

The Use of Pathology Services

A Comparison of Fee-for-Service and a Prepaid Group Practice

Willard G. Manning, Jr.

35th
Year

Rand

The project upon which this publication is based was performed pursuant to Contract No. 500-78-0048 with the Health Care Financing Administration, U.S. Department of Health and Human Services.

ISBN 0-8330-0482-4

The Rand Publications Series: The Report is the principal publication documenting and transmitting Rand's major research findings and final research results. The Rand Note reports other outputs of sponsored research for general distribution. Publications of The Rand Corporation do not necessarily reflect the opinions or policies of the sponsors of Rand research.

Published by The Rand Corporation

R-2919-HCFA

The Use of Pathology Services

A Comparison of Fee-for-Service and a Prepaid Group Practice

Willard G. Manning, Jr.

January 1983

Prepared for the Health Care Financing Administration,
U.S. Department of Health and Human Services

35th
Year



PREFACE

This report was prepared for the Health Care Financing Administration, U.S. Department of Health and Human Services, under Contract No. 500-78-0048. The research summarized here examines the differences between fee-for-service and a health maintenance organization (HMO) in the use of pathology (or "laboratory") services.

Related Rand reports on pathology services conducted under this contract are P. M. Danzon, Profits in Hospital Laboratories: The Effect of Reimbursement Policies on Health Costs and Charges, R-2582-HCFA, September 1980; P. M. Danzon, Economic Factors in the Use of Laboratory Tests by Office-Based Physicians, R-2525/1-HCFA, August 1982; M. S. Marquis, Laboratory Test Ordering by Physicians: The Effects of Reimbursement Policies, R-2901-HHS, August 1982; and P. M. Danzon, W. G. Manning, and M. S. Marquis, Factors Affecting Laboratory Test Use and Prices: Executive Summary, R-2987-HCFA, January 1983.

This report should be of interest to decisionmakers who formulate health policy and to members of the research community who are concerned about the choice between HMOs and fee-for-service.

SUMMARY

Health Maintenance Organizations (HMOs) have been widely discussed as an alternative to fee-for-service third-party reimbursement for containing medical expenses (Enthoven, 1978a,b, 1980). Arguments in favor of HMOs are that they provide incentives to reduce utilization, especially "unnecessary" hospitalization, while increasing access to ambulatory medical care and encouraging preventive care. Arguments against HMOs are that they encourage underuse of medical services, and that the evidence on differences in patterns of use between HMOs and fee-for-service coverage may reflect self-selection of people into (or out of) HMOs.

The debate over differences in patterns of use of medical services between HMO and fee-for-service plans has remained unresolved because, in past studies, insurance coverage (HMO versus fee-for-service) was self-selected. In this study we examine the HMO versus fee-for-service issues using data on the use of pathology services from a randomized controlled trial in health insurance, the Rand Health Insurance Study (HIS). The design of that study provides a unique opportunity to compare the use of medical services in an HMO and a free fee-for-service plan, and to assess the relative importance of different provider incentives as well as self-selection of enrollees.

This report compares the utilization of pathology services in three insurance plans with free care (no out-of-pocket costs for medical services in both ambulatory and inpatient settings). One plan is a free fee-for-service plan in Seattle, Washington. The other two plans are prepaid group practice plans at Group Health Cooperative (GHC) of Puget Sound, which serves the Seattle metropolitan area. The enrollees in the free fee-for-service plan and one of the GHC plans (the experimentals) are random samples of the Seattle fee-for-service population in 1976. Enrollees in the third plan, the GHC controls, are a random sample of the GHC population in 1976. All three plans exclude the aged, the Medicare disabled, the military and their dependents, and institutionalized persons.

By comparing the free fee-for-service plan and the GHC experimentals, we can assess the effects of different organizational structures on utilization. Unlike most comparisons of fee-for-service and prepaid group practices, this comparison of the use of pathology services will not be confounded by differences in cost-sharing (fee-for-service plans and HMOs often have different copayments) or self-selection into the alternative institutional settings. By comparing the GHC experimentals and controls, we can determine the effect of self-selection on utilization of pathology services.

The analysis of variance results indicate that, at a gross level, there appear to be no statistically significant differences in the utilization of pathology services among the three plans. A closer analysis, however, using more robust estimation techniques, produced three important results. First, there are no significant ambulatory differences between the free fee-for-service plan and the GHC plans. Second, the fee-for-service plan has a significantly greater use of inpatient pathology services than either GHC plan. These results confirm experimentally the results found in the literature for all medical services in a situation where the HMO versus fee-for-service comparison is not confounded by either self-selection or by differences in copayment.

The third major finding is that self-selection into (out of) an HMO does not appear to have an appreciable effect on the level of use of inpatient and outpatient pathology services. The GHC experimentals and GHC controls, the latter a self-selected population, did not behave differently. This was true whether one adjusted for all the measurable characteristics (including health status) or merely for age and gender. The GHC controls did have slightly higher use of ambulatory pathology services, but this was attributable to differences in the age and gender mix on the plans.

These findings may be useful in the policy debate over fee-for-service versus HMOs. HMOs do have lower use than fee-for-service, largely because of reduced inpatient use. This lower use appears to be the result of institutional differences instead of self-selection.

ACKNOWLEDGMENTS

Special thanks are due to my Rand colleagues Arleen Leibowitz, Patricia Danzon, George Goldberg, Susan Marquis, and Joseph Newhouse, for their comments and suggestions. Careful reviews by Richard Handschin, Michael Murray, and Frank Sloan markedly improved the report.

Ken Haber deserves a grateful thanks for his support as project officer during the research.

CONTENTS

PREFACE	iii
SUMMARY	v
ACKNOWLEDGMENTS	vii
FIGURES	xi
TABLES	xiii

Chapter

1. INTRODUCTION	1
2. THE DESIGN OF THE HEALTH INSURANCE STUDY, THE DATA AND THE SAMPLE	4
The Design	4
The Sample	8
Dependent Variables	9
Independent Variables	11
3. EMPIRICAL METHODOLOGY	15
Rationale for the Two-Part Model	15
Correlation in the Error Terms	20
4. EMPIRICAL FINDINGS	23
Adjusted Plan Comparisons	24
Alternative Reference Populations	25
Effect of Level of Adjustment	28
5. CONCLUSIONS	31

Appendix

A. ALTERNATIVE MODELS	35
B. SUPPORTING TABLES	37

REFERENCES	41
------------------	----

FIGURES

3.1. Normal Plot for Ambulatory Pathology Use	16
3.2. Normal Plot for Log (Positive Ambulatory Pathology Use)	19
3.3. Normal Plot for Residuals from Log (Positive Inpatient Use)	19

Chapter 1

INTRODUCTION

Cost control has been one of the major issues in the policy debate on health services. Both health maintenance organizations (HMOs) and cost sharing in fee-for-service have been offered as (partial) solutions to the problem of hemorrhaging health care costs. Proponents of HMOs argue that the provider will not overprescribe treatment because the HMO, as provider and insurer, bears the full costs of its actions (Enthoven, 1978a,b, 1980; Enthoven and Noll, 1979). In contrast, HMO advocates argue that, under third-party reimbursement, fee-for-service providers have an incentive to overutilize health services because the cost to the patient is smaller than the actual resource-cost of providing care. The insurance company pays the remainder. Supporters of this argument cite evidence from past studies that total medical costs and the use of complement services are lower in HMOs than in fee-for-service plans (Richardson et al., 1980; Luft, 1981). Arguments against HMOs are that they encourage underuse of medical services (Schwartz, 1978) and that the evidence on use between HMOs and fee-for-service coverage may reflect self-selection of people into (out of) HMOs (Luft, 1981; Richardson, et al., 1980; Berki et al., 1977, 1978).

The empirical evidence for these contentions is weak because of two major flaws in the available data. First, many comparisons of fee-for-service and HMOs are confounded by different levels of cost-sharing by patients. That is, it is not clear whether it is the cost-sharing differences or the institutional structures that are responsible for the results. Second, the empirical evidence frequently contains self-selection biases of unknown magnitude and direction. In many cases, an employee or private enrollee can choose between fee-for-service and the HMO. Sicker persons, who anticipate higher health expenditures, may choose the system with lower out-of-pocket costs (premiums plus copayments). Then, the higher utilization on

the fee-for-service plans may reflect either more sickness or an institutional difference in incentives for patients and providers.[1]

In this report, we will address part of the HMO versus fee-for-service debate by examining how different institutional settings affect the use of pathology (or "laboratory") services. The data available here do not have the two flaws mentioned above, which are often found in other studies. The data are from the Health Insurance Study, a social experiment in health insurance being conducted for the U.S. Department of Health and Human Services. In that study, there are three plans of interest to the fee-for-service versus HMO comparison. Random samples of the Seattle, Washington, population were enrolled in a free (no out-of-pocket cost) fee-for-service plan and a free HMO plan at Group Health Corporative of Puget Sound (GHC). Comparisons between these two plans provide estimates of the difference between fee-for-service and one HMO that are free of the confounding influences of both cost-sharing and self-selection. A third group, a random sample of persons who had already chosen GHC, acts as a control group. Differences between the GHC control and the GHC experimental groups provide a measure of the self-selection effect at GHC.

By comparing the utilization of pathology services on these three plans, we can disentangle the effects of self-selection and institutional arrangements. These data indicate that there are no significant differences in ambulatory (outpatient) pathology utilization among the three plans when age, gender, family characteristics, and health status are included as covariates. The self-selected control group does use insignificantly more ambulatory pathology services than age and gender alone would indicate, or than age, gender, and health and family characteristics would indicate. The free fee-for-service plan uses significantly more inpatient pathology services than either GHC plan. With or without adjustment for health characteristics, the two GHC plans are indistinguishable in their use of inpatient pathology services.

