A RAND NOTE

SAMPLE SELECTION IN THE HEALTH INSURANCE EXPERIMENT: COMPARING THE ENROLLED AND NONENROLLED POPULATIONS

Carl N. Morris

October 1985

N-2354-HHS

Prepared for

The U.S. Department of Health and Human Services

Rand
1700 MAIN STREET
P.O. BOX 2138
SANTA MONICA, CA 90406-2138

The research described in this report was supported by the U.S. Department of Health and Human Services under Grant No. 016B-80.

Published by The Rand Corporation

A RAND NOTE

SAMPLE SELECTION IN THE HEALTH INSURANCE EXPERIMENT: COMPARING THE ENROLLED AND NONENROLLED POPULATIONS

Carl N. Morris

October 1985

N-2354-HHS

Prepared for

The U.S. Department of Health and Human Services

Rand
1700 MAIN STREET
P.O. BOX 2138
SANTA MONICA, CA 90406-2138

PREFACE

The Rand Health Insurance Experiment (HIE) included a controlled trial of cost sharing in health insurance to determine the relative effects of various benefits on health service utilization, health status, and other factors. Such analyses require that the different cost-sharing benefits be applied to similar populations, so that any nonrandom differences in responses can be attributed solely to the differences in benefits. This Note establishes the similarity with respect to preexperimental baseline survey data of the HIE populations, at the time of initial enrollment in alternative cost-sharing medical insurance plans, to the nonenrolled population. These results encourage confidence in extrapolating estimates made from experimental samples to larger populations.

The Health Insurance Experiment is supported under a grant from the U.S. Department of Health and Human Services.

SUMMARY

Despite the inability to enroll some families selected for HIE participation, because of moving out of the area, refusals, etc., statistical tests described here reveal that Health Insurance Experiment samples at the time of enrollment are generally better matched with respect to key baseline survey dependent variables (number of physician visits, health status, worry about health, and pain experienced) than they would have been had the treatment assignments been made by simple random sampling with all selected families enrolling as planned. These results hold for: (A) comparisons between the entire enrolled sample and those not enrolled, but eligible for enrollment; and (B) comparisons between treatment groups for enrolled individuals only. Similar results hold for key preexperimental independent variables (age, income, insurance) in comparisons (A) and (B), except that for (A) only there is a higher percentage of children among enrolled individuals than among nonenrolled, eligible individuals. Thus, provided appropriate adjustments are made for age, and subject to no further distortions produced by attrition after initial enrollment, this analysis of baseline data suggests that the HIE samples will support inferences to the populations from which they were drawn at least as well as traditional random sampling methods would with a 100 percent enrollment rate.

ACKNOWLEDGMENTS

I am grateful to William Lisowski, who did all the programming for the analysis with great skill and care. I also thank Helene Mills, Sherry Trees, and William Fowler, who helped pull together the data requests and provided different components of the data. I especially appreciate Martha Cooper's excellent typing skills and the general assistance she provided for completing this task, and so many other tasks over the years.

CONTENTS

TABLES

TABLES (cont.)

I. INTRODUCTION AND SUMMARY

The experimental portion of the Health Insurance Experiment (Newhouse, 1974) focuses on a comparative analysis of numerous health insurance plans (Family Health Protection Plans, FHPP), having varying benefits for typical populations under the age of 65. A summary description of the experiment, its purposes, and its design, is available in the appendix. Such comparisons are made best if the sample used for the study meets two criteria:

A. The experimental sample is similar to the population about which generalizations are to be made; and
B. The samples on the various FHPP treatments are similar to one another.

This Note investigates the two preceding hypotheses in all six HIE sites (Dayton, Ohio; Seattle, Washington; Fitchburg and Franklin County, Massachusetts; and Charleston and Georgetown County, South Carolina) by comparing individual and family baseline data for the enrolled and nonenrolled samples, and also by comparing characteristics among those families actually enrolled on the various insurance plans at the time immediately after enrollment. The results are quite favorable. This section summarizes the main conclusions, which are discussed at greater length later:

A. Comparisons of differences between enrolled and eligible nonenrolled families show smaller than expected differences for the preexperimental measures of the response variables of the experiment (physician visits, health status, worry, pain). Among preexperimentally measured explanatory variables considered, significant differences were not found for income or insurance coverage. However, experimental families were found to have a higher proportion of children, necessitating adjustment for this variable when extrapolating experimental

results to the full population in the six experimental sites (particularly in Dayton, Seattle, and Massachusetts).

B. Among enrolled families, no statistically significant differences were found among the major groupings of FHPP insurance plans with respect to the same preexperimental response variables or preexperimental explanatory variables.

A. Tests for Differences Between Enrolled and Nonenrolled Populations

The differences between the enrolled population and the eligible nonenrolled populations are somewhat less than differences that would be expected at random in the six sites, with respect to all the key preexperimental measures of the dependent variables for the experiment (number of physician visits, health status, worry, and pain). These tests are determined by chi-square and t-test criteria applied to the variables in the six sites, using adjustments to these chi-square criteria to account for dependence among family members. Large chi-square values (values substantially exceeding their degrees of freedom) and large t-statistics (absolute value exceeding 2.0) would indicate significant departures between the enrolled and nonenrolled populations. Conversely, smaller values indicate satisfactory matching.

For example, the 30 chi-square tests of Table 1.1 (6 sites and 5 variables per site, physician visits being categorized in two different ways) reveal *no* statistically significant differences and the chi-square statistic exceeds its degrees of freedom, that is, its expected value, in only eight cases. The total chi-square for each variable of Table 1.1 is substantially less than would be expected if the sample had been divided at random between the enrolled and nonenrolled populations. (The expected value of a chi-square statistic is its degree-of-freedom, df, and the median is almost exactly df minus 0.66.) Those minor differences that do occur in Table 1.1, as when the chi-square statistic exceeds its expectation, are scattered across the sites. Thus, the HIE sites are at least as well balanced as one might expect from a simple random sampling assignment.

A variety of factors produced the final enrolled samples from those successfully baselined. During the entire experiment, 42 percent of all families that completed baselines, that were selected for enrollment,

Table 1.1

CHI-SQUARE VALUES MEASURING DIFFERENCES BETWEEN ENROLLED (GE) AND NONENROLLED (GN) ELIGIBLE POPULATIONS

Site	PV2	(df)	PV7	(df)	Health Status	(df)	Worry	(df)	Pain	(df)
Dayton	2.4	(3)	17.3	(18)	5.8	(9)	7.3	(9)	6.2	(9)
Seattle	0.11	(1)	2.41	(6)	5.48a	(3)	0.94	(3)	3.50a	(3)
Fitchburg	0.92	(1)	3.22	(6)	0.36	(3)	4.58a	(3)	5.66a	(3)
Franklin	2.20a	(1)	7.84a	(6)	1.22	(3)	1.20	(3)	1.81	(3)
Charleston	0.69	(1)	1.67	(6)	4.02a	(3)	0.66	(3)	0.36	(3)
Georgetown	0.01	(1)	4.77	(6)	2.64	(3)	3.67a	(3)	2.29	(3)
Totals[b]	6.3	(8)	37.2	(48)	19.5	(24)	18.4	(24)	19.8	(24)

NOTE: Smaller values of chi-square indicate greater matching of the two populations. PV 2 and PV7 indicate physician visit responses categorized into 2 or 7 categories.

[a]Exceeds expected value, but does not achieve statistical significance.

[b]Totals have a chi-square distribution with the indicated degrees of freedom.

and that were assigned to a treatment group, did not enroll for some reason (moved, not found, ineligibility, insurance nonverification, interview refusals, and enrollment refusals). This rate varied from 26 percent in Dayton to 49 percent in the South Carolina sites. The refusal of the offer to enroll was 17 percent overall (Morris, Newhouse, Archibald, 1980). Other nonenrolled individuals were deliberately selected for that status if more families were available than required. Section 2 discusses these factors and other definitions at greater length, but we summarize here. No differences between groups were introduced intentionally (other than the required oversampling of low-income families in Dayton at the enrollment stage, which has been accounted for in the statistics of this section).[1]

[1]Oversampling of low-income individuals was done at the screening interview stage in Massachusetts and South Carolina and was not attempted in Seattle.

The single feature of the HIE design used to obtain more closely matched samples than simple random assignment does is the Finite Selection Model (FSM). The FSM was developed for the HIE to provide controlled random sample selection with respect to selected baseline variables including those of this Note. Nonenrollment of families selected for HIE participation ordinarily would be expected to reduce matching of samples. The HIE design therefore attempted to minimize these distortions, when possible, which were caused by:

1. Out-of-area moves after the baseline interview;
2. Refusal to answer the preenrollment questionnaire;
3. Refusal of the FHPP offer (i.e., the offer to enroll);
4. Being found ineligible for FHPP participation after selection;
5. Inability to locate and enroll some families before the scheduled final enrollment deadline.

Thus, the results reported here indicate that the FSM was a stronger factor in favor of matching the experimental subjects to the nonexperimentals than were the negative factors. The tests used here are "omnibus" tests, in that the data used cannot identify which of the negative factors would have caused distortions, had any done so. Because of the favorable outcome, no such search has been attempted.

The preexperimental "dependent variables" just considered are the most important to control in the experimental design. Of less importance, because they can be controlled directly in subsequent analyses, are the three preexperimental "explanatory variables": *age; family income*; and possession of *preexperimental health insurance*. Our investigation of these variables reveals significant age differences, the experimental population being younger than the nonexperimental population, particularly so in Seattle and the Massachusetts sites (see Table 1.2). The preexperimental income and insurance variables are reasonably well-matched, however, nearly as well as would be expected at random.

