether the implicit assumptions about catastrophic insurance are correct, and the other would show how costly complete coverage (as opposed to, say, 80 percent coverage) would be. These questions should take high priority for health economics researchers studying demand for medical care.

Section II may give some insight into the second question. There I showed that the net price of medical care consists of a time price and a money price, so that NPH = $Cp_h + w \cdot t$. Further, the elasticity with respect to either C or p_h could be written approximately as

$$\eta_{hC} \approx \eta_{hh} \approx \frac{Cp_h}{Cp_h + w \cdot t} \eta_h^T ,$$

where η_h^T is the elasticity with respect to "total price" ($Cp_h + wt$). If η_h^T is roughly constant over all levels of total price, the elasticity with respect to C must decline as C approaches zero. Indeed, the higher the time price associated with a good or service, the lower will be the money-price elasticity, as a general phenomenon.

Hospitalization may be a good example of a high time-price good; the net price of medical care (even in these times of high money prices for hospitals) is composed primarily of time prices. Introduction of reimbursement insurance further reduces the effect of money prices, so we could expect to observe very low empirical estimates of the money price (or coinsurance) elasticity for hospital services.

Prescription drugs and dental care may be examples of low time-price goods (the fraction of total price accounted for by money price is high). In these goods, insured money price is close to total price, so that the elasticity with respect to money price does not change over levels of C (given my assumption that η_h^T is constant). With such goods, I have no way *a priori* to determine how η_{hC} will vary with C if we relax the assumption of constant η_h^T.

PRESENT AND POTENTIAL INSURANCE MARKETS

Reimbursement insurance is most probably provided through competitive markets. There are over 900 suppliers of health insurance in the United States, although not all of these operate in every state. In 1960, the minimum number of insurers operating in a given state was 83 (Alaska) and the maximum number of insurers in any state was 385 (Texas).[1] These numbers seem excessive to permit collusion among insurers. Yet there is some evidence to suggest that differential tax status of some insurers has altered the market away from a purely competitive equilibrium. A class of insurers, organized under non-profit laws in each state, has obtained exemption from state premium taxes, real estate taxes, corporate income taxes, and federal income taxes. If this group of insurers offered the same types of coverage as did other "commercial" insurers, the tax break would simply have provided a subsidy to some insurance purchasers. But typically these insurers (Blue Cross and Blue Shield, plus some "independent" health plans) offer first-dollar, first-day hospitalization coverage, whereas the "commercial" companies normally include some deductible amount and a coinsurance provision in their policies. Thus the differential tax treatment of one class of insurers alters the price line between types of insurance facing the consumer. The price shifts appear to alter relative market shares between these two classes of insurers. In another study, I have estimated that an increase of 1 percent in the state corporate income taxes would increase Blue Cross's market share by .2 percent (an elasticity of .20), with a similar effect arising from changes in premium taxes. I extrapolated these estimates over the range of average state corporate income and premium taxes, and over federal corporate income taxes, and concluded that the Blue Cross plans would maintain only 5 to 15 percent of the hospital insurance market if they did not receive favorable tax treatment *and if they maintained their present forms of insurance.*[2] This estimate seems to be well

[1] *Source Book of Health Insurance Data 1961*, Health Insurance Institute, New York, 1961, p. 47.

[2] Phelps, "Regulation of Non-Profit Health Insurers."

supported in the only two states that treat non-profit insurers identically to commercial insurers; in Nevada, the Blue Cross market share is 5 percent, and in South Dakota its share is 12 percent. The national average share for Blue Cross in 1968 was 30 percent.

There is additional evidence that the "commercial" insurers would like to avoid the premium tax when it is imposed; the differential tax treatment is economically as well as statistically significant. In New York State, one large insurer is currently suing for refund of over $1.5 million paid in premium taxes on certain group insurance plans. In these plans, the underwriter never actually collected 90 percent of the group insurance premiums; they were held by the companies purchasing the insurance. (The insurer was acting more as a reinsurance agent for a company-operated group insurance plan, and the premiums did not change hands.) The suit is attempting to establish the legality of this type of insuring activity and to establish the tax immunity of these transactions. I do not know the results of this suit at present, but the mere attempt in court of one insurer to achieve such a ruling is evidence that the differential tax status favoring Blue Cross and Blue Shield plans is disruptive to a purely competitive market.[1]

Other types of health insurance seem to be obstructed to some extent in this market. Prepaid group practice, one substitute for the combination of market insurance and fee for service medical care, has had long and frequent legal and political troubles. Kessel has argued that this mode of medical practice (which is also a substitute for insurance) has been obstructed on grounds that it eliminates opportunities for price discrimination in the practice of medicine.[2] MacColl has provided an excellent history of the legal and political opposition that accompanied

[1]Martin E. Segal, "News from Martin E. Segal Company," undated mimeograph.

[2]Reuben E. Kessel, "Price Discrimination in Medicine," *The Journal of Law and Economics*, I, 1958, pp. 20-53.

organization of these groups.[1] Indeed, it is still illegal in some states for medicine to be practiced under corporate auspices.

Most legal obstructions to prepaid group practice have now been removed, yet these plans still represent a small fraction of total insured persons. Some of the reasons for this may include the high time cost implied by having to travel to specific locations to obtain care, the long queues to obtain service,[2] or differences in the style of practice to which patients are not accustomed. These factors must be considered when weighing other advantages of prepayment plans and may help to explain why these plans have not grown as rapidly as their theoretically advantageous organization would imply.

In addition to consideration of the insurance market itself, it is necessary to consider the type of market in which medical care is delivered, if one is to fully evaluate the effectiveness of present reimbursement insurance. A major factor in this market is that outcomes in terms of health are highly uncertain, given an input of medical care. The production function that translates medical care into health has, in the real world, random elements that make the outcome highly uncertain, especially for some procedures. This may be one reason why reimbursement insurance is more common than cash indemnity insurance in this market; if the implicit value of the lost health could be estimated by insurers, there is no guarantee that the same amount of health could be obtained, even if the entire cash amount were invested in health care. With reimbursement insurance, the consumer can incrementally add medical care until he reaches the desired level of health, assessing the consequences of each medical procedure after it has been performed.

The general theory of demand for insurance tells us that, when loading fees are constant over all hazardous states of the world, optimal coverage produces identical final outcomes for all such states.[3]

---

[1]William A. MacColl, M.D., *Group Practice and Prepayment of Medical Care*, Public Affairs Press, Washington, 1966.

[2]Sidney R. Garfield, M.D., "The Delivery of Medical Care," *Scientific American*, 222, No. 4, April 1970, 15-23.

[3]Ehrlich and Becker, "Market Insurance, Self-Insurance, and Self-Protection," pp. 627-628.

To describe the competitive "optimal" insurance policy, one would need a theory of loading fees. I do not have such a theory and know of no attempts to develop one. If loading fees decrease as states of the world become more hazardous, it is clear that no upper bound would exist on the amount of insurance. Alternatively, if loadings increase as states of the world become more hazardous, then there would be some upper bound on insured states of the world. I infer that loadings *do* increase for more hazardous states, since I do not observe any unlimited health insurance coverage.[1]

Insurance, as set forth in the Ehrlich-Becker paper, guarantees outcomes (health) rather than inputs (expenditures for medical care). When the loss one faces is from a stock of health, it may be exceedingly difficult to insure outcomes, because of the random component in production of health in the medical sector. The costs to insurers of guaranteeing health levels would seem prohibitive, in the face of this uncertain production process. The implicit loading fees on serious illnesses intuitively may rise as severity increases. I do not expect to observe "full coverage" health insurance until medical technology itself becomes a certain process.

The concept of insuring outcomes also suggests what form "optimal" insurance may take. If medical production processes are uncertain, then it is clear that cash payments cannot produce uniform outcomes. Reimbursement insurance might produce such outcomes, if properly structured, and if medical care were supplied in a competitive market. An alternative way of guaranteeing optimal use of medical technology in production and "insuring" of health levels is to combine the functions of insurer and medical care provider. "Insuring" organizations such as this, if competitive, would offer insurance of outcomes, to the extent permitted by the state of medical technology and consumer information levels. Implicit loading fees on various states of the world would all be competitive. Production of health outcomes would be

---

[1]In 1963, the largest maximum payments observed in health insurance were around $40,000. In 1970, major medical policies with $100,000 maximum payments are not uncommon; recently one company offered a major medical policy with a $250,000 maximum payment.

carried out at minimum cost. Such organizations are known as prepaid group practice organizations. If they operate in a competitive environment, they should provide "ideal" insurance against health losses.

Similarly, there is every reason to believe that the same sorts of outcomes could be produced at minimum cost with reimbursement insurance, if insurance fully reflected all probabilities, and if physicians were competitive. Failure on either count would make this arrangement inferior to prepayment systems.

Any proposals for national health insurance schemes that include prepayment and group practice organizations should recognize the necessity of competition in these organizations. If prepayment organizations are monopolists, they should behave the same as any other monopolist; higher expenditures for inferior outcomes of health would be certain. A health industry so organized could well produce protection against health losses that was inferior to the present market systems.

Appendix A

COMPARATIVE STATICS OF DEMAND FOR MEDICAL CARE

	When Loss is known and h is Less Than h* Slope[a]	Partial Elasticity	Comments
1.	$\frac{\partial x}{\partial I} = \left(1 - \frac{\partial R}{\partial I}\right)\left[\frac{NPH \cdot U_{xH}g'(h) - p_x U_{hh}}{\lvert M\rvert}\right]$	η_{xI}	Assumed to be positive
2.	$\frac{\partial h}{\partial I} = \left(1 - \frac{\partial R}{\partial I}\right)\left[\frac{p_x U_{xH}g'(h) - NPH \cdot U_{xx}}{\lvert M\rvert}\right]$	η_{hI}	Assumed to be positive
3.	$\frac{\partial \lambda}{\partial I} = \left(1 - \frac{\partial R}{\partial I}\right)\left[\frac{U_{xx}U_{hh} - (U_{xH}g'(h))^2}{\lvert M\rvert}\right]$	$\eta_{\lambda I}$	Equal to $-r^*(I)$
4.	$\frac{\partial x}{\partial TP} = -h\,\frac{\partial x}{\partial I} + \lambda p_x \frac{NPH}{\lvert M\rvert}$	$\eta_{xTP} = \eta_x^T$	TP - NPH = "total price" facing consumer (no premium effects)
5.	$\frac{\partial h}{\partial TP} = -h\,\frac{\partial h}{\partial I} + \frac{\lambda p_x^2}{\lvert M\rvert}$	$\eta_{hTP} = \eta_h^T$	No premium effects
6.	$\frac{\partial \lambda}{\partial TP} = -h\,\frac{\partial \lambda}{\partial I} - \lambda\,\frac{\partial h}{\partial I}$	$\eta_{\lambda TP} = \eta_\lambda^T$	$= h(-\lambda)\left[r^*(I) - \eta_{hI}\right]$
7.	$\frac{\partial x}{\partial p_h} = \left(-Ch - \frac{\partial R}{\partial p_h}\right)\frac{\partial x}{\partial I} - \frac{C\lambda p_x NPH}{\lvert M\rvert}$	η_{xp_h}	$= \frac{Cp_h}{NPH}\,\eta_x^T - \frac{R}{I} \cdot \eta_{xI} \cdot \eta_{Rp_h}$
8.	$\frac{\partial h}{\partial p_h} = C\left(-ch - \frac{\partial R}{\partial p_h}\right)\frac{\partial h}{\partial I} + \frac{C\lambda p_x^2}{\lvert M\rvert}$	$\eta_{hp_h} = \eta_{hh}$	$= \frac{Cp_h}{NPH}\,\eta_h^T - \frac{R}{I}\,\eta_{hI}\,\eta_{Rp_h}$
9.	$\frac{\partial \lambda}{\partial p_h} = Ch\,\frac{\partial \lambda}{\partial I} - C\lambda\,\frac{\partial h}{\partial I} - \frac{\partial R}{\partial p_h}\,\frac{\partial \lambda}{\partial I}$	$\eta_{\lambda p_h}$	$= C\eta_{\lambda TP} + \frac{R}{I}\,r^*(I)\,\eta_{Rp_h}$