This study is one of a number of studies examining issues related to HMO versus fee-for-service comparisons and the utilization of

[1] Luft (1981) provides an excellent review of the issues and the literature.

pathology services. Work in progress on the Health Insurance Study is examining the HMO versus fee-for-service differences in aggregate medical utilization. Danzon (1980, 1982) and Marquis (1982) examine issues related to the effects of reimbursement practices, as well as the effects of the use of independent laboratories, on the use of pathology services. Many of the policy issues related to laboratory services are examined in those studies and not dealt with here.

This report is organized as follows. Chapter 2 describes the Health Insurance Study, the sample, and the data. Chapter 3 describes the empirical methodology for the comparisons. The nature of the data requires more careful modeling of the distribution of expenditures in order to obtain reliable estimates. Chapter 4 presents the empirical results, summarized above.

Chapter 2

THE DESIGN OF THE HEALTH INSURANCE STUDY, THE DATA AND THE SAMPLE

The Health Insurance Study (HIS) is a social experiment designed to study how health insurance affects the utilization of health services and the health status of individuals. Past studies in these areas have typically used nonexperimental data sources that suffer from several flaws:

- o Insurance policies are difficult to describe parametrically.
- o Utilization data are frequently based on recall or on limited claims data. Both sources are often subject to reporting biases.
- o Insurance coverage is frequently endogenous. For example, in many HMOs, the enrollees had a choice of several insurance plans. Sicker individuals may have chosen the plan with the lowest out-of-pocket cost.

The HIS was designed to avoid these problems.

THE DESIGN

The HIS enrolled a sample of nearly 7700 individuals in about 2700 families in six sites: Dayton, Ohio; Seattle, Washington; Fitchburg, Massachusetts; Franklin County, Massachusetts; Charleston, South Carolina; and Georgetown County, South Carolina.

Because our concern in this report is to estimate the effect of health insurance on demand for pathology services, we limit the description of the HIS design to its insurance plans. Newhouse (1974) and Brook et al. (1979) provide fuller descriptions of the study. The experiment contains fee-for-service and prepaid group plans. For this research, we have used data from the Seattle site because that is the only site with HMO participants in the experiment.

The HMO in this study is Group Health Cooperative of Puget Sound (GHC).[1] GHC is a nonprofit, prepaid group practice operating two

[1] These participants can also receive services outside GHC at their own expense.

hospitals, three mental health clinics, and 13 medical clinics in the Seattle area.

Families participating in the Seattle site were assigned to one of the two HMO plans or to one of 11 different fee-for-service insurance plans. In this analysis we use only the data for the two HMO plans and the free fee-for-service plan; comparisons with the other fee-for-service plans would confound the HMO versus fee-for-service comparison with cost sharing. (See Table 2.1 for the number of persons enrolled on each plan in Seattle.) About half of this sample was enrolled experimentally in GHC; this plan contains a random sample of the Seattle population. Participants in this plan received all outpatient and inpatient services at GHC for free (no out-of-pocket cost). Services covered for the GHC experimental population were designed to be as inclusive as those available on the HIS fee-for-service plans. The participants in the GHC experimental plan were not reimbursed for medical expenses incurred outside GHC unless these were GHC referrals or were for covered services not provided at GHC.[2]

Another third of this sample is the GHC control group, which is a random sample of families who were already enrolled in GHC. With three exceptions, the controls received all outpatient and inpatient care at GHC for free. First, there was some copayment for prescription drugs. Second, GHC control participants' plans limited free mental health outpatient visits to ten; beyond that number there was a copayment of \$8.00 to \$10.00 per visit, with no upper limit on the number of visits. Third, for most controls, benefits did not include inpatient psychiatric services. However, the hospitalization rate for inpatient mental health services is such a small percentage of total admissions that the conclusions about differences in hospitalizations should be only slightly affected.

[2] To encourage filing of claims for out-of-plan use not reimbursed by GHC, these participants were reimbursed a nominal five percent. These data on out-of-GHC use not reimbursed by GHC are not included in this analysis.

Table 2.1

SAMPLE IN HEALTH INSURANCE STUDY AT
INITIAL ENROLLMENT IN SEATTLE

Insurance Plan	Number	Percent ^a
Fee-for-service plan (free)	431	18.6
GHC plans		
Experimentals	1149	49.5
Controls	742	32.0
Total	2322	100.1

NOTE: Includes only Seattle plans used in this analysis.

^aDoes not sum to 100 because of rounding error.

The remainder of the people in the Seattle sample used in this analysis were enrolled in a fee-for-service plan with a zero coinsurance rate (they received free care).

The free fee-for-service plan and the GHC plans covered the same wide variety of services.[3] The only significant exclusion for this analysis was lack of reimbursement for inpatient mental health services in the GHC control group.[4]

Experimental GHC and free fee-for-service families were enrolled in the insurance plans as a unit, with only eligible members participating. No choice of plan was offered to experimentals; the family could either accept the experimental plan or choose not to participate.[5] To

[3] See Clasquin (1973) for a discussion of the reasons for the HIS structure of benefits.

[4] Nonpreventive orthodontia, cosmetic surgery (not related to preexisting conditions), and psychotherapy visits (in excess of 52 visits per person per year) were not covered on the experimental plans. In the case of each exclusion, it is questionable whether anything could have been learned about steady-state demand during the three-to-five-year lifetime of the experiment.

[5] The GHC controls are an exception because they had already selected a GHC plan. They also had the ability to choose which family members would be enrolled at GHC.

encourage acceptance, experimental families were given a lump-sum payment equal to their worst-case financial risk associated with the plan; thus, no family was worse off financially for being in the study.[6] The amount of the lump-sum payment was independent of use of health care services. It therefore should be considered a temporary change in income and should not affect the response to cost sharing.[7]

The families were assigned to experimental plans using the Finite Selection Model (Morris, 1979). This model is designed to achieve as much balance in the distribution of characteristics (e.g., age, sex) as is possible in a finite sample while retaining some randomization. The model ensures that experimental plans (e.g., all fee-for-service plans and the GHC experimental plan, but not the GHC control plan) are uncorrelated with the demographic covariates. Because the control group at GHC is a random sample of a self-selected population (GHC's 1976 membership), it may differ in its characteristics as well as its behavior from the other Seattle plans.

The sample in each site is a random sample of each site's population, but with the following groups ineligible for participation:

- (1) Those who were or would become eligible for "regular" Medicare at some time during the HIS;
- (2) Those with family incomes in excess of \$25,000 (in 1973 dollars);
- (3) Those eligible for the Medicare disability program;
- (4) Those in institutions (e.g., in jails);
- (5) Those in the military or their dependents; and
- (6) Those with military service-related disabilities.

[6] The family's nonexperimental coverage was maintained for the family by the HIS during the experimental period, with the benefits of the policy assigned to the HIS. If the family had no coverage, the HIS purchased a policy on their behalf. Thus, no family could become uninsurable as a result of their participation in the study.

[7] Some preliminary work on fee-for-service medical expenditures indicates that participants treat this payment as they would a transitory change in income. Only a small percentage of that payment appears to be used directly for medical expenses.

THE SAMPLE

The sample for this report consists of those individuals enrolled for a full year in either or both of the first two years of the Seattle site on the free fee-for-service plan or either of the Group Health Cooperative plans. The sample excludes individuals with partial years of participation in the Seattle area, participants who were suspended from the experiment because they entered a group from which we did not sample (e.g., for entering the military or being sentenced to jail), and participants who voluntarily quit the study, were involuntarily terminated for noncompliance during the year, died, or left the Seattle area.[8] For example, a person who was suspended in year I was included for Year II if he participated for all of year II. Similarly, an infant born in year I would be included only in year II, if he was eligible for benefits for all of the second year.

Partial-year enrollees were excluded because the models that we use in this report are not statistically well behaved for this group. The models rely on characteristics of the distribution of laboratory use, which can be observed only for a full year of data. For example, the positive use is approximately log normal. Partial-year data are not as well behaved because this distribution does not convolute.[9]

In other preliminary work on the HIS, we have checked whether partial-year enrollees behave like a prorated full-year enrollee. With the exception of individuals who died, we cannot reject the hypothesis that partial-year participants act like full-year participants. In this study, we excluded only one participant because of death. Consequently, we do not believe that our sample exclusions appreciably bias the results presented.

Table 2.2 contains the number of observations (persons) per plan by year. The exclusions would add about 18 percent to the sample. The change from year to year reflects death, attrition, termination,

[8] Individuals leaving the area could not receive services from GHC. To maintain comparability, we also deleted out-of-area moves in the free fee-for-service plan. Individuals in the GHC control plan who lost their GHC coverage (e.g., because of change of employment) were also deleted from the sample.

[9] The sum of two log normals is not log normal unless both are perfectly correlated.

suspension, and moves, net of births and adoptions. The GHC control plan has greater year-to-year losses than the other two plans. This occurs because, for some controls, change of employer can result in a loss of GHC coverage. The free fee-for-service and GHC experimentals maintain their insurance plan when they change employers.

DEPENDENT VARIABLES

In this analysis, we examined the inpatient and outpatient use of laboratory or pathology services. We measured annual laboratory services by expenses for ambulatory pathology (AMBPATH) and inpatient pathology (INPPATH) services, not visits. For the fee-for-service plans, the data are abstracted from insurance claims data submitted from hospitals, laboratories, and medical providers. For the GHC plans, the data are abstracted from the GHC medical record.