Table 1.2

CHI-SQUARE VALUES MEASURING DIFFERENCES BETWEEN ENROLLED (GE) AND NONENROLLED (GN) ELIGIBLE POPULATIONS FOR GROUPED DATA

Site	Age	(df)	Income	(df)	Insured	(df)
Dayton	24.30b	(9)	a		1.00	(2)
Seattle	17.02b	(3)	0.36	(2)	0.80	(1)
Fitchburg	21.54b	(3)	3.65	(2)	3.40	(1)
Franklin	25.13b	(3)	1.41	(2)	0.01	(1)
Charleston	1.42	(3)	4.81	(2)	0.03	(1)
Georgetown	3.94	(3)	2.08	(2)	2.30	(1)
Total	93.35a	(24)	12.30	(10)	7.62	(7)

NOTE: Age (4 categories: 0-5, 6-17, 18-44, 45-61), income (3 categories), and insurance (2 categories--covered or not covered). Full descriptions in Sec. II.

[a]Incomes in Dayton were deliberately unbalanced between GE and GN to produce a disproportionate number of low-income families in the experiment.

[b]Statistically significant (stronger than 0.05 level).

The average ages for the enrolled group (GE) and nonenrolled, but eligible group (GN), are shown in Table 1.3. FHPP individuals average 1.5 years younger than those eligible, but not enrolled. This difference is almost entirely due to the higher proportion of children in FHPP (41 percent children, 59 percent adults) than not in FHPP (35 percent children, 65 percent adults), age distributions within these categories being well-matched. Families chosen for FHPP participation who had children were more likely to be enrolled than those without children, perhaps because families with children are easier to locate. While this fact makes enrolled individuals slightly healthier than those not enrolled, it can be accounted for in extrapolations to any population by including an adult-child indicator variable in any analysis requiring extrapolation. However, age was well-matched *between* FHPP plans at enrollment, and therefore does not require adjustment for plan comparisons.

Table 1.3

AGE DATA FOR ENROLLED (GE) AND NON-ENROLLED (GN) GROUPS, BY SITE

Site	Average Age		t	P (Adults)[b]	
	GE	GN		GE	GN
Dayton	25.0	25.7	-1.2	0.59	0.66
Seattle	25.2	27.0	-3.2a	0.64	0.67
Fitchburg	24.6	27.9	-3.3a	0.57	0.66
Franklin	24.7	28.4	-4.3a	0.59	0.67
Charleston	25.1	24.7	0.5	0.59	0.61
Georgetown	24.9	24.7	0.2	0.57	0.60
All sites	24.9	26.4	-4.6a	0.59	0.65

[a]Significant at 0.01 level.

[b]Adults are those over age 18 at enrollment.

No test for income or insurance differences between GE and GN individuals, as shown in Table 1.2, is statistically significant (0.05 level) in any site, and the six site totals of these chi-square values are close to their expected values. The income tests shown are developed by forming three groups (baseline year family incomes of $0-$9,000, $9,001-$14,999, or $15,000 and up, categories that were used for income stratification in the experimental design). The two largest chi-square values for income in Fitchburg and Charleston both occur because of a lower proportion of higher-income families in GE. In the two most unbalanced sites, the percents of high-income families were: 40 percent in GN versus 32 percent in GE for Fitchburg, and 26 percent in GN versus 19 percent in GE for Charleston. Overall, incomes of GE families are 9 percent lower than GN in Fitchburg (significant at the 5 percent level), but this figure drops to 3 percent for Charleston GE families, which does not approach statistical significance.

B. Tests for Differences Among Families Assigned to Different Insurance Benefits

The insurance groupings considered here are listed in Table 1.4, with more complete descriptions in the appendix. The numbers 25, 50, 95, 100 are coinsurance rates, the percentage of each medical bill that the family pays before meeting its limit (thereafter expenditures are fully covered for the remainder of the year). The 25/50 notation indicates that the mental and dental coinsurance rate is 50 percent but that other medical expenditures are 25 percent.

Each site includes exactly six of the groups listed in Table 1.4, and therefore a chi-square test of independence of characteristics based on a c by 6 table, c = number of categories for a variable, will have 5(c - 1) degrees of freedom.

Table 1.5 contains the chi-square values for the preexperimental dependent variables. Small values indicate good matching between FHPP treatments. The table indicates there are no significant differences for any variable, or any site, and the totals, which have chi-square distributions under the hypothesis of no systematic differences, indicate much better matching than would be expected for four of the

Table 1.4

FHPP PLAN GROUPINGS USED IN THIS REPORT

o Free care	All sites
o 25/25[a]	All sites
o 25/50[a]	Not used in Dayton
o 50[a]	Not used in Seattle
o 95 or 100[a]	95 all sites, except 100 Dayton year 1
o ID 95 or 100	"Individual Deductible," except 100 Dayton year 1
o Control group	Dayton only
o Group Health	Seattle only

[a]These groups were not further separated by their maximum dollar expenditure (5, 10, and 15 percent of income, but not more than $1,000) because the resulting groups would be too small for this analysis.

Table 1.5

CHI-SQUARE MEASURING DIFFERENCES BETWEEN SAMPLES ENROLLED IN INSURANCE GROUPINGS OF TABLE 1.4

Site	PV2	(df)[a]	PV7	(df)[a]	Health Status	(df)[a]	Worry	(df)[a]	Pain	(df)[a]
Dayton	7.1	(5)	23.7	(30)	6.9	(10)	4.9	(10)	8.0	(10)
Seattle	4.2	(5)	28.2	(30)	4.3	(10)	3.5	(10)	2.7	(10)
Fitchburg	6.8	(5)	22.8	(30)	6.0	(10)	3.5	(10)	6.5	(10)
Franklin	2.9	(5)	19.1	(30)	2.6	(10)	8.8	(10)	5.4	(10)
Charleston	6.5	(5)	17.8	(30)	8.5	(10)	4.0	(10)	11.8	(10)
Georgetown	4.8	(5)	14.9	(30)	9.3	(10)	9.0	(10)	3.2	(10)
Total	32.2	(30)	126.5	(180)	37.6	(60)	33.7	(60)	37.6	(60)

[a]df = 5(c - 1), c = number of categories for the variable.

five variables, PV2 being the exception. Physician visits in two categories (PV2) have differences with magnitude about that expected of random sampling, $\chi^2_{30} = 32.2$. Fortunately, the probability of at least one physician visit does not have any pattern in the six sites.

The preexperimental versions of the independent variables also are well-balanced in each site with respect to the insurance groups of Table 1.4, as is shown in Table 1.6. There are no significant differences for any site or variable of Table 1.6, and the chi-square values totaled across sites indicate less variability among these insurance groups than would be expected from random sampling. In particular, the age variable, which was not balanced between the enrolled and nonenrolled samples, is nicely balanced among the plans for the enrolled populations.

Because there are no significant differences between insurance plans, direct comparisons between respondents to FHPP insurance plans can be made as safely as if all nonenrollment had occurred randomly. Thus, even though the sample assignment included nonrandom failure to enroll many selected families, differences between the insurance groups were relatively small at the time of enrollment.

Table 1.6

CHI-SQUARE VALUES INDICATING DIFFERENCES AMONG MAJOR INSURANCE GROUPS OF ENROLLED HIE SAMPLES

Site	Age	(df)	Income	(df)	Insured	(df)
Dayton	15.9	(15)	(a)		4.3	(5)
Seattle	14.1	(15)	8.6	(10)	4.4	(5)
Fitchburg	5.9	(15)	10.1	(10)	3.6	(5)
Franklin	8.1	(15)	6.2	(10)	5.1	(5)
Charleston	7.1	(15)	4.6	(10)	2.4	(5)
Georgetown	12.1	(15)	12.4	(10)	4.7	(5)
Total	63.2	(90)	41.9	(60)	24.5	(30)

NOTE: Age (4 categories); income (3 categories); insurance (2 categories, covered and not covered).

[a]No analysis made.

The following sections spell out more fully the assumptions and definitions (Sec. II), and provide greater detail for the six geographical areas (Sec. III). Some statistical theory is included in Sec. IV. The Appendix contains a brief description of the HIE experiment.

II. DESCRIPTION AND PHILOSOPHY OF THE EVALUATION OF HIE SAMPLE ASSIGNMENTS

The analyses of Sec. I, and their justification, are described more fully here. This section also treats results not in Sec. I. The precise statistical methods used, which require certain technical adjustments to account for observed positive correlation of responses among family members, are postponed until Sec. IV following presentation of the numerical results in Sec. III.

VARIABLES AND GROUP ASSIGNMENTS USED FOR THIS COMPARISON

With exception of random differences, HIE samples in each site are intended to match the population of individuals eligible for the HIE, the only exception being the deliberate oversampling of low-income families in all sites except Seattle. However, no experiment can maintain full control over human subjects, and so there will be some nonenrollment. This inevitably provides some potential for bias--i.e., the enrolled sample may not match the population in some important way.

If the sample differs from the population in a known way, compensating adjustments can be made in the analysis to remove the distortion. The intentional oversampling of low-income families in the HIE provides such an example, with reweighting as an acceptable adjustment. On the other hand, when sample and population differences occur in an unknown way, conclusions drawn from experimental data may be biased with respect to the full population. Such biases may be greater for direct extrapolation to the population than for comparisons between responses on different treatments. For example, when estimating the effect of coinsurance, even from distorted samples, significant biases would not be expected unless there exists a significant *interaction* between an out-of-balance variable (e.g., age in Seattle) and the explanatory variable of interest (e.g., coinsurance) with respect to the response. Thus, an age adjustment to population age levels would be necessary only if one must estimate how coinsurance affects the health expenditures of all Seattle families and if the effect of coinsurance for Seattle children differs from that of adults.

Serious biases, however, could occur if the health levels of enrolled families differ significantly among the enrolled and nonenrolled families (though analysis has not revealed any important interactions between health levels and plan). As Sec. I showed, this did not occur, based on evidence from preexperimental values of the dependent variables, i.e., measures of health services utilization and of health status.

Preexperimental values of the experimental responses are not available in most experiments. They were in the HIE because almost all families completed baseline survey interviews before being considered for enrollment on an FHPP insurance plan. This includes over 20,000 individuals in the six HIE sites who met age and income constraints for eligibility. These HIE baseline data include preexperimental values of the experimental responses: the number of physician visits in the preceding year, and several measures of health status (self-perceived health status, worry about health, and pain experienced). These data form the basis of this Note.

