
Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment

Author(s): Willard G. Manning, Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler and Arleen Leibowitz

Source: *The American Economic Review*, Jun., 1987, Vol. 77, No. 3 (Jun., 1987), pp. 251-277

Published by: American Economic Association

Stable URL: <https://www.jstor.org/stable/1804094>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



is collaborating with JSTOR to digitize, preserve and extend access to *The American Economic Review*

JSTOR

Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment

By WILLARD G. MANNING, JOSEPH P. NEWHOUSE, NAIHUA DUAN,
EMMETT B. KEELER, ARLEEN LEIBOWITZ, AND M. SUSAN MARQUIS*

We estimate how cost sharing, the portion of the bill the patient pays, affects the demand for medical services. The data come from a randomized experiment. A catastrophic insurance plan reduces expenditures 31 percent relative to zero out-of-pocket price. The price elasticity is approximately -0.2 . We reject the hypothesis that less favorable coverage of outpatient services increases total expenditure (for example, by deterring preventive care or inducing hospitalization).

Over the past four decades medical care costs have grown about 4 percent per year in real terms, and the share of GNP devoted to medical care has increased from 4.4 percent in 1950 to 10.7 percent in 1985 (Daniel Waldo, Katherine Levit, and Helen Lazenby, 1986). A prominent explanation of this rapid increase has emphasized the spread of health insurance, which has generated demand for both a higher quality and an increased quantity of medical services (Martin Feldstein, 1971, 1977). In turn, the spread of health insurance has been linked to the exemption of employer-paid health insurance premiums from the individual income tax (Feldstein and Elizabeth Allison, 1974; Feldstein and Bernard Friedman, 1977; Mark Pauly, 1986). Thus, the increase in expenditure is often

portrayed as a type of market failure induced by public policy, although such an argument is not universally accepted (Morris Barer, Robert Evans, and Gregory Stoddart, 1979; Robert Evans, 1984; John Goddeeris and Burton Weisbrod, 1985).

No one has shown, however, that the spread of health insurance can quantitatively account for most of the sustained rise in health expenditure (Pauly, 1986). If it cannot, the widespread presumption that distorted prices (because of insurance) are inducing excess resources in medical care is not necessarily correct. Central to appraising the quantitative role of insurance, of course, is the magnitude of the demand response to changes in insurance. The literature exhibits substantial disagreement, by a factor of 10 or more, about the price elasticity, or coinsurance elasticity, of demand (Richard Rosett and Lien Fu Huang, 1973; Karen Davis and Louise Russell, 1972; Charles Phelps and Newhouse, 1974; Fred Goldman and Michael Grossman, 1978; Ann Colle and Grossman, 1978; Newhouse and Phelps, 1974, 1976).¹

Such disagreement is not surprising in light of the problems of using nonexperimental data to estimate elasticities (Newhouse,

*The Rand Corporation, 1700 Main Street, Santa Monica, CA 90406. This research was conducted under grant 016B80 from the Department of Health and Human Services. Neither the Department nor the Rand Corporation necessarily agrees with the conclusions presented here. We acknowledge the comments of three very detailed anonymous referee reports. We, of course, are responsible for any remaining errors. Bernadette Benjamin and Jack Zwanziger provided invaluable programming support and assistance throughout the analysis underlying this paper. Individuals too numerous to mention provided advice, assistance, and support throughout the project; however, two of our government project officers, Larry Orr and James Schuttinga, deserve special mention; so too do our former colleagues Charles Phelps and Carl Morris.

¹The elasticity estimates at the mean vary from around -0.1 to -2.1 .

Phelps, and Marquis, 1980). In cross-sectional data, insurance is endogenous; those who expect to demand more services have a clear incentive to obtain more complete insurance, either by selecting a more generous option at the place of employment, by working for an employer with a generous insurance plan, or by purchasing privately more generous coverage.

Ignoring this selection issue (i.e., treating insurance as exogenous) has generally produced results showing that demand for medical care responds to insurance-induced variation in price. Treating insurance as endogenous, however, has generally led to coefficients with confidence intervals that are insignificantly different from zero at conventional levels (Newhouse and Phelps, 1976).²

That upward bias may be present is suggested by results from several natural experiments that compared demands of the same individuals before and after their group insurance changed (Anne Scitovsky and Nelda Snyder, 1972; Scitovsky and Nelda McCall, 1977; Phelps and Newhouse, 1972; R. G. Beck, 1974). In these cases the change in insurance is presumptively exogenous, and the elasticity estimates cluster near the low end of those cited above. But natural experiments have no control group, so that any other factor that changed over time is perfectly confounded with the insurance change. Moreover, the samples available in such studies are not necessarily representative of the general population, and the changes in insurance that could be studied were limited to those that occurred in the natural experiment. Hence, these results too have been suspect.³

²Although many believe this failure to reject the null hypothesis when insurance is treated as endogenous occurs because the insurance variable is only weakly identified, the magnitude of any upward bias in elasticity estimates from treating insurance as exogenous remains unknown. Hausman (or Wu) type tests have not been used to test for endogeneity, but if they failed to reject the null hypothesis of exogeneity, it could be for lack of power because of a lack of a useful set of instruments.

³For reviews of the nonexperimental demand literature and a discussion of its methodological problems, see Newhouse (1978; 1981).

In light of the uncertainty about how demand responds to insurance-induced changes in price, and the importance for both public and private decisions of quantifying that response, the federal government initiated the Rand Health Insurance Experiment (HIE) in 1974, one aim of which was to narrow uncertainty about this issue (Newhouse, 1974). In this article we report the results of that experiment. Our findings have implications for the role of insurance in explaining the post-war increase in medical expenditure, as well as for the magnitude of the welfare loss from health insurance.

The HIE had several objectives other than improved estimates of how demand responds to insurance. Four such objectives merit mention here:

1) Many poor individuals are insured through public programs; whether the demand response differs for the poor is therefore an issue in decisions on the scope of these programs.

2) Insurance need not be uniform across various medical services. In fact, second-best pricing implies that coverage should be more generous for less price elastic (or less insurance elastic) services (Frank Ramsey, 1927; Richard Zeckhauser, 1970; William Baumol and David Bradford, 1970). We therefore wished to learn if insurance elasticities differed for various types of medical services. In particular, are demand elasticities greater for outpatient physician services, psychotherapy, and preventive services, which would accord with the observed lesser coverage of these services?⁴

3) The public financing of medical care has been justified by its status as a merit good (Richard Musgrave, 1959) and in particular the claim that the consumption of medical services leads to improved health, which can generate externalities (Cotton Lindsay, 1969; Anthony Culyer, 1971, 1976, 1978; Pauly, 1971; Evans, 1984). Thus, we

⁴Other explanations, not mutually exclusive, for the lower coverage of these services include greater loading charges and asymmetric information between insurer and insured.

sought to quantify how the change in the consumption of medical services at the margin might affect health. The answer to this question would inform the political debate about the benefits of public financing of medical care services for the indigent and would also inform the insurance decisions of private agents such as employers and unions.

4) For the past decade, public policy has promoted Health Maintenance Organizations (HMOs) on the groups that such organizations were more efficient in the delivery of services. Almost all evidence of lower cost, however, came from uncontrolled settings, leaving unresolved the question of whether selection of healthier members or more efficient treatment was responsible for lower costs in HMOs (Harold Luft, 1981). Also unresolved was the question of whether any true reduction in services at HMOs might adversely affect health status. Therefore, we sought to decompose the observed lower use of services at one HMO into the pure effect of the HMO, on the one hand, and treating a possibly less sickly group of enrollees on the other. Moreover, we sought to determine whether any reduced use of services affected health status and satisfaction.

This article considers the first two questions in some detail and summarizes the findings on the latter two.

I. Data and Sample

A. *The Design of the Rand Health Insurance Experiment*⁵

Between November 1974 and February 1977, the HIE enrolled families in six sites: Dayton, Ohio; Seattle, Washington; Fitchburg, Massachusetts; Franklin County, Massachusetts; Charleston, South Carolina; and Georgetown County, South Carolina.⁶

⁵Newhouse (1974) and Robert Brook et al. (1979), provide fuller descriptions of the design. Newhouse et al. (1979) discuss the measurement issues for the second generation of social experiments, to which the HIE belongs. John Ware et al. (1980) discuss many aspects of data collection and measurement for health status.

⁶The sites were selected to represent the four census regions; to represent the range of city sizes (a proxy for

Families participating in the experiment were assigned to one of 14 different fee-for-service insurance plans or to a prepaid group practice; additionally, some members already enrolled in the prepaid group practice were enrolled as a separate group. The fee-for-service insurance plans, the main focus of this article, had different levels of cost sharing, which varied over two dimensions: the coinsurance rate (percentage paid out-of-pocket) and an upper limit on annual out-of-pocket expenses. The coinsurance rates were 0, 25, 50, or 95 percent. Each plan had an upper limit (the Maximum Dollar Expenditure or MDE) on annual out-of-pocket expenses of 5, 10, or 15 percent of family income, up to a maximum of \$1,000.⁷ Beyond the MDE, the insurance plan reimbursed all covered expenses in full.

Covered expenses included virtually all medical services.⁸ One plan had different coinsurance rates for inpatient and ambulatory medical services (25 percent) than for dental and ambulatory mental health services (50 percent). And on one plan, the families faced a 95 percent coinsurance rate for outpatient services, subject to a \$150 annual limit on out-of-pocket expenses per person (\$450 per family). In this plan, all inpatient services were free; in effect, this plan had approximately an outpatient individual deductible.⁹

the complexity of the medical delivery system); to cover a range of waiting times to appointment and physician per capita ratios (to test for the sensitivity of demand elasticities to nonprice rationing); and to include both urban and rural sites in the North and the South.

⁷The maximum was \$750 in some site-years for the 25 percent coinsurance plans. The \$1000 was kept fixed in nominal dollars from 1974 to 1981. During this time the medical care component of the CPI rose by 96 percent.

⁸See Lorraine Clasquin (1973) for a discussion of the rationale for the HIE structure of benefits. Nonpreventive orthodontia and cosmetic surgery (related to pre-existing conditions) were not covered. Also excluded were outpatient psychotherapy services in excess of 52 visits per year per person. In the case of each exclusion, it is questionable whether anything could have been learned about steady-state demand during the 3- to 5-year lifetime of the experiment.

⁹The coinsurance rate for the 95 percent and individual deductible plans was actually 100 percent in the first

Families were assigned to these insurance plans using the Finite Selection Model (Carl Morris, 1979). This model was used to achieve as much balance across plans as possible while retaining randomization; that is, it minimizes the correlation between the experimental treatments and health, demographic, and economic covariates.

To study methods effects, the HIE employed four randomized subexperiments (Newhouse et al., 1979). We describe two here. To test for transitory demand effects (Charles Metcalf, 1973; Kenneth Arrow, 1975), 70 percent of the households were enrolled for three years; the remainder for five years. Also, to ensure that no one was worse off financially from participating in the study, families were paid a lump sum payment.¹⁰ To test for a possible stimulus to utilization, 40 percent of the families were given an unanticipated increase in their lump sum payment during the next to the last year of the study.

B. The Sample

The enrolled sample is for the most part a random sample of each site's nonaged population, but some groups were not eligible.¹¹

year of Dayton, the first site. The rate was changed to 95 percent for all other site-years of the experiment, in order to increase the incentive to file claims, although there was no statistical evidence at that time of underfiling. Subsequent analysis has shown that the mean outpatient physician expenditure on the 95 percent coinsurance plans relative to the free-care plan is understated by about 5 to 10 percent because of a lower propensity to file claims (William Rogers and Newhouse, 1985).

¹⁰Because of size of the lump sum payment, there is a theoretical presumption of no bias from refusal or attrition. Although refusal and attrition occurred at higher rates on higher coinsurance plans, refusal and attrition appear to have been random within plan. More precisely, we detect no differences by plan at enrollment in pre-experimental use or health status, nor do we detect differences in the rate of spending between those who withdrew from the experiment and those who did not. More detailed data on issues of refusal and attrition can be found in Brook et al. (1983, 1984); Kevin O'Grady et al. (1985); Newhouse et al. (1987). The details of the lump sum payment rules can be found in Clasquin and Marie Brown (1977).