	When Loss is known and h is Less Than h* Slope[a]	Partial Elasticity	Comments
10.	$\frac{\partial x}{\partial w} = \left(-th - \frac{\partial R}{\partial w}\right)\frac{\partial x}{\partial I} - \frac{\lambda t p_x NPH}{\lvert M\rvert}$	η_{xw}	$= \frac{w \cdot t}{NPH}\,\eta_x^T - \frac{R}{I}\,\eta_{xI}\,\eta_{Rw}$
11.	$\frac{\partial h}{\partial w} = \left(-th - \frac{\partial R}{\partial w}\right)\frac{\partial h}{\partial I} + \frac{\lambda t p_x^2}{\lvert M\rvert}$	η_{hw}	$= \frac{w \cdot t}{NPH}\,\eta_h^T - \frac{R}{I}\,\eta_{hI}\,\eta_{Rw}$
12.	$\frac{\partial \lambda}{\partial w} = \left(-th - \frac{\partial R}{\partial w}\right)\frac{\partial \lambda}{\partial I} - t\lambda\,\frac{\partial h}{\partial I}$	$\eta_{\lambda w}$	$= \frac{w \cdot t}{NPH}\,\eta_\lambda^T + \frac{R}{I}\,r^*(I)\,\eta_{Rw}$
13.	$\frac{\partial x}{\partial t} = \left(-wh - \frac{\partial R}{\partial t}\right)\frac{\partial x}{\partial I} - \frac{\lambda w p_x NPH}{\lvert M\rvert}$	η_{xt}	$= \frac{w \cdot t}{NPH}\,\eta_x^T - \frac{R}{I}\,\eta_{xI}\,\eta_{Rt}$
14.	$\frac{\partial h}{\partial t} = \left(-wh - \frac{\partial R}{\partial t}\right)\frac{\partial h}{\partial I} + \frac{\lambda w p_x^2}{\lvert M\rvert}$	η_{ht}	$= \frac{w \cdot t}{NPH}\,\eta_h^T - \frac{R}{I}\,\eta_{hI}\,\eta_{Rt}$
15.	$\frac{\partial \lambda}{\partial t} = \left(-wh - \frac{\partial R}{\partial t}\right)\frac{\partial \lambda}{\partial I} - w\lambda\,\frac{\partial h}{\partial I}$	$\eta_{\lambda t}$	$= \frac{w \cdot t}{NPH}\,\eta_\lambda^T + \frac{R}{I}\,\eta_{Rt}\,r^*(I)$
16.	$\frac{\partial x}{\partial \ell} = \frac{NPH}{g'(h)}\,\frac{\partial x}{\partial I}$	$\eta_{x\ell}$	$= -NPH\,\eta_{xI}\,\frac{\ell}{g'(h)I}$
17.	$\frac{\partial h}{\partial \ell} = \frac{P_x}{g'(h)}\,\frac{\partial x}{\partial I}$	$\eta_{h\ell}$	$= \omega_x\,\eta_{xI}\,\frac{\ell}{g'(h) \cdot h}$
18.	$\frac{\partial \lambda}{\partial \ell} = -\frac{NPH}{g'(h)}\,\frac{\partial x}{\partial I}$	$\eta_{\lambda\ell}$	$= r^*(I)\,\frac{NPH \cdot \ell}{g'(h) \cdot I}$
19.	$\frac{\partial x}{\partial H_o} = \frac{NPH}{g'(h)}\,\frac{\partial x}{\partial I} - \frac{\partial R}{\partial H_o}\,\frac{\partial x}{\partial I}$	η_{xH_o}	$= -\eta_{x\ell} - \frac{R}{I}\,\eta_{xI}\,\eta_{RH_o}$

	When Loss is known and h is Less Than h* Slope[a]	Partial Elasticity	Comments
20.	$\frac{\partial h}{\partial H_o} = - \frac{p_x}{g'(h)} \frac{\partial x}{\partial I} - \frac{\partial R}{\partial H_o} \frac{\partial h}{\partial I}$	η_{hH_o}	$= - \eta_{h\ell} - \frac{R}{I} \eta_{hI} \eta_{RH_o}$
21.	$\frac{\partial \lambda}{\partial H_o} = \frac{\partial \lambda}{\partial I} \left(\frac{NPH}{g'(h)} - \frac{\partial R}{\partial H_o} \right)$	$\eta_{\lambda H_o}$	$= - \eta_{\lambda \ell} + \frac{R}{I} r^*(I) \eta_{RH_o}$
22.	$\frac{\partial x}{\partial C} = (-R_C - p_h h) \frac{\partial x}{\partial I} - \frac{\lambda p_h p_x NPH}{\lvert M \rvert}$	η_{xC}	$= \frac{Cp_h}{NPH} \eta_x^T - \frac{CR_c}{I} \eta_{xI}$
23.	$\frac{\partial h}{\partial C} = (-R_C - p_h h) \frac{\partial h}{\partial I} + \frac{\lambda p_x^2 p_h}{\lvert M \rvert}$	η_{hC}	$= \frac{Cp_h}{NPH} \eta_h^T - \frac{CR_c}{I} \eta_{hI}$
24.	$\frac{\partial \lambda}{\partial C} = (-R_C - p_h h) \frac{\partial \lambda}{\partial I} - \lambda p_h \frac{\partial h}{\partial I}$	$\eta_{\lambda C}$	$= \frac{Cp_h}{NPH} \eta_\lambda^I + r^*(I) \frac{CR_c}{I}$
25.	$\frac{\partial x}{\partial h^*} = -R_{h^*} \frac{\partial x}{\partial I}$	η_{xh^*}	$= - \frac{R}{I} \eta_{xI} \eta_{Rh^*}$
26.	$\frac{\partial h}{\partial h^*} = -R_{h^*} \frac{\partial h}{\partial I}$	η_{hh^*}	$= - \frac{R}{I} \eta_{hI} \eta_{Rh^*}$
27.	$\frac{\partial \lambda}{\partial h^*} = -R_{h^*} \frac{\partial \lambda}{\partial I}$	$\eta_{\lambda h^*}$	$= \frac{R}{I} \eta_{Rh^*} r^*(I)$
28.	$\frac{\partial x}{\partial \theta} = -R_\theta \frac{\partial x}{\partial I}$	$\eta_{x\theta}$	$= - \frac{R}{I} \eta_{xI} \eta_{R\theta}$
29.	$\frac{\partial h}{\partial \theta} = -R_\theta \frac{\partial h}{\partial I}$	$\eta_{h\theta}$	$= - \frac{R}{I} \eta_{hI} \eta_{R\theta}$
30.	$\frac{\partial \lambda}{\partial \theta} = -R_\theta \frac{\partial \lambda}{\partial I}$	$\eta_{\lambda\theta}$	$= \frac{R}{I} r^*(I) \eta_{R\theta}$

	When Loss is Known and h exceeds h* Slope[a]	Partial Elasticity	Comments
31.	$\frac{\partial x}{\partial p_h} = \left[-h+(1-C)h^* - \frac{\partial R}{\partial p_h}\right]\frac{\partial x}{\partial I} + \frac{\lambda p_x^2}{\lvert M'\rvert}$	η_{xp_h}	Income effect negative since h > (1-C)h*
32.	$\frac{\partial h}{\partial p_h} = \left[-h+(1-C)h^* - \frac{\partial R}{\partial p_h}\right]\frac{\partial h}{\partial I} + \frac{\lambda p_x NPH'}{\lvert M'\rvert}$	η_{hp_h}	Income effect and substitution effect negative
33.	$\frac{\partial \lambda}{\partial p_h} = \left[-h+(1-C)h^* - \frac{\partial R}{\partial p_h}\right]\frac{\partial \lambda}{\partial I} - \lambda \frac{\partial h}{\partial I}$	η_{xp_h}	$\gtrless 0$
34.	$\frac{\partial x}{\partial C} = (-R_c - p_h h^*) \frac{\partial x}{\partial I}$	η_{xc}	Income effect only (always negative)
35.	$\frac{\partial h}{\partial C} \quad (-R_c - p_h h^*) \frac{\partial h}{\partial I}$	η_{hc}	Income effect only (always negative)
36.	$\frac{\partial \lambda}{\partial C} \quad (-R_c - p_h h^*) \frac{\partial \lambda}{\partial I}$	$\eta_{\lambda c}$	Income effect only (always negative)
37.	$\frac{\partial x}{\partial h^*} = (-R_{h^*} + (1-C)p_h) \frac{\partial x}{\partial I}$	η_{xh^*}	Income effect always positive
38.	$\frac{\partial h}{\partial h^*} = (-R_{h^*} + (1-C)p_h) \frac{\partial h}{\partial I}$	η_{hh^*}	Income effect always positive
39.	$\frac{\partial \lambda}{\partial h^*} = (-R_{h^*} + (1-C)p_h) \frac{\partial \lambda}{\partial I}$	$\eta_{\lambda h^*}$	Income effect always positive

[a]In this table, $(-\lambda) = \frac{\partial U}{\partial I}$, $NPH = TP = (Cp_h + w \cdot t)$, $NPH' = (p_h + w \cdot t)$, $\lvert M\rvert = -p_x^2 U_{hh} - (NPH)^2 U_{xx} + 2p_x NPH\, U_{xH} g'(h)$ and $\lvert M'\rvert = -p_x^2 U_{hh} - p_h^2 U_{xx} + 2p_x NPH' U_{xH} g'(h)$. The derivations are available in Charles E. Phelps, "Demand for Health Insurance: A Theoretical and Empirical Investigation," Ph.D. dissertation, University of Chicago, June 1973.

Appendix B

THE PREMIUM FUNCTION R AND ITS ASSOCIATED DERIVATIVES

Let the total insurance premium be defined by

$$R = (1 + \theta)(1 - C) \int_{o}^{\ell^*} p_h h f(\ell) d\ell, \tag{B.1}$$

where θ is a constant "loading fee," C is the percentage of all payments for medical care made by the consumer, and $f(\ell)$ is a known illness distribution facing the consumer.

The derivatives $(-R_c)$ and R_{h*} are:

$$-R_c = \frac{-\partial}{\partial C}(R) = (1 + \theta) \int_{o}^{\ell^*} \left[p_h h - (1 - C) \frac{\partial h}{p_h \partial C} \right] f(\ell) d\ell. \tag{B.2a}$$

From Appendix A, it is known that

$$(1 - C)p_h \frac{\partial h}{\partial C} = (1 - C)p_h\left(\frac{\partial h}{\partial C} \cdot \frac{C}{h}\right)\frac{h}{C} = \left\{(-R_c) \frac{\partial h}{\partial I} + h \frac{p_h}{Cp_h + w \cdot t} \eta_h^T\right\}(1 - C)p_h,$$

where η_h^T is the elasticity with respect to total price $(Cp_h + w \cdot t)$. The ratio $k = \dfrac{p_h}{Cp_h + w \cdot t}$ may be greater or less than unity, depending on the values of C and $w \cdot t$. We now have:

$$(1 - C)p_h \frac{\partial h}{\partial C} = \left\{(-R_c) \frac{\partial h}{\partial I} + (h)k\eta_h^T\right\}(1 - C)p_h.$$

Collecting terms from (B.2a) involving $(-R_c)$, we have

$$\text{(B.2b)} \qquad -R_c = \frac{(1+\theta)\int_o^{\ell^*} p_h h(1-(1-C)k\eta_h^T)f(\ell)d\ell}{1+(1+\theta)(1-C)\int_o^{\ell^*} p_h \frac{\partial h}{\partial I} f(\ell)d\ell} \cdot$$

$$= \frac{(1+\theta\int_o^{\ell^*} p_h h(1-(1-C)k\eta_h^T)f(\ell)d\ell}{1+(1+\theta)(1-C)\overline{\omega_h\eta_{hI}}} > 0 .$$