In comparing laboratory or pathology use across plans, we must compare different bundles of pathology services. To facilitate such a comparison, we need a single measure of utilization. In the fee-for-service, there exists such a measure to summarize utilization--actual expenditures. In an HMO, no such ready measure exists.

Fortunately, both fee-for-service and GHC assign California Relative Value Studies (CRVS) units to ambulatory services rendered. These data can be used to construct a single measure, an estimated expenditure, which allows us to compare use in different settings. In the case of ambulatory care, and for pathologists who bill separately

Table 2.2

SAMPLES BY PLAN YEAR

Plan	Year I	Year II
Free fee-for-service	383	358
Group Health Cooperative experimentals	1070	1011
Group Health Cooperative controls	667	593

NOTE: Samples include only full-year participants.

from hospitals, the measure of pathology use is based on the 1974 CRVS units assigned to each procedure in the 80000-89999 procedure code range. For procedures that have no units, a physician reviewed the procedure and assigned a unit value. To convert CRVS units to an estimated expenditure, all CRVS units were multiplied by a factor of \$.90 for 1976 services, \$1.00 for 1977 services, and \$1.10 for 1978 services. These are the factors that GHC used to bill for any pathology services rendered to non-GHC participants.[10]

For inpatient stays if the pathology services were rendered by the hospital instead of by a pathologist who billed separately, we did not have CRVS unit information to create our estimated expenditure measure. In fee-for-service cases, we used actual charges for inpatient laboratory services. In the GHC cases, the data give the amount that GHC would have charged a non-GHC participant for the service. Thus, the GHC inpatient and all ambulatory use are comparable with each other. However, the fee-for-service inpatient amounts may not be completely comparable with the GHC and CRVS imputed amounts. The models proposed below allow us to see how sensitive the results would be to any underlying incomparability between fee-for-service inpatient and the fee-for-service ambulatory and GHC-estimated expenditures.

This analysis uses expenses rather than the more traditional visits for two reasons. First, using visits would lose all of the information about the intensity of utilization. Some visits are for minor complaints that require a limited use (if any) of ancillary services, such as pathology, while other visits are for more serious conditions that may require extensive lab work. An expenses measure captures this intensity, but a visit measure will not.

Second, comparisons based on a count of visits would be quite sensitive to data-collection techniques. Use-data for fee-for-service is based on claims data, while those for GHC are derived from medical records. In fee-for-service, multiple providers may bill for parts of the same service. For example, a hospital laboratory, the hospital emergency room, and the private physician may all bill separately for

[10] We use GHC weights in order to keep inpatient use at GHC and all ambulatory use in comparable terms.

services rendered during one emergency room visit. Unless all of the utilizations are perfectly linked, we will count more than one fee-for-service visit for this use. In contrast, the GHC data would provide only one visit. Thus, although the content of the visit is identical for both systems, the measures will differ.[11] Expenses measures avoid this data-methods effect.

INDEPENDENT VARIABLES

There are two plan indicator variables, one each for the free fee-for-service plan (P00) and the GHC controls (GHCC). The GHC experimentals (GHCE) are the omitted group. Thus, the coefficients are measures of (1) the difference between free care in an HMO and in a free fee-for-service plan (the P00 coefficient) when there are no selection effects, and (2) the selection effect from people who voluntarily join a prepaid group practice as compared with a random sample of the population (the GHCC coefficient).

The model specification includes covariates for other experimental treatments, age, gender, race, family size and income, and self- or parent-reported health status. Table 2.3 lists these other covariates.

The other experimental treatments are given by: whether one received a physical screening examination at the beginning of the study (TOOKPHYS) or not; how often one was required to file a diary on sick-loss days and utilization (HRTYPE); and whether one was enrolled for three years (TERM3) or for five years. The first two treatments were introduced to test for Hawthorne effects in the study associated with

[11] In the preliminary analysis for this report, a visit approach was used, because we wanted to comment on the use of pathology services per visit. Although overall ambulatory pathology use was the same on all the plans, the visits-level data showed differences. A systematic examination of the discrepancy revealed that the source of the problem was the lower quality of linking for fee-for-service data. A particular problem was linking data for pathology use billed for by independent labs, because the date on the record was frequently the date of billing, rather than the date of service. Such a problem did not exist at GHC. Rather than report results which were probably methods effects, we used an approach that avoided the potential bias.

Table 2.3

INDEPENDENT VARIABLES

Treatment Variables

Indicator variables

GHCC = 1 if GHC control group
 GHCE = 1 if GHC experimental group
 P00 = 1 if free fee-for-service plan
 TOOKPHYS = 1 if received physical examination at enrollment^a
 TERM3 = 1 if enrolled for 3 years^b
 HRTYPE = 0 if did not file health diary
 = 2 if filed biweekly

Sociodemographic Variables

Indicator variables

BLACK = if person is black
 AFDC = if any family member on AFDC
 FEMALE = if female
 CHILD = if age <18
 FADULT = if adult woman
 YR2 = if year 2^c

Continuous variables

LINC = ln (family income in 1973 dollars)^d
 LFAM = ln (family size)
 AGE, SQAGE = age, square root of age. Newborns have an age of zero.

Health Status Measures

Indicator variables

PHYSLM = 1 if physically limited
 POSDIS = 1 if any chronic disease
 GHINMIS = 1 if no value for GHINDX
 FLAG0 = 1 if no health questionnaire at enrollment
 FLAG1 = 1 infant form of health care questionnaire,
 no mental health scale
 FLAG2 = 1 pediatric form of health care questionnaire,
 mental health scale with limited number of items^e

Continuous variables

LNDIS = log (max (number of chronic diseases, 1))
 MHI = mental health index
 GHINDX = general health index

NOTE: Indicator values = 0 or 1.

^aNo examination is the omitted category.

^bFive years is the omitted category.

^cYear 1 is the omitted category.

^dIncome is set equal to \$1000, if reported to be less.
 The year is 1976 in Seattle.

^eThe omitted group is the adult form of the Enrollment Medical
 History Questionnaire.

data-collection methods for initial health status and reporting sick-loss activity. The three- versus five-year split was introduced to test the transitory effect of the study. If transitory effects are prominent, there should be behavioral differences between the two groups in the third year as the three-year group exits.

The resources available to family members are measured by LINC, LFAM, and AFDC. LINC is the natural logarithm of the maximum of 1976 family income or \$1,000. LFAM is the natural logarithm of the number of individuals in the family during the accounting year. AFDC is an indicator for whether the family head or any adult in the family received Aid to Families with Dependent Children.

The age and gender variables include indicators for being a child (age <18 at the beginning of the year), for being female, and for being a female adult. In addition, we include age and the square root of age; these two age variables give an age function that parallels age differences in utilization of ambulatory services, other than those for maternity care.[12]

The measures of health status are based on self-reported or parent-reported answers to questions asked at enrollment. The variables include measures of general health, physical limitation, mental health, and the number of chronic disease complaints (applies to ages 14+ only).[13] Those measures are coded so that higher scores indicate better general health, lower physical functioning, better mental health and more chronic complaints.[14]

[12] For this analysis, maternity care is treated as inpatient because fee-for-service providers typically bill for maternity care in a single fee which covers both prenatal care and delivery.

[13] See Davies and Ware (1981) for the general health and disease measures; Stewart, Ware, and Brook (1981) for the physical functioning measure; Veit and Ware (draft) for the mental health measure; and Eisen, Donald, Ware, and Brook (1980) for the health measures for children.

[14] Missing data are replaced with the site mean. Flags were introduced for different sources of missing data. Practically all missing data are due to form nonresponse or to particular groups (e.g., children aged <5) not being targeted for the measure.

The transformations for income, family size, age, and chronic disease were selected to provide a better model that would not be overly influenced by a few large outlying values for these variables, and to provide a homoscedastic error. For the analysis of inpatient use, the independent variables include plan variables, age and gender indicators, and the general health measure and its missing-value indicator. The more parsimonious specification was dictated by the small numbers of inpatient cases and a fear that we would overfit the model.

Chapter 3

EMPIRICAL METHODOLOGY

In this analysis we have used a two-part model to estimate the demand for pathology services, rather than the more common analysis of variance (ANOVA) and analysis of covariance (ANOCOVA) techniques.

The distribution of pathology use has three characteristics that make it difficult to obtain reliable estimates of the effect of covariates on the demand for pathology services using either ANOVA or ANOCOVA. First, a substantial fraction of the population does not use any pathology services during the year. Second, the remaining positive use is very skewed. Through much of its range, positive use is approximately log normal. Third, both tails of the distribution of positive use are heavier than even the log normal.

These characteristics imply that ANOVA and ANOCOVA techniques will yield very imprecise (though unbiased) estimates of the effects of health insurance on the use of pathology services, even for a sample as large as the one used here. As Duan, Manning, Morris, and Newhouse (1982) have shown, a statistical model that exploits the characteristics of the expenditure distribution can provide more precise estimates.

This Chapter briefly describes the statistical methods used in this analysis. The topics include the rationale for the two-part model, the formal statistical model for estimation and prediction, and the correction of inferences for correlations in the error.