Three groups are created for the first analyses of this Note, which aim to determine how closely HIE samples matched the nonsampled populations. In all cases below, a family is considered to be "enrolled" if they were enrolled initially, whether or not they later dropped out of the study. The main two groups, denoted GE and GN are defined as follows.

GE: The Group of Enrolled Individuals

These individuals met: age (under the age of 62 at the time of enrollment); income (family income not exceeding $25,000 in 1973 dollars for the year before the baseline); and other minor constraints (described in GI below).

GN: Group of Nonenrolled Individuals

These were eligible to be in the HIE but were not selected for enrollment (because HIE budget constraints limited the number of enrolled families to less than the number available), or were selected but were not contacted (moved, couldn't be located, etc.), or were

selected and were or would have been asked to participate, but refused at some stage in the enrollment to comply with the interview procedures or to be enrolled.

A third group, GI, includes ineligible individuals who met age and income constraints, but who were ineligible for other reasons. It also is considered in the comparisons of this Note. However, GI was not expected to match the GE or GN groups, and therefore was not discussed in Sec. I.

GI: The Group of Ineligible Individuals

These individuals met age and income ineligibility requirements on the baseline survey, but were ineligible for one of the following reasons. They were in the military or institutionalized, their insurance could not be verified, their insurance provider would not cooperate with the HIE, they were eligible to use student health services, or they were dependent on others in such categories.

The relative sizes of these groups vary in the six cities. Considering all sites, about 34 percent are in GE, 50 percent in GN, and 16 percent in GI.

The main objective of this Note is to compare GE with respect to baseline preexperimental utilization and health status. All responses are self-reported or reported by one of the family heads. Descriptions of these variables follow.

1. *Preexperimental utilization* in this Note is measured by the number of annual physician visits (PV), other than in hospitals, as determined in response to the question:

> "In the last 12 months--since (DATE OF INTERVIEW), a year ago, how many times altogether did (PERSON) see (PROVIDER) about (his/her) own health? This should *not* include visits while (PERSON) was a patient in a hospital."

This question was repeated for each individual for as many physicians (PROVIDERS) as he/she visited, then summed over providers to yield the measurement used here. Thus, PV = 0, 1, 2, ... is a nonnegative integer. For the analysis, it was useful to define several derivatives of PV:

a. Physician visits in 2 categories (PV2):

$$PV2 = 0 \text{ if } PV = 0 \text{ and } PV2 = 1 \text{ if } PV \geq 1.$$

b. Physician visits in 7 categories (PV7). The categories were chosen as

0, 1, 2, 3-4, 5-7, 8-17, 18 and higher,

because they have roughly equal probabilities.

c. Log(2 + PV), a scale chosen (and used in the Finite Selection Model) because it has approximately a normal distribution.

2. *Self-perceived health status* (HS) is indicated as excellent, good, fair, or poor in response to the baseline question:

"Would you say (PERSON's) health, in general, is excellent, good, fair, or poor?"

For comparisons within GE among insurance plan groups, which causes very small groups in "poor" health, these four HS categories are collapsed to three by grouping the fair and poor categories.

3. *Worry about health* is indicated as "great deal," "some," "a little," or "none at all," in response to the question:

"Over the *past year* has (PERSON)'s health caused you a great deal of worry, some worry, a little worry, or no worry at all?"

Worry about health in 3 categories, used for comparisons within GE among insurance groups, combines the "great deal of worry" and "some worry" categories, because few individuals Note a great deal of worry.

4. *The four categories for pain* (PAIN) are determined by the response to the question:

"In the *past year* would you say (PERSON) has experienced a great deal of pain, some pain, a little pain, or no pain at all?"

Pain in 3 categories, which is used for comparison within GE among insurance groups, combines the "great deal of pain" and "some pain" categories.

Although these are imperfect health measures, systematic utilization and health differences should emerge in these large samples if important health differences are present among various HIE groups.

Several key explanatory variables, also are analyzed.

5. *Insured*, in 2 categories is coded as 0 if the individual has no coverage, 1 if there is coverage.

6. *Age* is analyzed in both its natural scale--an integer, in years, and as a categorical variable, the categories being 0-5, 6-17, 18-44, 45-61 (individuals over 61 years of age at the outset of the experiment were ineligible). "Child" here refers to ages 0-17, "Adult" to ages 18-61.

7. *Income* refers to total self-reported family income in dollars for the year preceding the baseline interview. When analyzed as a continuous variable, the logarithm of income is used, being more symmetrically distributed than income. When analyzed in three categories, the categories are $0-$9,000, $9,001-$14,999, and $15,000-$25,000. Families above $25,000 in 1973 dollars were ineligible, although because of inflation this limit was higher in sites enrolled after 1973. These categories were used because they defined the "low", "medium", and "high" categories used to stratify the Dayton sample when eligible families were assigned to insurance groups.

HOW THE HIE DESIGN LED TO GROUPS GE, GN, GI

No differences were introduced intentionally between GE and GN, except in Dayton where the low-income families ($0 to $9,000) were oversampled, medium-income families ($9,001 to $14,999) were undersampled, and high-income families ($15,000 and above, subject to the upper limit of $25,000 in Dayton) were proportionately sampled. Because of this sampling plan in Dayton, analysis for that site

necessarily proceeds differently than in other sites. In Dayton, we make comparisons within income class to account for the deliberate distortion between the GE and GN samples. This is unnecessary in the other five sites because no oversampling was used in Seattle, and oversampling in the Massachusetts and South Carolina sites was accomplished before the baseline survey was administered, so that GE and GN comparisons using baseline data are unaffected. The income class breakdown in Dayton provides a weaker test than for the other sites. However, when the Dayton results were combined for the different income groups in a direction-preserving way that made the Dayton degrees-of-freedom equal to the other sites, the results were much the same, and still insignificant.

Differences between GE and GN were deliberately minimized by using a design algorithm to assign families to insurance plans based on a variety of independent variables, including the health and utilization variables of this study. This model, the "Finite Selection Model" (FSM), is described in Morris (1979), and Seda and Keesey (1980). After *each* assignment was made for each site by the FSM, we verified with the data that the assignments to GE and GN were better balanced (the covariates of the assigned individuals to the insurance plans were more equally distributed between the group selected as experimental subjects and those who were not) than stratified random sampling permitted. The FSM does this by a process of controlled randomization, partly based on stratification of a few key variables, but mainly by using a multitude of other variables. Had all selected individuals been located and enrolled, this Note would be unnecessary, because the variables used in this Note also were used by the FSM for assigning subjects to treatments, and always were substantially better matched than would be expected from random sampling.

While use of the FSM to make assignments of experimental subjects to treatments (FHPP insurance plans, in the HIE) permits improving upon random sampling, operational effects counter this, reducing statistical control over assignments. However, the net result still will dominate random sampling if these effects act randomly.) In fact, the gains made by the FSM over simple random sampling are reduced in approximate proportion to the nonenrollment rate if nonenrollment is random (Morris, Newhouse, and Archibald, 1979).

One potential GE-GN difference arose from out-of-area moves by families selected for FHPP (moving families were followed before enrollment if they moved within the site). The HIE guarded against this mobility bias caused by moves after the baseline survey by enrolling the families that moved into dwellings vacated by families selected for FHPP enrollment that moved out of the area. Thus, the GN and GE groups are more closely matched by this procedure than they otherwise would have been even if the moving population's health levels differ from that of the more stationary population.

Families selected for FHPP that were contacted, but who refused the preenrollment questionnaire (in order to reduce refusal bias, knowledge of selection for participation and the type of FHPP plan was withheld until after completion of the questionnaire), or who completed the questionnaire and then refused the FHPP offer also are assigned in this analysis to GN. Those found ineligible in the preenrollment questionnaire are assigned to GI.

Finally, GN includes families that were selected for enrollment, but never were enrolled because the enrollment period terminated before their enrollment could be completed. Such delays were caused for hard-to-locate families, because of protracted time to verify insurance coverage, or because the benefits of the family's existing insurance were not assigned to the HIE in time. (Without this assignment, some families could have continued to use their existing insurance in addition to the FHPP, thereby invalidating their experimental responses.)

In summary, many processes, some deliberate (FSM) and others beyond HIE control, led to placement of families in GN.

EVALUATION PHILOSOPHY AND QUALITATIVE RESULTS

The evaluation of this Note is based on the idea that if the two groups GE and GN differ with respect to key preexperimental dependent variables by no more than random assignment suggests they should, then differences between the groups can be ignored and the experimental data adequately reflect the eligible population in the sites. No significant differences were found. This is not just a failure to reject the null

hypothesis, but differences are generally less than expected, with a plausible reason for that--the FSM. We do not show that no distortions were created by nonrandom losses, but only that, with respect to preexperimental experimental variables, the net distortions are less than distortions that would be expected from random sampling.

These tests are "omnibus" tests, in that they are sensitive only to distortions caused by the net effect of all factors operating between the baseline survey and the enrollment period. Had significant differences occurred, it may have been necessary to determine which factors caused the distortion. But as it turned out, this was not required.

The tests are intended only to determine whether the tested groups are as well-matched as would be expected of random sampling. A more powerful test might have been conducted to determine if a specific subgroup, e.g., those refusing the enrollment offer, differs from those enrolling. However, such tests are not appropriate for our purposes here. To see this, consider a simple extreme example. Suppose that the groups GE* and GN* were perfectly matched, as assigned before enrollment offers were made, and suppose that just one individual in the GE* group refused the enrollment offer. Then GE is just the group GE* omitting that individual, and GN is GN* plus that individual. Also, imagine that that individual's physician visits exceeded the 97.5th percentile in the baseline year. Then that person (refusal group) is statistically significantly different from GN at the 5 percent (two-tailed) level of significance. However, GE and GN still are nicely balanced, and the GE experience would still be very appropriate for extrapolation of results to GE combined with GN, or to the entire population. (A few other obvious assumptions are needed here, namely that GE and GN are both sizable, and that GE and GN, combined, are reflective of the general population.) Thus, the most powerful test that the refusal group differs from GE is not the appropriate test for asking whether GE is as adequate for inference to the full population as is a similarly sized group chosen by randomly splitting the members of GE and GN. Random sampling also would lead to differences between GE and GN. Given that in the HIE GE* and GN* were better matched than would be expected of random sampling, we are only asking if nonenrollment degraded the

matching of GE and GN enough so that the GE-GN differences exceed random sampling differences. Our analyses say no.