Note that in the expression for $(-R_c)$, the integral over $p_h h$ is multiplied by $(1-(1-C)k\eta_h^T)$, rather than the traditional expression involving price changes of $(1+\eta_h^T)$. Since η_h^T is negative, $(-R_c$ will in general be larger than $(1-C)p_h\overline{h}$. In effect, $(-R_c)$ is reflecting the dual increase in R as C decreases; more of "the pie" is being covered by the insurance company, and the size of "the pie" is increasing because of the subsidy effect of C on demand for h. A percentage change in C will alter $(-R_c)$ in the following fashion:

$$\text{(B.2c)} \qquad -\frac{\partial R}{(\partial C/C)} = \frac{(1+\theta)\int_o^{\ell^*} p_h h\left(C-(1-C)\frac{Cp_h}{NPM}\eta_h^T\right)f(\ell)d\ell}{1+(1+\theta)(1-C)\overline{\omega_h\eta_{hI}}} .$$

If η_h^T is approximately constant over losses, then the elasticity of R with respect to C can be written (assuming $\overline{\omega_h\eta_{hI}} \approx 0$) as:

$$\text{(B.2d)} \qquad -\frac{\partial R}{\partial C}\cdot\frac{C}{R} \approx \frac{\left[(1+\theta)\int_o^{\ell^*} p_h h f(\ell)d\ell\right]\left(C-(1-C)\frac{Cp_h}{NPM}\eta_h^T\right)}{(1-C)(1+\theta)\int_o^{\ell^*} p_h h f(\ell)d\ell} \cdot$$

$$= \frac{C-(1-C)\eta_h^T\left(\frac{Cp_h}{Cp_h+w\cdot t}\right)}{(1-C)} \approx \frac{C-(1-C)\eta_{hh}}{(1-C)} ,$$

where η_{hh} = the observed own-price elasticity of demand for h. Thus, the average own-price elasticity of demand for C may be calculated

directly from observed premium shifts (as C changes) if the insurance is priced on expected costs.

The derivative R_{h*} is:

(B.3a)

$$R_{h*} = \frac{\partial}{\partial h*}\left[(1+\theta)(1-C)\int_{o}^{\ell*} p_h h f(\ell)d\ell\right]$$

$$= (1+\theta)(1-C)\int_{o}^{\ell*}\left[p_h \frac{\partial h}{\partial h*}\right]f(\ell)d\ell + (1+\theta)(1-C)p_h h*g'(h*)f(\ell*)$$

$$= (1+\theta)(1-C)\int_{o}^{\ell*}\left(p_h(-R_{h*})\frac{\partial h}{\partial I}\right)f(\ell)d\ell + (1+\theta)(1-C)p_h h*\gamma(h*),$$

where $\gamma(h*) \equiv g'(h*)f(\ell*)$.

Collecting terms involving R_{h*},

(B.3b)

$$R_{h*} = \frac{(1+\theta)(1-C)p_h h*\gamma(h*)}{1+(1+\theta)(1-C)\int_{o}^{\ell*}\left[p_h \frac{\partial h}{\partial I}\right]f(\ell)d\ell}$$

$$= \frac{(1+\theta)(1-C)p_h h*\gamma(h*)}{1+(1-C)(1+\theta)\overline{\omega_h \eta_{hI}}} > 0.$$

The partial derivative with respect to p_h (holding C and h* constant) is:

(B.4a)

$$R_{p_h} = \frac{\partial R}{\partial p_h} = (1+\theta)(1-C)\int_{o}^{\ell*}\left[h + p_h \frac{\partial h}{\partial p_h}\right]f(\ell)d\ell$$

$$= (1+\theta)(1-C)\int_{o}^{\ell*}\left(h + p_h\left(C \cdot \frac{\partial h}{\partial TP} - \frac{\partial R}{\partial p_h}\frac{\partial h}{\partial I}\right)f(\ell)d\ell.\right.$$

Collecting terms involving $\partial R/\partial p_h$, and converting to elasticities (using terms from Appendix A),

$$\text{(B.4b)} \qquad \frac{\partial R}{\partial p_h} = \frac{(1+\theta)(1-C)\int_o^{\ell^*} h\left(1 + \frac{Cp_h}{Cp_h + w \cdot t}\,\eta_h^T\right) f(\ell)d\ell}{1 + (1+\theta)(1-C)\int_o^{\ell^*} p_h \frac{\partial h}{\partial I} f(\ell)d\ell}$$

$$= \frac{(1+\theta)(1-C)\overline{h(1+k'\eta_h^T)}}{1 + (1-C)(1+\theta)\overline{\omega_h\eta_{hI}}} \gtrless 0 \quad \text{as} \quad k'\eta_h^T \gtrless (-1).$$

Since $k'\eta_h^T \approx \eta_{hh}$ (the observed own-price elasticity of demand for h), I can show that $\frac{\partial R}{\partial p_h} > 0$ if demand for h, "on average," is inelastic.

The partial derivative $\frac{\partial R}{\partial I}$ (holding C, h* constant) is:

$$\text{(B.5a)} \qquad \frac{\partial R}{\partial I} = \frac{\partial}{\partial I}\left[(1-C)(1+\theta)\int_o^{\ell^*} p_h h f(\ell)d\ell\right]$$

$$= (1-C)(1+\theta)\int_o^{\ell^*}\left(p_h \frac{\partial h}{\partial I}\right)\left(1 - \frac{\partial R}{\partial I}\right) f(\ell)d\ell,$$

using terms from Appendix A.

Collecting terms,

$$\text{(B.5b)} \qquad \frac{\partial R}{\partial I} = \frac{(1-C)(1+\theta)\int_o^{\ell^*} p_h \frac{\partial h}{\partial I} f(\ell)d\ell}{1 + (1-C)(1+\theta)\int_o^{\ell^*} p_h \frac{\partial h}{\partial I} f(\ell)d\ell}$$

$$\equiv \frac{(1-C)(1+\theta)\overline{\omega_h\eta_{hI}}}{1 + (1-C)(1+\theta)\overline{\omega_h\eta_{hI}}} > 0.$$

The partial derivative $\frac{\partial R}{\partial \theta}$ (holding C, h* constant) is:

$$\frac{\partial R}{\partial \theta} = \frac{\partial}{\partial \theta}\left[(1 - C)(1 + \theta)\int_{o}^{\ell^*} p_h h f(\ell)d\ell\right]$$

(B.6a)

$$= (1 - C)(1 + \theta)\int_{o}^{\ell^*} p_h \frac{\partial h}{\partial I} (-R_\theta) f(\ell)d\ell + (1 - C)\int_{o}^{\ell^*} p_h h f(\ell)d\ell .$$

Collecting terms involving R_θ,

(B.6b) $$\frac{\partial R}{\partial \theta} = \frac{(1 - C)\int_{o}^{\ell^*} p_h h f(\ell)d\ell}{1 + (1 - C)(1 + \theta)\int_{o}^{\ell^*} p_h \frac{\partial h}{\partial I} f(\ell)d\ell}$$

$$= \frac{(1 - C)p_h\overline{h}}{1 + (1 - C)(1 + \theta)\overline{\omega_h \eta_{hI}}} > 0 .$$

A change in w acts much as a change in p_h, as seen here. Again holding C and h* constant (terms from Appendix A):

(B.7a) $$\frac{\partial R}{\partial w} = \frac{\partial}{\partial w}\left((1 - C)(1 + \theta)\int_{o}^{\ell^*} p_h h f(\ell)d\ell\right)$$

$$= (1 - C)(1 + \theta)\int_{o}^{\ell^*} p_h\left[t\frac{\partial h}{\partial TP} - \frac{\partial R}{\partial w}\frac{\partial h}{\partial I}\right] f(\ell)d\ell .$$

Collecting terms involving $\frac{\partial R}{\partial w}$, and converting to elasticities,

(B.7b) $$\frac{\partial R}{\partial w} = \frac{(1 - C)(1 + \theta)\int_{o}^{\ell^*} \frac{t}{Cp_h + w \cdot t}(p_h h)\eta_h^T f(\ell)d\ell}{1 + (1 + \theta)(1 - C)\int_{o}^{\ell^*} p_h \frac{\partial h}{\partial I} f(\ell)d\ell} ,$$

and

$$\text{(B.7c)} \qquad \frac{\partial R}{\partial(\%w)} = \frac{\partial R}{(\partial w/w)} \frac{(1-C)(1+\theta)\left(\frac{w \cdot t}{Cp_h + w \cdot t}\right)\overline{p_h h \eta_h^T}}{1 + (1+\theta)(1-C)\overline{\omega_h \eta_{hI}}} < 0 .$$

Without completing the proof, I note that

$$\text{(B.8)} \qquad \frac{\partial R}{\partial(\%t)} \approx \frac{\partial R}{(\partial t/t)} \approx \frac{\partial R}{(\partial w/w)} .$$

In other words, the premium R is always reduced as w or t increases, since there is an own-price effect on demand for h (given any loss), but the insurance company never partakes of any of the time costs of medical care.

Total Derivatives of R

Regressions using total premiums as the dependent variable may be used to infer effects of explanatory variables on C, using the following relationships.

(B.9) Let $R = (1 - C)(1 + \theta)\int_o^{\ell^*} p_h h f(\ell) d\ell$. Then

$$\eta_{RI} = \frac{dR}{dI} \cdot \frac{I}{R} = \frac{I}{R}\left[(1+\theta)(1-C)\int_o^{\ell^*} p_h \frac{dh}{dI} f(\ell) d\ell \right.$$

$$\left. - \frac{\partial C}{\partial I}(1+\theta)\int_o^{\ell^*} p_h h f(\ell) d\ell + (1-C)(1+\theta)\int_o^{\ell^*} p_h h \frac{\partial h}{\partial C} \cdot \frac{\partial C}{\partial I}\right]$$

$$= \frac{(1+\theta)(1-C)\int_o^{\ell^*} p_h h}{R} \eta_{hI} f(\ell) d\ell - \frac{C}{R}(1+\theta)\int_o^{\ell^*} p_h h f(\ell) d\ell \eta_{CI}$$

$$+ \frac{(1-C)(1+\theta)\int_o^{\ell^*} p_h h}{R} \eta_{hC}\eta_{CI} .$$

Assuming that η_{hC} and η_{hI} are "approximately" constant over all ℓ, then

(B.10a) $$\eta_{RI} \approx \eta_{hI} - \frac{C}{1 - C}\eta_{CI} + \eta_{hC}\eta_{CI} ,$$

and

(B.10b) $$\eta_{CI} \approx \frac{\eta_{RI} - \eta_{hI}}{-\frac{C}{1 - C} + \eta_{hC}} \lessgtr 0 \quad \text{as} \quad \eta_{RI} \gtrless \eta_{hI} .$$

If we assume $\eta_{hC} = -.20$, $\eta_{hI} = .25$, and $C = .25$, then an observed $\eta_{RI} = .30$ implies $\eta_{CI} = \frac{.05}{-.53} = -.095$. Defining coverage (K) as $K = 1 - C$, then $\partial K/\partial I \cdot I/K = -(1 - C)/C\ (\eta_{CI}) = -3(-.095) = .285$.

The effects of a change in p_h are similar:

(B.11) $$\eta_{Rp_h} = \frac{p_h}{R}\Bigg[(1 + \theta)(1 - C)\int_0^{\ell^*} h(1 + \eta_{hh})f(\ell)d\ell - (1 + \theta)\frac{\partial C}{\partial p_h}\int_0^{\ell^*} p_h h f(\ell)d\ell + (1 - C)(1 + \theta)\int_0^{\ell^*} p_h \frac{\partial h}{\partial C}\cdot\frac{\partial C}{\partial p_h} f(\ell)d\ell\Bigg].$$

Again, assuming η_{hh} and η_{hC} are constant over all losses, then

(B.12a) $$\eta_{Rp_h} = (1 + \eta_{hh}) - \frac{C}{1 - C}\eta_{C,p_h} + \eta_{hC}\eta_{C,p_h} ,$$

and

(B.12b) $$\eta_{Cp_h} = \frac{\eta_{R,p_h} - (1 + \eta_{hh})}{-\frac{C}{1 - C} + \eta_{h,C}} .$$

If I assume $C = .25$, and $\eta_{hh} = \eta_{hC} = -.20$, then an observed $\eta_{Rp_h} = 0$ implies $\eta_{Cp_h} = -.8/-.53 = 1.51$ and $\eta_{Kp_h} = -4.53$.