¹¹The ineligible groups include: 1) those 62 years of age and older at the time of enrollment; 2) those with

Table 1 gives the sample by plan and site; it excludes the 1,982 persons in the HMO experiment. Note that plans are not perfectly balanced by site; in particular, no one was enrolled in the 50 percent plan in Seattle, and about half of those in the 50 percent plan are in Dayton, whereas only 20 percent of all participants are in Dayton.¹²

1. *Dependent Variables.* In the interest of brevity, we focus primarily on the use of medical services other than outpatient psychotherapy and dental services.¹³ We do, however, summarize results for dental services below.

incomes in excess of \$25,000 in 1973 dollars or \$58,000 in 1984 dollars); this excluded 3 percent of the families contacted; 3) those eligible for the Medicare disability program, 4) those in jails or institutionalized for indefinite periods; 5) those in the military or their dependents; and 6) veterans with service-connected disabilities.

¹²About 3 percent of the actual participant-years are truncated because the participant withdrew partway through an accounting year. With the exception of deaths, we do not use such participants in the estimation sample because the 4-part model (see below) requires equal time periods for each observation. If a person is only observed for one quarter and the expenditure distribution is lognormal, the annual distribution is not simply the quarterly distribution scaled up by a factor of 4; i.e., the lognormal distribution does not convolute. The sample used in this analysis more specifically includes enrollees during each full year that they participated, and the last accounting year in the study for those who died. We excluded data on partial years of participation by newborns. (Their expenses in the hospital at the time of birth, however, are attributed to the mother.) We tested the legitimacy of excluding those with partial years by comparing expenditure rates of part-year persons, adjusted for time at risk, with what they would have spent if they behaved like full-year people. Specifically we regressed actual expenditure minus (time at risk times the 4-part model prediction) on plan dummy variables. We could not reject the null hypothesis of no difference by plan ($\chi^2(4) = 2.67, p > .50$). The estimated effect of including part-year participants is to negligibly increase the estimated response to plan.

¹³See Manning et al. (1984b, 1986b) and Kenneth Wells et al. (1982) for additional results on the use of mental health care, and Manning et al. (1985) for additional results on dental use. Mental health care use is on the order of 4 percent of the expenditures discussed here.

TABLE 1—NUMBER OF PERSONS AT ENROLLMENT AND NUMBER OF PERSON-YEARS IN ESTIMATION SAMPLE

Plan	Site						Enrollment Total ^a	Estimation Sample Total ^b
	Dayton	Seattle	Fitch- burg	Frank- lin County	Charles- ton	George- town		
Free	301	431	241	297	264	359	1893	6822
25 Percent ^c	260	253	125	152	146	201	1137	4065
50 Percent	191	0	56	58	26	52	383	1401
95 Percent	280	253	113	162	146	166	1120	3727
Individual								
Deductible	105	285	188	220	196	282	1276	4175
Total	1137	1222	723	889	778	1060	5809	20190

^a Persons.^b Person-years.^c Includes those with 50 percent coinsurance for dental and mental health and 25 percent coinsurance for all other services.

2. *Independent Variables.* Although we present sample means by plan, we also present results controlling for site, health status, sociodemographic, and economic variables.

Insurance Plan Variables. Rather than impose a functional form, we have conservatively used dummy variables for insurance plans. We have grouped the insurance plans into five groups: 1) the free plan (no out-of-pocket cost to the family); 2) 25 percent coinsurance rate plans for medical services; 3) 50 percent coinsurance rate plans for medical services; 4) 95 percent coinsurance rate plans for medical services; and 5) the plan with a 95 percent coinsurance rate for outpatient services (subject to a limit of \$150 per person or \$450 per family per year) and free inpatient care.¹⁴ The middle three groups we call the family-pay plans.

Other Covariates. In addition to dummy variables for each plan group, we also included covariates for age, sex, race, family income, health status, family size, and site (Manning et al., 1987). With the exception of family size and income, the data were col-

lected before or at enrollment in the study. The value for family size varies by year. Family income data are from 1975 in Dayton, 1978 for the three-year group in South Carolina, and 1976 for all other participants.¹⁵ Health status measures are described more fully in Brook et al. (1983, 1984), R. Burciaga Valdez et al. (1985), and Valdez (1986).

Although we have not tested for all possible interactions among covariates, we did examine some that are important for policy purposes (for example, income and plan). As a result, we have included interactions between being a child and plan in the inpatient and outpatient use equations (see below), between plan and income in the probabilities of any use of medical and of any inpatient use (see below), and between sex and age in all equations. The remaining interactions

¹⁴ Differences among plans with 5, 10, and 15 percent upper limits are too small to detect at the level of annual expenditure. Hence, we have pooled across these different expenditure ceilings. See Keeler et al. (1987) for further discussion of how a varying ceiling affects demand.

¹⁵ The first year of participation was 1975 for the Dayton participants; the South Carolina 3-year group began participation in late 1978 (about a quarter participated for two months and another quarter for one month of 1978); and the remainder of the sample enrolled in 1976 or early 1977. Most of the enrollment was in the latter half of 1976. We used these data because we believed the income measure was more reliable than the data on pre-experimental income. The data we used were collected on forms keyed to income tax returns, whereas data on pre-experimental income were responses to a personal interview.

were neither significant nor appreciable, and have been omitted.

C. Unit of Analysis

The unit of analysis is a person-year. We use the year as the time frame for ease of interpretation and because the upper limit on out-of-pocket expenses is an annual limit. We use the person as the unit of observation because most major determinants of the use of services are individual (for example, age, sex, and health status) rather than family (for example, insurance coverage, and family income).

II. Statistical Methods

We estimated two types of models. In addition to simple means (ANOVA), we present more robust estimates based on a four-equation model developed by Duan et al. (1982 and 1983). This model gains over ANOVA (and ANOCOVA) by exploiting three characteristics of the distribution of medical expenses. First, a large proportion of the participants use no medical services during the year. Second, the distribution of expenses among users is highly skewed. Third, the distribution of medical expenses is different for individuals with only outpatient use than for individuals with inpatient use.

Because of these three characteristics, ANOVA (and ANOCOVA) yields imprecise, though consistent, estimates of the effects of health insurance, health status, and socioeconomic status on the use of medical services, even for a sample size on the order of 20,000 (not all independent) observations. As Duan et al. (1982 and 1983) and Manning et al. (1987) show, a four-equation model that exploits the characteristics of the medical expense distribution yields consistent estimates with lower mean square error than ANOVA.

A. The Four-Equation Model

We partition the participants into three groups: nonusers, users of only outpatient services, and users of any inpatient services.

We examine the expenses of the last two groups of users separately.

The first equation of the model is a probit equation for the probability that a person will receive any medical service during the year—from either inpatient or outpatient sources. Thus, this equation separates users from nonusers, and addresses the first characteristic described above, a large proportion of the population does not use medical services during the year. The second equation is a probit equation for the conditional probability that a user will have at least one inpatient stay, given that he has some medical use. This equation separates the two user groups, and thus addresses the third characteristic noted above, different distributions of medical expenses for inpatient and outpatient users.

The third equation is a linear regression for the logarithm of total annual medical expenses of the outpatient-only users. The fourth equation is a linear regression for the logarithm of total annual medical expenses for the users of any inpatient service. This last equation includes both outpatient and inpatient expenses for users of any inpatient services.¹⁶

The logarithmic transformation of annual expenses practically eliminates the undesirable skewness in the distribution of expenses among users, the second characteristic noted above. In particular, the logarithmic transform yields nearly symmetric and roughly normal error distributions. Further details are available in Duan et al. (1982 and 1983) and Manning et al. (1987).

While our use of the four-equation model is motivated by our desire to have the stochastic term approximate the normal assumption as closely as possible (to obtain robust estimates), the error distributions for the two levels of expense equations still deviate from the normal assumption. As a re-

¹⁶Grouping expenses by person rather than the more natural all-inpatient and all-outpatient expenditure eliminates the need to account for across-equation correlation in calculating standard errors of total expenditure.

sult, if we were to use the normal theory retransformation from the logarithmic scale to the raw dollar scale ($\exp(\sigma^2/2)$), the predictions would be inconsistent. Instead we use a nonparametric estimate of the retransformation factors, the smearing estimate, developed by Duan (1983), which in this application is the sample average of the exponentiated least squares residuals:

$$(1) \quad \hat{\phi}_j = \sum_i \exp(\hat{\varepsilon}_{ij})/n_j, \quad j = 3, 4,$$

where n_j = sample size for equation j ,

$$\hat{\varepsilon}_{ij} = \ln(y_{ij}) - x_{ij}\hat{\beta}_j,$$

$$\hat{\beta}_j = \text{OLS estimate of } \beta_j,$$

and i indexes the person. The smearing estimate is weakly consistent (asymptotically unbiased) for the retransformation factor if the error distribution does not depend on the characteristics x_i .¹⁷

A consistent estimate of the expected medical expense for person i based on the four-equation model is given by

$$(2) \quad E(\text{Medical Expenditure}_i) \\ = \hat{p}_i [(1 - \hat{\pi}_i) \exp(x_i \hat{\beta}_3) \hat{\phi}_3 + \hat{\pi}_i \exp(x_i \hat{\beta}_4) \hat{\phi}_4]$$

where $\hat{p}_i = \Phi(x_i \hat{\beta}_1)$
= estimated probability of any medical use,
 $\hat{\pi}_i = \Phi(x_i \hat{\beta}_2)$ = estimated conditional probability for a medical user to have any inpatient use,
 $\exp(x_i \hat{\beta}_3) \hat{\phi}_3$ = estimate of the conditional expense for medical services if outpatient only,

$\exp(x_i \hat{\beta}_4) \hat{\phi}_4$ = estimate of the conditional expense for medical services if any inpatient,

$\hat{\phi}_3, \hat{\phi}_4$ = estimated retransformation ("smearing") factors of the error terms for level of outpatient only and any inpatient expenditure equations.

Our estimates of predicted expenditure presented below are based on equation (2). We use equation (2) to predict medical expenditure for each person we enrolled, alternatively placing that person on each plan (by successively turning on plan dummy variables). We then average within plans over each predicted value to obtain a mean value for each plan. Standard errors of the predicted values are obtained by the delta method (see Duan et al., 1983, pp. 40, 48). The regression equations underlying our predicted values are presented in Manning et al. (1987).

B. Correlation in the Error Terms

Although we have over 20,000 observations, we do not have the same number of independent observations, because of substantial positive correlations in the error terms among family members and over time among observations on the same person. These correlations exist in all four equations. Failure to account for them in the analysis would yield inefficient estimates of the coefficients and inconsistent estimates of the standard errors. In the results presented below we have corrected the inference statistics (t , F , and χ^2) for this positive correlation using a nonparametric approach.¹⁸

C. Selection Modes

The econometric literature provides an additional class of models for continuous but limited dependent variables such as medical

¹⁷Moreover, when the normal assumption does hold, the smearing factor has high efficiency (90 percent or more) relative to the normal retransformation for a wide range of parameter values, including those in this analysis (see Duan, 1983, Section 5; and F. Mehran, 1973). In the results presented below, the smearing factors for the log level of expense for outpatient only users are estimated separately by plan and year to allow for heteroscedasticity. For the log level of expenses for users of any inpatient services, the smearing factor is a constant. See Duan et al. (1983) and Appendix C of Manning et al. (1987) for a comparison of normal theory and nonparametric retransformations.

¹⁸The correction is similar to that for the random effects least squares model, or equivalently the intracluster correlation model (S. R. Searle, 1971). The model is described in Brook et al. (1984), based on prior work by P. J. Huber (1967) on the variance of a robust regression.

expenditure. These models include the Tobit model (James Tobin, 1958), the Adjusted Tobit model (Wynand van de Ven and Bernard van Praag, 1981a,b), and sample selection models (G. S. Maddala, 1983). Like our four-equation model, these are multi-equation models, with an equation (often a probit) for whether there is a positive amount, and another equation for the level of the positive amount. These models differ from ours in that they explicitly model the correlation between the probability of any use and the level of use. Although they may appear to be more general, in fact for this problem they are not (Duan et al., 1984). In particular, the four-equation model just described is not nested within the sample selection model. Manning et al. (1987) provides a fuller discussion of these models and, using a split-sample validation, show that the four-equation model has significantly less bias than the sample selection model and is statistically indistinguishable on the basis of mean square error.¹⁹ In a separate Monte Carlo study, Manning, Duan, and William Rogers (forthcoming) show that models such as the four-equation model can be more robust, and are no worse than selection models when the data are truly generated by a selection model.