RATIONALE FOR THE TWO-PART MODEL

With the sample sizes available, the nature of the distribution of pathology use causes an ordinary least-squares (OLS) regression of pathology use on the covariates (either ANOVA or ANOCOVA) to yield very imprecise estimates. OLS is the best linear unbiased estimator if the error is independent of the covariates. OLS is a maximum likelihood estimator if the error is normally distributed. However, if the data are not normally distributed, then OLS estimates will be quite sensitive to extreme values. In the case of pathology use, the distribution is

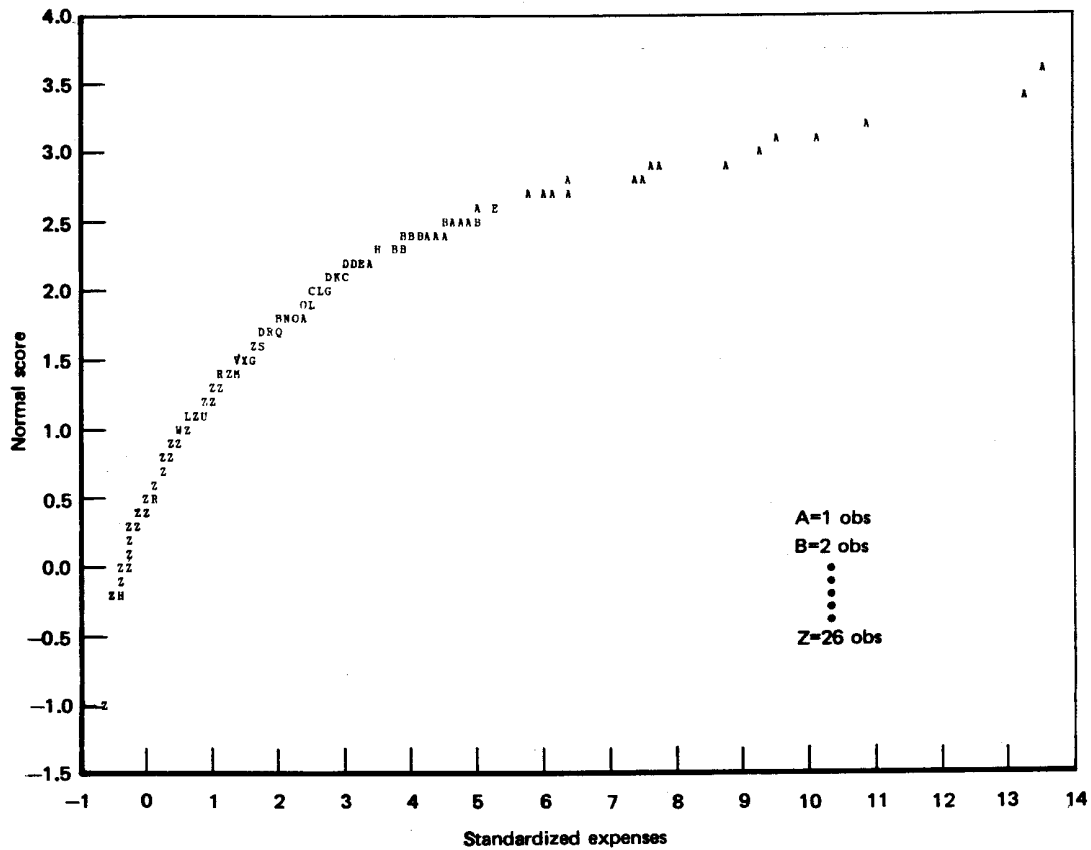


Fig. 3.1 -- Normal plot for ambulatory pathology use

highly skewed toward the positive side. Figure 3.1 provides a normal plot of the ambulatory use of pathology services for both years, unadjusted for covariates.[1] The data clearly deviate from a straight line, as required for normality. The distribution of residuals after fitting an OLS model are also highly skewed.

[1] The normal plot is sometimes called a Q-Q (Q for quantiles) plot. These figures plot the quantiles (percentiles) $F^{-1}(p)$ of the empirical distribution against the quantiles of a normal distribution with the same mean and variance. If the empirical distribution is normal, the quantiles will have the same values, and a plot of the quantiles will fall on a 45-degree line. In these plots, the axes are measured in \log units as deviations from the mean.

Although ANOVA is unduly influenced by the extreme values found in pathology services data,[2] we will examine ANOVA results as a first rough cut at the analysis.

The two-part model uses two equations to explain the utilization of pathology services. We have modeled pathology expenses by separating the enrolled population into two groups: nonusers and users of pathology services; the model also examines the expenses of the users. The first equation, the ANY USE equation, is a probit equation for the probability that a person will receive any pathology service during the year.[3] The equation therefore separates the users from the nonusers, and thus addresses the first characteristic described above. The second equation, the LEVEL OF USE equation, is a linear regression for the logarithm of annual pathology expenses, given that there is any use.[4] The log transformation of expenses practically eliminates the undesirable skewness in the distribution of expenses among users, described above as the second characteristic. The log transformation yields nearly symmetric and roughly normal error distributions, for which the least-squares estimate is efficient. Figures 3.2 and 3.3 are the normal plots for the residuals from the level of outpatient and inpatient use equations; the residuals are clearly more normal than the

[2] In the split-sample, cross-validation work on medical data, Duan, Manning, Morris, and Newhouse (1982), hereafter referred to as DMMN, did find that ANOVA was more reliable than ANOCOVA on untransformed data.

[3] More formally, the first equation is a probit equation for the decision to have positive use. For an individual with characteristics x_i , a column vector, we have

$$I_i = x_i' \beta_1 + \varepsilon_{1i}, \quad (3.1)$$

where I is an unobserved propensity to have positive use, which if positive implies that pathology use is positive, and if negative implies zero use. The error ε_1 is assumed to be independently and normally distributed with zero mean and unit variance.

[4] The second equation is:

$$\ln(\text{PATH} | \text{PATH} > 0)_i = x_i' \beta_2 + \varepsilon_{2i}, \quad (3.2)$$

where the error is independently distributed with mean zero and variance σ_2^2 .

data in Fig. 3.1. Therefore the two-part model estimates should be more precise than those obtained from ANOVA or ANOCOVA.[5]

This approach allows us to model the behavior of patients with no use differently from those with nonzero expenses. Although the model was developed to solve a statistical problem, the separation of users from nonusers does have a behavioral interpretation. The decision to have any care depends on patients' decisions to seek medical care and providers' decisions to order tests. The provider alone largely influences the decisions about the level of care.

A consistent estimate of the expected pathology expense for an enrollee is

$$E(\text{Pathology } \$) = \text{probability that Pathology } \$ > 0$$

* Pathology \$ if a user,

which, for someone with characteristics x_i , is

$$E(\text{pathology } \$) = \hat{p}_i \exp(x_i' \hat{\beta}_2) \hat{\phi},$$

where

$$\begin{aligned} \hat{p}_i &= \Phi(x_i' \hat{\beta}_1) = \text{estimated probability of any pathology use;} \\ \exp(x_i' \hat{\beta}_2) \hat{\phi} &= \text{estimated expense for pathology, given that there is use;} \\ \hat{\phi} &= \text{estimated retransformation of the error term in the} \\ &\quad \text{level-of-use equation. [6]} \end{aligned}$$

[5] The two equations of this model can be estimated separately because the log likelihood is separable in the parameters of Eqs. (3.1) and (3.2). This separability does not depend on any independence assumption for the error structure. Instead it is a result of the way the conditional estimates are calculated. For a more extensive discussion, see DMMN (1982) or Manning et al. (1981).

[6] This analysis uses the smearing estimate of ϕ . While the use of the two-part model is motivated by the need to approximate the normal assumption as closely as possible, the error distribution for the level-

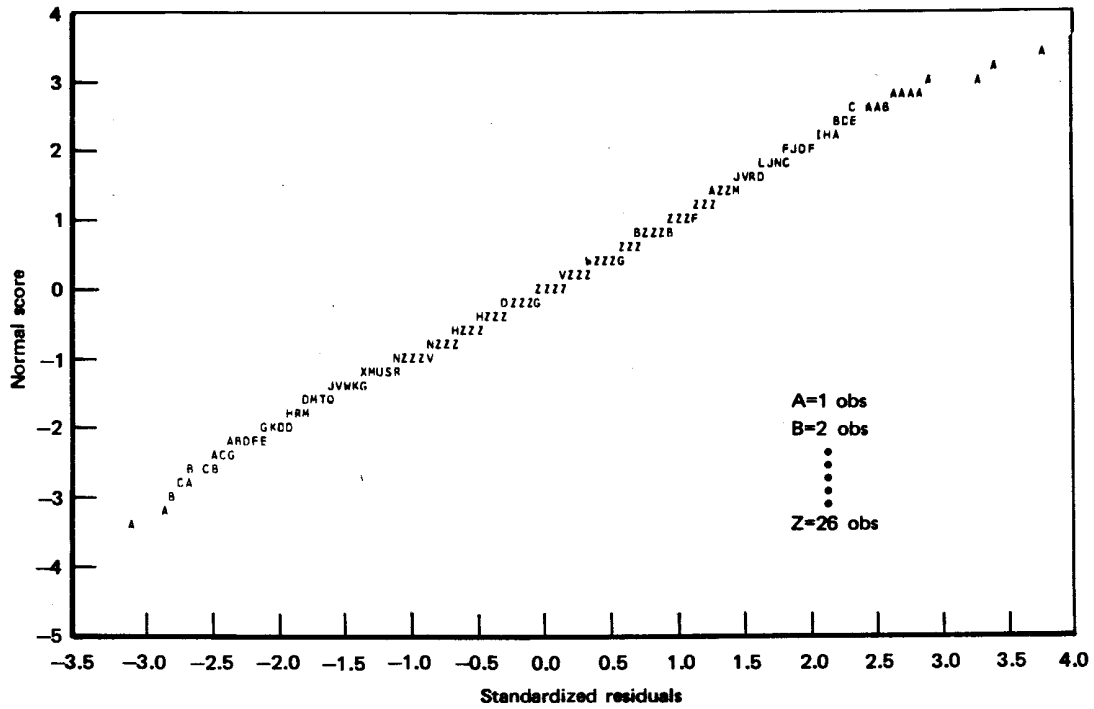


Fig. 3.2 -- Normal plot for log (positive ambulatory pathology use)