A change in w or t produces the same type of result:

$$\text{(B.13)}\quad \frac{dR}{dw}\cdot\frac{w}{R}=\frac{w}{R}\left[(1+\theta)(1-C)\int_{o}^{\ell^*} p_h\,\frac{\partial h}{\partial w}\,f(\ell)d\ell-(1+\theta)\,\frac{\partial C}{\partial w}\int_{o}^{\ell^*} p_h h f(\ell)d\ell + (1-C)(1+\theta)\int_{o}^{\ell^*} p_h\,\frac{\partial h}{\partial w}\cdot\frac{\partial w}{\partial C}\right].$$

Making similar assumptions about η_{hC} and η_{hw}, this becomes

$$\text{(B.14a)}\qquad \eta_{Rw}=\eta_{hw}-\frac{C}{1-C}\,\eta_{Cw}+\eta_{hC}\eta_{Cw}\ ,$$

and

$$\text{(B.14b)}\qquad \eta_{Cw}=\frac{\eta_{Rw}-\eta_{hw}}{\frac{-C}{1-C}+\eta_{hC}}\ .$$

Assuming $C = .25$, $\eta_{hC} = -.20$, $\eta_{hw} = -.40$, then an observed $\eta_{Rw} = .30$ implies

$$\eta_{Cw}=\frac{.70}{-.53}=-1.31,$$

and

$$\eta_{Kw}=(-3)\ \eta_{Cw}=3.96.$$

Finally, a change in θ results in the following:

$$\frac{dR}{d\theta}=(1-C)\int_{o}^{\ell^*} p_h h f(\ell)d\ell-(1+\theta)\,\frac{\partial C}{\partial\theta}\int_{o}^{\ell^*} p_h h f(\ell)d\ell$$

$$\text{(B.15)}\qquad +(1+\theta)(1-C)\int_{o}^{\ell^*} p_h\,\frac{\partial h}{\partial\theta}\,f(\ell)d\ell$$

$$+(1+\theta)(1-C)\int_{o}^{\ell^*} p_h\,\frac{\partial h}{\partial C}\cdot\frac{\partial C}{\partial\theta}\,f(\ell)d\ell$$

where $\frac{dh}{d\theta} = -\frac{dR}{d\theta} \cdot \frac{d\theta}{\partial I}$

Collecting terms and converting to elasticities (again assuming η_{hC}, and η_{hI} are constant over all losses, then

$$\text{(B.16a)}\quad \eta_{R\theta} = \frac{dR}{d\theta} \cdot \frac{\theta}{R} = \left[\frac{\theta}{1+\theta} - \frac{C}{1-C}\,\eta_{C\theta} + \eta_{hC}\eta_{C\theta}\right] \frac{1}{1+\bar{\omega}_h\eta_{hI}},$$

and

$$\text{(B.16b)}\quad \eta_{C\theta} = -\frac{(1+\bar{\omega}_h\eta_{hI})\eta_{R\theta} - \frac{\theta}{1+\theta}}{-\frac{C}{1-C} + \eta_{hC}}.$$

Let $1 + \bar{\omega}\eta_{hI} = 1.02$ ($\eta_{hI} = .25$, $\bar{\omega} = .08$), $C = .25$, $\eta_{hC} = -.20$, and $\theta = .20$. One might be tempted to report that demand for coverage was "elastic" if $\eta_{R\theta}$ were negative ("expenditure increases as price decreases").

Given these data, an elasticity of premiums with respect to theta of -.30 means that

$$\eta_{C\theta} = \frac{(1.02)(-.30) - \frac{.2}{1.2}}{\frac{-.25}{.75} - .20} = \frac{-.46}{-.53} \approx .87,$$

and

$$\eta_{K\theta} = -\frac{(1-C)}{C}\,\eta_{C\theta} \approx -2.60,$$

which is certainly "elastic." However, the same value for $\eta_{R\theta} = -.30$ has entirely different meaning at (say) $\theta = .10$ and $C = .5$. Then

$$\eta_{C\theta} = \frac{(1.02)(-.30) - \frac{1}{1.1}}{-1. - .2} \approx \frac{-.40}{-1.2} = .33,$$

and

$$\eta_{K\theta} = \frac{-.5}{.5}\,(.33) = -.33,$$

which is "inelastic."

Appendix C

SECOND-ORDER CONDITIONS OF THE INSURANCE PURCHASE DECISION

The second-order conditions for a maximum in expected utility with respect to C and h* are:

$$\text{(C.1)} \qquad a_{11} = \frac{\partial^2 EU}{\partial C^2} < 0 \qquad\qquad a_{22} = \frac{\partial^2 EU}{\partial h^{*2}} < 0,$$

and

$$\text{(C.2)} \qquad |D| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0,$$

where $a_{12} = \partial^2 EU/\partial C \partial h^*$.

The derivative a_{11} is given as:

$$\text{(C.3)} \qquad \frac{\partial^2 EU}{\partial C^2} = s_o \left\{ (-\lambda)\left(\frac{-\partial^2 R}{\partial C^2} - p_h \frac{\partial h}{\partial C}\right) + (-R_C - p_h h)\left(\frac{-\partial \lambda}{\partial C}\right) \right\}$$

$$+ \int_o^{\ell^*} \left\{ (-\lambda)\left(\frac{-\partial^2 R}{\partial C^2} - p_h \frac{\partial h}{\partial C}\right) + (-R_C - p_h h)\left(\frac{-\partial \lambda}{\partial C}\right) \right\} f(\ell)\, d\ell$$

$$+ \int_{\ell^*}^{\infty} \left\{ (-\lambda)\left(\frac{-\partial^2 R}{\partial C^2}\right) + (-R_c - p_h h^*)\left(\frac{-\partial \lambda}{\partial C}\right) \right\} f(\ell)\, d\ell \ .$$

Making the appropriate substitutions for $\partial h/\partial C$ and $\partial \lambda/\partial C$, this becomes:

$$\text{(C.4)}\quad \frac{\partial^2 EU}{\partial C^2} = s_o\left\{(-R_c-p_h h)^2\left(\frac{-\partial\lambda}{\partial I}\right) + 2p_h\lambda\,\frac{\partial h}{\partial I}(-R_C-p_h h) + \left(\frac{\lambda p_h p_x}{|M|}\right)^2\right\}$$

$$+\int_o^{\ell^*}\left\{(-R_C-p_h h)^2\left(\frac{-\partial\lambda}{\partial I}\right) + 2\lambda p_h\,\frac{\partial h}{\partial I}(-R_C-p_h h) + \left(\frac{\lambda p_h p_x}{|M|}\right)^2\right\}f(\ell)d\ell$$

$$+\int_{\ell^*}^{\infty}(-R_C-p_h h^*)\left(\frac{-\partial\lambda}{\partial I}\right)f(\ell)d\ell + E(-\lambda)(-\partial^2 R/\partial C^2)\ .$$

The derivative a_{22} is found by differentiating the first-order condition with respect to h*:

$$\text{(C.5)}\quad \frac{\partial^2 EU}{\partial h^{*2}} = s_o(-R_{h^*})\left(\frac{-\partial\lambda_o}{\partial h^*}\right) + (-\lambda_o)\left(\frac{-\partial^2 R}{\partial h^{*2}}\right)$$

$$+\int_o^{\ell^*}\left\{(-R_{h^*})\left(\frac{-\partial\lambda}{dh^*}\right) + (-\lambda)\left(\frac{\partial^2 R}{\partial h^{*2}}\right)\right\}f(\ell)d\ell$$

$$+\int_{\ell^*}^{\infty}\left\{(-R_{h^*} + (1-C)p_h)\left(\frac{-\partial\lambda}{\partial h^*}\right) + (-\lambda)\left(\frac{-\partial^2 R}{\partial h^{*2}}\right)\right\}f(\ell)d\ell$$

$$-\ (1-C)p_h f(\ell^*)g'(h^*)\ ,$$

where the last term arises from the terms involving limits of integration from the Leibnitz equation for differentiation of integrals. Substituting for $(-\partial\lambda/\partial h^*)$ this becomes:

$$\text{(C.6)}\quad \frac{\partial^2 EU}{\partial h^{*2}} = s_o(-R_{h^*})^2\left(\frac{\partial\lambda}{\partial I}\right) + \int_o^{\ell^*}\left\{(-R_{h^*})^2\left(\frac{\partial\lambda}{\partial I}\right)\right\}f(\ell)d\ell$$

$$+\int_{\ell^*}^{\infty}\left\{(-R_{h^*} + (1-C)p_h)^2\ \frac{\partial\lambda}{\partial I}\right\}f(\ell)d\ell$$

$$-\ E(\lambda)\left(\frac{\partial^2 R}{\partial h^{*2}}\right) - (1-C)p_h f(\ell^*)g'(h^*)\ .$$

This expression is always negative as long as $\partial^2 R/\partial h^{*2}$ is non-negative, and $g'(h)$ is positive when evaluated at h^*.

The cross-product derivative $a_{12} = a_{21} = (\partial^2 EU)/(\partial C \partial h^*)$ is found either by differentiation of the first-order conditions with respect to h^* or to C. (The results are identical under each solution.) That expression is given by:

$$\text{(C.7)}\quad \frac{\partial^2 EU}{\partial C \partial h^*} = a_{12} = a_{21} = s_o(-R_c - p_h h^*)\left(\frac{-\partial\lambda}{\partial h^*}\right) + \lambda\left(p_h \frac{\partial h}{\partial h^*} + \partial^2 R/\partial C \partial h^*\right)$$

$$+ \int_o^{\ell^*} \left\{(-R_C - p_h h)\left(\frac{-\partial\lambda}{\partial h^*}\right) + \lambda\left[p_h \frac{\partial h}{\partial h^*} \frac{+\partial^2 R}{\partial C \partial h^*}\right]\right\} f(\ell)d\ell$$

$$+ \int_{\ell^*}^{\infty} \left\{(-R_c - p_h h^*)\left(\frac{-\partial\lambda}{\partial h^*}\right) + \lambda\left(p_h \frac{+\partial^2 R}{\partial C \partial h^*}\right)\right\} f(\ell)d\ell\ .$$

Making the appropriate substitutions this becomes:

$$\text{(C.8)}\quad \frac{\partial^2 EU}{\partial C \partial h^*} = s_o\left(\lambda p_h \frac{\partial h}{\partial I}\right)(-R_{h*}) + \int_o^{\ell^*} (-R_{h*})(\lambda p_h) \frac{\partial h}{\partial I} f(\ell)d\ell$$

$$+ \int_{\ell^*}^{\infty} (\lambda p_h) f(\ell)d\ell - \left\{s_o(-R_{h*})(-R_c - p_h h) r(I)(-\lambda)\right.$$

$$+ \int_o^{\ell^*} (-R_{h*})(-R_C - p_h h) r(I)(-\lambda) f(\ell)d\ell$$

$$\left.+ \int_{\ell^*}^{\infty} (-R_C - p_h h^*)(-R_{h*} + (1-C)p_h)(-\lambda) r(I) f(\ell)d\ell\right\}$$

$$+ E(-\lambda)\left(\frac{-\partial^2 R}{\partial C \partial h^*}\right).$$

This expression has made use of the Pratt formula for risk aversion,

$$r(I) = -U_{II}/U_I = (\partial\lambda/\partial I)/(-\lambda).$$

The Leibnitz formula terms involving the limits of integration all cancel.