III. Empirical Results

A. Main Effects of Insurance Plan: ANOVA Estimates

The data from the Health Insurance Experiment (HIE) clearly show that the use of medical services responds to changes in the amount paid out-of-pocket. Table 2 provides the sample means and standard errors by plan for several measures of use of services—the probability of being treated, visit and admission rates, and total expenses. The per capita expenses on the free plan (no out-of-

pocket costs) are 45 percent higher than those on the plan with a 95 percent coinsurance rate, subject to an upper limit on out-of-pocket expenses. Spending rates on plans with an intermediate level of cost sharing lie between these two extremes. The right-most column shows that adjusting for the site imbalance in plan assignments (see Table 1) makes little difference.

Cost sharing affects primarily the number of medical contacts, rather than the intensity of each of those contacts. In other words, the differences in expenditures across plans reflect real variation in the number of contacts rather than an increase in the intensity or charge per service.²⁰ For example, outpatient expenses on the free plan are 67 percent higher than those on the 95 percent plan, while outpatient visit rates are 66 percent higher.

The largest decreases in the use of outpatient services occurs between the free and 25 percent plans, with smaller but statistically significant differences between the 25 percent and other family coinsurance (pay) plans ($\chi^2(2) = 9.48$, $p < .01$).

There are no significant differences among the family coinsurance (25, 50, and 95 percent) plans in the use of inpatient services. For the probability of any inpatient use, total admission rates, and inpatient expenses, the contrasts between the 25, 50, and 95 percent plans have p values greater than 0.50. As noted above, this lack of a significant difference is probably due to the effect of the upper limit on out-of-pocket expenses. Seventy percent of people with inpatient care exceeded their upper limit. Hence, the out-of-pocket cost of a hospitalization was at most \$1000 (in current dollars), and did not vary much among the pay plans (other than the individual deductible).²¹

¹⁹The bias in the selection models in the forecast sample was appreciable, on the order of 10–25 percent of the mean in the two replications we made ($p < .10$). In contrast, the bias for the 4-part model was 2 percent ($t = .50$).

²⁰Keeler and John Rolph (1982) found that cost sharing affected the number of episodes of treatment, rather than the size of the episode. They used data from the first three years of the Dayton site. Kathleen Lohr et al. (1986) found a similar result for diagnosis-specific episodes.

²¹This is a good example of the difference between the response to a marginal price or coinsurance and the response to plan.

TABLE 2—SAMPLE MEANS FOR ANNUAL USE OF MEDICAL SERVICES PER CAPITA

Plan	Face-to-Face Visits	Outpatient Expenses (1984 \$)	Admissions	Inpatient Dollars (1984 \$)	Prob. Any Medical (%)	Prob. Any Inpatient (%)	Total Expenses (1984 \$)	Adjusted Total Expenses (1984 \$) ^a
Free	4.55 (.168)	340 (10.9)	.128 (.0070)	409 (32.0)	86.8 (.817)	10.3 (.45)	749 (39)	750 (39)
25 Percent	3.33 (.190)	260 (14.70)	.105 (.0090)	373 (43.1)	78.8 (1.38)	8.4 (0.61)	634 (53)	617 (49)
50 Percent	3.03 (.221)	224 (16.8)	.092 (.0116)	450 (139)	77.2 (2.26)	7.2 (0.77)	674 (144)	573 (100)
95 Percent	2.73 (.177)	203 (12.0)	.099 (.0078)	315 (36.7)	67.7 (1.76)	7.9 (0.55)	518 (44.8)	540 (47)
Individual Deductible	3.02 (.171)	235 (11.9)	.115 (.0076)	373 (41.5)	72.3 (1.54)	9.6 (0.55)	608 (46)	630 (56)
Chi-Squared (4) ^b	68.8	85.3	11.7	4.1	144.7	19.5	15.9	17.0
P Value for chi-Squared (4)	<.0001	<.0001	.02	n.s.	<.0001	.0006	.003	.002

Note: All standard errors (shown in parentheses) are corrected for intertemporal and intrafamily correlations. Dollars are expressed in June 1984 dollars. Visits are face-to-face contacts with MD, DO, or other health providers; excludes visits for only radiology, anesthesiology or pathology services. Visits and expenses exclude dental care and outpatient psychotherapy.

^aThe figures in this column are adjusted for the imbalance of plans across sites as follows: the site-specific responses on each plan (simple means by site) are weighted by the fraction of the sample in each site and summed across sites. In the case of the 50 percent plan, which has no observations in Seattle, the weights are renormalized excluding Seattle.

^bThe *chi*-square statistic with 4 d.f. tests the null hypothesis of no difference among the five plan means. The *chi*-square statistic is a Wald test from the robust estimate of the information matrix (see Brook et al., 1984, for further details). It is used in lieu of the usual *F*-statistic because of the difficulty of computing such a statistic while allowing for intertemporal and interfamily correlation.

The Individual Deductible plan exhibits a somewhat different pattern from the other cost sharing plans. Recall that this plan has free inpatient care, but a 95 percent coinsurance rate (up to a \$150 per person, or \$450 per family annual maximum) for outpatient services. Total expenditures on this plan are significantly less than the free plan ($t = -2.34$, $p < .02$). This overall response is the sum of a one-third reduction in outpatient expenses ($t = -6.67$), and a less than one-tenth reduction in inpatient expenses ($t = -0.68$). Thus, this plan looks like a combination of the 50 or 95 percent plans for outpatient care and the free of 25 percent plan for inpatient care. The admission rate for the Individual Deductible plan lies roughly midway between the free plan and family coinsurance plan rates, suggesting a nontrivial cross-price elasticity between inpatient and outpatient services.

B. Main Effects of Insurance Plan:
Four-Equation Estimates

Because sample means are quite sensitive to the presence of catastrophic cases, we used the four-equation model to provide more robust estimates of the plan responses.²² The use of covariates in these equations further enhances precision and re-

²²For example, the ANOVA estimates of the response to cost sharing for total expenses (not adjusted for site) show a statistically insignificant reversal between the 50 and 25 percent plans. Although such a reversal is compatible with theory (due to the MDE) the reversal is almost certainly due to chance. One participant on the 50 percent plan had a very expensive hospitalization (total medical expenses of \$148,000 in one year); that single observation, which was the largest observation in the entire sample, adds \$106 dollars to the 50 percent plan mean (16 percent of that plan's mean).

TABLE 3—VARIOUS MEASURES OF PREDICTED MEAN ANNUAL USE OF MEDICAL SERVICES, BY PLAN

Plan	Likelihood of Any Use (%)	One or More Admissions (%)	Medical Expenses (1984 \$)
Free	86.7 (0.67)	10.37 (0.420)	777 (32.8)
Family Pay			
25 Percent	78.8 (0.99)	8.83 (0.379)	630 (29.0)
50 Percent	74.3 (1.86)	8.31 (0.400)	583 (32.6)
95 Percent	68.0 (1.48)	7.75 (0.354)	534 (27.4)
Individual			
Deductible	72.6 (1.14)	9.52 (0.529)	623 (34.6)

Note: Standard errors are shown in parentheses. Medical services exclude dental and outpatient psychotherapy. The predictions are for the enrollment population carried forward through each year of the study. The standard errors are corrected for intertemporal and intrafamily correlation. The *t*-statistics for the contrasts with the free plan are -6.69, -6.33, -11.57, and -10.69 for the last four rows of the first col., respectively; -2.74, -3.57, -4.80, and -1.28 for the last four rows of the second col., respectively, and -4.05, -4.91, -6.74, and -3.78 for the last four rows of the third col., respectively. These *t*-statistics are larger than those one would compute from the standard errors shown in the table because use of the standard errors ignores the positive covariance between the two predicted plan means from the shared $X\beta$ terms. The differences in expenses between the 25 and 50 percent plans are significant at the 5 percent level ($t = 1.97$), and between the 50 and 95 percent plans are significant at the 6 percent level ($t = 1.93$). The parameter estimates underlying these predictions are available in Manning et al. (1987).

moves the relatively minor imbalances across plan, including the site imbalance. Table 3 presents estimates from this model of plan response for the probability of any use of medical services, the unconditional probability of any inpatient use, and total medical expenses. Figure 1 displays the expenditure results.

Mean predicted expenditure in the free care plan is 46 percent higher than in the 95 percent plan ($p < .001$), almost exactly the difference found in the sample means.²³ Like

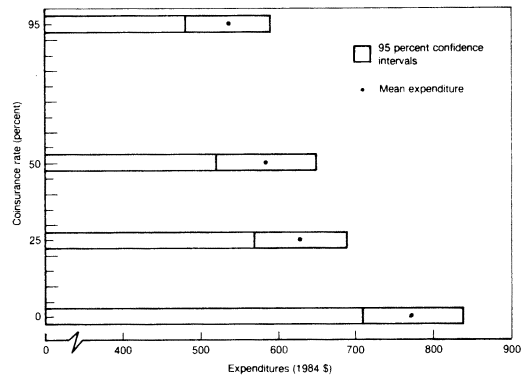


FIGURE 1. DEMAND AND 95 PERCENT CONFIDENCE INTERVALS BY COINSURANCE RATE

the sample means, these more robust estimates also indicate that the largest response to plan occurs between free care and the 25 percent plan, with smaller decreases thereafter.

Not surprisingly, given the approximate orthogonality of plan and covariates, adding covariates does not change the estimated probability of any use of medical services—87 percent of the free plan participants are predicted to use any service during the course of the year, while only 68 percent of the 95 percent plan participants are. These differences in the likelihood of receiving any care account for over three-fifths of the overall response to cost sharing. Virtually all the remaining response is attributable to the effect of cost sharing on hospital admissions.

Cost sharing for outpatient services only (the individual deductible plan) produces a different pattern of utilization than cost sharing for all services. Outpatient-only cost sharing reduces total expenditures relative to free care ($p < .0001$), largely by reducing the likelihood of any use ($p < .0001$). Outpatient-only cost sharing also reduces inpatient use, but by an insignificant amount ($p = .20$ for the probability of any inpatient use). This last result is the only important change

²³ It may seem that this is a trivial result that follows from the orthogonality of plan and covariates. Such is not the case because of the nonlinear transformations in the 4-part model. Using the logarithm of expenditure

plus \$5, for example, as a dependent variable instead of the 4-part model would lead to a much larger estimate of plan response, one that would be biased upward. (See Duan et al., 1983; Manning et al., 1987.)

TABLE 4—VARIOUS MEASURES OF PREDICTED ANNUAL USE OF MEDICAL SERVICES, BY INCOME GROUP

Plan	Income			Significance Tests <i>t</i> on Contrast of:	
	Lowest Third Mean	Middle Third Mean	Highest Third Mean	Middle vs. Lowest Thirds ^a	Highest vs. Lowest Thirds ^a
Likelihood of Any Use (Percent)					
Free	82.8	87.4	90.1	4.91	5.90
Family Pay					
25 Percent	71.8	80.1	84.8	5.45	6.28
50 Percent	64.7	76.2	82.3	4.35	4.86
95 Percent	61.7	68.9	73.8	3.96	4.64
Individual					
Deductible	65.3	73.9	79.1	6.09	7.09
Likelihood of One or More Admissions (Percent)					
Free	10.63	10.14	10.35	-0.91	-0.35
Family Pay					
25 Percent	10.03	8.44	7.97	-2.95	-2.75
50 Percent	9.08	8.06	7.77	-1.78	-1.66
95 Percent	8.77	7.38	7.07	-2.79	-2.46
Individual					
Deductible	9.26	9.44	9.88	0.31	0.68
Expenses (1984 \$)					
Free	788	736	809	-1.78	0.53
Family Pay					
25 Percent	680	588	623	-3.17	-1.47
50 Percent	610	550	590	-1.89	-0.49
95 Percent	581	494	527	-3.09	-1.41
Individual					
Deductible	609	594	670	-0.57	1.38

Note: Excludes dental and outpatient psychotherapy. Predictions for enrollment population carried forward for all years of the study.