Finally, two other analyses provide reassurance concerning inferences from HIE samples.

(1) While our data did not provide the means to consider the more powerful test, such an exercise, with different objectives, was undertaken by W. Rogers (1983) with respect to the refusal of the FHPP offer. He found that refusals were significantly less likely for the free plan. However, while some personal characteristics were found to be associated with refusal, none of the variables used in this study were, not even the age of head variable.

(2) Even if certain variables should differ between GE and the population, they bias comparisons between insurance plans only if they interact with plan on the response. Analysis of experimental data has thus far revealed only one quantitatively important interaction of demographic characteristics with plan: hospital admissions among adults varies with plan (Newhouse et al., 1981). Hence, estimated plan responses from the HIE are unlikely to vary in any quantitatively important way from random sampling.

III. DETAILED RESULTS FOR SIX SITES

Tables 3.1-3.24 provide more detailed information about the data and tests summarized in Sec. I. In particular, the ineligible group (GI) is included in these comparisons.

Each of the six sites has four tables, two for comparisons of the enrolled group (GE) versus GN or versus GN and GI, and two to test differences between the FHPP plan groups among the GE category. Thus, there are 24 tables.

There are two chi-square values reported in each table. These include the test statistics reported in Sec. I that compare GE and GN, and another set that simultaneously compares GE, and GN, and GI in three groups. Thus, the latter will have twice as many degrees-of-freedom as the former (Dayton is an exception, for which GN and GI were grouped when making this test). The latter chi-square values were expected to be larger because GI is acknowledged to be a different population than GE or GN (see Sec. II).

Each table provides "deflation factors" D derived to adjust the chi-square values and thereby to properly account for correlation among family members (see Sec. IV and Table 4.1). Chi-square statistics computed before division by the deflator are too large, as is easily seen if all members of a family have the same value for a variable, as with insurance and family income.

Extra information is provided in each table, yielding the numbers of *individuals* analyzed in each group, and the proportions or averages for the variables analyzed. These latter data permit one to consider the magnitudes and directions of differences. A key point is that, with the exception of the age variable, no consistent directions of effects were established for GE versus GN across the six sites. It follows that combining tests that are sensitive to direction of imbalances between GE and GN across sites would not lead to significant effects.

The ineligible group (GI), which represents less than 20 percent of all baseline families, except in Dayton, does differ importantly from GE and GN. The GI group, which meets age and income restrictions, has

lower income, and is much less likely to be insured than other baseline families. In Seattle, GI is much older than the other groups, but this effect is limited to that site. The health status of the GI group is substantially lower than GE and GN in the first four sites, but this does not happen in the South Carolina sites.

The analysis for the Dayton site differs somewhat from the others, and hence the Dayton tables (3.1-3.4) differ somewhat from those in the last 5 sites (3.5-3.24), which all have the same structure. This happened partly because Dayton alone had a deliberate oversampling of low-income groups *after the baseline survey*, necessitating an adjustment for income. The adjustment was accomplished by conducting separate analyses in Dayton for each income group, and summing the independent chi-square values. Hence, the degrees-of-freedom for GE-GN tests in Tables 3.1 and 3.2 are several times as large as in the corresponding tables for the other sites. The Dayton analysis also was prepared before the others, and the GI comparisons were made by lumping GI and GN and then comparing this combined (GI, GN) group with GE. Thus, the degrees-of-freedom is the same for the GE versus GN and the GE versus (GI, GN) tests in Dayton. In all sites other than Dayton, the GI test kept all three groups separate, so that the chi-square statistic (listed in the column between GN and GI) is sensitive to any departure between the three groups for those sites.

The analysis of GE families by FHPP plan status offers some differences for Dayton (which has the only control group, but no 25/50 plan), and Seattle (which has the GHC plan but no 50-percent coinsurance plan). The Dayton control group, which was large, did generally have individuals in poorer health than the FHPP plan there, and had more children (see Tables 3.3, 3.4). This group did not stay intact during the experiment, however, and hence later was dropped--thus its contributions to any imbalances between the plans should be disregarded.

Table 3.1

DAYTON: BY EXPERIMENTAL STATUS AND DEPENDENT VARIABLES

Variables	Group Enrolled (GE)		Group Nonen-rolled (GN)		Group Ineli-gible (GI)
1. Physician visits (2 categories, D=1.74)					
A. test statistics		$\chi^2_3 = 2.4$		$\chi^2_3 = 5.9$	
B. number individuals	1,584		1,046		1,342
C. P(visits > 0)	0.806		0.762		0.760
2. Physician visits (7 categories, D=1.74)					
A. test statistics		$\chi^2_{18} = 17.3$		$\chi^2_{18} = 19.8$	
B. number individuals	1,584		1,046		1,342
C. P(visits ≥ 5)	0.347		0.291		0.323
3. Physician visits (log(2+PV), D=1.74)					
A. test statistics		t = 1.4		t = 1.3	
B. number individuals	1,584		1,046		1,342
C. geometric mean	3.070		2.710		2.920
4. Health status (4 categories, D=2.50)					
A. test statistics		$\chi^2_9 = 5.8$		$\chi^2_9 = 14.2$	
B. number individuals	1,584		1,046		1,550
C. P(H.S. = excellent)	0.509		0.463		0.397
5. Worry about health (4 categories, D=2.05)					
A. test statistics		$\chi^2_9 = 7.3$		$\chi^2_9 = 13.3$	
B. number individuals	1,583		1,046		1,550
C. P(worry = never)	0.600		0.611		0.555
6. Pain (4 categories, D=2.23)					
A. test statistic		$\chi^2_9 = 6.2$		$\chi^2_9 = 5.0$	
B. number individuals	1,583		1,046		1,550
C. P(pain = none)	0.438		0.427		0.419

NOTE: Income not analyzed in Dayton because income was stratified and deliberately required to differ from nonenrolled families by oversampling those earning less than $9,000 and undersampling those between $9,001 and $14,999.

Table 3.2

DAYTON: BY EXPERIMENTAL STATUS AND INDEPENDENT VARIABLES

Variables	Group Enrolled (GE)		Group Nonen-rolled (GN)		Group Ineli-gible (GI)
1. Insured (2 categories, D=3.83)					
A. test statistics		$\chi^2_2 = 1.0$		$\chi^2_2 = 0.6$	
B. number individuals	1,584		1,046		1,342
C. P(insured)	0.900		0.858		0.907
2. Age (4 categories, D=1.5)					
A. test statistics		$\chi^2_9 = 24.3$		$\chi^2_9 = 20.0$	
B. number individuals	1,582		1,038		1,548
C. P(adult)	0.585		0.663		0.603
3. Age (as a numerical variable: D=1.5)					
A. test statistics		t = -1.2		t = -1.4	
B. number individuals	1,582		1,038		1,548
C. mean age	25.00		25.70		25.70
4. Income*					

NOTE: Income not analyzed in Dayton because income was stratified and deliberately required to differ from nonenrolled families by oversampling those earning less than $9,000 and undersampling those between $9,001 and $14,999.

Table 3.3

DAYTON: COMPARISONS BY INSURANCE PLANS AND DEPENDENT VARIABLES

	Plans					
Variables	Free	25/25	50	95/100	ID95	Control
1. Physician visits (2 categories, D=1.74)						
A. test statistics	$\chi^2_5 = 7.1$					
B. number individuals	315	263	197	283	109	644
C. P(visits > 0)	0.803	0.859	0.848	0.784	0.807	0.772
2. Physician visits (7 categories, D=1.74)						
A. test statistics	$\chi^2_{30} = 23.7$					
B. number individuals	315	263	197	283	109	644
C. P(visits ≥ 5)	0.330	0.379	0.361	0.383	0.385	0.337
3. Physician visits (log(2+PV), D=1.74)						
A. test statistics	$\chi^2_5 = 3.8$					
B. number individuals	315	263	197	283	109	644
C. geometric mean	3.290	3.450	3.270	3.230	3.030	2.830
4. Health status (3 categories, D=2.50)						
A. test statistics	$\chi^2_{10} = 6.9$					
B. number individuals	316	265	199	286	111	645
C. P(H.S. = excellent)	0.582	0.479	0.497	0.462	0.477	0.474
5. Worry about health (3 categories, D=2.05)						
A. test statistics	$\chi^2_{10} = 4.9$					
B. number individuals	316	265	199	286	111	644
C. P(worry = never)	0.630	0.570	0.628	0.605	0.622	0.573
6. Pain (3 categories, D=2.23)						
A. test statistic	$\chi^2_{10} = 8.0$					
B. number individuals	316	265	199	286	111	644
C. P(pain = none)	0.487	0.460	0.467	0.427	0.396	0.391

NOTE: Income not analyzed in Dayton because income was stratified and deliberately required to differ from nonenrolled families by oversampling those earning less than $9,000 and undersampling those between $9,001 and $14,999.

Table 3.4

DAYTON: COMPARISONS BY INSURANCE PLANS AND INDEPENDENT VARIABLES

	Plans					
Variables	Free	25	50	95/100	ID95	Control
1. Insured (2 categories, D=3.83)						
A. test statistics	$\chi^2_5 = 4.3$					
B. number individuals	315	263	197	283	109	644
C. P(insured)	0.922	0.966	0.929	0.902	0.964	0.955
2. Age (4 categories, D=1.5)						
A. test statistics	$\chi^2_{15} = 15.9$					
B. number individuals	312	263	197	285	111	642
C. P(adult)	0.596	0.590	0.650	0.593	0.666	0.570
3. Age (as a numerical variable; D=1.5)						
A. test statistics	$\chi^2_5 = 4.3$					
B. number individuals	312	263	197	285	111	642
C. mean age	26.58	25.05	26.56	25.93	27.20	24.31
4. Income*						

NOTE: Income is not analyzed in Dayton.