Evaluation of this expression in braces is best done by noting that r(I) is essentially a weighting function on a series of integrated values, the total of which we knew to be zero (from the first-order conditions). Decreasing risk aversion in income implies increasing risk aversion as ℓ increases. That is, r(I) increases as ℓ increases under those conditions. Note also that $(-R_{h*})$ and $(-R_{h*} + (1-C)p_h)$ may be treated as weights on the same terms, and progressing in the same direction (from smaller to larger algebraically) as a decreasing risk aversion function r(I). *Both* of these "weighting functions" place more weight on negative terms. Hence, the expression in braces is negative, and makes a positive contribution to the value of a_{12}.

The remaining expressions in a_{12} all make positive contributions to a_{12}, with the exception of the integral term $\int_{\ell*}^{\infty}(\lambda p_h)f(\ell)d\ell$, which must be negative. For that reason, a_{12} is unsigned.

EFFECTS OF INCOME ON DEMAND FOR C

To determine the effect of income on demand for C, we must find the value of:

$$\text{(C.9)} \qquad b_1 = \frac{\partial}{\partial I}\left(s_o(-R_C-p_h h)(-\lambda) + \int_o^{\ell*} (-R_C-p_h h)(-\lambda)f(\ell)d\ell + \int_{\ell*}^{\infty} (-R_C-p_h h^*)(-\lambda)f(\ell)d\ell\right),$$

which is given as:

$$\text{(C.10)} \qquad b_1 = s_o\left((-R_C-p_h h)\left(\frac{-\partial\lambda}{\partial I}\right) + \lambda p_h \frac{\partial h}{\partial I}\right) + \int_o^{\ell^*}\left((-R_C-p_h h)\left(\frac{-\partial\lambda}{\partial I}\right) + \lambda p_h \frac{\partial h}{\partial I}\right)f(\ell)d\ell + \int_{\ell^*}^{\infty}\left((-R_C-p_h h^*)\left(\frac{-\partial\lambda}{\partial I}\right)\right)f(\ell)d\ell + E(-\lambda)\left(\frac{\partial R_C}{\partial I}\right).$$

This may be written as:

$$\text{(C.11)} \qquad -\left\{s_o(-R_C-p_h h)r(I)(-\lambda) + \int_o^{\ell^*}(-R_C-p_h h)r(I)(-\lambda)f(\ell)d\ell + \int_{\ell^*}^{\infty}(-R_C-p_h h^*)r(I)(-\lambda)f(\ell)d\ell\right\}$$

$$-\left\{s_o(-\lambda p_h)\frac{\partial h}{\partial I} + \int_o^{\ell^*}\left(-\lambda p_h \frac{\partial h}{\partial I}\right)f(\ell)d\ell + E(-\lambda)\cdot\left(\frac{-R_C}{\partial I}\right)\right\}.$$

The first expression in braces of (C.11) is simply the first-order condition for maximizing expected utility with respect to C, except that each term is weighted by $r(I)$. When $r(I)$ is decreasing in income, it increases with ℓ. Therefore, the original first-order condition is weighted more heavily on negative terms and less heavily on positive terms, and becomes negative. The expression in braces makes a positive contribution to b_1 (since that entire expression is multiplied by (-1)). The remaining portion of (C.11) is best evaluated by expanding the full expression for $\partial(-R_C)/\partial I$. From Appendix B,

$$-R_C = (1+\theta)\int_o^{\ell^*} p_h h\left(1-\left(\frac{1-C}{C}\right)\eta_{hC}\right)f(\ell)d\ell \ .$$

So
$$\frac{\partial(-R_C)}{\partial I} = (1+\theta)\int_o^{\ell^*}\left[p_h\frac{\partial h}{\partial I}\left(1-\left(\frac{1-C}{C}\right)\eta_{hC}\right)-p_h h\left(\frac{1-C}{C}\right)\frac{\partial\eta_{hC}}{\partial I}\right]f(\ell)d\ell \ .$$

Obviously if η_{hC} does not interact with income, then

$$\frac{\partial(-R_C)}{\partial I} = (1+\theta)\int_o^{\ell^*} p_h\frac{\partial h}{\partial I}f(\ell)d\ell = (1+\theta)\overline{\omega_h^T\eta_{hI}} \ ,$$

where $\overline{\omega_h\eta_{hI}}$ is an average value of the product of η_{hI} and the budget-share ω_h^T. Assuming that η_{hC} does not interact with income, the remainder of (D.11) is:

$$s_o(-\lambda_o)\left[\frac{-p_h h_o}{I}\eta_{h_oI}+\overline{\omega_h^T\eta_{hI}}\right]+\int_o^{\ell^*}(-\lambda)\left\{\overline{\omega_h^T\eta_{hI}}-\omega_h\eta_{hI}\right\}f(\ell)d\ell$$

$$+\int_{\ell^*}^{\infty}(-\lambda)\overline{\omega_h^T\eta_{hI}}\ f(\ell)d\ell \ .$$

Clearly, if $\eta_h^T\omega_{hI}$ is constant over all ℓ, the first two terms here vanish, leaving an expression that is necessarily positive (causing $\partial C/\partial I$ to become more positive--less insurance purchased with increasing income). However, if the income elasticity-budget share product *increases* with losses, then the expression may be negative. Similarly, if an interaction between η_{hC} and income is allowed, that will further alter the results, depending upon the sign of that interaction. Since the results are ambiguous, the response of demand for C to changes in income is dependent not only upon the behavior of risk aversion, but upon the behavior of $\partial h/\partial I$ as the size of ℓ changes, *and*

upon possible interactions between η_{hC} and income. Hence, one cannot make inferences about risk aversion characteristics by observing income elasticity of demand for reimbursement insurance, and one cannot predict whether $\partial C/\partial I$ will be positive or negative in this model without specifying the various elasticities to a considerable degree.

EFFECTS OF INCOME ON DEMAND FOR h*

To determine the income effects on demand for h*, we need to evaluate the expression $b_2 = \partial^2 EU/\partial h^*\partial I$. That expression is given as:

$$(C.12) \qquad \frac{\partial^2 EU}{\partial h^*\partial I} = \frac{\partial}{\partial I}\left[s_o(-R_{h^*})(-\lambda) + \int_o^{\ell^*} (-R_{h^*})(-\lambda)f(\ell)d\ell + \int_{\ell^*}^{\infty} (-R_{h^*}+(1-C)p_h)(-\lambda)f(\ell)d\ell\right]$$

$$= s_o(-R_{h^*})\left(\frac{-\partial\lambda}{\partial I}\right) + \int_o^{\ell^*} (-R_{h^*})\left(\frac{-\partial\lambda}{I}\right)f(\ell)d\ell + \int_{\ell^*}^{\infty} (-R_{h^*}+(1-C)p_h\left(\frac{-\partial\lambda}{\partial I}\right)f(\ell)d\ell$$

Making the appropriate substitutions for $(-\partial\lambda/\partial I)$ this becomes:

$$(C.13) \qquad b_2 = -\left\{s_o(-R_{h^*})^2 r(I)(-\lambda) + \int_o^{\ell^*} (-R_{h^*})^2 r(I)(-\lambda)f(\ell)d\ell + \int_{\ell^*}^{\infty} (-R_{h^*}+(1-C)p_h)^2 r(I)(-\lambda)f(\ell)d\ell\right\},$$

It can be proved that $(\partial^2 R)/(\partial h^*\partial I) = 0$

which is seen to be the first-order condition for maximization in h*, weighted by r(I). For decreasing risk aversion, the weights become larger for higher values of ℓ, and the expression b_2 becomes negative. If r(I) is constant, (C.13) reduces to the first-order condition, and is equal to zero. If risk aversion is increasing in wealth, then b_2 is positive.

EFFECTS ON DEMAND FOR C IF p_h CHANGES

To determine the effects on demand for C if p_h changes, we need to evaluate the expression $\partial^2 EU/\partial C \partial p_h$. That expression is:

$$\text{(C.14)} \quad \frac{\partial^2 EU}{\partial C \partial p_h} = \frac{\partial}{p_h}\left[s_o(-R_C - p_h h)(-\lambda) + \int_o^{\ell^*} (-R_C - p_h h)(-\lambda) f(\ell) d\ell + \int_{\ell^*}^{\infty} (-R_C - p_h h^*)(-\lambda) f(\ell) d\ell\right]$$

$$= s_o\left[(-R_C - p_h h)\left(Ch\frac{\partial\lambda}{\partial I} + C\lambda\frac{\partial h}{\partial I}\right) + \lambda\left(\frac{\partial^2 R}{\partial C \partial p_h} + p_h\frac{\partial h}{\partial p_h} + h\right)\right]$$

$$+ \int_o^{\ell^*}\left[(-R_C - p_h h)\left(Ch\frac{\partial\lambda}{\partial I} + C\lambda\frac{\partial h}{\partial I}\right) + \lambda\left(\frac{\partial^2 R}{\partial C \partial p_h} + p_h\frac{\partial h}{\partial p_h} + h\right)\right] f(\ell) d\ell$$

$$+ \int_{\ell^*}^{\infty}\left[(-R_C - p_h h^*)\left((h - (1-C)h^*)\frac{\partial\lambda}{\partial I} + \lambda\frac{\partial h}{\partial I} + \lambda\frac{\partial^2 R}{\partial C \partial p_h}\right)\right] f(\ell) d\ell .$$

Noting that $[h + (p_h)(\partial h/\partial p_h)]$ may be written as $h(1 + \eta_{hh})$, and that $h[(\partial\lambda/\partial I) + \lambda(\partial h/\partial I)]$ may be written as $(-\lambda)h/I[r^*(I) - \eta_{hI}]$, these expressions may be rewritten in elasticity form as:

$$(C.15)\quad \frac{\partial^2 EU}{\partial C \partial p_h} = s_o\left[\left(\frac{Ch}{I}\right)(-R_C - p_h h)(-\lambda)(r^*(I) - \eta_{hi}) + p_h(1 + \eta_{hh})\right]$$

$$+ \int_o^{\ell^*} \left[\frac{Ch}{I}(-R_C - p_h h)(-\lambda)(r^*(I) - \eta_{hI}) + p_h(1+\eta_{hh})\right] f(\ell)d\ell$$

$$+ \int_{\ell^*}^{\infty} \left[(-R_C - p_h h^*)\left(\frac{h}{I}\right)(r^*(I) - \eta_{hI})(-\lambda)\right.$$

$$\left. - (-R_C - p_h h^*)(-\lambda)r^*(I)\left(\frac{(1-C)h^*}{I}\right)\right] f(\ell)d\ell + E(-\lambda)\left(\frac{-\partial^2 R}{\partial C \partial p_h}\right).$$

This expression cannot be evaluated in general and is therefore technically unsigned. Section II discusses evaluation attempts.