^aThe *t*-statistics are corrected for intertemporal and intrafamily correlation. The statistics test the null hypothesis that the mean of middle (highest) third equals the mean of the lowest third; for example, the 4.91 figure implies we can reject at the .001 level the hypothesis that in the free plan the likelihood of any use for the lowest and middle thirds of the income distribution are equal.

from the previously published analysis of the first 40 percent of the data (Newhouse et al., 1981). In that analysis, inpatient use was less on the deductible plan, and one could reject at the 5 percent level the hypothesis that the free plan and individual deductible plan means for inpatient use were the same. This difference may have occurred because inflation in the late 1970's reduced the real value of the deductible, which was kept fixed at \$150 (i.e., in nominal dollars), or may have simply been due to chance.

C. Use by Subgroups

An important goal of the HIE was to study how the response to cost sharing varied

across subgroups. These included differences in responses across income groups, differences between adults and children, differences between the sickly and healthy, as well as differences across time (for example, any transitory surges in use as insurance changed), and differences across medical markets (for example, urban vs. rural).

1. *Across Income Groups.* Different aspects of the use of medical services exhibit different responses to income (Table 4).²⁴ In Ta-

²⁴Recall that the income measure comes from the first partial year of enrollment.) The division into thirds is site specific (for example, the lowest third is the lowest third of each site's income distribution), because

ble 4 we observe differences in use that are due to both income directly and the effects of variables correlated with income; that is, these are not partial effects.

Within each of the five plans the probability of any use of medical services increases with income, with larger increases for the family pay (25, 50, and 95 percent) and individual deductible plans than the free plan.²⁵ In contrast, the (unconditional) probability of any use of inpatient services declines with income for the family pay plans, but is not significantly different across income groups for the two plans with free inpatient care (the free and individual deductible plans). Because of these two conflicting effects of income—positive on outpatient use but negative on inpatient use—the net result on total expenditure is a shallow U-shaped response.

Our estimate of the differences by income group within the family-pay plans is influenced by the income-related upper limit in out-of-pocket expenses. The observed response in a combination of the direct response to income, and the fact that families with lower incomes are more likely to exceed their (lower) limit and receive free care for part of the year.²⁶ If medical care is a normal good, then any positive direct effect of greater income would be reduced by the decreased likelihood of going over the limit. In the case of the positive effect of income on the probability of any use, the direct income effect is probably more important, and in the case of the negative effect on the probability of any inpatient use, the limit has relatively more influence.²⁷

1) expenses are not corrected for cross-sectional differences in prices, and 2) we did not want to confound income and site; the sites were chosen to represent a spectrum of medical market characteristics. See Manning et al. (1987, Table 1, Appendix D) for the ANOVA estimates by plan income group (as well as by other subgroups).

²⁵Note that this is not a *ceteris paribus* statement, so there is no contradiction with standard theory, which would suggest no income effect in the free plan.

²⁶See Manning et al. (1987, Appendix B) for data on the proportion exceeding the upper limit on out-of-pocket expenses.

²⁷Some may argue that income is endogenous with respect to inpatient expenditure. This may well be true,

The Individual Deductible plan provides a cleaner test of the differences by income group of use of medical services, because the deductible in that plan is not income related. We observe an insignificant 10 percent increase in medical expenses between the bottom and top third of the income distribution. The effect of income is limited to an increased likelihood of using outpatient services, probably because inpatient services are free on this plan.

Thus far we have compared response among income groups rather than examining the partial effect of income. Although income has a statistically significant positive partial effect on use of service, the magnitude is small enough to be swamped by other factors correlated with income (for details see Manning et al., 1987, Appendix A, Tables 2–4 and 6).²⁸

2. *Across Age Groups.* We found about the same outpatient response to insurance plans for children (ages less than 18) as for adults, but children are less plan responsive for inpatient care (Table 5).²⁹ As we observed with a subset of these data (see Newhouse et al., 1981 and 1982; Leibowitz et al., 1985), we cannot reject the hypothesis that admission rates for children show no response to insurance coverage.³⁰ By contrast, adults

but is not likely to account for our result because only a few months of data are “tainted.”

²⁸Income has a moderately significant (at $p < .10$) and positive partial effect on use in all but the inpatient expenditure equation; in the level of outpatient-only expenditures, however, the income coefficients are of mixed sign. The probabilities with which we can reject the null hypothesis that the income coefficients are zero are: $p < .001$ for any use of medical services, $p < .10$ for the probability of any inpatient use given any medical use, $p < .001$ for the (log) level of outpatient-only use, and $p > .10$ for (log) level of medical expenditure if any inpatient use. The test statistics include plan income interactions and missing value replacement dummy variables.

²⁹Recall that children are overrepresented in the study relative to the population of our sites. Hence, our estimates understate (modestly) the population responsiveness in our sites.

³⁰ $\chi^2(4) = 5.19$ using ANOVA estimates for the probability of any inpatient use, and $\chi^2(4) = 5.36$ for the admission rate. Another possible hypothesis is no differential plan response for children relative to adults.

TABLE 5—VARIOUS MEASURES OF PREDICTED ANNUAL USE OF MEDICAL SERVICES, BY AGE GROUP AND PLAN

Plan	Likelihood of Any Use (%) Mean	One or More Admissions (%) Mean	Medical Expenses (1984 \$) Mean
Children			
Free	84.0	5.33	346
Family Pay			
25 Percent	75.1	4.98	287
50 Percent	70.3	4.62	279
95 Percent	63.5	4.23	236
Individual			
Deductible	68.5	5.86	299
Adults			
Free	88.6	13.9	1080
Family Pay			
25 Percent	81.4	11.5	872
50 Percent	77.1	10.9	797
95 Percent	71.2	10.2	744
Individual			
Deductible	75.6	12.1	852

Note: Excludes dental and outpatient psychotherapy services. The eight *t*-statistics for the contrasts between the free plan and the pay plans for the likelihood of any use all exceed 6. For one or more admissions, the *t*-statistics for children for contrasts with the free plan (rows 2–5) are 0.55, 1.13, 1.81, and –0.63, respectively, and for adults are 2.92, 3.64, 4.69, and 1.89, respectively (for example, the *t*-statistic on the difference between 13.9 and 12.1 is 1.89). For medical expenses the *t*-statistics on contrasts with the free plan for children are 2.16, 2.20, 4.10, and 1.42, respectively, and for adults are 3.70, 4.80, 6.07, and 3.63, respectively.

have significantly lower use of inpatient services on the family-pay plans than they do on the free plan.³¹ For outpatient services, we observe a very similar pattern of plan responses for children and adults.

D. Other Subgroups

1. *Health Status.* Although health status was a strong predictor of expenditure levels, we

observed no differential response to health insurance coverage between the healthy and the sickly (Manning et al., 1987). This null result is striking because of the upper limit feature. If anything, the presence of an upper limit on out-of-pocket expenses would lead to less plan response for the sickly; all other things equal, sicker individuals are more likely to exceed their upper limit and receive some free care—especially on the 95 percent plan, where care is free after gross expenditures of \$1050 or more. Furthermore, some might expect the sickly to be less responsive to insurance coverage than the healthy, on the supposition that their use of services is less discretionary. If, in fact, there is no interaction between plan and health status, one can infer that the opposite is true at the margin; that is, at the margin the sickly exhibit more discretion.

2. *Sites.* The six sites in the HIE were selected to reflect a spectrum of city sizes, waiting times to appointment, and physician to

We can reject this hypothesis; the test statistics are $\chi^2(4) = 16.49$ for the probability of any inpatient use and $\chi^2(4) = 14.08$ for total admissions. Hence, it appears that children and adults respond differently and that children do not respond to cost sharing for inpatient care.

³¹ $\chi^2(3) = 24.22$ for the probability of any inpatient use and 16.31 for the admission rate. By contrast, there are no significant differences among the family pay plans for adults. $\chi^2(2) = 1.69$ for expenditures, 0.73 for total admissions, and 1.39 for the probability of any inpatient use, again based on ANOVA (see Manning et al., 1987, Table 2, Appendix D for the ANOVA estimates).

TABLE 6—USE OF DENTAL SERVICES BY DENTAL PLAN: SAMPLE MEANS

Dental Insurance Plan	Year 1 of Dental Coverage			Year 2 of Dental Coverage		
	Probability (%)	Visits	Expenses per Enrollee (\$)	Probability (%)	Visits	Expenses per Enrollee (\$)
Free	68.7 (1.19)	2.50 (.065)	380 (18.0)	66.8 (1.18)	1.93 (.049)	261 (12.5)
25 Percent	53.6 (3.39)	1.73 (.138)	224 (32.8)	52.6 (3.34)	1.51 (.111)	190 (28.0)
50 Percent	54.1 (2.41)	1.80 (.118)	219 (31.3)	53.0 (2.55)	1.50 (.103)	177 (32.3)
95 Percent	47.1 (2.59)	1.39 (.098)	147 (18.7)	48.3 (2.62)	1.44 (.099)	179 (24.9)
Individual Deductible	48.9 (2.12)	1.70 (.104)	242 (24.1)	48.1 (2.12)	1.33 (0.080)	158 (20.4)

Note: Expenses were converted to January 1984 dollars using the dental fee component of the Consumer Price Index. There has been no adjustment for regional differences in prices, or differences in population characteristics across plans and years. Standard errors (shown in parentheses) are corrected for intrafamily correlations.

population ratios (Newhouse, 1974).³² Our concern was that the response to insurance coverage could vary according to the complexity of the medical market or to the excess demand in the medical delivery system. Yet we found no differences among the sites in the response to insurance coverage, $\chi^2(19) = 14.96$ ($p > .50$). The uniformity of response across the sites gives some reason to believe the results may be representative of the United States, and we have so used them below.

Interestingly, the site with the longest delay to appointment and lowest physician to population ratio (Fitchburg) had the second highest probability of any use, the second highest expenditures per enrollee, and the highest probability of any inpatient use. The latter two phenomena may represent substitution of inpatient for outpatient care (Jeffrey McCombs, 1984), and the first may indicate that the presence of emergency rooms removes the constraint of the queue

(Stephen Long, Russell Settle, and Bruce Stuart, 1986).³³

2. *Period of Enrollment.* As noted above, we enrolled families for three or five years to see if the response to insurance changed over time and if the duration of enrollment mattered. The free plan might generate transitorily high demand; the 95 percent plan might generate postponement of demand at the end of the experiment (Arrow, 1975; Metcalf, 1973). Neither effect was found; see Manning et al. (1987) for further details.³⁴ Nor did duration of enrollment matter to either the absolute level of spending or the responsiveness to plan.

3. *Subexperiments.* As described above, the Health Insurance Experiment contained a number of subexperiments to study methods effects. None of the subexperiments had a measurable effect on expenditure (Manning et al., 1987).

³² For example, city sizes in 1970 ranged from 34,000 (Georgetown County) to 1.2 million (Seattle), waiting times for nonemergent care in 1973–74 ranged from 4.1 days (Seattle) to 25.0 days (Fitchburg), and physicians per capita in 1972 ranged from 30 per 100,000 (Fitchburg) to 59 per 100,000 (Seattle).

³³ Length of waiting time to an appointment with a primary care physician is associated positively with the use of emergency rooms (O'Grady et al.).

³⁴ A transitory effect was found for dental services; see Manning et al. (1985, 1986a) for details.

E. Dental Results

These results are reported in greater detail elsewhere (Manning et al., 1985, 1986a). Dental services do show greater responsiveness to plan in Year 1 than in subsequent years ($p < .001$) (Table 6). This would be expected if dental services were more durable than other medical services, as is plausible. The responsiveness of demand by plan in Year 2, which is typical of the middle years, is of the same general magnitude as that for other medical services.