Table 3.5

SEATTLE: BY EXPERIMENTAL STATUS AND DEPENDENT VARIABLES

Variables	Group Enrolled (GE)		Group Nonen-rolled (GN)		Group Ineli-gible (GI)
1. Physician visits (2 categories, D=1.375)					
A. test statistics		$\chi^2_1 = 0.11$		$\chi^2_2 = 0.41$	
B. number individuals	2,251		2,824		383
C. P(visits > 0)	0.787		0.782		0.770
2. Physician visits (7 categories, D=1.375)					
A. test statistics		$\chi^2_6 = 2.41$		$\chi^2_{12} = 11.29$	
B. number individuals	2,251		2,824		383
C. P(visits ≥ 5)	0.260		0.254		0.295
3. Physician visits (log(2+PV), D=1.371)					
A. test statistics		t = 0.154		$\chi^2_2 = 3.37$	
B. number individuals	2,251		2,824		383
C. geometric mean	2.480		2.460		2.830
4. Health status (4 categories, D=1.871)					
A. test statistics		$\chi^2_3 = 5.48$		$\chi^2_6 = 122.1$	
B. number individuals	2,259		2,835		384
C. P(H.S. = excellent)	0.592		0.555		0.367
5. Worry about health (4 categories, D=1.629)					
A. test statistics		$\chi^2_3 = 0.94$		$\chi^2_6 = 45.6$	
B. number individuals	2,259		2,834		384
C. P(worry = never)	0.539		0.526		0.388
6. Pain (4 categories, D=1.858)					
A. test statistic		$\chi^2_3 = 3.50$		$\chi^2_6 = 37.91$	
B. number individuals	2,252		2,832		383
C. P(pain = none)	0.408		0.434		0.306

Table 3.6

SEATTLE: BY EXPERIMENTAL STATUS AND INDEPENDENT VARIABLES

Variables	Group Enrolled (GE)		Group Nonen-rolled (GN)		Group Ineli-gible (GI)
1. Insured (2 categories, D=3.32)					
A. test statistics		$\chi^2_1 = 0.80$		$\chi^2_2 = 21.02$	
B. number individuals	2,261		2,848		385
C. P(insured)	0.750		0.730		0.546
2. Ages (4 categories, D=1.449)					
A. test statistics		$\chi^2_3 = 17.02$		$\chi^2_6 = 48.25$	
B. number individuals	2,261		2,848		361
C. P(adults)	0.637		0.671		0.800
3. Age (as a numerical variable: D=1.449)					
A. test statistics		t = -3.248		$\chi^2_2 = 38.40$	
B. number individuals	2,261		2,848		385
C. mean age	25.18		26.97		31.82
4. Income (3 categories, D=3.32)					
A. test statistics		$\chi^2_2 = 0.36$		$\chi^2_4 = 19.92$	
B. number individuals	2,190		2,655		362
C. P(income > $15,000)	0.396		0.403		0.271
5. Income (as log(income): D=3.32)					
A. test statistics		t = 0.432		$\chi^2_2 = 21.50$	
B. number individuals	2,190		2,655		362
C. geometric mean	$10,775		$10,572		$7,187

Table 3.7

SEATTLE: COMPARISONS BY INSURANCE PLANS AND DEPENDENT VARIABLES

	Plans					
Variables	Free	25/25	25/50	95	ID95	GHC
1. Physician visits (2 categories, D=1.375)						
A. test statistics	$\chi^2_{5} = 4.16$					
B. number individuals	412	127	109	244	274	1,085
C. P(visits > 0)	0.782	0.803	0.752	0.803	0.741	0.798
2. Physician visits (7 categories, D=1.375)						
A. test statistics	$\chi^2_{30} = 28.20$					
B. number individuals	412	127	109	244	274	1,085
C. P(visits ≥ 5)	0.282	0.252	0.174	0.217	0.252	0.273
3. Physician visits (log(2+PV))						
A. number individuals	412	127	109	244	274	1,085
B. geometric mean	2.490	2.410	2.070	2.250	2.380	2.620
4. Health status (3 categories, D=1.871)						
A. test statistics	$\chi^2_{10} = 4.32$					
B. number individuals	412	128	110	244	275	1,090
C. P(H.S. = excellent)	0.575	0.609	0.482	0.627	0.596	0.598
5. Worry about health (3 categories, D=1.629)						
A. test statistics	$\chi^2_{10} = 3.51$					
B. number individuals	412	128	110	244	275	1,090
C. P(worry = never)	0.536	0.563	0.482	0.566	0.556	0.532
6. Pain (3 categories, D=1.858)						
A. test statistic	$\chi^2_{10} = 2.73$					
B. number individuals	412	125	110	244	275	1,086
C. P(pain = none)	0.410	0.408	0.346	0.430	0.422	0.405

Table 3.8

SEATTLE: COMPARISONS BY INSURANCE PLANS AND INDEPENDENT VARIABLES

	Plans					
Variables	Free	25/25	25/50	95	ID95	GHC
1. Insured (2 categories, D=3.32)						
A. test statistics	$\chi^2_{5} = 4.36$					
B. number individuals	412	129	110	244	275	1,091
C. P(insured)	0.752	0.775	0.609	0.742	0.789	0.753
2. Age (4 categories, D=1.449)						
A. test statistics	$\chi^2_{15} = 14.05$					
B. number individuals	412	129	110	244	275	1,091
C. P(adults)	0.624	0.659	0.645	0.635	0.587	0.627
3. Age (as a numerical variable)						
A. number individuals	412	129	110	244	275	1,091
B. mean age	25.36	26.20	26.80	25.31	26.29	24.52
4. Income (3 categories, D=3.32)						
A. test statistics	$\chi^2_{10} = 8.64$					
B. number individuals	393	126	103	231	266	1,071
C. P(income > $15,000)	0.387	0.325	0.388	0.407	0.402	0.404
5. Income (as log(income))						
A. number individuals	393	126	103	231	266	1,071
B. geometric mean	$10,340	$10,930	$9,029	$10,417	$11,643	$10,979

Table 3.9

FITCHBURG: BY EXPERIMENTAL STATUS AND DEPENDENT VARIABLES

Variables	Group Enrolled (GE)		Group Nonen-rolled (GN)		Group Ineli-gible (GI)
1. Physicians visits (2 categories, D=1.619)					
A. test statistics		$\chi^2_1 = 0.92$		$\chi^2_2 = 1.45$	
B. number individuals	699		1,246		489
C. P(visits > 0)	0.808		0.785		0.812
2. Physicians visits (7 categories, D=1.619)					
A. test statistics		$\chi^2_6 = 3.22$		$\chi^2_{12} = 9.14$	
B. number individuals	699		1,246		489
C. P(visits ≥ 5)	0.281		0.256		0.317
3. Physicians visits (log(2+PV): D=1.441)					
A. test statistics		t = 0.916		$\chi^2_2 = 3.74$	
B. number individuals	699		1,246		489
C. geometric mean	2.64		2.48		2.87
4. Health status (4 categories, D=2.071)					
A. test statistics		$\chi^2_3 = 0.36$		$\chi^2_6 = 5.37$	
B. number individuals	699		1,240		487
C. P(H.S. = excellent)	0.567		0.561		0.493
5. Worry about health (4 categories, D=1.634)					
A. test statistics		$\chi^2_3 = 4.58$		$\chi^2_6 = 6.55$	
B. number individuals	699		1,240		487
C. P(worry = never)	0.587		0.644		0.600
6. Pain (4 categories, D=1.866)					
A. test statistics		$\chi^2_3 = 5.66$		$\chi^2_6 = 7.14$	
B. number individuals	699		1,238		486
C. P(Pain = none)	0.486		0.557		0.494

Table 3.10

FITCHBURG: BY EXPERIMENTAL STATUS AND INDEPENDENT VARIABLES

Variables	Group Enrolled (GE)		Group Nonen-rolled (GN)		Group Ineli-gible (GI)
1. Insured (2 categories, D=3.575)					
A. test statistics		$\chi^2_{1} = 3.40$		$\chi^2_{2} = 36.51$	
B. number individuals	699		1,230		483
C. P(insured)	0.751		0.818		0.551
2. Ages (4 categories, D=1.514)					
A. test statistics		$\chi^2_{3} = 21.54$		$\chi^2_{6} = 31.23$	
B. number individuals	699		1,246		478
C. P(adults)	0.572		0.663		0.567
3. Age (as a numerical variable: D=1.514)					
A. test statistics		t = -3.322		$\chi^2_{2} = 19.79$	
B. number individuals	699		1,246		478
C. mean age	24.57		27.86		23.50
4. Income (3 categories, D=3.575)					
A. test statistics		$\chi^2_{2} = 3.65$		$\chi^2_{4} = 34.54$	
B. number individuals	681		1,163		466
C. P(income > $15,000)	0.323		0.404		0.204
5. Income (as log(income): D=1.512)					
A. test statistics		t = -2.128		$\chi^2_{2} = 91.51$	
B. number individuals	681		1,163		466
C. geometric mean	$10,257		$11,324		$6,711