EFFECTS OF A CHANGE IN p_h ON DEMAND FOR h*

To determine the effects of a change in demand for h* as p_h changes, we need to evaluate $\partial^2 EU/\partial h^* \partial p_h$. That expression is:

$$(C.16)\quad \frac{\partial^2 EU}{\partial h^* p_h} = \frac{\partial}{p_h}\left[s_o(-R_{h^*})(-\lambda) + \int_o^{\ell^*} (-R_{h^*})(-\lambda)f(\ell)d\ell\right.$$

$$\left. + \int_{\ell^*}^{\infty} (-R_{h^*} + (1-C)p_h)(-\lambda)f(\ell)d\ell\right]$$

$$= s_o(-R_{h^*})\left(\frac{-\partial\lambda}{\partial p_h}\right) + \int_o^{\ell^*} (-R_{h^*})\left(\frac{-\partial\lambda}{\partial p_h}\right) f(\ell)d\ell$$

$$+ \int_{\ell^*}^{\infty} (-R_{h^*} + (1-C)p_h)\left(\frac{-\partial\lambda}{\partial p_h}\right) f(\ell)d\ell$$

$$+ E(-\lambda)(-\partial^2 R/\partial h^* \partial p_h) + \int_{\ell^*}^{\infty} (-\lambda)(1-C)f(\ell)d\ell\ .$$

Making the appropriate substitutions for $-(\partial\lambda/\partial p_h)$, we find:

$$\text{(C.17)} \quad \frac{\partial^2 EU}{\partial h^* \partial p_h} = s_o(-R_{h*})\left(Ch\,\frac{\partial\lambda}{\partial I} + C\lambda\,\frac{\partial h}{\partial I}\right) + \int_o^{\ell*} (-R_{h*})\left(Ch\,\frac{\partial\lambda}{\partial I} + C\lambda\,\frac{\partial h}{\partial I}\right) f(\ell)d\ell$$

$$+ \int_{\ell*}^{\infty} (-R_{h*}+(1-C)p_h)((h-h^*)+Ch^*)\,\frac{\partial\lambda}{\partial I}\,f(\ell)d\ell$$

$$+ \int_{\ell*}^{\infty} (-\lambda)(1-C)f(\ell)d\ell + E(-\lambda)(-\partial^2R/\partial h^*\partial p_h)\ .$$

This equation may be rewritten similarly to (C.17) (in elasticity form), with a similar conclusion:

$$\text{(C.18)} \quad \frac{\partial^2 EU}{\partial h^* \partial p_h} = \left\{ s_o\left(\frac{-R_{h*}}{I}\right)(-\lambda)(r^*(I) - \eta_{hI}) + \int_o^{\ell*} \left(\frac{-R_{h*}}{I}\right)(-\lambda)(r^*(I) - \eta_{hI})f(\ell)d\ell \right.$$

$$\left. + \int_{\ell*}^{\infty} (-R_{h*}+(1-C)p_h)((h-h^*)+Ch^*)(-\lambda)(r^*(I)) - \eta_{hI})f(\ell)d\ell \right\}$$

$$+ \int_{\ell*}^{\infty} (-\lambda)(1-C)f(\ell)d\ell + E(-\lambda)(-\partial^2R/\partial h^*\partial p_h)\ .$$

Again we see that this expression is simply the first-order condition for maximization with respect to h*, with each term weighted by $[r^*(I) - \eta_{hI}]$ and by the terms Ch or $[(h-h^*) + Ch^*]$. The latter set of terms *must* increase as ℓ increases. Therefore, if the weighting function $[r^*(I) - \eta_{hI}]$ is constant, or increases with ℓ, then the expression in braces becomes more positive. Technically, this expression is unsigned unless one specifies that the income elasticity of

demand for h is zero (the random wealth loss model), in which case (C.17) is always positive.

EFFECTS OF A CHANGE IN H_0 ON DEMAND FOR C

To determine the effects on demand for C if the consumer enters the period with a different H_0, we need to evaluate the expression

$$\text{(C.19)}\quad \frac{\partial^2 EU}{\partial C \partial H_0} = \frac{\partial}{H_0}\left[(-R_C-p_h h)(-\lambda) + \int_0^{\ell^*} (-R_C-p_h h)(-\lambda) f(\ell)d\ell + \int_{\ell^*}^{\infty} (-\lambda)(-R_C-p_h h^*) f(\ell)d\ell\right]$$

$$= s_o\left[(-R_C-p_h h)\left(\frac{-\partial\lambda}{\partial H_0}\right) + \left(p_h \frac{\partial h}{\partial H_0} + \frac{\partial^2 R}{\partial C\partial H_0}\right)\right]$$

$$+ \int_0^{\ell^*}\left[(-R_C-p_h h)\left(\frac{-\partial\lambda}{\partial H_0}\right) + \left(p_h \frac{\partial h}{\partial H_0} + \frac{\partial^2 R}{C\partial H_0}\right)\right] f(\ell)d\ell$$

$$+ \int_{\ell^*}^{\infty}\left[(-R_C-p_h h^*)\left(\frac{-\partial\lambda}{\partial H_0}\right) + \left(\frac{\partial^2 R}{\partial C\partial H_0}\right)\right] f(\ell)d\ell\ .$$

Substituting for $(-\partial\lambda/\partial H_0)$ we find:

$$\text{(C.20)}\quad \frac{\partial^2 EU}{\partial C\partial H_0} = \left[s_o(-R_C-p_h h)\ \frac{NPH}{g'}\ r(I)(-\lambda) + (-\lambda)p_h p_x \frac{\partial x}{\partial I} + \partial^2 R/\partial C\partial H_0\right]$$

$$+ \int_0^{\ell^*}\left[(-R_C-p_h h)\frac{NPH}{g'}\ r(I)(-\lambda) + (-\lambda)p_h p_x \frac{\partial x}{\partial I} + \lambda(\partial^2 R/\partial C\partial H_0)\right] f(\ell)d\ell$$

$$+ \int_{\ell^*}^{\infty}\left[(-R_C-p_h h^*)\ \frac{NPH}{g'}\ (-\lambda) + (\partial^2 R/\partial C\partial H_0)\right] f(\ell)d\ell\ .$$

If NPH is increasing over ℓ (as it assuredly is), or if g'(h) is decreasing as ℓ increases (if there are diminishing returns to h), then the function NPH/g' acts as a weighting function with the same behavior as a risk aversion function r(I), when r(I) is decreasing in income. In that case, those terms that form the first-order condition for maximization with respect to C (as weighted by NPH/g' and r(I)) act to make (C.20) positive. The remaining terms, involving $\partial x/\partial I$, also tend to make (C.20) positive, as long as "x" is a normal good. The remaining terms, involving the second derivative $\partial^2 R/\partial C\partial H_0$ act in an offsetting direction. If community rating is practiced (so that $\partial^2 R/\partial C\partial H_0 = 0$), then (C.20) is uniformly positive.

EFFECT OF CHANGES IN H_0 ON DEMAND FOR h*

In assessing demand for h* as H_0 changes, we need to evaluate the derivative $\partial^2 EU/\partial h^*\partial H_0$, which is given as:

$$\text{(C.21)} \quad \frac{\partial^2 EU}{\partial h^*\partial H_0} = \frac{\partial}{H_0}\left[s_o(-R_{h^*})(-\lambda) + \int_o^{\ell^*} (-R_{h^*})(-\lambda)f(\ell)d\ell + \int_{\ell^*}^{\infty} (-\lambda)(-R_{h^*}+(1-C)p_h)f(\ell)d\ell\right]$$

$$= \left[s_o(-R_{h^*})\left(\frac{-\partial\lambda}{\partial H_0}\right) + \int_o^{\ell^*} (-R_{h^*})\left(\frac{-\partial\lambda}{\partial H_0}\right)f(\ell)d\ell + \int_{\ell^*}^{\infty} (-R_{h^*}+(1-C)ph)\left(\frac{-\partial\lambda}{\partial H_0}\right)f(\ell)d\ell + E(-\lambda)(-\partial^2 R/\partial h^*\partial H_0)\right] .$$

Making the appropriate substitutions for $(-\partial\lambda/\partial H_0)$ we find:

$$\text{(C.22)} \quad \frac{\partial^2 EU}{\partial h^* \partial H_0} = -\left[s_o(-R_{h^*})\frac{NPH}{g'(h)} r(I)(-\lambda) + \int_o^{\ell^*} (-R_{h^*})\frac{NPH}{g'(h)} r(I)(-\lambda) f(\ell) d\ell \right.$$

$$\left. + \int_{\ell^*}^{\infty} (-R_{h^*} + (1-C)p_h)\frac{NPH}{g'(h)} r(I)(-\lambda) f(\ell) d\ell + E(-\lambda)(-\partial^2 R/\partial h^* \partial H_0) \right].$$

The same analysis may be used here as for equation (C.20). If the ratio NPH/g'(h) increases as ℓ increases, and if r(I) is decreasing as income increases, then (C.22) is negative, with the exception of the second derivative of the function R (the offsetting price reduction of a unit of h* of coverage). That term equals zero if premiums do not change as H_0 changes (a common state of the world in actual markets).

EFFECTS OF A CHANGE IN THE LOADING FEE ON DEMAND FOR C

When determining the effects of a change in the loading fee θ on demand for C, we need to evaluate the expression $\partial^2 EU/\partial C \partial \theta$, which is:

$$\text{(C.23)} \quad \frac{\partial^2 EU}{\partial C \partial \theta} = \frac{\partial}{\partial \theta} \left[s_o(-R_C - p_h h)(-\lambda) + \int_o^{\ell^*} (-R_C - p_h h)(-\lambda) f(\ell) d\ell \right.$$

$$\left. + \int_{\ell^*}^{\infty} (-R_C - p_h h^*)(-\lambda) f(\ell) d\ell \right]$$

$$= s_o \left[(-R_C - p_h h)\left(\frac{-\partial \lambda}{\partial \theta}\right) + \lambda\left(ph \frac{\partial h}{\partial \theta} + \frac{\partial^2 R}{\partial C \partial \theta}\right) \right]$$

$$+ \int_o^{\ell^*} \left[(-R_C - p_h h)\left(\frac{-\partial \lambda}{\partial \theta}\right) + \lambda\left(p_h \frac{\partial h}{\partial \theta} + \frac{\partial^2 R}{\partial C \partial \theta}\right) \right] f(\ell) d\ell$$

$$+ \int_{\ell^*}^{\infty} \left[(-R_C - p_h h^*)\left(\frac{-\partial \lambda}{\partial \theta}\right) + \left(\frac{\partial^2 R}{\partial C \partial \theta}\right) \right] f(\ell) d\ell \; .$$

Substituting for $(-\partial\lambda/\partial\theta)$ (where R_θ is defined in Appendix B), this becomes:

$$\text{(C.24)}\quad \frac{\partial^2 EU}{\partial C\partial\theta} = s_o\left[(-R_C-p_h h)R_\theta\ r(I)(-\lambda) + (-\lambda)p_h \frac{\partial h}{\partial I} R_\theta\right]$$

$$+ \int_o^{\ell*} \left[(-R_C-p_h h)R_\theta\ r(I) \cdot (-\lambda) + (-\lambda)p_h R_\theta \frac{\partial h}{\partial I}\right] f(\ell)d\ell$$

$$+ \int_{\ell*}^{\infty} (-R_\theta-p_h h^*) \cdot R_\theta \cdot r(I)(-\lambda)f(\ell)d\ell + E(-\lambda)(-\partial^2 R/\partial C\partial\theta) \ .$$

For utility functions with decreasing risk aversion, this expression is of indeterminant value. One set of the terms is seen to be the first-order conditions for maximization with respect to C, with a weighting function of $r(I)$, and all multiplied by R_θ (which must be positive). When risk aversion is decreasing, those terms make a negative contribution to the value of (C.24). The terms involving $\partial h/\partial I$ make a positive contribution, as do the terms involving the second derivative of the rate function $\partial^2 R/\partial C\partial\theta$.

EFFECTS ON DEMAND FOR h* as θ CHANGES

In assessing the effects of a change in θ on demand for h*, we must evaluate the expression $\partial^2 EU/\partial h^*\partial\theta$, which is:

$$\text{(C.25)}\quad \frac{\partial^2 EU}{\partial h^*\partial\theta} = \frac{\partial}{\partial\theta}\left[s_o(-R_{h*})(-\lambda) + \int_o^{\ell*} (-R_{h*})(-\lambda)f(\ell)d\ell \right.$$

$$\left. + \int_{\ell*}^{\infty} (-R_{h*}+(1-C)p_h)(-\lambda)f(\ell)d\ell\right]$$

$$= \left[s_o(-R_{h*})\left(\frac{-\partial\lambda}{\partial\theta}\right) + \int_o^{\ell*} (-R_{h*})\left(\frac{-\partial\lambda}{\partial\theta}\right)f(\ell)d\ell\right.$$

$$\left. + \int_{\ell*}^{\infty} (-R_{h*}+(1-C)p_h)\left(\frac{-\partial\lambda}{\partial\theta}\right)f(\ell)d\ell \quad + E(-\lambda)(-\partial^2R/\partial h*\partial\theta)\right] .$$

Substituting for $(-\partial\lambda/\partial\theta)$ we find:

$$\text{(C.26)} \qquad \frac{\partial^2 EU}{\partial h*\partial\theta} = R_\theta\left[s_o(-R_{h*})r(I)(-\lambda) + \int_o^{\ell*} (-R_{h*})r(I)(-\lambda)f(\ell)d\ell\right.$$

$$\left. + \int_{\ell*}^{\infty} (-R_{h*}+(1-C)p_h)r(I)(-\lambda)f(\ell)d\ell\right] + E(-\lambda)(-\partial^2R/\partial h*\partial\theta) \ .$$

For utility functions with decreasing risk aversion, the expression (C.26) is seen to be positive, with the exception of the terms involving $\partial^2R/\partial h*\partial\theta$, which make a negative contribution to the value of (C.26). The series of terms (all multiplied by R_θ) are seen to be the first-order condition for maximizing expected utility with respect to h*, weighted by the risk aversion measure r(I). When r(I) decreases with income, those terms become positive overall.