F. Health Status Outcome Results

These results also are reported in greater detail elsewhere (Brook et al., 1983, 1984; Valdez et al.; Valdez; Howard Bailit et al., 1985). For the person with mean characteristics, we can rule out clinically significant benefits from the additional services in the free fee-for-service plan relative to either the cost-sharing plans or the HMO experimental group. For poor adults (the lowest 20 percent of the income distribution) who began the experiment with high blood pressure (specifically, who were in the upper 20 percent of the diastolic blood pressure distribution) there was a clinically significant reduction in blood pressure in the free fee-for-service plan compared to the plans with cost sharing. Epidemiologic data imply that the magnitude of this reduction would lower mortality about 10 percent each year among this group, about 6 percent of the population. (The sample size is much too small to test this prediction with actual mortality among the experimental population.) For poor adults who began the experiment with vision problems that were correctable with eyeglasses, there was a modest improvement in corrected vision. Individuals on the free care plan between the ages of 12 and 35 showed a modest improvement in the health of the gums; caries (decayed teeth) were also more likely to be filled on the free care plan.

The specific gains in health just described, for high blood pressure, myopia, and dental care, were all for relatively prevalent chronic problems (of course, we had difficulty detecting effects for rare problems) that are rela-

tively inexpensive to diagnose and remedy. One can infer that programs targeted at these problems would be much more cost effective in achieving these gains in health than free care for all services. For example, more than half the benefit of free care for high blood pressure (and presumably for risk of dying) was available from a one-time screening examination, whose cost is a small fraction of free care for all services (Keeler et al., 1985).

G. Health Maintenance Organization Results

We also randomized a group of participants into an HMO, the Group Health Cooperative of Puget Sound in Seattle.³⁵ This group, whom we call the HMO Experimentals, was given a plan of benefits identical to the free fee-for-service (FFS) plan. In addition, we enrolled a random sample of existing HMO enrollees, the HMO Controls. Thus, a comparison of the experimentals and the free fee-for-service plan establishes the "pure" HMO effect on use; a comparison of the experimentals and controls establishes the extent, if any, of selection with respect to the HMO.³⁶

Our results (Table 7) show no evidence of selection in the single HMO that we studied; those previously enrolled at the HMO (the Controls) used services at approximately the same rate as those who were not previously enrolled (the Experimentals). By contrast, the percentage of Experimental plan participants with one or more hospital admissions was only two-thirds as great as the percentage on the free fee-for-service plan. Because outpatient use was approximately similar on the two plans, the expenditure difference between the HMO Experimentals and free fee-for-service participants was

³⁵An HMO is reimbursed a fixed amount per month, in return for which it agrees to provide medical care. Thus, unlike fee-for-service medicine, the approximate marginal revenue from delivering additional services is zero. Of course, there are market constraints on the HMO's behavior because it competes with fee-for-service medicine for patients.

³⁶The fee-for-service sample in this comparison is from Seattle, in order to keep the population sampled the same between the two groups.

TABLE 7—ANNUAL USE OF MEDICAL SERVICES PER CAPITA, SEATTLE SAMPLE, BY HMO AND FFS STATUS^a

Plan	Likelihood of Any Use (%)	One or More Admissions (%)	Imputed Expenditures ANOVA ^b (1983 \$)	Imputed Expenditures with Age-Sex Covariates ^b (1983 \$)	Person Years
HMO Experimental	87.0 (1.0)	7.1 (0.50)	434 (28)	426 (23)	3687
HMO Control	91.1 (0.8)	6.4 (0.55)	432 (34)	465 (47)	2596
Free Fee-for-Service	85.3 (1.6)	11.2 (1.17)	640 (81)	612 (66)	1221
<i>t</i> -Statistic on Free-Experimental Difference ^c	-0.88	3.24	2.44	2.69	
<i>p</i> Value for <i>t</i> -Statistic, 2 tail	n.s.	.0012	.016	.007	

^aStandard errors are shown in parentheses. The sample includes participants while they remained in the Seattle area. The sample excludes children born into the study and excludes partial years except for deaths, similar to Tables 1 and 2 above. For HMO Controls and Experimentals, the data include both in- and out-of-plan use. The standard errors are corrected for intertemporal and intrafamily correlation using an approach due to Huber in a similar fashion to Tables 1 and 2 above. The numbers differ slightly from those in Manning et al. (1984), because of minor corrections in the data, as well as the use of a less precise, but more robust method of calculating standard errors. The method is the same as that described in Table 2.

^bSee Manning et al. (1984) for details of imputation method.

^cTesting null hypothesis of no difference between HMO Experimental and Free Fee-for-Service plan.

somewhat narrower; expenditures per person among the HMO Experimentals were only 72 percent of expenditures on the free fee-for-service plan.

These findings demonstrate that a markedly less hospital-intensive style of medicine than is commonly practiced in the fee-for-service system is technically feasible. Whether the technical style will be attractive to consumers, and, if it is, whether a market of competing HMOs is economically feasible—or whether adverse selection problems will prove insurmountable (Michael Rothschild and Joseph Stiglitz, 1976)—are still somewhat open questions, although the size and history of large HMOs such as Group Health Cooperative of Puget Sound suggest that the style is attractive to some consumers.

In projecting the effect of the growing HMO market share on hospital admissions and medical expenditure, one must keep in mind that the above comparisons have been made against the free care plan. Because virtually all private fee-for-service health insurance plans include some cost sharing, one should compare the reduction in hospital

admissions at the HMO, some 35 percent, with the reduction caused by cost sharing, some 15 to 25 percent depending on plan. The values presented above, however, do represent the *ceteris paribus* HMO effect; if an HMO were to use cost sharing, its observed rates of use might be even lower.

Consumers contemplating enrollment in an HMO will weigh the cost savings against any effect of the reduction in services upon health status and consumer satisfaction. Our findings on health status of the HMO are analogous to those in the free fee-for-service system; the mean person in the fee-for-service plan appeared to derive few or no benefits from the additional hospital services (Ware et al., 1986; Elizabeth Sloss et al. 1987). Those who are both in poor health and of low income who were in the HMO exhibited a higher rate of bed-days and serious symptoms (relative to those in the free fee-for-service plan). There is thus some suggestive evidence that special programs to facilitate access for Medicaid enrollees in HMOs may be worthwhile, but we caution that this result comes only from one HMO

(albeit a well-established and well-regarded HMO) and that the precision with which we could measure results among the poor, sick group makes this result less than definitive, even in the case of this HMO.

Those who had self-selected the HMO (the Controls) were on average as satisfied with their care as those in the fee-for-service system (Allyson Davies et al., 1986). Theory would suggest the marginal person would be equally satisfied in both systems, and it is not surprising that we detected no difference for the average person. By contrast, the HMO Experimentals were less satisfied overall with their care than those in the fee-for-service system, although on certain dimensions they were as satisfied or even more satisfied.

IV. Conclusions

A. *On Comparing our Estimates of Demand with those in the Literature*

Our results leave little doubt that demand elasticities for medical care are nonzero and indeed that the response to cost sharing is nontrivial. How do our estimates compare with those in the nonexperimental literature?

This question is difficult to answer, because most prior empirical work has parameterized cost sharing as a constant coinsurance rate (for example, Feldstein, 1971, 1977) or has examined particular changes in insurance plans (for example, an imposition of a \$3 per visit copayment: Scitovsky and Snyder; Phelps and Newhouse, 1972; Scitovsky and McCall). By contrast, experimental policies were from a two-parameter family (coinsurance rate and maximum dollar expenditure). We make no apologies for this intentional noncomparability; a constant coinsurance rate, while convenient for obtaining comparative statics results, is not an insurance policy that theory suggests would be optimal, assuming risk aversion (Arrow, 1963, 1971, 1973, 1975). Indeed, an optimal policy would almost certainly contain a stop-loss feature, exactly as the experimental plans did.³⁷

³⁷A stop-loss feature means there is a maximum out-of-pocket loss that the insured can sustain. In ad-

dition to its risk-reduction properties, no worst-case payment would have been possible without a stop-loss feature, and hence selection effects might have been introduced into the experiment.

One could, of course, attempt to estimate the functional response of demand to variation in the two parameters; one can view the values presented above as selected points in the response surface generated by varying coinsurance at given maximum dollar expenditure levels. In order to compare our results with those in the literature, however, we must extrapolate to another part of the response surface, namely, the response to coinsurance variation when there is no maximum dollar expenditure. Although any such extrapolation is hazardous (and of little practical relevance given the considerable departure from optimality of such an insurance policy), we have undertaken such an extrapolation rather than forego entirely any comparison with the literature. Specifically, we have used three different methods to estimate a price elasticity comparable to the estimates in the literature:

1) One can estimate a pure coinsurance elasticity by analyzing variation in the demand for episodes of care rather than annual expenditure per person (Keeler and John Rolph, 1982; Keeler et al., 1987). The theory of demand suggests that individuals who have not yet exceeded the upper limit on out-of-pocket expenses, when making a marginal medical consumption decision, will discount the nominal price by the probability of exceeding the limit (because with that probability the true price is zero) (Keeler, Newhouse, and Phelps, 1977; Randall Ellis, 1986).³⁸ We therefore examine demand for episodes of treatment by individuals who are more than \$400 from their limit. This gives an approximation of the pure price effect if such people treat the true probability of exceeding their limit as nearly zero.³⁹ The

dition to its risk-reduction properties, no worst-case payment would have been possible without a stop-loss feature, and hence selection effects might have been introduced into the experiment.

³⁸The specific result requires risk neutrality and separability of the utility function in health and money, but the qualitative results does not.

³⁹Because there was no appreciable difference between demand for outpatient episodes when the MDE remaining was between \$1 and \$400 and when it was more than \$400, this assumption seems reasonable for

TABLE 8—ARC ELASTICITIES FOR VARIOUS TYPES OF CARE
CALCULATED FROM EPISODES^a

Range of Nominal Coinsurance Variation	Type of Care					
	Outpatient ^b				Hospital	All Care ^c
	Acute	Chronic	Well	All ^c		
0–25 Percent	.16 (.02)	.20 (.04)	.14 (.02)	.17 (.02)	.17 (.04)	.17 (.02)
25–95 Percent	.32 (.05)	.23 (.07)	.43 (.05)	.31 (.04)	.14 (.10)	.22 (.06)

^a The method of calculating standard errors (shown in parentheses) is described in Keeler et al. (1987).

^b Acute conditions are unforeseen and treatment opportunities are nondeferrable. Chronic episodes comprise foreseen and continuing expenditure; treatment is designed to ameliorate the consequences of the disease rather than cure. Flare-up of chronic conditions, which are unforeseen, we treat as acute. Well care episodes are medically deferrable without great loss and can occur when the patient is not considered sick.

^c Estimate derived by weighting elasticities for various types of care by budget shares.

estimation method controls for unobserved propensities to have episodes, as well as other observed covariates by looking at experience before and after the MDE is exceeded; see Keeler and Rolph for a description of the methodology. We have computed arc elasticities for the 0–25 and 25–95 percent ranges of coinsurance; those elasticities are shown in Table 8.

2) A second estimate comes from using an indirect utility function and applying it to total expenditure in the 25–95 percent range. This estimate is very close to the first, -0.18 (Manning, 1986).

3) A third estimate comes from a similar calculation to those in the literature, that is, it uses average coinsurance rates (Table 9). The usual proof of an upward bias in the elasticity estimate from using the average coinsurance rate (Newhouse, Phelps, and Marquis) does not apply here because of the balance across plans. The amount of bias, if any, depends on two effects that work in opposite directions. For small expenditures the experimental plans will exhibit smaller expenditure than would a pure coinsurance

TABLE 9—ARC ELASTICITIES FOR VARIOUS
TYPES OF CARE CALCULATED FROM AVERAGE
COINSURANCE RATES

Range of Nominal Coinsurance Variation	Range of Average Coinsurance Variation	All Care	Outpatient Care
0–25 Percent	0–16	.10	.13
25–95 Percent	16–31	.14	.21

Source: Calculated from data in Table 2 (outpatient) and Table 3 (total). For those who wish to calculate arc elasticities with the 50 percent plan, from the data in Tables 2 or 3, the average coinsurance rate in the 50 percent plan is 24 percent.

rate plan of 16 or 31 percent (because the effective coinsurance rate is likely to be higher); for large expenditures exceeding the MDE the opposite will be true (because the marginal coinsurance rate will be zero, not positive). Which effect predominates is an empirical question the experimental data cannot resolve; empirically, this method yields values that are somewhat lower but still close to those of the other two methods. (The lower value suggests the first bias predominates.)