Table 3.11

FITCHBURG: COMPARISONS BY INSURANCE PLANS AND DEPENDENT VARIABLES

	Plans					
Variables	Free	25/25	25/50	50	95	ID95
1. Physician visits						
(2 categories, D=1.619)						
A. test statistics	$\chi^2_{5} = 6.84$					
B. number individuals	229	37	88	55	106	184
C. P(visits > 0)	0.843	0.811	0.807	0.909	0.802	0.739
2. Physician visits						
(7 categories, D=1.619)						
A. test statistics	$\chi^2_{30} = 22.76$					
B. number individuals	229	37	88	55	106	184
C. P(visits ≥ 5)	0.293	0.243	0.296	0.273	0.255	0.283
3. Physician visits						
(log(2+PV)						
A. number individuals	229	37	88	55	106	184
B. geometric mean	2.900	2.520	2.690	2.570	2.480	2.440
4. Health status						
(3 categories, D=2.071)						
A. test statistics	$\chi^2_{10} = 6.04$					
B. number individuals	229	37	88	55	106	184
C. P(H.S. = excellent)	0.550	0.703	0.511	0.673	0.632	0.516
5. Worry about health						
(3 categories, D=1.634)						
A. test statistics	$\chi^2_{10} = 3.49$					
B. number individuals	229	37	88	55	106	184
C. P(worry = never)	0.572	0.514	0.534	0.673	0.585	0.620
6. Pain						
(3 categories, D=1.866)						
A. test statistic	$\chi^2_{10} = 6.47$					
B. number individuals	229	37	88	55	106	184
C. P(pain = none)	0.467	0.595	0.466	0.564	0.491	0.473

Table 3.12

FITCHBURG: COMPARISONS BY INSURANCE PLANS AND INDEPENDENT VARIABLES

	Plans					
Variables	Free	25/25	25/50	50	95	ID95
1. Insured (2 categories, D=3.575)						
A. test statistics	$\chi^2_{5} = 3.583$					
B. number individuals	229	37	88	55	106	184
C. P(insured)	0.790	0.649	0.670	0.782	0.830	0.706
2. Age (4 categories, D=1.514)						
A. test statistics	$\chi^2_{15} = 5.898$					
B. number individuals	229	37	88	55	106	184
C. P(adults)	0.563	0.568	0.557	0.545	0.613	0.576
3. Age (as a numerical variable)						
A. number individuals	229	37	88	55	106	184
B. mean age	24.16	25.57	25.44	24.24	24.67	24.50
4. Income (3 categories, D=3.575)						
A. test statistics	$\chi^2_{10} = 10.055$					
B. number individuals	218	37	88	55	106	177
C. P(income > $15,000)	0.312	0.460	0.159	0.527	0.274	0.356
5. Income (as log(income))						
A. number individuals	218	37	88	55	106	177
B. geometric mean	$10,169	$12,332	$9,020	$11,410	$10,238	$10,300

Table 3.13

FRANKLIN COUNTY: BY EXPERIMENTAL STATUS AND DEPENDENT VARIABLES

Variables	Group Enrolled (GE)		Group Nonen-rolled (GN)		Group Ineli-gible (GI)
1. Physician visits (2 categories, D=1.383)					
A. test statistics		$\chi^2_{1} = 2.20$		$\chi^2_{2} = 3.46$	
B. number individuals	878		1,513		533
C. P(visits > 0)	0.841		0.812		0.846
2. Physician visits (7 categories, D=1.383)					
A. test statistics		$\chi^2_{6} = 7.84$		$\chi^2_{12} = 17.71$	
B. number individuals	878		1,513		533
C. P(visits ≥ 5)	0.267		0.288		0.357
3. Physician visits (log(2+PV), D=1.546)					
A. test statistics		t = -0.517		$\chi^2_{2} = 7.45$	
B. number individuals	878		1,513		533
C. geometric mean	2.700		2.780		3.290
4. Health status (4 categories, D=2.050)					
A. test statistics		$\chi^2_{3} = 1.22$		$\chi^2_{6} = 13.86$	
B. number individuals	878		1,512		532
C. P(H.S. = excellent)	0.591		0.560		0.470
5. Worry about health (4 categories, D=1.453)					
A. test statistics		$\chi^2_{3} = 1.20$		$\chi^2_{6} = 6.03$	
B. number individuals	878		1,513		532
C. P(worry = never)	0.600		0.597		0.592
6. Pain (4 categories, D=1.792)					
A. test statistic		$\chi^2_{3} = 1.81$		$\chi^2_{6} = 6.23$	
B. number individuals	878		1,510		532
C. P(pain = none)	0.526		0.488		0.485

Table 3.14

FRANKLIN COUNTY: BY EXPERIMENTAL STATUS AND INDEPENDENT VARIABLES

Variables	Group Enrolled (GE)		Group Nonen-rolled (GN)		Group Ineli-gible (GI)
1. Insured (2 categories, D=3.413)					
A. test statistics		$\chi^2_1 = 0.01$		$\chi^2_2 = 86.40$	
B. number individuals	877		1,510		524
C. P(insured)	0.794		0.794		0.428
2. Age (4 categories, D=1.455)					
A. test statistics		$\chi^2_3 = 25.13$		$\chi^2_6 = 27.22$	
B. number individuals	878		1,513		498
C. P(adults)	0.590		0.673		0.602
3. Age (as a numerical variable: D=1.455)					
A. test statistics		t = -4.329		$\chi^2_2 = 20.89$	
B. number individuals	878		1,513		498
C. mean age	24.65		28.43		25.47
4. Income (3 categories, D=3.4133)					
A. test statistics		$\chi^2_2 = 1.41$		$\chi^2_4 = 61.85$	
B. number individuals	854		1,430		505
C. P(income > $15,000)	0.367		0.339		0.141
5. Income (as log(income): D=3.4133)					
A. test statistics		t = -0.348		$\chi^2_2 = 74.65$	
B. number individuals	854		1,430		505
C. geometric mean	$10,604		$10,788		$6,426

Table 3.15

FRANKLIN COUNTY: COMPARISONS BY INSURANCE PLANS AND DEPENDENT VARIABLES

	Plans					
Variables	Free	25/25	25/50	50	95	ID95
1. Physician visits (2 categories, D=1.383)						
A. test statistics	$\chi^2_{5} = 2.94$					
B. number individuals	290	60	90	58	162	218
C. P(visits > 0)	0.841	0.900	0.856	0.776	0.852	0.826
2. Physician visits (7 categories, D=1.383)						
A. test statistics	$\chi^2_{30} = 19.13$					
B. number individuals	290	60	90	58	162	218
C. P(visits ≥ 5)	0.283	0.183	0.167	0.241	0.253	0.326
3. Physician visits (log(2+PV))						
A. number individuals	290	60	90	58	162	218
B. geometric mean	2.860	2.700	2.240	2.150	2.660	2.880
4. Health status (3 categories, D=2.050)						
A. test statistics	$\chi^2_{10} = 2.56$					
B. number individuals	290	60	90	58	162	218
C. P(H.S. = excellent)	0.569	0.600	0.589	0.655	0.617	0.583
5. Worry about health (3 categories, D=1.453)						
A. test statistics	$\chi^2_{10} = 8.76$					
B. number individuals	290	60	90	58	162	218
C. P(worry = never)	0.586	0.567	0.656	0.603	0.605	0.601
6. Pain (3 categories, D=1.792)						
A. test statistic	$\chi^2_{10} = 5.35$					
B. number individuals	290	60	90	58	162	218
C. P(pain = none)	0.517	0.467	0.589	0.603	0.537	0.500

Table 3.16

FRANKLIN COUNTY: COMPARISONS BY INSURANCE PLANS AND INDEPENDENT VARIABLES

	Plans					
Variables	Free	25/25	25/50	50	95	ID95
1. Insured (2 categories, D=3.413)						
A. test statistics	$\chi^2_{5} = 5.11$					
B. number individuals	290	60	90	58	161	218
C. P(insured)	0.731	0.833	0.756	0.810	0.888	0.807
2. Age (4 categories, D=1.455)						
A. test statistics	$\chi^2_{15} = 8.07$					
B. number individuals	290	60	90	58	161	218
C. P(adults)	0.579	0.667	0.533	0.552	0.599	0.610
3. Age (as a numerical variable)						
A. number individuals	290	60	90	58	161	218
B. mean age	24.28	26.15	23.77	24.64	24.71	25.06
4. Income (3 categories, D=3.413)						
A. test statistics	$\chi^2_{10} = 6.24$					
B. number individuals	282	60	85	58	155	214
C. P(income > $15,000)	0.305	0.400	0.282	0.483	0.458	0.374
5. Income (as log(income))						
A. number individuals	282	60	85	58	155	214
B. geometric mean	$9,736	$10,390	$10,647	$11,918	$11,477	$10,902

Table 3.17

CHARLESTON: BY EXPERIMENTAL STATUS AND DEPENDENT VARIABLES

Variables	Group Enrolled (GE)		Group Nonen-rolled (GN)		Group Ineli-gible (GI)
1. Physician visits (2 categories, D=1.735)					
A. test statistics		$\chi^2_1 = 0.69$		$\chi^2_2 = 0.89$	
B. number individuals	717		1,996		243
C. P(visits > 0)	0.739		0.762		0.774
2. Physician visits (7 categories, D=1.735)					
A. test statistics		$\chi^2_6 = 1.67$		$\chi^2_{12} = 8.54$	
B. number individuals	717		1,996		243
C. P(visits ≥ 5)	0.280		0.294		0.313
3. Physician visits (log(2+PV), D=1.538)					
A. test statistics		t = -0.843		$\chi^2_2 = 1.96$	
B. number individuals	717		1,996		243
C. geometric mean	2.450		2.610		2.920
4. Health status (4 categories, D=2.176)					
A. test statistics		$\chi^2_3 = 4.02$		$\chi^2_6 = 5.80$	
B. number individuals	716		1,986		243
C. P(H.S. = excellent)	0.412		0.476		0.444
5. Worry about health (4 categories, D=1.825)					
A. test statistics		$\chi^2_3 = 0.66$		$\chi^2_6 = 4.08$	
B. number individuals	716		1,984		242
C. P(worry = never)	0.594		0.575		0.546
6. Pain (4 categories, D=1.939)					
A. test statistic		$\chi^2_3 = 0.36$		$\chi^2_6 = 8.27$	
B. number individuals	715		1,984		242
C. P(pain = none)	0.476		0.482		0.364