Appendix D

SUPPLEMENTARY TABLES OF EXPECTED VALUE LOCI DEMAND FOR VARIOUS TYPES OF INSURANCE

Table D-1

EXPECTED VALUE LOCUS OF MAXPRM AS INCOME CHANGES

Level of Income	Probability that y>0	E(y)	E(y|y>0)
2,000	.497	53.05	106.69
4,000	.537	59.94	111.65
6,000	.576	67.35	116.92
8,000	.614	75.27	122.50
10,000	.652	83.70	128.41
12,000	.688	92.62	134.67
14,000	.722	102.01	141.29
16,000	.754	111.84	148.27
18,000	.784	122.08	155.63
20,000	.812	132.72	163.37
22,000	.838	143.70	171.51
24,000	.861	155.02	180.03
26,000	.882	166.62	188.94
28,000	.900	178.49	198.24
30,000	.917	190.59	207.92

Table D-2

EXPECTED VALUE LOCUS OF MAXPRM AS EDUCATION CHANGES

Years of Education	Probability that y>0	E(y)	E(y|y>0)
0	.353	30.40	86.19
1-4	.479	47.58	99.41
5-6	.580	64.92	111.86
7-8	.652	79.74	122.28
9-11	.696	90.10	129.55
12	.714	94.82	132.87
13-15	.708	93.41	131.88
≥16	.679	86.01	126.67

Table D-3

EXPECTED VALUE LOCUS OF MAXPRM AS GROUP SIZE CHANGES

Work Group Size	Probability that y>0	E(y)	E(y|y>0)
1	.551	59.50	108.01
2	.621	73.02	117.56
3	.688	88.12	128.15
4	.748	104.70	139.87
5	.802	122.60	152.77
6	.849	141.65	166.90
7	.887	161.69	182.25
8	.918	182.53	198.80

Table D-4

EXPECTED VALUE LOCUS OF HOSMAX AS INCOME CHANGES

Level of Income	Probability that y>0	E(y)	E(y|y>0)
2,000	.651	9.15	14.06
4,000	.680	9.94	14.62
6,000	.709	10.78	15.21
8,000	.736	11.64	15.82
10,000	.762	12.54	16.46
12,000	.787	13.47	17.13
14,000	.810	14.43	17.83
16,000	.831	15.42	18.55
18,000	.851	16.42	19.31
20,000	.869	17.46	20.09
22,000	.885	18.51	20.90
24,000	.900	19.58	21.75
26,000	.914	20.67	22.61
28,000	.926	21.77	23.51
30,000	.937	22.89	24.43

Table D-5

EXPECTED VALUE LOCUS FOR HOSMAX AS EDUCATION CHANGES

Years of Education	Probability that y>0	E(y)	E(y|y>0)
0	.209	1.74	8.31
1-4	.316	3.01	9.51
5-6	.415	4.43	10.66
7-8	.494	5.75	11.65
9-11	.546	6.76	12.38
12	.572	7.30	12.76
13-15	.571	7.29	12.75
≥16	.545	6.73	12.36

Table D-6

EXPECTED VALUE LOCUS OF HOSMAX AS GROUP SIZE CHANGES

Work Group Size	Probability that y>0	E(y)	E(y|y>0)
1	.368	3.71	10.10
2	.454	5.05	11.13
3	.542	6.67	12.31
4	.628	8.57	13.66
5	.708	10.75	15.18
6	.779	13.17	16.91
7	.839	15.80	18.84
8	.887	18.61	20.98

Table D-7

EXPECTED VALUE LOCUS OF MM-MAX AS INCOME CHANGES [a]

Level of Income	Probability that y>0	E(y)	E(y|y>0)
2,000	.214	2,190.94	10,243.25
4,000	.253	2,723.23	10,777.49
6,000	.295	3,348.09	11,354.10
8,000	.340	4,072.87	11,976.72
10,000	.388	4,903.77	12,649.16
12,000	.437	5,845.59	13,375.42
14,000	.487	6,901.47	14,159.63
16,000	.538	8,072.80	15,006.00
18,000	.588	9,359.12	15,918.73
20,000	.637	10,758.13	15,901.91
22,000	.683	12,265.84	17,959.46
24,000	.727	13,876.72	19,094.91
26,000	.767	15,583.95	20,311.32
28,000	.804	17,379.75	21,611.09
30,000	.837	19,255.65	22,995.84

[a]In these estimates, work-group size is set at 7 (i.e., the loading fee is approximately .10).

Table D-8

EXPECTED VALUE LOCUS OF MM-MAX AS EDUCATION CHANGES

Years of Education	Probability that y>0	E(y)	E(y|y>0)
0	.347	4,182.61	12,067.71
1-4	.345	4,158.86	12,048.08
5-6	.351	4,254.20	12,126.68
7-8	.364	4,474.45	12,306.25
9-11	.384	4,832.89	12,593.13
12	.412	5,350.64	12,997.64
13-15	.448	6,057.06	13,534.61
≥16	.491	6,989.95	14,224.29

Table D-9

EXPECTED VALUE LOCUS OF MM-MAX AS GROUP SIZE CHANGES

Work Group Size	Probability that y>0	E(y)	E(y|y>0)
1	.059	453.37	7,723.57
2	.083	683.59	8,219.82
3	.115	1,004.98	8,769.61
4	.154	1,441.57	9,380.02
5	.201	2,018.96	10,059.02
6	.255	2,762.92	10,815.53
7	.317	3,697.57	11,659.43
8	.384	4,843.46	12,601.50

Table D-10

EXPECTED VALUE LOCUS OF SURMAX AS EDUCATION CHANGES

Years of Education	Probability that y>0	E(y)	E(y|y>0)
0	.319	78.06	244.99
1-4	.395	105.65	267.50
5-6	.464	134.13	289.16
7-8	.522	161.39	309.05
9-11	.569	185.52	326.26
12	.603	204.95	339.98
13-15	.625	218.53	349.52
≥16	.636	225.51	354.41

Table D-11

EXPECTED VALUE LOCUS OF SURMAX AS GROUP SIZE CHANGES

Work Group Size	Probability that y>0	E(y)	E(y|y>0)
1	.419	115.04	274.78
2	.492	147.08	298.69
3	.566	184.35	325.43
4	.638	226.75	355.28
5	.705	274.06	388.50
6	.766	325.88	425.31
7	.819	381.72	465.89
8	.864	441.00	510.33

Appendix E

STATISTICAL SUMMARY OF CROSS-SECTION DATA

Sample: Entire data set; n = 2367

Variable	Mean	Standard Deviation Individual Data	Standard Deviation Grouped Data
MAXPRM	116.84	111.55	39.15
VERPRM	83.20	106.13	34.34
FAMPRM	56.77	84.31	26.67
EMPPRM	26.67	61.87	17.27
HOSMAX	10.86	12.38	6.05
HOSDAY	65.93	106.93	34.15
SURMAX	171.18	230.34	77.62
MM-MAX	1776.85	5501.22	2022.64
PERINC	6310.50	3060.69	--
INCOME (observed)	6370.35	5080.76	2730.19
EDUC.	5.31	1.62	1.45
AGE	4.67	1.85	.91
TOTPRI	459.19	60.66	35.19
MD-PRI	85.68	17.75	10.08
HOSPRI	221.46	38.07	29.24
ESTLOS	3.40	.73	.55
GRPSIZ	3.77	3.04	1.31
FREE CARE	.058	.23	.09
WELFARE CARE	.045	.207	.08
BLACK	.125	.33	--
FEMALE	.195	.39	--

Sample: Families with no unverified policies; n = 1579

Variable	Mean	Standard Deviation Individual Data Set	Standard Deviation Grouped Data Set
MAXPRM	109.02	111.13	48.48
VERPRM	102.08	110.32	47.22
FAMPRM	70.96	90.45	36.93
EMPPRM	31.41	65.72	22.11
HOSMAX	13.15	12.51	8.15
HOSDAY	79.15	111.36	41.63
SURMAX	203.26	242.15	101.12
MM-MAX	2108.17	5709.57	2403.83
PERINC	5897.19	3100.88	--
INCOME (observed)	5719.45	5188.08	2277.97
EDUC.	5.16	1.67	1.45
AGE	4.74	1.91	.977
TOTPRI	458.13	61.63	38.09
MD-PRI	85.14	17.88	10.67
HOSPRI	220.25	37.94	29.60
ESTLOS	3.34	.752	.70
GRPSIZ	3.13	2.92	1.38
FREE CARE	.068	.25	.10
WELFARE CARE	.059	.24	.10
BLACK	.139	.34	--
FEMALE	.217	.41	--

This study uses an instrumental variable technique to achieve a measure of permanent income free of transitory errors. A multiple regression fitted observed income to a set of independent variables, which should systematically affect permanent income. The fitted values are used in the study as the measure of permanent income (PERINC). The exact equation fitted was:

$$\begin{aligned}
\text{Income} = {} & \underset{(t=1.90)}{(125.22)}(\text{FAMILY SIZE}) + \underset{(t=.297)}{69.867}\ (\text{RURAL}) + \underset{(t=7.00)}{7.03}\ (\text{COUNTY INCOME}) \\
& + \underset{(t=1.38)}{608.5}\ (\text{HOME OWNER}) + \underset{(t=2.21)}{546.1}\ (\text{SUBURB}) - \underset{(t=2.47)}{978.5}\ (\text{NEGRO}) - \underset{(t=4.73)}{1860}\ (\text{FEMALE}) \\
& + \underset{(t=.04)}{18.59}\ (\text{EDUC}=6) + \underset{(t=1.73)}{763.43}\ (\text{EDUC}=8) + \underset{(t=4.06)}{1379}\ (\text{EDUC}=11) + \underset{(t=4.06)}{1857}\ (\text{EDUC}=12) \\
& + \underset{(t=5.93)}{2980}\ (\text{EDUC}=15) + \underset{(t=9.45)}{5066}\ (\text{EDUC}=16+) + \underset{(t=1.17)}{611}\ (\text{MOUNTAIN}) - \underset{(t=.67)}{336}\ (\text{NORTHEAST}) \\
& - \underset{(t=.11)}{37.4}\ (\text{MIDATLANTIC}) + \underset{(t=1.18)}{378.6}\ (\text{E.N. CENTRAL}) - \underset{(t=0)}{3.2}\ (\text{W.N. CENTRAL}) \\
& + \underset{(t=2.31)}{850}\ (\text{S. ATLANTIC}) + \underset{(t=1.04)}{547}\ (\text{E.S. CENTRAL}) + \underset{(t=.99)}{414.9}\ (\text{W.S. CENTRAL}) \\
& + \underset{(t=.71)}{237.6}\ (\text{PROFESSIONAL PERSON}) + \underset{(t=.01)}{10.1}\ (\text{MINING}) - \underset{(t=\ 27)}{152}\ (\text{FARMER}) \\
& - \underset{(t=1.24)}{575}\ (\text{CRAFTS}) - \underset{(t=.03)}{9.43}\ (\text{SERVICE}) + \underset{(t=.95)}{456.7}\ (\text{CONSTRUCTION}) \\
& + \underset{(t=3.50)}{1450}\ (\text{MANUFACTURING}) + \underset{(t=2.23)}{1103}\ (\text{COMMUNICATION}) + \underset{(t=2.38)}{1103}\ (\text{RETAIL}) \\
& + \underset{(t=4.00)}{2377}\ (\text{FINANCE}) + \underset{(t=.96)}{445}\ (\text{PUBLIC ADMIN.}) - \underset{(t=.97)}{1215}\ (\text{ENTERTAINMENT}) \\
& - \underset{(t=.21)}{99.4}\ (\text{PERSONAL SERVICES}) + \underset{(t=1.42)}{687}\ (\text{PROFESSIONAL SERVICES}) \\
& - \underset{(t=6.66)}{4327}\ (\text{AGE-24}) - \underset{(t=2.52)}{2675}\ (\text{AGE-29}) - \underset{(t=1.66)}{1523}\ (\text{AGE-34}) - \underset{(t=2.12)}{1385}\ (\text{AGE-44}) \\
& - \underset{(t=1.70)}{1162}\ (\text{AGE-64}) - \underset{(t=4.72)}{2916}\ (\text{AGE-74}) - \underset{(t=5.07)}{3409}\ (\text{AGE-75+}) - \underset{(t=5.57)}{2297}\ (\text{ALL MEN}) \\
& + \underset{(t=1.41)}{366}\ (\text{MORE MEN}) + \underset{(t=.40)}{105}\ (\text{MORE FEMALE}) - \underset{(t=2.74)}{1272}\ (\text{ALL FEMALE}) + \underset{(t=.88)}{990.34}\ .
\end{aligned}$$

In this regression, the constant term embodies the omitted variables, as well as any real "intercept" term. Omitted as standardizations were: Education less than six years; Pacific Coast Region; Age of head 45-54; Family of equal sex ratio; Managers (occupation); Agriculture (industry). The equation had uncorrected $R^2 = .3625$, with an F-ratio of 22.78. The standard deviation of observed income is 5055; the standard error of the regression is 4086.