In sum, these three methods suggest that price elasticities for a constant coinsurance policy are in the -0.1 to -0.2 range, values that are consistent with those in the lower range of the nonexperimental literature.

outpatient episodes. It may cause some bias in the estimated hospital elasticity; if the true MDE were, say \$10,000 rather than \$1000, we might observe fewer hospitalizations.

B. *On the Explanation of the Sustained Rise in Medical Expenditure*

At first blush, our estimates of demand response imply that the spread of health insurance can account for only a modest portion of the postwar rise in medical expenditure, contrary to the commonly held view described in the introduction. Between 1950 and 1984, real medical expenditure rose by a factor of 7,⁴⁰ but our estimates of insurance elasticity do not begin to imply this degree of increase. To demonstrate this point, we use the average coinsurance rate. Despite its imperfect measure of the generosity of insurance, it is a gross measure of how much insurance changed over the post-1950 period and therefore indicative of the role insurance might have played in this increase. Table 10 shows the average coinsurance rate by type of service (see Table 9 for comparable values from the 25, 50, and 95 percent plans). Although the figures by service are based on an arbitrary accounting convention, they suggest that the change in insurance in the postwar period was of roughly the same absolute magnitude as the difference between the 95 percent coinsurance and free care plans.⁴¹

Because the free plan demand was only around 1.5 times that of the 95 percent plan, it appears that the change in insurance can explain only a small part, perhaps a tenth, of the factor of 7 change in health expenditure in the post-World War II period.

Nor can changes in real income (around a factor of 3 during this period) directly account for much of the rise. Income elasticities estimated from the experimental data (the partial response, not the one shown in Table 5) are at most 0.2—much too small to account for anything like a factor of 7 change.⁴²

⁴⁰Nominal expenditure data from Katherine Levit et al. (1985) deflated by the GNP deflator.

⁴¹The accounting convention used by the Health Care Financing Administration allocates a common deductible to services in proportion to gross expenditure. We have followed the same convention in calculating comparable figures from the experimental data.

⁴²Real GNP increased between 1950 and 1983 by a factor of 2.9. Even allowing for the usual downward

TABLE 10—CHANGE IN AVERAGE COINSURANCE RATE, 1950–84, BY TYPE OF SERVICE

Year	Hospital	Physician	Other	Total
1950	.30	.83	.86	.66
1984	.09	.28	.56	.28

Source: Levit et al. (1985).

Thus, we still must account for the bulk of the expenditure increase. The rather obvious “accounting” explanation of the expenditure increase is technological change; there are a host of new medical products and procedures today that did not exist in 1950. For example, those with kidney failure are now treated with renal dialysis or kidney transplantation; in 1950 these individuals died rather quickly. This merely pushes the puzzle back one stage, however; what role, if any, did insurance (and income growth) play in inducing the technological change? Unfortunately that question cannot be answered from experimental data.⁴³

Thus, if insurance is playing a role in inducing a welfare loss, given the rate of increase in medical expenditure, the bulk of that loss must come from its having induced innovation for which unsubsidized consumers would not be willing to pay.⁴⁴ Given that most countries in the world have also experienced a long-term sustained increase in expenditure despite widely varying institutional arrangements, it is at least arguable that consumers would be willing to pay for much of the increase, but there clearly

bias from using measured income to estimate income elasticities, it is clear that changes in income can only explain a modest portion of the expenditure increase.

⁴³Because most consumers have been insured for inpatient services throughout the relevant time period, it is an extremely difficult question to answer from non-experimental data. Moreover, one does not observe insurance policies that do and do not cover new procedures, so there is no straightforward test of willingness to pay for new technology. Although virtually all policies do not cover “experimental” procedures, once efficacy and “safety” are demonstrated, insurance plans tend to cover all procedures.

⁴⁴The willingness-to-pay calculation should include any willingness to pay for others’ care.

has been no pure market test (Newhouse, 1977, 1984).

C. *On the Magnitude of Welfare Loss from Health Insurance*

Setting aside the issue of possible welfare loss from induced technological change, one can estimate the welfare loss in the usual static framework. Under a number of strong assumptions (including that gross medical care prices are competitive and there are no externalities), our estimates imply a nontrivial welfare loss from first-dollar health insurance coverage. An approximation to the loss from moving from a universal 95 percent plan (with a \$1000 MDE) to the free care plan is \$37 to \$60 billion, as against an expenditure around \$200 billion on these services in 1984 by the under 65 population.⁴⁵

From the \$37–60 billion figure must be deducted some amount for the reduced risk in the free plan relative to the 95 percent plan. Usual values for risk aversion, however, would suggest the deduction is small in the presence of a \$1000 cap (Feldstein, 1973; Keeler, Morrow, and Newhouse, 1977). Although the \$37–60 billion figure is probably overstated by ignoring externalities and assuming medical care prices are competitive, it ignores any welfare loss from induced technological change.⁴⁶

D. *On the Existing Insurance Coverage of Various Medical Services*

One can find several economic reasons for the traditionally more generous coverage of inpatient services relative to outpatient services (Table 10). Loading charges (as a per-

⁴⁵ The \$37 and \$60 billion figures are calculated in the usual Harberger fashion by taking the \$325 per capita difference in spending between the 95 percent and free plans from Tables 3 and 6 (Year 2 values) and adding \$19 for mental health services (Wells et al., inflated by the change in the CPI Medical Services prices index between 1977 and 1984). We then multiply by 207 million, the number of resident civilians under 65. This yields a figure of \$71 billion. One then multiplies by 0.525 or 0.845. Both fractions are larger than the usual 0.5 because we do not start at an unsubsidized point. Our 95 percent \$1000 MDE plan had an average coinsurance rate of 0.31. An upper bound on the welfare loss comes from assuming that individuals valued the last dollar at 0.31. A lower bound on the welfare loss comes from assuming that the extra spending is all from individuals who valued the last dollar of spending at 0.95, the nominal coinsurance rate. The 0.525 figure equals $1 - .95/2$, and the 0.845 figure equals $1 - .31/2$.

The \$200 billion figure can be estimated in two ways: 1) Data from Levit et al. show expenditure on personal health care services of \$342 billion in 1984. Waldo and Lazenby (1984, Table 11) estimate that \$120 billion of this is for the over 65, leaving \$222 billion for the under 65. Some of this, however, is for noncovered services, such as nonprescription drugs, and some other part is for ineligible populations, such as the institutionalized. Adjusting for these noncomparabilities is necessarily somewhat imprecise, but would probably leave a final figure around \$200 billion. 2) Data from Tables 3 and 6 (Year 2 values) plus data on outpatient mental health spending from Wells et al. inflated to 1984 and scaled up by 207 million population imply an expenditure of \$224 billion on the free care plan in our sites and \$178 billion on the 25 percent coinsurance plans. Adjusting

for price and usage levels in our sites relative to the nation is necessarily imprecise, but these two values probably bracket the true national figure.

⁴⁶ The induced technological change is clearly only a welfare loss if patent protection is at the level to induce the appropriate investment in new products in an unsubsidized market. If there is not enough patent protection, there is no necessary welfare loss from insurance's inducing a too rapid rate of innovation. There appears to be one estimate in the literature of the welfare loss from induced change; Feldstein (1973) attempted to adjust for the willingness of consumers to pay for "higher quality care." There is no empirical way to do this, however, so the magnitude of the true welfare loss is highly problematic. Feldstein's method, although not explicit on the point, in effect ignores true technological change. He implicitly assumes that consumers in earlier years could have purchased "higher quality" medical care, but they chose not to because they faced a higher coinsurance rate and/or had lower incomes. (Alternatively, physician "norms of care" were lower because of the higher coinsurance rate and lower income.) As the renal dialysis example makes clear, however, consumers were simply unable to purchase some medical services in earlier years because they did not exist. In many cases their subsequent existence depended on fundamental scientific advance such as the discovery of DNA and would not have occurred without that advance, despite lower coinsurance or higher incomes. Whether consumers in the 1950's and early 1960's would have purchased such services if they had existed then obviously cannot be answered from actual expenditure data. Feldstein's method also yields an upper bound for the same reason our \$60 billion estimate is an upper bound.

centage of premium) are less, and the risk of a large loss is greater. For children, price elasticities for inpatient services are not measurably different from zero, and hence for them there is no measurable moral hazard.

This structure of more extensive insurance for inpatient services has been attacked as misguided, however (Milton Roemer et al., 1975), on the grounds that lack of insurance for outpatient services deters ignorant individuals from seeking care at a time in their illness when they can be treated relatively cheaply. Others have also asserted that the more generous coverage of inpatient services leads physicians to hospitalize patients who could be treated on an outpatient basis, thereby minimizing private but increasing social expenditure.

Analysis of a natural (not randomized) experiment supported the claim that more complete coverage of outpatient expenditure reduced total expenditure (Roemer et al.; L. Jay Helms, Newhouse, and Phelps, 1978), but a prior controlled experimental study testing this hypothesis rejected it (Charles Lewis and Harold Keairnes, 1970; D. B. Hill and James Veney, 1980). At issue is whether outpatient and inpatient services are substitutes or complements.

Our findings decisively reject the hypothesis that increased coverage of outpatient services, holding constant the coverage of inpatient services, will reduce expenditure. As Table 3 shows, the mean expenditure on the individual deductible plan (free inpatient, costly outpatient care) is 20 percent less than the mean on the free care plan (free inpatient, free outpatient care), and the difference is statistically significant ($p < .001$).⁴⁷ Disaggregation shows that the outpatient deductible not only reduces outpatient expenditure (Table 2) but, if anything, decreases hospital admissions for adults as well (Table 5). The (possibly) decreased admissions for adults suggests that outpatient and inpatient services are, if anything, complements not substitutes.

⁴⁷In the ANOVA results (Table 2), the estimated reduction is 19 percent and the t -statistic is 2.34 ($p < .02$, two-tailed test).

In the interests of brevity we summarize four other implications for health insurance coverage (these are discussed at greater length in Manning et al., 1987):

There appears to be little justification for the common practice of group insurance policies' treating emergency room services more generously than physician office visits, because emergency room services are as responsive to plan as physician office visits.⁴⁸

There is no support for the so-called offset hypothesis, namely that more complete coverage of psychotherapy services will reduce total medical costs (or at least not increase them) (W. Follette and Nicholas Cummings, 1967, 1968). The experimental data, however, are not very precise on this question.

The observed lesser coverage of outpatient mental health care relative to all outpatient care would be consistent with a greater plan response for mental health care. Although the estimated plan response is in fact substantially larger for mental health care, the difference with all outpatient care is statistically insignificant.⁴⁹

Well-care services are about as price responsive as other medical services. Although there are other reasons for the common practice of not covering well-care

⁴⁸We assume that a presumed lower response to insurance is the reason for greater coverage of emergency room services. The alternative explanations, differential loading charges or asymmetric information, are not particularly plausible as explanations of the better coverage of emergency room services. Asymmetric information (differential knowledge of insurer and insured) is not very relevant to a single insurance plan offered in a group setting unless the service is costly enough to motivate an employment change (which might apply to psychotherapy or certainly costly dental services such as orthodontia). Routine office visits do not match this description. Moreover, asymmetric information may apply to both office and emergency room services. An individual may know that his use of office visits differs from average (whereas the insurer does not) but may also know that his likelihood of an accident differs from average, and the insurer may not.

⁴⁹The estimated ratio of the free to 95 plan expenditures is 233 percent, compared with a 169 percent estimate for medical outpatient care (Manning et al., 1986b).

services as generously as other outpatient services (primarily there is little or no uncertainty and loadings are relatively high), greater price responsiveness is not a reason.

E. *Was It Worth It?*

One question frequently raised about social experimentation is whether its benefits are worth its costs (for example, Orley Ashenfelter, 1986; Robert Haveman, 1986). Because the question concerns the value of information, and because the links from this type of information to actual behavior are generally impossible to establish with any rigor, the question admits of no easy answer (save for the trivial case in which the experiment was so poorly designed or conducted that it produced no information). In other words, any attempt to justify the cost of an experiment is necessarily speculative.