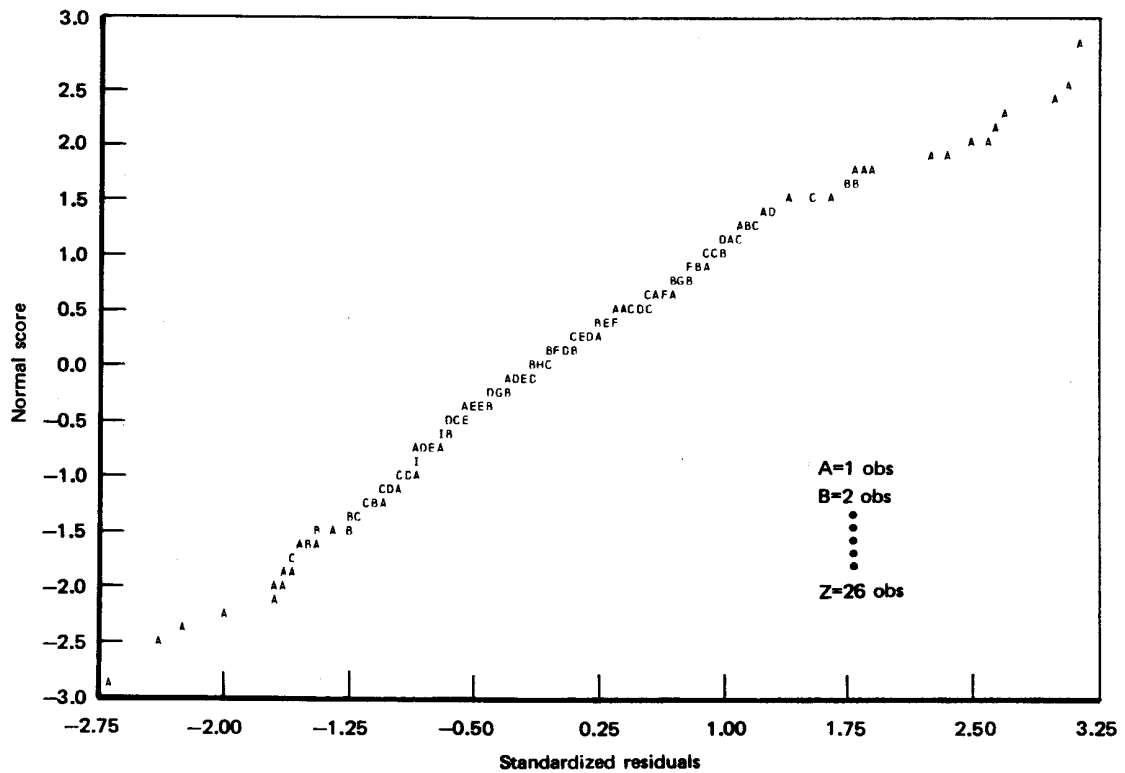


Fig. 3.3 -- Normal plot for residuals from log (positive inpatient use)

In this report we analyze the inpatient and outpatient use of pathology services separately, for two reasons. First, much of the literature on the HMO versus fee-for-service comparison suggests different responses for outpatient and inpatient care. Second, a single two-part model for all pathology use would yield biased estimates because of the differences in the probability of any use, and mean and variance of use between the two services. The use of two two-part models solves this problem.[7,8]

CORRELATION IN THE ERROR TERMS

Although there are observations on several thousand person-years of data, there are not that number of independent observations because of substantial positive correlations in the error terms among family

of-expense equations still deviates appreciably from the normal assumption. If we used the normal theory retransformation ($\exp(\sigma_2^2/2)$), the predictions would be inconsistent. Instead we use a nonparametric estimate, the smearing estimate developed by Duan (1982b), which is the sample average of the exponentiated residuals. The smearing estimate is statistically consistent for the retransformation factor ϕ if the error distribution does not depend on the characteristics x_1^i .

[7] With HIS data, a different model has been used to analyze medical utilization. DMMN (1982) used a four-part model that had separate equations for the decision to have any medical expenditures, the decision to have inpatient utilization conditional on positive medical use, a decision about the level of log-positive ambulatory use conditional on no inpatient use, and a decision about the level of log-positive medical use conditional on positive inpatient use. The four-part model of DMMN (1982) and separate two-part models for inpatient and ambulatory use are different "accounting" methods for looking at the same data, each with its own advantages and disadvantages. The four-part model is designed to predict aggregate use and to obtain standard errors for that aggregate prediction. Its disadvantage is that the composition of the response cannot be examined because of the way the conditioning in the last three parts works; the conditional equations refer to total use, not ambulatory and inpatient use separately. The use of two two-part models allows us to examine the ambulatory and inpatient pathology response separately. Unfortunately, the presence of correlation between ambulatory and inpatient use does not allow us to obtain either efficient estimates of aggregate use or to obtain easily the correct standard errors for the aggregate prediction by using two-part models. However, the predictions from two two-part models are consistent. Because the predictions in Chap. 4 are correlated, we must be careful not to interpret them as independent realizations.

[8] We also rejected the use of the tobit and self-selection (Heckman) models. For further detail see App. A.

members and over time among observations on the same person. These correlations exist in both equations: in the equation for any use of pathology services, and the equation for the level of pathology expenses if any use. Failure to account for these correlations in the analysis yields inefficient estimates of the coefficients and statistically inconsistent estimates of the standard errors. As a result, the inference statistics (e.g., t and χ^2) calculated in the usual way (without adjusting for these correlations) can be too large. In the probit equation, for any use of ambulatory pathology services, failure to account for these correlations leads to t statistics that can be as much as 1.09 times too large for individual characteristics (e.g., age, sex), and 1.39 times too large for family characteristics (e.g., plan and family income).

For the level-of-expense equations, we have used a nested variance components model to approximate the pattern of correlations in our data.[9] The error term in each equation is assumed to be the sum of three error components: a random, unobserved, family-specific error component, a random, unobserved, individual-specific error component (both of the preceding two are assumed to be constant over time), and a random, unobserved, time-specific error component.[10] The regression coefficients and standard errors for the expense equations are actually estimated using generalized least squares. Estimates of the family and individual correlations are obtained using OLS residuals. The regression coefficients are then reestimated by generalized least squares.

We considered a similar estimation model for the probit equation, but with the unbalanced design (unequal family sizes and unequal numbers

[9] See Balestra and Nerlove (1966), and Searle (1971, Chaps. 9-11), for a discussion of this model in a regression context.

[10] We cannot estimate the family and individual components as fixed effects, because we would then be unable to estimate the effects of insurance plan and other covariates. For example, there would be no variation left in insurance plan variables because insurance coverage is constant within families and individuals.

of person-years), the computation for the multivariate probit model would have been prohibitively expensive. Instead, we estimated those equations with univariate probit equations, treating observations as if they were stochastically independent; we then estimated an upper bound for the standard error. Because the precision associated with the probit equations contributes only a small fraction to the overall variance of the prediction for total pathology expenditures, we lose little by bounding standard errors in this manner.

The method for bounding the standard errors is described in Duan (1982a) for the single variance component case. With two variance components in our probit equations, we judge that an adaptation of the correction in Duan (1982a) applies. For family-level variables, such as plan, we bound the error covariance matrix by

$$\text{Cov}(\varepsilon_j) \leq (1 - \rho_{\text{FAM}} - \rho_{\text{IND}})I_N + (\rho_{\text{FAM}} + \rho_{\text{IND}})D_{\text{FAM}}, j = 1, 2$$

ρ_{IND} is a pure intraperson correlation (net of family effect), and D_{FAM} is a block diagonal matrix with a block of 1's for each family. We then apply Duan's correction to the upper-bound error covariance matrix (which is equivalent to a single variance component model). For individual-level variables such as CHILD and AGE, we apply Duan's correction only to the intraperson correlation structure, neglecting the intrafamily correlation among persons in the same family. (Such a modification is plausible because most of our person-level variables are nearly orthogonal to the family clusters.)

All inference statistics (e.g., t , χ^2) reported below have been corrected for correlation.

Chapter 4

EMPIRICAL FINDINGS

The aggregate utilization of pathology services does vary slightly across the three insurance plans: free fee-for-service, GHC experimentals, and GHC controls. Table 4.1 contains the analysis of variance results: means and standard errors for total (ambulatory plus inpatient) pathology use in both Year I and Year II. The standard errors are conventionally calculated, which is to assume (incorrectly here) that the observations are independent. The free fee-for-service plan uses more than its experimental GHC counterpart, but the difference is not significantly different from zero ($t = 0.86$) using a test that overstates the true precision of the comparison because it ignores intrafamily and intrapersonal correlation. The GHC controls also use more than their experimental GHC counterparts. This difference is also not significantly different from zero ($t = -0.18$), using a test that overstates the precision for the comparison.