A similar equation was used to predict the expected illness level facing the family. The dependent variable used was "disability days," a summary measure for the family. Fitted values from this equation were used as the Expected Illness measure (ESTLOS). The equation fitted was:

Disability days =

.1984 (FAMSIZ) + .107 (RURAL) - .1275E-04 (INCOME) + .00125 (COUNTY INCOME)
(t=5.98) (t=.91) (t=1.11) (t=2.51)

\+ .6369 (HOME OWNER) + .2294 (FEMALE) - .1317 (NEGRO) + .1635 (AGE-24)
(t=2.87) (t=1.15) (t=.66) (t=.23)

\- .30 (AGE-29) - .281 (AGE-34) + .054 (AGE-44) + .023 (AGE-64) - 1.07 (AGE-75)
(t=.57) (t=.61) (t=.16) (t=.06) (t=3.38)

\- 1.11 (AGE-75+) + .11 (EDUC-6) - .18 (EDUC-8) - .12 (EDUC-11) - .14 (EDUC-12)
(t=3.22) (t=.46) (t=.85) (t=.57) (t=.62)

\- .229 (EDUC-15) - .12 (EDUC-16+) + .07 (MOUNTAIN) - .07 (N. EAST) - .21 (MIDATL.)
(t=.90) (t=.47) (t=.28) (t=.27) (t=1.33)

\- .17 (E.N. CENT) - .53 (W.N. CENT) + .14 (S. ATLAN.) - .13 (S. CENT)
(t=1.10) (t=2.81) (t=.77) (t = .55)

\- .72 (ALL MALE) - .16 (MORE MALE) + .01 (MORE FEMALE) - .84 (ALL FEMALE)
(t=3.48) (t=1.26) (t=.07) (t=3.59)

\- 1.34 (OLDEST-24) - .89 (OLDEST-29) - .31 (OLDEST-34) - .78 (OLDEST-44)
(t=1.83) (t=1.67) (t=.70) (t=2.47)

\- .87 (OLDEST-54) - 1.00 (OLDEST-64) + .52E-03 (DENTAL) + 2.76
(t=3.00) (t=3.27) (t=1.09) (t=4.97)

In this equation, the "standardizing variables" were: Education under six years; Age over 65; Pacific Region; Equal sex ratio in family; Oldest in family

over 65; Urban; Non-Negro; Male Household Head; apartment dweller.

The equation had R^2 = .116 and F = 5.70.

BIBLIOGRAPHY

I. BOOKS AND REPORTS

Andersen, Ronald and Odin W. Anderson, *A Decade of Health Services*, University of Chicago Press, 1967.

Arrow, Kenneth J., *Essays in the Theory of Risk Bearing*, Markham Publishing Company, Chicago, 1971.

Becker, Gary S., *Economic Theory*, Alfred A. Knopf, Inc., New York, 1971.

Dickerson, O. D., *Health Insurance*, Richard D. Irwin & Co., Homewood, Illinois, 1959.

Faulkner, E. J., *Health Insurance*, McGraw-Hill, New York, 1960.

Friedman, Milton, *A Theory of the Consumption Function*, Princeton University Press, 1957.

Goldberger, Arthur S., *Econometric Theory*, John Wiley & Sons, Inc., New York, 1964.

Grossman, Michael, *The Demand for Health: A Theoretical and Empirical Investigation*, Columbia University Press, New York, 1972.

Klarman, Herbert E. with Helen H. Jaszi, *Empirical Studies in Health Economics*, The Johns Hopkins Press, Baltimore, 1970.

MacColl, William A., *Group Practice and Prepayment of Medical Care*, Public Affairs Press, Washington, D.C., 1966.

MacIntyre, Duncan M., *Voluntary Health Insurance and Rate Making*, Cornell University Press, Ithaca, 1962.

Newhouse, Joseph P. and Charles E. Phelps, with the assistance of David J. Weinschrott and Mark S. Thompson, *Price and Income Elasticities of Demand for Medical Care Services*, R-1197-NC, The Rand Corporation, Santa Monica, 1973.

Newhouse, Joseph P. and Charles E. Phelps, *On Having Your Cake and Eating it Too: An Analysis of Estimated Effects of Insurance on Demand for Medical Care*, R-1149-NC, The Rand Corporation, Santa Monica, 1973.

Phelps, Charles E. and Joseph P. Newhouse, *Coinsurance and the Demand for Medical Services*, R-964-OEO, The Rand Corporation, Santa Monica, 1973.

Somers, Herman Miles and Anne Ramsay Somers, *Doctors, Patients, and Health Insurance*, The Brookings Institution, Washington, D.C., 1961.

Theil, Henri, *Principles of Econometrics*, John Wiley & Sons, Inc., New York, 1971.

Von Neumann, John and Oscar Morgenstern, *The Theory of Games and Economic Behavior*, Princeton University Press, 1947.

Zellner, Arnold, *An Introduction to Bayesian Inference in Econometrics*, John Wiley & Sons, Inc., New York, 1971.

II. JOURNAL ARTICLES AND ARTICLES FROM COLLECTED ESSAYS

Andersen, Ronald and Lee Benham, "Factors Affecting the Relationship between Family Income and Medical Care Consumption," in Herbert J. Klarman, ed., *Empirical Studies in Health Economics*, The Johns Hopkins Press, Baltimore, 1970, 73-95.

Arrow, Kenneth J., "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, 53, 1963, 941-973.

Becker, Gary S., "A Theory of the Allocation of Time," *The Economic Journal*, 75, September 1965, 493-517.

Bernoulli, Daniel, trans. by Louise Sommer, *Econometrica*, 22, 1, January 1954, 22-36.

Crew, Michael, "Coinsurance and the Welfare Economics of Medical Care," *American Economic Review*, 54, No. 5, December 1969, 406-408.

Ehrlich, Isaac and Gary S. Becker, "Market Insurance, Self-Insurance, and Self-Protection," *Journal of Political Economy*, 80, No. 4, July-August 1972, 623-648.

Eilers, Robert D., "Post-Payment Health Coverage: Contributions and Limitations," *The Journal of Risk and Insurance*, 38, No. 3, 1970, 381-396.

Feldstein, Martin S., "An Econometric Model of the Medicare System," *The Quarterly Journal of Economics*, No. 338, February 1971, 1-20.

-----, "The Welfare Loss of Excess Health Insurance," *Journal of Political Economy*, 81, No. 2, Part 1, March-April 1973, 251-280.

Friedman, Milton, "The 'Welfare' Effects of Taxes," *Essays in Positive Economics*, The University of Chicago Press, 1953, 100-116.

Friedman, Milton and L. J. Savage, "The Utility Analysis of Choices Involving Risks," *The Journal of Political Economy*, 56, 1948, 279-304.

Gould, John P., "The Expected Utility Hypothesis and the Selection of Optimal Deductibles for a Given Insurance Policy," *The Journal of Business*, 42, No. 2, April 1969, 143-151.

Greenlich, Merwyn R. and Benjamin J. Darsky, "A Comparison of General Drug Utilization in a Metropolitan Community with Utilization Under a Drug Prepayment Plan," *The American Journal of Public Health*, 58, No. 11, November 1968, 2121-2135.

Hedinger, Fredric R., "The Social Role of Blue Cross: Progress and Problems," *Inquiry*, 5, No. 2, June 1968, 3-12.

Kessel, Reuben A., "Price Discrimination in Medicine," *The Journal of Law and Economics*, 1, No. 1, October 1958, 20-53.

Langford, Elizabeth A., "Medical Care in the Consumer Price Index, 1936-56," *Monthly Labor Review*, September 1958, 1053-1058.

Mossin, Jan, "Aspects of Rational Insurance Purchasing," *The Journal of Political Economy*, 76, August 1968, 553-568.

Pashigian, B. Peter, Laurence L. Schkade, and George H. Menefee, "The Selection of an Optimal Deductible for a Given Insurance Policy," *The Journal of Business*, 29, 1966, 35-44.

Pauly, Mark V., "The Economics of Moral Hazard: Comment," *American Economic Review*, 58, No. 3, 531-537.

-----, "The Welfare Economics of Community Rating," *The Journal of Risk and Insurance*, 38, No. 3, 407-418.

Phelps, Charles E. and Joseph P. Newhouse, "The Effects of Coinsurance on the Demand for Physician Services," *Social Security Bulletin*, 35, No. 6, June 1972, 20-28.

Pratt, John W., "Risk Aversion in the Large and in the Small," *Econometrica*, 32, Nos. 1-2, January-April 1964, 122-136.

Reed, Louis, "Private Health Insurance Coverage and Financial Experience," *Social Security Bulletin*, 30, No. 11, November 1967, 3-22.

Rosenthal, Gerald, "Price Elasticity of Demand for Short-Term General Hospital Services," in Herbert E. Klarman, ed., *Empirical Studies in Health Economics*, The Johns Hopkins Press, Baltimore, 1970, 101-117.

Rosett, Richard N. and Lien-Fu Huang, "The Effect of Health Insurance on the Demand for Medical Care," *Journal of Political Economy*, 81, No. 2, Part 1, March-April 1973, 281-305.

Smith, Vernon L., "Optimal Insurance Coverage," *Journal of Political Economy*, 76, No. 1, January-February 1968, 68-77.

Tobin, James, "Estimation of Relationships for Limited Dependent Variables," *Econometrica*, 26, No. 1, 1958, 26-36.

Zeckhauser, Richard, "Medical Insurance: A Case Study of the Tradeoff between Risk Spreading and Appropriate Incentives," *Journal of Economic Theory*, 2, No. 1, March 1970, 14-26.

Zellner, Arnold, "Estimation of Cross-Section Relations: Analysis of a Common Specification Error," *Metroeconomica*, 14, Nos. 1, 2, 3, April, August, December 1962, 111-117.

-----, "On the Aggregation Problem: A New Approach to a Troublesome Problem," in K. A. Fox, ed., *Econometric Models, Estimation and Risk Programming: Essays in Honor of Gerhard Tintner*, Springer-Verlag, Inc., New York, 1969, 365-373.

----- and Tong Hun Lee, "Joint Estimation of Relationships Involving Discrete Random Variables," *Econometrica*, 33, No. 2, April 1965, 382-393.

III. DUPLICATED AND UNPUBLISHED MATERIAL

Cooper, Richard V. and Joseph P. Newhouse, "Further Results on the Errors in the Variables Problem," P-4715, The Rand Corporation, Santa Monica, October 1971.