Despite the circumstantial nature of the evidence, we believe that the benefits of this particular experiment greatly exceeded the (current dollar, undiscounted) costs of a little over \$80 million (\$136 million if put in 1984 dollars, and brought forward to 1984 using a 3 percent real discount rate.⁵⁰ Between 1982 and 1984, there was a remarkable increase in initial cost sharing in the United States, at least for hospital services. For example, the number of major companies with first-dollar charges for hospital care rose from 30 to 63 percent in those two years, and the number of such firms with an annual deductible of \$200 per person or more rose from 4 percent to 21 percent (Jeff Goldsmith, 1984). Although it is impossible to know how much of this change can be attributed to the experimental results, the initial findings of the experiment were published in December 1981 (Newhouse et al., 1981) and December 1983 (Brook et al., 1983) and given wide publicity in both the general and trade press. In certain instances a direct link between changes in cost sharing and the experimental results can be made.⁵¹

⁵⁰ We have used the GNP deflators to inflate costs.

⁵¹ For example, the Xerox Corporation in 1983 announced an increase in its deductible from \$100 per

person or \$200 per family to 1 percent of earnings per family. It raised coinsurance on hospital and surgical services from 0 to 20 percent. Additionally, it lowered its cap on out-of-pocket expenditures (analogous to the MDE) from 6 percent of earnings to 4 percent of earnings. In a brochure distributed to its employees it said: "According to a study by the Rand Corporation, when consumers are required to increase their share of medical costs, there is a significant decrease in the total amount spent for these services. Furthermore, this study—and other similar studies—does not indicate that the health of the employees was affected adversely by the decrease in costs." Despite the large increase in initial cost sharing, the average coinsurance rate for hospital services nationally only rose from 7.6 to 8.7 percent between 1982 and 1984. This modest change in the average rate may reflect both the lowering of ceilings on out-of-pocket expenditure, as in the Xerox case and the highly skewed distribution of hospital expenditure, which means most expenditure exceeds the initial cost sharing.

person or \$200 per family to 1 percent of earnings per family. It raised coinsurance on hospital and surgical services from 0 to 20 percent. Additionally, it lowered its cap on out-of-pocket expenditures (analogous to the MDE) from 6 percent of earnings to 4 percent of earnings. In a brochure distributed to its employees it said: "According to a study by the Rand Corporation, when consumers are required to increase their share of medical costs, there is a significant decrease in the total amount spent for these services. Furthermore, this study—and other similar studies—does not indicate that the health of the employees was affected adversely by the decrease in costs." Despite the large increase in initial cost sharing, the average coinsurance rate for hospital services nationally only rose from 7.6 to 8.7 percent between 1982 and 1984. This modest change in the average rate may reflect both the lowering of ceilings on out-of-pocket expenditure, as in the Xerox case and the highly skewed distribution of hospital expenditure, which means most expenditure exceeds the initial cost sharing.

⁵² The average cost per hospital day in 1984 was \$417. This uses the 1983 \$368 figure from the American Hospital Association (1984) inflated by 13.3 percent, the change in per day inpatient costs from 1983 to 1984 (American Hospital Association, 1985). Bernard Friedman and Pauly (1981, 1983) have argued that the marginal cost/average cost ratio for hospital services is near one. Hence, a *ceteris paribus* estimate of the savings from decreased use, assuming a marginal cost/average cost ratio of 0.9, is around \$7 billion ($19 \text{ million} \times 417 \times 0.9$). The American Hospital Association cost per day figure includes the over 65; however, cost per day is not very different for the over 65.

⁵³ In part, we do not account for such a change because the physician visit rate rose in 1985 to its 1982 value. Thus, the decrease from 1982 to 1984 could have been attributable to chance; alternatively the continued decrease in hospital care in 1985 (another 7.1 percent decrease in patient-days, USDHHS, 1987) may have led to a substitution of outpatient use. Data on physician visits are from the *National Health Interview Survey* (USDHHS, Series 10, 1985; 1986).

If all the changes in patient-days were attributable to the increased cost sharing, and if all the increase in cost sharing is due to the publication of experimental results, and if the benefits of the foregone use were negligible, as our results suggest, the experiment paid for itself in about a week $(.136/7)(52)$.⁵⁴ It is clear that these assumptions overstate the benefits of the experiment, yet it is equally clear that the assumptions can be greatly relaxed and still yield the result that the experiment was worth it. Moreover, we have ignored any benefits to countries other than the United States, and any benefits from the decrease in physician visits or changes in dental or mental health coverage or emergency room coverage. We have also ignored any benefits from the results of the HMO portion of the experiment, although HMO's market share has been expanding rapidly from a period just before and subsequent to our first article describing the HMO results (Manning et al., 1984a). Finally, we have ignored the value of the public use files to future research efforts.⁵⁵

Implicit in our conclusions is the assumption that one could not reduce uncertainty with nonexperimental data to the satisfaction of those making decisions about cost sharing. We believe this is likely to be true, because of the wide range of nonexperimental estimates of insurance elasticity cited in the introduction, the difficulty of inferring health status effects from nonexperimental data, and the temporal proximity of the changes in cost sharing to the publication of the experimental results (many of the nonexperimental results had been in the literature for a decade, during which time cost sharing had, if anything, decreased). Thus, we think it highly plausible that the benefits of this endeavor were indeed worth its costs.

⁵⁴ The negligible benefits assumption relies on the observation that cost sharing for hospital services was near zero in 1982 and that there were no measurable health benefits outside the dental area for the middle-class employees who would have been the dominant group for whom the cost sharing changed.

⁵⁵ The public use files can be ordered from Publications Department, The Rand Corporation, 1700 Main Street, Santa Monica, CA 90406-2138.

F. On Experimentation in Economics

Econometric and economics texts often have a statement near the beginning that experimentation is not nearly as possible in economics as it is in the physical sciences. Perhaps the degree of difference is not as great as many think. Well-designed and executed field and laboratory experiments are feasible and can add substantially to the body of knowledge (Walter Heller, 1975; Charles Plott, 1982).⁵⁶ We hope this example will encourage others to ask whether an experiment is practical or feasible when approaching empirical questions.

⁵⁶ For other views of field experiments see Jerry Hausman and David Wise (1985) and Robert Ferber and Werner Hirsch (1978).

REFERENCES

- Arrow, Kenneth J., "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, December 1963, 53, 941-73.
- , *Essays in the Theory of Risk-Bearing*, Chicago: Markham, 1971.
- , *Optimal Insurance and Generalized Deductibles*, Publ. No. R-1108-OEO, Santa Monica: Rand Corporation, 1973.
- , "Two Notes on Inferring Long Run Behavior from Social Experiments," Publ. No. P-5546, Rand Corporation, 1975.
- Ashenfelter, Orley, "Book Review of *The National Supported Work Demonstration*," *Journal of Economic Literature*, September 1986, 24, 1268-1270.
- Bailit, Howard et al., "Does More Generous Dental Insurance Coverage Improve Oral Health?," *Journal of the American Dental Association*, May 1985, 110, 701-07.
- Barer, Morris L., Evans, Robert G. and Stoddart, Gregory L., "Controlling Health Care Costs by Direct Charges to Patients: Snare or Delusion?," Occasional Paper No. 10, Toronto: Ontario Economic Council, 1979.
- Baumol, William J. and Bradford, David F., "Optimal Departures from Marginal Cost Pricing," *American Economic Review*, June 1970, 60, 263-83.

- Beck, R. G.**, "The Effects of Co-Payment on the Poor," *Journal of Human Resources*, Winter 1974, 9, 129-42.
- Brook, Robert H. et al.**, "Overview of Adult Health Status Measures Fielded in Rand's Health Insurance Study," *Medical Care*, 1979, Suppl., 17, 1-131.
- _____, **et al.**, "Does Free Care Improve Adults' Health? Results from a Randomized Controlled Trial," *New England Journal of Medicine*, December 8, 1983, 309, 1426-34.
- _____, **et al.**, *The Effect of Coinsurance on the Health of Adults*, Publ. No. R-3055-HHS, Santa Monica: Rand Corporation, 1984.
- Clasquin, Lorraine A.**, *Mental Health, Dental Services, and Other Coverage in the Health Insurance Study*, Publ. No. R-1216-OEO, Santa Monica: Rand Corporation, 1973.
- _____, **and Brown, Marie E.**, *Rules of Operation for the Rand Health Insurance Study*, Publ. No. R-1602-HEW, Santa Monica: Rand Corporation, 1977.
- Colle, Ann D. and Grossman, Michael**, "Determinants of Pediatric Care Utilization," *Journal of Human Resources*, 1978, Suppl., 13, 115-58.
- Culyer, Anthony J.**, "The Nature of the Commodity 'Health Care' and Its Efficient Allocation," *Oxford Economic Papers*, July 1971, 23, 189-211.
- _____, *Need and the National Health Service* Totowa: Rowman and Littlefield, 1976.
- _____, "Measuring Health: Lessons for Ontario," Ontario Economic Council Research Studies No. 14, University of Toronto, 1978.
- Davies, Allyson et al.**, "Consumer Acceptance of Prepaid and Fee-for-Service Medical Care: Results from a Randomized Trial," *Health Services Research*, August 1986, 21, 429-52.
- Davis, Karen and Russell, Louise B.**, "The Substitution of Hospital Outpatient Care for Inpatient Care," *Review of Economics and Statistics*, May 1972, 54, 109-20.
- Duan, Naihua**, "Smearing Estimate: A Non-parametric Retransformation Method," *Journal of the American Statistical Association*, September 1983, 78, 605-10.
- _____, **et al.**, "A Comparison of Alternative Models for the Demand for Medical Care," *Journal of Economic and Business Statistics*, April 1983, 1, 115-26 [also R-2754-HHS, Rand Corporation, 1982].
- _____, **et al.**, "Choosing Between the Sample-Selection Model and the Multi-Part Model," *Journal of Business and Economic Statistics*, July 1984, 2, 283-89.
- Ellis, Randall P.**, "Rational Behavior in the Presence of Coverage Ceilings and Deductibles," *Rand Journal of Economics*, Summer 1986, 17, 158-75.
- Evans, Robert G.**, *Strained Mercy: The Economics of Canadian Health Care*, Toronto: Butterworths, 1984.
- Feldstein, Martin S.**, "Hospital Cost Inflation: A Study of Nonprofit Price Dynamics," *American Economic Review*, December 1971, 61, 853-72.
- _____, "The Welfare Loss of Excess Health Insurance," *Journal of Political Economy*, March/April 1973, 81, 251-80.
- _____, "Quality Change and the Demand for Hospital Care," *Econometrica*, October 1977, 45, 1681-702.
- _____, **and Allison, Elizabeth**, "Tax Subsidies of Private Health Insurance," in *The Economics of Federal Subsidy Programs*, Pt. 8, papers submitted to the Subcommittee on Priorities and Economy in Government of the Joint Economic Committee, 93rd Congress, 2nd session, July 1974.
- _____, **and Friedman, Bernard**, "Tax Subsidies, the Rational Demand for Insurance, and the Health Care Crisis," *Journal of Public Economics*, April 1977, 7, 155-78.
- Ferber, Robert and Hirsch, Werner**, "Social Experimentation and Economic Policy: A Survey," *Journal of Economic Literature*, December 1978, 16, 1379-404.
- Follette, W. and Cummings, Nicholas A.**, "Psychiatric Services and Medical Utilization in a Prepaid Health Plan Setting," *Medical Care*, January 1967, 5, 26-35.
- _____, **and _____**, "Psychiatric Services and Medical Utilization in a Prepaid Health Plan Setting, Part 2," *Medical Care*, January 1968, 6, 31-41.
- Friedman, Bernard and Pauly, Mark**, "Cost Functions for a Service Firm with Variable Quality and Stochastic Demand," *Review of Economics and Statistics*, November 1981, 63, 610-24.
- _____, **and _____**, "A New Approach to Hospital Cost Functions and Some Issues