Table 4.1

ANALYSIS OF VARIANCE: MEANS AND STANDARD
ERRORS FOR TOTAL ANNUAL PATHOLOGY
IMPUTED EXPENSES BY PLAN

Plan	Mean ^a	Standard Error ^b	t on Difference from GHC Experimental
GHC experimental	42.18	1.95	----
Free fee-for-service	45.84	3.80	0.86
GHC control	42.73	2.28	-0.18

^aSample means for full-year enrollees with one observation for each person-year. Users and non-users are both included.

^bUncorrected for intrafamily and intrapersonal correlation. Calculated standard errors are too small, while t-statistics are too large.

Thus, at a gross aggregate level, it appears that there is no major (significant) utilization difference between a free-care plan in fee-for-service and in a prepaid group practice, when the two groups are random samples of the same population. Similarly, at the aggregate level, it appears that there is no major (significant) effect on utilization that can be attributed to self-selection. The self-selected GHC controls are not "significantly" different from the GHC experimentals.

In this chapter, we will examine the data more carefully to see if these two null findings are the result of inefficient estimation techniques or of other underlying factors. The aggregate utilization response will be disaggregated into ambulatory and inpatient responses. With this disaggregation we will examine whether the response by plan is due to true plan differences or due to differences in observable population characteristics. Specifically, we will examine the response to the three plans by:

- (1) Examining utilization by plan when the data have been adjusted to remove any differences in the nonplan characteristics of the three plans by using a standard reference population;
- (2) Examining utilization for the GHC controls when alternative reference populations are used. This provides a partial measure of the effect of self-selection on the use by GHC controls; and
- (3) Examining utilization by plan under two alternative levels of adjustment. One alternative adjusts for age and gender but not for health and other nonplan characteristics. The second adjusts for all observable characteristics, including self- or parent-reported health status.

ADJUSTED PLAN COMPARISONS

The lack of difference in the comparison of the GHC experimentals and controls may be due to underlying differences in individual and family characteristics that may partially offset any pure plan response. To isolate the plan responsiveness, we will examine what would happen if

the three plans had the same distribution of nonplan covariates, i.e., if we adjust for family and individual characteristics. Later we will examine the effect of different nonplan characteristics on utilization.

To measure plan response alone, we will assume that each plan has the same measured characteristics as the pooled free fee-for-service and GHC experimentals in Year I. The logic is the same as that used for age and sex adjustments in epidemiological studies. The only difference is that here we carry out the adjustment for all variables. (We pool the free and experimentals because they are a random sample of the same population.[1]) Table 4.2 contains the plan comparisons for ambulatory and inpatient pathology use, adjusting for a full set of covariates; the underlying coefficient and smearing estimates can be found in App. B. Table 4.3 contains the t-statistics for plan comparisons.

The numbers in Tables 4.2 and 4.3 indicate three interesting results. First, when a full set of covariates are used, there is a negligible (at most 9 percent) and statistically insignificant difference in ambulatory pathology utilization among the plans.[2] Second, when a full set of covariates are used, there is no statistically significant difference in inpatient pathology use between the self-selected GHC controls and the GHC experimentals. Third, however, there is a significantly higher rate of inpatient pathology utilization on the free fee-for-service plan than on either the GHC experimentals or the GHC controls; the GHC control result is not shown.

ALTERNATIVE REFERENCE POPULATIONS

The GHC controls might be expected to have different use from either the free or the GHC experimentals because they are a self-selected population. Table 4.4 contains the predictions for the GHC controls under two alternative assumptions. The first is that the

[1] When making predictions for nonlinear models such as the two-part model, we have to correct for any possible differences in the distribution of characteristics. Otherwise different predictions might be due to minor differences in the distribution, especially of rare characteristics. The free fee-for-service plan and GHC experimental plan samples were designed to be balanced. However, actual balance is unachievable ex post.

[2] $(34.68 - 31.81)/34.68 = 9$ percent.

Table 4.2

PREDICTED PATHOLOGY USE BY ASSUMING THAT
THE PLAN POPULATIONS HAVE THE SAME
OBSERVED CHARACTERISTICS

(Standard Error in Parentheses)

Plan	Ambulatory ^a	Inpatient ^b
GHC experimental	34.04 (1.69)	6.91 (1.34)
Free fee-for-service	31.81 (2.55)	12.26 (2.62)
GHC controls	34.68 (2.54)	5.96 (1.33)

NOTE: Reference group is pooled free fee-for-service and GHC experimental populations.

^aStandard errors are calculated using the correction for correlation in the probit. See Chap. 3.

^bStandard errors assume independent error in probit, because the correlations are not significantly different from zero.

Table 4.3

t-STATISTICS FOR PLAN DIFFERENCES ASSUMING THAT THE PLAN
POPULATIONS HAVE THE SAME OBSERVED CHARACTERISTICS

Plan	Ambulatory ^a	Inpatient ^b
Free fee-for-service vs. GHC experimental	-0.97	2.30
GHC controls vs. GHC experimental	+0.28	-0.70

NOTE: Reference group is pooled free fee-for-service and GHC experimental populations.

^at-statistics are calculated using the correction for correlation in the probit. See Chap. 3.

^bt-statistics assume independent error in the probit, because the correlations are not significantly different from zero.

Table 4.4

PREDICTION FOR THE GHC CONTROLS YEAR I
USING DIFFERENT REFERENCE POPULATIONS
(Standard Error)

Reference Population Characteristics	Ambulatory ^a	Inpatient ^b
Free fee-for-service and GHC experimentals	34.68 (2.54)	5.96 (1.33)
GHC controls	37.25 (2.23)	6.10 (1.36)

^aStandard errors are calculated using the correction for correlation in the probit. See Chap. 3.

^bStandard errors are calculated assuming independent error terms in the probit, because the correlations are not significantly different from zero.

GHC controls have the same distribution of observed nonplan characteristics as the free fee-for-service and the GHC experimentals plans combined. The second is that the controls have their actual distribution of observed characteristics. The difference between the two numbers provides a measure of the effect of self-selection on observable characteristics.

The effect of adjusting for population characteristics on predicted pathology use is large for ambulatory use and negligible for inpatient pathology use. Predictions for ambulatory pathology use from the two-part model are 7 percent higher for the GHC controls than they would have been if the controls had the same characteristics as the other two plans. The sample differences observed are 5 percent, which is imprecisely measured. As the plan coefficients in App. B indicate, all of the differences between the GHC controls and experimentals are in nonplan characteristics (e.g., age, sex, or health status). In Table 4.4, the data have been adjusted both ways, with the free fee-for-

service plus GHCE plans and the GHC control plan acting as the reference population.

EFFECT OF LEVEL OF ADJUSTMENT

In the preceding comparisons between the GHC controls and both experimental plans, we have used a more extensive set of characteristics, especially health status characteristics, than are usually available. Tables 4.5 and 4.6 contain the plan comparisons using a far more parsimonious specification, which includes only plan, age, gender, and year variables. The plan coefficient in the sample specification captures the effect of any omitted variables (e.g., sickness) that may be correlated with plan.

Table 4.5

PREDICTIONS OF AMBULATORY PATHOLOGY USE YEAR I:
PLAN, AGE, GENDER, AND YEAR SPECIFICATION

Reference Population Characteristics	Plan	Mean Prediction (Standard Error)	t on Difference from GHC Experimentals
Free fee-for-service + GHC experimentals	GHC experimentals	34.10 (1.73)	----
	Free fee-for- service	31.47 (2.57)	-1.11
	GHC controls	35.48 (2.73)	+0.59
GHC control	GHC experimentals	36.20 (2.37)	----
	Free fee-for- service	33.39 (3.30)	-1.12
	GHC controls	37.71 (2.46)	+0.60

If we are limited to plan, age, gender, and year variables, then none of the results presented earlier for the contrast between free fee-for-service and GHC experimentals change. Ambulatory use is still not significantly different on the two plans. Inpatient use is still dramatically different.

Similarly, the results of the comparison between GHC experimentals and controls do not change. The GHC controls have an insignificantly higher rate of ambulatory utilization than that of the experimentals. Instead of a 7 percent difference in ambulatory use, adjusted to the experimental population, there is now a 4 percent difference. There are still no significant inpatient differences. In fact, if one makes no adjustment for population characteristics other than year, the controls and experimentals are not significantly different in their inpatient use.

Table 4.6

PREDICTIONS OF INPATIENT PATHOLOGY USE YEAR I:
PLAN, AGE, GENDER, AND YEAR SPECIFICATION

Reference Population Characteristics	Plan	Mean Prediction (Standard Error)	t on Difference from GHC Experimentals
Free fee-for-service + GHC experimentals	GHC experimentals	7.29 (1.45)	----
	Free fee-for service	13.31 (2.93)	2.32
	GHC controls	6.12 (1.40)	-0.80
GHC Control	GHC Experimentals	7.56 (1.50)	----
	Free fee-for- service	13.76 (3.02)	2.31
	GHC controls	6.35 (1.45)	-0.80

Thus, we find that, adjusting for measured characteristics, the GHC controls have an insignificantly higher rate of utilization of ambulatory pathology services than that of the GHC experimentals. In addition, the controls have higher utilization because their measured characteristics are different, whether the characteristics include age and sex, or age, sex, family characteristics, and health status.

Chapter 5

CONCLUSIONS

This report compares the utilization of pathology services in three insurance plans with free care (no out-of-pocket costs) for medical services in both ambulatory and inpatient settings. One plan is a free fee-for-service plan in Seattle, Washington. The other two plans are prepaid group practice plans at Group Health Cooperative (GHC) of Puget Sound, which serves the Seattle metropolitan area. The enrollees in the free fee-for-service plan and one of the GHC plans (the experimentals) are random samples of the Seattle fee-for-service population in 1976. Enrollees in the third plan, the GHC controls, are a random sample of the GHC population in 1976. In all three plans, the aged, the Medicare disabled, the military and their dependents, and the institutionalized are excluded.