Table 3.18

CHARLESTON: BY EXPERIMENTAL STATUS AND INDEPENDENT VARIABLES

Variables	Group Enrolled (GE)		Group Nonen-rolled (GN)		Group Ineli-gible (GI)
1. Insured (2 categories, D=3.613)					
A. test statistics		$\chi^2_1 = 0.03$		$\chi^2_2 = 14.14$	
B. number individuals	717		1,974		243
C. P(insured)	0.685		0.692		0.465
2. Age (4 categories, D=1.495)					
A. test statistics		$\chi^2_3 = 1.42$		$\chi^2_6 = 5.47$	
B. number individuals	717		1,996		243
C. P(adults)	0.590		0.607		0.675
3. Age (as a numerical variable: D=1.495)					
A. test statistics		t = 0.53		$\chi^2_2 = 3.39$	
B. number individuals	717		1,996		243
C. mean age	25.12		24.67		26.02
4. Income (3 categories, D=3.613)					
A. test statistics		$\chi^2_2 = 4.81$		$\chi^2_4 = 12.89$	
B. number individuals	640		1,701		201
C. P(income > $15,000)	0.186		0.263		0.169
5. Income (as log(income): D=3.613)					
A. test statistics		t = -0.365		$\chi^2_2 = 15.53$	
B. number individuals	640		1,701		201
C. geometric mean	$7,587		$7,790		$4,843

Table 3.19

CHARLESTON: COMPARISONS BY INSURANCE PLANS AND DEPENDENT VARIABLES

	Plans					
Variables	Free	25/25	25/50	50	95	ID95
1. Physician visits (2 categories, D=1.735)						
A. test statistics	$\chi^2_{5} = 6.48$					
B. number individuals	239	64	77	25	132	180
C. P(visits > 0)	0.762	0.594	0.831	0.680	0.773	0.706
2. Physician visits (7 categories, D=1.735)						
A. test statistics	$\chi^2_{30} = 17.84$					
B. number individuals	239	64	77	25	132	180
C. P(visits ≥ 5)	0.297	0.219	0.286	0.200	0.295	0.278
3. Physician visits (log(2+PV))						
A. number individuals	239	64	77	25	132	180
B. geometric mean	2.450	1.570	2.890	2.050	2.760	2.450
4. Health status (3 categories, D=2.176)						
A. test statistics	$\chi^2_{10} = 8.53$					
B. number individuals	239	64	77	25	132	179
C. P(H.S. = excellent)	0.423	0.563	0.390	0.560	0.311	0.408
5. Worry about health (3 categories, D=1.825)						
A. test statistics	$\chi^2_{10} = 3.97$					
B. number individuals	239	64	77	25	132	179
C. P(worry = never)	0.590	0.625	0.623	0.640	0.538	0.609
6. Pain (3 categories, D=1.939)						
A. test statistic	$\chi^2_{10} = 11.76$					
B. number individuals	239	64	76	25	132	179
C. P(pain = none)	0.515	0.469	0.592	0.560	0.296	0.497

Table 3.20

CHARLESTON: COMPARISONS BY INSURANCE PLANS AND INDEPENDENT VARIABLES

	Plans					
Variables	Free	25/25	25/50	50	95	ID95
1. Insured (2 categories, D=3.613)						
A. test statistics	$\chi^2_{5} = 2.36$					
B. number individuals	239	64	77	25	132	180
C. P(insured)	0.682	0.828	0.662	0.640	0.629	0.694
2. Age (4 categories, D=1.495)						
A. test statistics	$\chi^2_{15} = 7.11$					
B. number individuals	239	64	77	25	132	180
C. P(adults)	0.611	0.563	0.532	0.520	0.576	0.617
3. Age (as a numerical variable)						
A. number individuals	239	64	77	25	132	180
B. mean age	25.71	26.81	23.41	24.80	23.69	25.57
4. Income (3 categories, D=3.613)						
A. test statistics	$\chi^2_{10} = 4.63$					
B. number individuals	220	59	70	25	107	159
C. P(income > $15,000)	0.173	0.322	0.257	0.080	0.187	0.157
5. Income (as log(income))						
A. number individuals	220	59	70	25	107	159
B. geometric mean	$7,909	$10,066	$8,419	$6,978	$6,092	$7,237

Table 3.21

GEORGETOWN COUNTY: BY EXPERIMENTAL STATUS AND DEPENDENT VARIABLES

Variables	Group Enrolled (GE)		Group Nonen-rolled (GN)		Group Ineli-gible (GI)
1. Physician visits (2 categories, D=1.830)					
A. test statistics		$\chi^2_1 = 0.01$		$\chi^2_2 = 0.21$	
B. number individuals	941		1,814		213
C. P(visits > 0)	0.668		0.671		0.648
2. Physician visits (7 categories, D=1.830)					
A. test statistics		$\chi^2_6 = 4.77$		$\chi^2_{12} = 11.31$	
B. number individuals	941		1,814		213
C. P(visits ≥ 5)	0.223		0.267		0.329
3. Physician visits (log(2+PV), D=1.740)					
A. test statistics		t = -1.045		$\chi^2_2 = 1.78$	
B. number individuals	941		1,814		213
C. geometric mean	2.030		2.200		2.400
4. Health status (4 categories, D=2.490)					
A. test statistics		$\chi^2_3 = 2.64$		$\chi^2_6 = 7.05$	
B. number individuals	936		1,810		211
C. P(H.S. = excellent)	0.346		0.365		0.332
5. Worry about health (4 categories, D=1.962)					
A. test statistics		$\chi^2_3 = 3.67$		$\chi^2_6 = 5.72$	
B. number individuals	937		1,805		212
C. P(worry = never)	0.584		0.614		0.594
6. Pain (4 categories, D=2.199)					
A. test statistic		$\chi^2_3 = 2.29$		$\chi^2_6 = 3.14$	
B. number individuals	938		1,807		211
C. P(pain = none)	0.472		0.500		0.441

Table 3.22

GEORGETOWN COUNTY: BY EXPERIMENTAL STATUS AND INDEPENDENT VARIABLES

Variables	Group Enrolled (GE)		Group Nonen-rolled (GN)		Group Ineli-gible (GI)
1. Insured (2 categories, D=3.954)					
A. test statistics		$\chi^2_1 = 2.30$		$\chi^2_2 = 10.72$	
B. number individuals	904		1,766		210
C. P(insured)	0.731		0.674		0.500
2. Age (4 categories, D=1.436)					
A. test statistics		$\chi^2_3 = 3.94$		$\chi^2_6 = 5.14$	
B. number individuals	941		1,814		213
C. P(adults)	0.571		0.599		0.629
3. Age (as a numerical variable: D=1.436)					
A. test statistics		t = 0.209		$\chi^2_2 = 0.05$	
B. number individuals	941		1,814		213
C. mean age	24.90		24.73		24.91
4. Income (3 categories, D=3.954)					
A. test statistics		$\chi^2_2 = 2.08$		$\chi^2_4 = 5.71$	
B. number individuals	788		1,482		143
C. P(income > $15,000)	0.207		0.215		0.119
5. Income (as log(income): D=3.954)					
A. test statistics		t = 0.343		$\chi^2_2 = 7.23$	
B. number individuals	788		1,482		143
C. geometric mean	$7,889		$7,699		$5,357

Table 3.23

GEORGETOWN COUNTY: COMPARISONS BY INSURANCE PLANS AND DEPENDENT VARIABLES

	Plans					
Variables	Free	25/25	25/50	50	95	ID95
1. Physician visits						
(2 categories, D=1.830)						
A. test statistics	$\chi^2_{5} = 4.83$					
B. number individuals	317	77	107	48	146	246
C. P(visits > 0)	0.609	0.675	0.720	0.646	0.664	0.728
2. Physician visits						
(7 categories, D=1.830)						
A. test statistics	$\chi^2_{30} = 14.93$					
B. number individuals	317	77	107	48	146	246
C. P(visits ≥ 5)	0.208	0.234	0.252	0.188	0.178	0.260
3. Physician visits						
(log(2+PV))						
A. number individuals	317	77	107	48	146	246
B. geometric mean	1.780	2.050	2.200	1.730	1.860	2.440
4. Health status						
(3 categories, D=2.49)						
A. test statistics	$\chi^2_{10} = 9.33$					
B. number individuals	317	77	107	46	145	244
C. P(H.S. = excellent)	0.306	0.390	0.467	0.304	0.338	0.344
5. Worry about health						
(3 categories, D=1.96)						
A. test statistics	$\chi^2_{10} = 9.04$					
B. number individuals	317	77	107	46	145	245
C. P(worry = never)	0.539	0.610	0.608	0.522	0.655	0.592
6. Pain						
(3 categories, D=2.20)						
A. test statistic	$\chi^2_{10} = 3.15$					
B. number individuals	317	77	107	46	145	246
C. P(pain = none)	0.461	0.571	0.430	0.478	0.510	0.451

Table 3.24

GEORGETOWN COUNTY: COMPARISONS BY INSURANCE PLANS
AND INDEPENDENT VARIABLES

	Plans					
Variables	Free	25/25	25/50	50	95	ID95
1. Insured (2 categories, D=3.954)						
A. test statistics	$\chi^2_5 = 4.67$					
B. number individuals	297	67	100	48	146	246
C. P(insured)	0.788	0.582	0.790	0.750	0.658	0.720
2. Age (4 categories, D=1.436)						
A. test statistics	$\chi^2_{15} = 12.07$					
B. number individuals	317	77	107	48	146	246
C. P(adults)	0.574	0.520	0.542	0.614	0.555	0.598
3. Age (as a numerical variable)						
A. number individuals	317	77	107	48	146	246
B. mean age	25.78	23.95	24.02	25.85	23.40	25.13
4. Income (3 categories, D=3.954)						
A. test statistics	$\chi^2_{10} = 12.42$					
B. number individuals	263	55	89	42	133	206
C. P(income > $15,000)	0.255	0.236	0.202	0.286	0.053	0.223
5. Income (as log(income))						
A. number individuals	263	55	89	42	133	206
B. geometric mean	$8,121	$8,000	$8,641	$8,258	$5,845	$8,754

IV. STATISTICAL THEORY: DEFLATORS FOR TEST STATISTICS

All chi-square values in the preceding sections have been deflated by a factor $D = 1 + \rho(\bar{F} - 1)$ (ρ = correlation among individuals in the family for the given variable, and F = "person-weighted average family size"), from the standard chi-square value to account for within-family correlation when using person-level data. This is required because most variables used are individual level (physician visits, health status, worry, pain, and age). Strong correlations are estimated among individuals within a family, and so chi-square test values that fail to acknowledge this correlation are too large.