- in Revenue Regulation," *Health Care Financing Review*, March 1983, 4, 105-14.
- Goddeeris, John and Weisbrod, Burton, "What We Don't Know about Why Health Expenditures Have Soared: Interaction of Insurance and Technology," *Mount Sinai Journal of Medicine*, November 1985, 52, 685-91.
- Goldman, Fred and Grossman, Michael, "The Demand for Pediatric Care: An Hedonic Approach," *Journal of Political Economy*, April 1978, 86, 259-80.
- Goldsmith, Jeff, "Death of a Paradigm: The Challenge of Competition," *Health Affairs*, Fall 1984, 3, 5-19.
- Hausman, Jerry and Wise, David, *Social Experimentation*, Chicago: University of Chicago Press, 1985.
- Haveman, Robert, "Social Experimentation and Social Experimentation," *Journal of Human Resources*, Fall 1986, 21, 586-605.
- Helms, L. Jay, Newhouse, Joseph, P. and Phelps, Charles E., "Copayments and the Demand for Medical Care: The California Medicaid Experience," *Bell Journal of Economics*, Spring 1978, 9, 192-208.
- Heller, Walter, W., "What's Right with Economics," *American Economic Review*, March 1975, 65, 1-26.
- Hill, D. B. and Veney, James E., "Kansas Blue Cross/Blue Shield Outpatient Benefits Experiment," *Medical Care*, March-April 1980, 8, 143-58.
- Huber, P. J., "The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions," *Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, 1967, 221-33.
- Keeler, Emmett B. and Rolph, John E., *The Demand for Episodes of Medical Services: Interim Results from the Health Insurance Experiment*, R-2829-HHS, Santa Monica: Rand Corporation, December 1982.
- _____, Morrow, Daniel and Newhouse, Joseph P., "The Demand for Supplementary Health Insurance, or Do Deductibles Matter?," *Journal of Political Economy*, August 1977, 85, 789-802.
- _____, Newhouse, Joseph, P. and Phelps, Charles E., "Deductibles and the Demand for Medical Care Services: The Theory of a Consumer Facing a Variable Price Schedule under Uncertainty," *Econometrica*, April 1977, 45, 641-56.
- _____, et al., "How Free Care Reduced Hypertension of Participants in the Rand Health Insurance Experiment," *Journal of the American Medical Association*, October 11, 1985, 154, 1926-31.
- _____, et al., *The Demand for Episodes of Medical Treatment*, Publ. No. R-3454-HHS, Santa Monica: Rand Corporation, 1987.
- Leibowitz, Arleen et al., "Effect of Cost Sharing on the Use of Medical Services by Children: Interim Results from a Randomized Controlled Trial," *Pediatrics*, May 1985, 75, 942-51.
- Levit, Katherine R. et al., "National Health Expenditures, 1984," *Health Care Financing Review*, Fall 1985, 7, 1-35.
- Lewis, Charles E. and Keairnes, Harold, "Controlling Costs of Medical Care by Expanding Insurance Coverage," *New England Journal of Medicine*, June 18, 1970, 292, 1405-12.
- Lindsay, Cotton M., "Medical Care and the Economics of Sharing," *Economica*, November 1969, 36, 351-62.
- Lohr, Kathleen N. et al., "Use of Medical Care in the Rand Care in the Rand Health Insurance Experiment: Diagnosis- and Service-Specific Analyses," *Medical Care*, September 1986, Suppl.
- Long, Stephen H., Settle, Russell F. and Stuart, Bruce C., "Reimbursement and Access to Physicians' Services under Medicaid," *Journal of Health Economics*, September 1986, 5, 235-52.
- Luft, Harold S., *Health Maintenance Organizations*, New York: Wiley & Sons, 1981.
- McCombs, Jeffrey S., "Physician Treatment Decisions in a Multiple Treatment Model: The Effects of Physician Supply," *Journal of Health Economics*, August 1984, 3, 155-71.
- Maddala, G. S., *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press, 1983.
- Manning, Willard G., "Estimating Demand Functions with Data from Declining Block Tariffs," draft, June 1986.
- _____, Duan, Naihua and Rogers, William H., *Monte Carlo Evidence on the Choice between Sample Selection and Two-Part Models*, Santa Monica: Rand Corpora-

- tion, forthcoming.
- _____, et al., (1984a) "A Controlled Trial of the Effect of a Prepaid Group Practice on Use of Services," *New England Journal of Medicine*, June 7, 1984, 310, 1505–10 [also R-3029-HHS, Rand Corporation, December 1985].
- _____, et al., (1984b) "Cost Sharing and the Demand for Ambulatory Mental Health Services," *American Psychologist*, October 1984, 39, 1090–100.
- _____, et al., "The Demand for Dental Care: Evidence from a Randomized Trial in Health Insurance," *Journal of the American Dental Association*, June 1985, 110, 895–902.
- _____, et al., (1986a) *Health Insurance and the Demand for Dental Care: Evidence from a Randomized Experiment*, Publ. No. R-3225-HHS, Santa Monica: Rand Corporation, 1986.
- _____, et al., (1986b) "How Cost Sharing Affects the Use of Ambulatory Mental Health Services," *Journal of the American Medical Association*, 1986, 256, 1930–34.
- _____, et al., *Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment*, Publ. No. R-3476-HHS, Santa Monica: Rand Corporation, 1987.
- Mehran, F., "Variance of the MVUE for the Log Normal Mean," *Journal of the American Statistical Association*, 1973, 68, 726–27.
- Metcalfe, Charles E., "Making Inferences from Controlled Income Maintenance Experiments," *American Economic Review*, June 1973, 63, 478–483.
- Morris, Carl N., "A Finite Selection Model for Experimental Design of the Health Insurance Study," *Journal of Econometrics*, September 1979, 11, 43–61.
- Musgrave, Richard, *The Theory of Public Finance*, New York: McGraw-Hill, 1959.
- Newhouse, Joseph P., "A Design for a Health Insurance Experiment," *Inquiry*, March 1974, 11, 5–27.
- _____, "Medical Care Expenditure: A Cross-National Survey," *Journal of Human Resources*, Winter 1977, 12, 115–25.
- _____, "Insurance Benefits, Out-of-Pocket Payments, and the Demand for Medical Care: A Review of the Literature," *Health and Medical Care Services Review*, July-August 1978, 1, 3–15.
- _____, "The Demand for Medical Care Services: A Retrospect and Prospect," in J. van der Gaag and M. Perlman, eds., *Health, Economics, and Health Economics*, Amsterdam: North-Holland, 1981.
- _____, "Are Medical Care Costs Too High?," *Journal of Dental Education*, November 1984, 48, 587–90.
- _____, and Phelps, Charles E., "Price and Income Elasticities for Medical Care Services," in M. Perlman, ed., *The Economics of Health and Medical Care*, New York: Wiley & Sons, 1974.
- _____, and _____, "New Estimates of Price and Income Elasticities of Medical Care Services," in R. N. Rosett, ed., *The Role of Health Insurance in the Health Services Sector*, Universities-National Bureau Conference Series No. 27, New York, 1976.
- _____, _____, and Marquis, M. Susan, "On Having Your Cake and Eating It Too: Econometric Problems in Estimating the Demand for Health Services," *Journal of Econometrics*, August 1980, 13, 365–90.
- _____, et al., "Design Improvements in the Second Generation of Social Experiments: The Health Insurance Study," *Journal of Econometrics*, September 1979, 11, 117–29.
- _____, et al., "Some Interim Results from a Controlled Trial of Cost Sharing in Health Insurance," *New England Journal of Medicine*, December 17, 1981, 305, 1501–07 [see R-2847-HHS, Rand Corporation, 1982].
- _____, et al., "The Findings of the Rand Health Insurance Experiment: A Response to Welch et al.," *Medical Care*, February 1987, 25, 157–79.
- O'Grady, Kevin F. et al., "The Impact of Cost Sharing on Emergency Department Use," *New England Journal of Medicine*, August 22, 1985, 313, 484–90.
- Pauly, Mark V., *Medical Care at Public Expense*, New York: Praeger, 1971.
- _____, "Taxation, Health Insurance, and Market Failure," *Journal of Economic Literature*, June 1986, 24, 629–75.
- Phelps, Charles E. and Newhouse, Joseph P., "Effects of Coinsurance: A Multivariate Analysis," *Social Security Bulletin*, June 1972, 35, 20–29.

- _____, and _____, "Coinsurance, the Price of Time, and the Demand for Medical Services," *Review of Economics and Statistics*, August 1974, 56, 334-42.
- Plott, Charles R.**, "Industrial Organization Theory and Experimental Economics," *Journal of Economic Literature*, December 1982, 20, 1435-527.
- Ramsey, Frank**, "A Contribution to the Theory of Taxation," *Economic Journal*, March 1927, 37, 47-61.
- Roemer, Milton J. et al.**, "Copayments for Ambulatory Care: Penny-wise and Pound-foolish," *Medical Care*, 1975, 13, 475-66.
- Rogers, William H., and Newhouse, Joseph P.**, "Measuring Unfiled Claims in the Health Insurance Experiment," in L. Burstein et al., eds., *Collecting Evaluation Data: Problems and Solutions*, Beverly Hills: Sage, 1985, 121-33.
- Rosett, Richard N. and Huang, Lien Fu**, "The Effect of Health Insurance on the Demand for Medical Care," *Journal of Political Economy*, March/April 1973, 81, 281-305.
- Rothschild, Michael and Stiglitz, Joseph**, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, November 1976, 90, 629-50.
- Scitovsky, Anne A. and McCall, Nelda M.**, "Coinsurance and the Demand for Physician Services: Four Years Later," *Social Security Bulletin*, May 1977, 40, 19-27.
- _____, and **Snyder, Nelda M.**, "Effect of Coinsurance on Use of Physician Services," *Social Security Bulletin*, June 1972, 35, 3-19.
- Searle, S. R.**, *Linear Models*, New York: Wiley & Sons, 1971.
- Sloss, Elizabeth M. et al.**, "Effect of a Health Maintenance Organization on Physiologic Health: Results from a Randomized Trial," *Annals of Internal Medicine*, January 1987, 106, 130-38.
- Tobin, James**, "Estimation of Relationships for Limited Dependent Variables," *Econometrica*, January 1958, 26, 24-36.
- Valdez, R. Burciaga**, *The Effects of Cost Sharing on the Health of Children*, Publ. No. R-3270, Santa Monica: Rand Corporation, 1986.
- _____, et al., "The Consequences of Cost Sharing for Children's Health," *Pediatrics*, May 1985, 75, 957-61.
- Van de Ven, Wynand P. and van Praag, Bernard M.**, (1981a) "The Demand for Deductibles in Private Health Insurance: A Probit Model with Sample Selection," *Journal of Economics*, November 1981, 17, 229-52.
- _____, and _____, (1981b) "Risk Aversion of Deductibles in Private Health Insurance," in J. van der Gaag and M. Perlman, eds., *Health, Economics, and Health Economics*, Amsterdam: North-Holland, 1981.
- Waldo, Daniel and Lazenby, Helen C.**, "Demographic Characteristics and Health Care Use and Expenditures by the Aged in the United States, 1977-1984," *Health Care Financing Review*, Fall 1984, 6, 1-29.
- _____, **Levit, Katherine R. and Lazenby, Helen C.**, "National Health Expenditures, 1985," *Health Care Financing Review*, Fall 1986, 8, 1-21.
- Ware, John E. et al.**, *Conceptualization and Measurement of Health*, Vol. I, R-1987/1-HEW, Santa Monica: Rand Corporation, May 1980.
- _____, et al., "Comparison of Health Outcomes at a Health Maintenance Organization with Those of Fee-for-Service Care," *The Lancet*, May 3, 1986, 848, 1017-22.
- Wells, Kenneth B. et al.**, *Cost Sharing and the Demand for Ambulatory Mental Health Services*, Publ. No. R-2960-HHS, Santa Monica: Rand Corporation, 1982.
- Zeckhauser, Richard J.**, "Medical Insurance: A Case Study of the Trade-Off between Risk Spreading and Appropriate Incentives," *Journal of Economic Theory*, March 1970, 2, 10-26.
- American Hospital Association**, "1984 Hospital Cost and Utilization Trends," *Economic Trends*, Spring 1985, 1.
- _____, *Hospital Statistics*, 1984 Edition, Chicago: AH Association, 1984.
- U.S. Department of Health and Human Services**, "Current Estimates from the *National Interview Health Survey*," Series 10: 150, 1985; 156, 1986, Washington: USDHHS.
- _____, "Utilization of Short Stay Hospitals," Series 13: 78, 1984; 84, 1986; 90, 1987; Washington: USDHHS.