By comparing the free fee-for-service plan and the GHC experimentals, we can assess the effect of different organizational structures on utilization. Unlike most comparisons of fee-for-service and prepaid group practices, this comparison of the use of pathology services will not be confounded by differences in cost-sharing (the fee-for-service plans and HMOs often have different copayments) or self-selection into the alternative institutional settings. By comparing the GHC experimentals and controls, we can determine the effect of self-selection on utilization of pathology services.

The analysis of variance results indicate that, at a gross level, there appear to be no statistically significant differences in the utilization of pathology services among the three plans. However, a careful analysis using more robust estimation techniques (the two-part model) indicates three important results. First, there are no significant ambulatory differences between the free fee-for-service plan and the GHC plans. Second, the fee-for-service plan has a significantly greater use of inpatient pathology services than either GHC plan. These results confirm experimentally the results found in the literature for all medical services in a situation where the HMO versus fee-for-service

comparison is not confounded by either self-selection or by differences in copayment. The results also parallel the findings for total medical use being observed on the HIS.

The third major finding is that self-selection into (out of) an HMO does not appear to have an appreciable effect on the level of use of inpatient and outpatient pathology services. The GHC experimentals and GHC controls, the latter a self-selected population, did not behave differently. This was true whether one adjusted for all the measurable characteristics (including health status) or merely for age and gender. The GHC controls did have slightly higher use of ambulatory pathology services, but this was attributable to differences in the age and gender mix on the plans.

The differences between the GHC experimentals and free fee-for-service on the use of pathology services can be explained by differences in incentives for providers and patients. In an outpatient setting, patients' decisions should largely determine utilization patterns because it is the patient who initiates visits. For patients under both plans, the out-of-pocket cost of additional visits is zero. Patients under the two plans therefore have equal incentives to seek care. In contrast to the outpatient case, it is the physician who is largely responsible for decisions about inpatient use. In a free fee-for-service plan, the provider has little financial incentive to restrict inpatient use and will admit patients on the basis of medical need alone rather than cost and medical need. HMO providers, who are paid on a capitation basis, must keep inpatient use down if they are to keep their premium costs down.

The lack of a self-selection effect is at first surprising. The literature has been quite concerned about this issue. Nevertheless, there is a plausible explanation for the results. Most members of GHC join it through employer or group plans. The major source of selection may be between the employed and the not employed, and their families. In a non-aged population, there may be little room for additional selection effects after removing those who are too ill to work. It is important to remember that the sample populations in all of these plans, both experimentals and controls, exclude the aged, the Medicare disabled, and those with service-related disabilities.

These findings may be useful in the policy debate over fee-for-service versus HMOs. HMOs do have lower use than fee-for-service, largely because of reduced inpatient use. This lower use appears to be the result of institutional differences instead of self-selection.

Whether this lower use is desirable or not is an open question. Until we know the effects of different health insurance arrangements on health status, we cannot make a cost-benefit calculation.

These conclusions are preliminary. Additional HIS data from subsequent years in Seattle may provide enough additional precision to enable the detection of effects that are not now easily observed. Additional data may also lead to refinements in the estimation methodology.

Moreover, these results are for only one site, Seattle, and for one prepaid group practice, GHC. It is not clear how well these results will generalize to other areas or to other prepaid group practices. Seattle is a large metropolitan area with a well-developed medical system. It has a high ratio of providers to population and a short waiting-time for appointments. Nonprice rationing may be less important in Seattle than in other areas of the country.

GHC is a prepaid group practice with salaried providers, and about 80 percent of its enrollees enter through group policies (e.g., work-related policies). Alternative reimbursement arrangements or prepaid group practices that have a significant fraction of nongroup enrollees may behave differently.

Appendix A

ALTERNATIVE MODELS

The economic literature has two alternative models--the Tobit, and the selection models--for similar data. The two-part model can be considered a less restrictive generalization of the Tobit model; see Cragg (1971). The two-part model does not have the Tobit's assumption that the decision to use and the decision about level of use are the same.

The selection model, described by Heckman (1974, 1976, 1979), is also a two-equation model. The first is a censoring function that determines whether positive use is observed:

$$POS_i = x_i' \delta_1 + v_{1i}, \quad (A.1)$$

where v_{1i} is i.i.d $N(0, 1)$. If POS_i is positive, then positive use is observed; otherwise 0 is observed. This equation is a probit equation that is indistinguishable from the first equation of the two-part model.

The second equation in the selection model is an unconditional linear model of some function of use;

$$f(PATH_i) = x_i' \delta_2 + v_{2i}, \quad (A.2)$$

where

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \sim N(0, \Sigma)$$

$$\Sigma = \begin{pmatrix} 1 & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$$

The expected use is given by the unconditional expectation $x_i' \delta_2$.

The selection model must have the two equations estimated jointly if the errors are correlated ($\sigma_{12} \neq 0$). If the errors are correlated, then estimating Eq. (A.2) on the observed (positive) users will yield biased estimates of the unconditional use.[1]

Although the first equation is identical in the two-part and the selection model, the second equation is quite different. In the two-part model, the second equation is a linear model for conditional (positive) use. In the selection model, the second equation is a linear model for unconditional use.

We have rejected the selection model for four reasons.

First, the interpretation of the selection model is inappropriate for this application. The selection model estimates the unconditional use for all individuals, including nonusers, as if they were all users. In this case, we know that some persons have no use and we want the estimates to reflect this lack of use.

Second, the selection model assumes that the functional form is such that the unconditional error is bivariate normal. That assumption is not testable because the "uncensored" data are never observed.

Third, the selection model assumes that the conditional error is not symmetric (if $\sigma_{12} \neq 0$). Because the underlying error term v_2 has a normal error structure, the observed error must be skewed, with more of the population missing from one tail than the other. Hence, the conditional error will be shorter than normal in one tail. The data on pathology are longer in both tails than (log) normal. Thus, using the selection model would be a misspecification that could yield biased estimates.

Fourth, the selection model is not very well behaved numerically or statistically. The selection model has multiple local maxima, while the two-part model has a single maximum. It does not appear to be robust to minor departures from the underlying assumptions.

For a fuller discussion of these issues see DMMN (1982).

[1] Equation (A.1) can be estimated separately. If it is, then Eq. (A.2) must be corrected for selection by using a procedure such as Amemiya's (1973) or Heckman's (1974).

Appendix B

SUPPORTING TABLES

Table B.1

PROBIT EQUATION FOR POSITIVE
AMBULATORY PATHOLOGY USE

Variable	Coefficient	Standard Error	t Statistic
INTERCEPT	0.7475	0.5001	1.4938
P00	-0.0817	0.0579	-1.4107
GHCC	0.1327	0.0587	2.2606
TOOKPHYS	0.0130	0.0421	0.3095
TERM3	0.0528	0.0514	1.0273
HRTYPE	0.0375	0.0448	0.8364
CHILD	0.0484	0.1055	0.4592
FEMALE	0.2330	0.0657	3.5476
FADULT	0.6254	0.0861	7.2615
AGE	0.0278	0.0090	3.0822
SQAGE	-0.3648	0.1001	-3.6445
YR2	-0.0502	0.0414	-1.2114
LINC	0.0555	0.0425	1.3072
LFAM	-0.0335	0.0463	-0.7238
BLACK	-0.2176	0.1399	-1.5554
AFDC	-0.2608	0.0995	-2.6215
GHINDX	-0.0032	0.0016	-1.9912
MHI	-0.0026	0.0019	-1.3348
PHYSLM	0.0963	0.0669	1.4386
POSDISB	-0.1867	0.1146	-1.6286
LNDISB	0.1844	0.0537	3.4341
GHINMIS	-0.3050	0.1844	-1.6542
FLAG1	-0.2091	0.1503	-1.3910
FLAG2	-0.3009	0.0938	-3.2082

NOTE: The standard errors and t-statistics here are uncorrected for intrafamily and intraperson correlation. Upper bound for correcting t-statistics is divided by 1.39 for family variables and 1.09 for person variables. The person variables are age, sex, health and flag variables.

Table B.2

CONDITIONAL EQUATION FOR LOG POSITIVE
AMBULATORY PATHOLOGY USE

Variable	Coefficient	Standard Error	t Statistic
INTERCEPT	3.7992	0.4555	8.3406
P00	-0.0728	0.0559	-1.3022
GHCC	-0.0750	0.0540	-1.3894
TOOKPHYS	-0.0788	0.0390	-2.0196
TERM3	-0.0379	0.0489	-0.7749
HRTYPE	-0.0426	0.0410	-1.0405
CHILD	-0.2025	0.1015	-1.9958
FEMALE	0.3153	0.0617	5.1134
FADULT	0.0738	0.0786	0.9395
AGE	-0.0064	0.0081	-0.7886
SQAGE	0.0844	0.0904	0.9334
YR2	0.0467	0.0329	1.4189
LINC	-0.0145	0.0390	-0.3722
LFAM	-0.1098	0.0422	-2.5986
BLACK	0.0828	0.1337	0.6190
AFDC	-0.0234	0.0948	-0.2473
GHINDX	-0.0042	0.0014	-3.0399
MHI	0.0019	0.0017	1.1125
PHYSLM	0.1707	0.0554	3.0821
POSDISB	-0.2412	0.1011	-2.3869
LNDISB	0.1200	0.0437	2.7461
GHINMIS	-0.1757	0.1800	-0.9761
FLAG1	0.0514	0.1386	0.3710
FLAG2	-0.0888	0.0885	-1.0027

NOTE: The standard errors and t-statistics here are corrected for intrafamily and intraperson correlation. The estimated value for intrafamily correlation is 0.0462 and 0.1236 for intraperson correlation.

Table B.3

PROBIT EQUATION FOR POSITIVE INPATIENT
PATHOLOGY USE

Variable	Coefficient	Standard Error	t Statistic
INTERCEPT	-1.0019	0.1547	-6.4776
P00	0.2866	0.0782	3.6651
GHCC	-0.0469	0.0739	-0.6348
CHILD	-0.0590	0.0857	-0.6882
FADULT	0.4251	0.0765	5.5581
YR2	0.0531	0.0623	0.8516
GHINDX	0.0103	0.0020	-5.2703
GHINMIS	0.1694	0.2152	0.7871

NOTE: Standard errors and t-statistics are uncorrected for intrafamily and intraperson correlation. Intrafamily correlation is not significantly different from zero. Correction for intraperson correlation is done by dividing t-statistic by 1.04.

Table B.4

CONDITIONAL EQUATION FOR LOG POSITIVE
INPATIENT PATHOLOGY USE

Variable	Coefficient	Standard Error	t Statistic
INTERCEPT	5.4074	0.3032	17.8361
P00	0.0627	0.1565	0.4010
GHCC	-0.0585	0.1554	-0.3763
CHILD	-0.9541	0.1912	-4.9894
FADULT	-0.5829	0.1624	-3.5895
YR2	-0.1937	0.1237	-1.5662
GHINDX	-0.0122	0.0038	-3.2167
GHINMIS	-0.1884	0.4610	-0.4086

NOTE: Standard errors and t-statistics are corrected for intraperson and intrafamily correlation. The estimated value for intraperson correlation is 0.3970 and for intrafamily is 0.0024.

Table B.5

SMEARING FACTORS FOR PATHOLOGY USE

Use and Plan	Smearing Factor	Standard Error
Ambulatory use		
Free fee-for-service	1.4764	0.0682
GHC experimental	1.4057	0.0387
GHC control	1.4462	0.0548
Inpatient use	1.9225	0.2029

REFERENCES

- Amemiya, T., "Regression Analysis When the Dependent Variable is Truncated Normal," Econometrica, Vol. 41, 1973, pp. 997-1017.
- Balestra, P., and M. Nerlove, "Pooling Cross-Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas," Econometrica, Vol. 34, 1966, pp. 585-612.
- Berki, S. E., "Enrollment Choice in a Multi-HMO Setting: The Roles of Health Risk, Financial Vulnerability, and Access to Care," Medical Care, Vol. 15, 1977, pp. 95-114.
- , et al., "Enrollment Choices in Different Types of HMOs: A Multivariate Analysis," Medical Care, Vol. 16, 1978, pp. 682-697.
- Brook, R. H., et al., "Overview of Adult Health Status Measures Fielded in Rand's Health Insurance Study," Medical Care (Supplement), Vol. 17, 1979, pp. 1-131.
- Clasquin, L. A., Mental Health, Dental Services, and Other Coverage in the Health Insurance Study, The Rand Corporation, R-1216-OEO, 1973.
- Cragg, John G., "Some Statistical Models for Limited Dependent Variables, with Applications to the Demand for Durable Goods," Econometrica, Vol. 39, 1971, pp. 829-844.
- Danzon, P. M., Profits in Hospital Laboratories: The Effects of Reimbursement Policies on Hospital Costs and Charges, The Rand Corporation, R-2582-HCFA, September 1980.
- , Economic Factors in the Use of Laboratory Tests by Office-Based Physicians, The Rand Corporation, R-2525/1-HCFA, August 1982.
- , W. G. Manning, and M. S. Marquis, Factors Affecting Laboratory Test Use and Prices: Executive Summary, The Rand Corporation, R-2987-HCFA, January 1983.
- Davies, A. R., and J. E. Ware, Measuring Health Perceptions in the Health Insurance Experiment, The Rand Corporation, R-2711-HHS, October 1981.
- Duan, N., "Correcting for Intraclass Correlation in Probit Regression Models," App. C in N. Duan, W. G. Manning, Jr., C. N. Morris, and J. P. Newhouse, A Comparison of Alternative Models for the Demand for Medical Care, The Rand Corporation, R-2754-HHS, January 1982a.
- , "Smearing Estimate: A Nonparametric Retransformation Method," App. B in Duan, Manning, Morris, and Newhouse, 1982b.

- , W. G. Manning, Jr., C. N. Morris, and J. P. Newhouse, A Comparison of Alternative Models for the Demand for Medical Care, The Rand Corporation, R-2754-HHS, January 1982.
- Eisen, M., C. A. Donald, J. E. Ware, and R. H. Brook, Conceptualization and Measurement for Children in the Health Insurance Study, The Rand Corporation, R-2313, May 1980.
- Enthoven, A. C., "Consumer-Choice Health Plan": Part I, "Inflation and Inequity in Health Care Today: Alternative for Cost Control and Analysis of Proposals for National Health Insurance," New England Journal of Medicine, Vol. 298, 1978a, pp. 650-658.
- , Part II "A National Health Insurance Proposal Based on Regulated Competition in the Private Sector," New England Journal of Medicine, Vol. 298, 1978b, pp. 709-720.
- , Health Plan: The Only Practical Solution to the Soaring Cost of Medical Care, Addison-Wesley Publishing Company, Reading, Massachusetts, 1980.
- , and R. Noll, "Regulatory and Nonregulatory Strategies for Controlling Health Care Costs," in S. H. Altman and R. Blendon (eds.), Medical Technology: The Culprit behind Health Care Costs?, U. S. Department of Health, Education, and Welfare, DHEW No. (PHS) 79-3216, 1979, pp. 213-234.
- Gronau, R., "Wage Comparisons--A Selectivity Bias," Journal of Political Economy, Vol. 82, 1974, pp. 1119-1144.
- Heckman, J., "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables, and a Sample Estimator for such Models," Annals of Economic and Social Measurement, Vol. 5, 1976, pp. 475-592.
- , "Sample Bias as a Specification Error," Econometrica, Vol. 47, 1979, pp. 153-167.
- , "Shadow Prices, Market Waves, and Labor Supply," Econometrica, Vol. 42, 1974, pp. 679-694.
- Luft, H. S., Health Maintenance Organizations: Dimensions of Performance, John Wiley & Sons, New York, 1981.
- Maddala, G. S., "The Use of Variance Components Models in Pooling Cross-Section and Time Series Data," Econometrica, Vol. 39, 1971, pp. 341-370.
- Manning, W. G., et al., "A Two-Part Model of the Demand for Medical Care: Preliminary Results from the Health Insurance Study," in J. Van der Gaag and M. Perlman (eds.), Health, Economics, and Health Economics, North Holland Publishing Co., Amsterdam, 1981.

- Marquis, M. S., Laboratory Test-Ordering by Physicians: The Effect of Reimbursement Policies, The Rand Corporation, R-2901-HHS, August 1982.
- Morris, C. N., "A Finite Selection Model for Experimental Design of the Health Insurance Study," in D. J. Aigner and C. N. Morris (eds.), "Experimental Design in Econometrics," Journal of Econometrics, Vol. 11, No. 1, September 1979, pp. 43-62.
- Mundlak, Y., "On the Pooling of Time Series and Cross-Section Data," Econometrica, Vol. 46, 1978, pp. 69-86.
- Newhouse, J. P., "A Design for a Health Insurance Experiment," Inquiry, Vol. 11, 1974, pp. 5-27.
- Olsen, R., "The Analysis of Two-Variable Models When One of the Variables Is Dichotomous," Economics Department, Yale University, 1975 (unpublished manuscript).
- Poirier, D. J., and P. A. Ruud, "On the Appropriateness of Endogenous Switching," Journal of Econometrics, Vol. 16, 1981, pp. 249-256.
- Richardson, W. C., et al., The Seattle Prepaid Health Care Project: Comparison of Health Services Delivery: Chap.1, "Introduction to the Project, the Study, and the Enrollees," National Technical Information Service, PB 267-488-SET, Springfield, Va., 1976.
- , Comparisons of Prepaid Health Plans in a Competitive Market: The Seattle Prepaid Health Care Project. National Technical Information Service, PB-267-489, Springfield, Va., 1980.
- Schwartz, H., "Conflicts of Interest in Fee for Service and in HMOs," The New England Journal of Medicine, Vol. 299, 1978, pp. 1071-1073.
- Searle, S. R., Linear Models, John Wiley & Sons, New York, 1971.
- Stewart, A. L., J. E. Ware, and R. H. Brook, Construction and Scoring of Aggregate Functional Status Indexes, Vol. I, The Rand Corporation, R-2551-HHS, January 1981.
- Tobin, J., "Estimation of Relationships for Limited Dependent Variables," Econometrica, Vol. 26, 1958, pp. 24-36.
- Van de Ven, W.P.M.M., and B.M.S. van Praag, "Risk Aversion and Deductibles in Private Health Insurance: Applications of an Adjusted Tobit Model to Family Health Care Expenditures," in J. van der Gaag and M. Perlman (eds.), Health, Economics, and Health Economics, North Holland Publishing Co., Amsterdam, 1981.
- Viet, C., and J. E. Ware, Jr. Refinements in the Measurement of Mental Health for Adults in the Health Insurance Study, The Rand Corporation (forthcoming).