The easiest way to see the need for deflation is to consider $\rho = 1$, the case of insurance and income, in which every family member has the same value. The correct number of independent observations is the number of families, but the standard analysis of individuals assumes there are as many independent observations as there are individuals. The chi-square statistic is inflated by a factor $\bar{F}$ in this case. In Dayton, for example, $\bar{F} = 3.83$, and while Dayton averages 3.0 persons per family, the person-weighted average is 3.83 people.

When $\rho < 1$, as with the other variables, then the statistics should be deflated by dividing by $D = 1 + \rho(\bar{F} - 1) = 1 + 2.83\,\rho$ for Dayton. When $\rho = 0$ or $\rho = 1$, $D = 1$ or $D = \bar{F}$ is consistent with the proper use of the chi-square statistic applied to independent individuals ($\rho = 0$) or to independent families ($\rho = 1$).

Here ρ is the ratio of the variance between families to the total variance;

$$\rho = \frac{\tau^2}{\tau^2 + \sigma^2} \qquad (4.1)$$

with τ^2 = variance between family means and σ^2 = within-family variance. Thus, $\tau^2 + \sigma^2$ = total variance, or the variance of a randomly selected individual. These values could vary with family size, but they do not

do so for the variables of this study, except for the age variable, for which ρ drops sharply as family size increases. All relevant values of ρ have been estimated for this study for each site and variable and are given in Table 4.1. The t-statistics are properly deflated by dividing by the square root of the deflator: $\sqrt{1 + \rho(\bar{\bar{F}} - 1)} = D^{0.5}$.

The largest values of $\rho < 1$ occur for health status, worry, and pain. This may be due, in part, to the use of proxy respondents (one person who answers for the entire family). If such a respondent answered, for example, "I never worry about any of them" without carefully considering individual cases, then a spuriously large correlation would ensue.

Table 4.1

DEFLATORS AND CORRELATIONS: $D = 1 + \rho(\bar{\bar{F}} - 1)$
ESTIMATED FROM FORMULA (4.15)

Site	Physician Visits	Log (Visits + 2)	Health Status	Worry	Pain	Insurance[a]	Age	Income[a]
Dayton	1.74 0.26	1.74 0.26	2.50 0.53	2.05 0.37	2.23 0.43	3.83 1.00	1.5 0.16	3.83 1.00
Seattle	1.375 0.16	1.371 0.16	1.871 0.38	1.629 0.27	1.858 0.37	3.32 1.00	1.449 0.19	3.32 1.00
Fitchburg	1.619 0.24	1.441 0.17	2.071 0.42	1.634 0.25	1.866 0.34	3.575 1.00	1.514 0.20	3.575 1.00
Franklin Co.	1.383 0.16	1.546 0.23	2.050 0.44	1.453 0.19	1.792 0.33	3.413 1.00	1.455 0.19	3.143 1.00
Charleston	1.735 0.28	1.538 0.21	2.176 0.45	1.825 0.32	1.939 0.36	3.613 1.00	1.495 0.19	3.613 1.00
Georgetown Co.	1.830 0.28	1.740 0.25	2.490 0.50	1.962 0.32	2.199 0.40	3.954 1.00	1.436 0.15	3.954 1.00

[a]Deflators for (family level) insurance and income variables are $D = \bar{\bar{F}}$.

DEFLATOR MODEL AND ANALYSIS

The deflators and correlations were estimated using the following variance components model. Let y_{if} be the observation for individual i in family f. Let s be the family size, i = 1, ..., s and f = 1, ..., N_s with N_s the number of families of size s. Assume for families of size s that

$$y_{if} = \mu_s + u_f + e_{if} \tag{4.2}$$

with μ_s = mean response, u_f = family-specific component with mean zero and variance $\sigma^2{}_s$, and e_{if} has mean zero and variance $\tau^2{}_s$. Assume $e_{1f}, \ldots, e_{sf}, u_f$ are all independent. Then the variance of an individual selected at random from the family is $v_s = \tau^2{}_s + \sigma^2{}_s$. Because of correlation $\rho_s \equiv \tau^2{}_s/v_s = \mathrm{Corr}(y_{1f}, y_{2f})$ between members of a family, the variance of the average response for family f is

$$\mathrm{Var}(\frac{1}{s}\Sigma_i y_{if}) \equiv D_s(\frac{1}{s^2}\Sigma_i \ \mathrm{Var}\ y_{if}) \tag{4.3}$$

with D_s, the "deflator", defined by this relationship. The right-hand side of (4.3) is calculated as if family members are independent, which they are not, and multiplication by D_s produces the correct variance. Using (4.2), (4.3) simplifies to

$$\mathrm{Var}(u_f + (1/s)\Sigma\varepsilon_{if}) = D_s((1/s)v_s)$$

or

$$D_s = \frac{s(\tau^2{}_s + \sigma^2{}_s/s)}{v_s} = 1 + (s - 1)\rho_s \tag{4.4}$$

with

$$\rho_s \equiv \tau^2{}_s/v_s. \tag{4.5}$$

If family members behave independently, then there is no intrafamily component, so $\tau^2{}_s = 0 = \rho_s$ and hence $D_s = 1$, meaning that the variance computed assuming independence is correct. On the other hand, if $\rho_s = 1$, so that $\sigma^2{}_s = 0$, no variation exists between family members (e.g., the income variable) and $D_s = s$, the family size.

Next, we let family sizes vary, supposing the largest family size is denoted m. Let $n = \Sigma_1^m N_s$ be the total number of families. Assume the model (4.2) and independence of all observations among different families. We compute the variance of the total response of individuals across families as (letting f_s be a family of size s below)

$$\text{Var}(\sum_{f=1}^{n} \sum_{i=1}^{s} y_{if}) \tag{4.6}$$

$$= \sum_{s=1}^{m} N_s \text{Var}(\sum_{i=1}^{s} y_{ifs})$$

$$= \Sigma_1^m \; N_s D_s \Sigma_1^s \; \text{Var}(y_{ifs})$$

$$= \Sigma_1^m \; sN_s D_s v_s. \tag{4.7}$$

The value of (4.6), computed by assuming no within family component u_f, but with marginal variance v_s yields

$$\sum_{f=1}^{n} \sum_{i=1}^{s} \text{Var}(y_{if}) = \Sigma_1^m \; N_s s v_s. \tag{4.8}$$

Thus, with $W_s \equiv sN_s v_s$ the deflator is the ratio of (4.7) to (4.8),

$$D \equiv \frac{\Sigma_1^m \; W_s D_s}{\Sigma_1^m \; W_s} = \frac{\text{Var}(\Sigma_f \Sigma_i y_{if})}{\Sigma_f \Sigma_i \text{Var}(y_{if})}. \tag{4.9}$$

The deflator D is to be used when the variance of a linear statistic is computed as in (4.8) assuming independence, if the correct variance is given by (4.7). We divide the statistic by $D \geq 1$ (deflate the statistic) because the variance appears in the denominator of most statistics.

Brier (1980) gives a similar result for tests based on binomial and multinomial distributions. The deflators D_s here agree with his formulas in the binomial case, but he replaces W_s when computing D by $W'_s \equiv sN_s$. The derivation here shows that for the model (4.2) that the W_s weights are better and consequently they are used in this study. However, Brier's weights were computed for this application and were found to be almost always within five percent of the values of Table 4.1, with no systematic bias.

One must estimate σ^2_s and τ^2_s to implement this model. We estimate σ^2_s by

$$\hat{\sigma}^2_s = \frac{1}{N_s} \sum_{f=1}^{N_s} \hat{\sigma}^2_{s,f}, \qquad (4.10)$$

where $\hat{\sigma}^2_{s,f}$ is the (unbiased) sample variable of observations within the family. The sample between group variance for families of size s is

$$MSB_s = \frac{1}{N_s - 1} \sum_{f=1}^{N_s} (m_f - \bar{m})^2 \qquad (4.11)$$

with $m_f = (1/s)\Sigma y_{if}$ the sample mean for family f and m the average of all family sample means. Then

$$E(MSB_s) = \tau^2_s + \sigma^2_s/s,$$

and so we estimate τ^2_s by

$$\hat{\tau}^2_s = (MSB)_s - \sigma^2_s/s. \qquad (4.12)$$

The estimates in (4.10) and (4.12) were used in this paper to estimate D_s by

$$\hat{D}_s = 1 + (s - 1)\hat{\rho}_s, \tag{4.13}$$

with

$$\hat{\rho}_s \equiv \hat{\tau}^2{}_s/(\hat{\tau}^2{}_s + \hat{\sigma}^2{}_s). \tag{4.14}$$

Then

$$\hat{D} = \Sigma\hat{W}_s\hat{D}_s/\Sigma\hat{W}_s \tag{4.15}$$

with

$$\hat{W}_s = sN_s(\hat{\sigma}^2{}_s + \hat{\tau}^2{}_s). \tag{4.16}$$

We define the aggregate intrafamily correlation as

$$\hat{\rho} = \frac{\Sigma W_s(s - 1)\hat{\rho}_s}{\Sigma W_s(s - 1)}, \tag{4.17}$$

and the average family size of a randomly selected person (a value weighted toward the large families) as

$$\bar{s} = \frac{\Sigma W_s s}{\Sigma W_s} = \bar{F}, \tag{4.18}$$

in the notation of the paper. Then it follows that

$$\hat{D} = 1 + (\bar{s} - 1)\hat{\rho}. \tag{4.19}$$

Values of the deflators $\hat{D}$ and correlations $\hat{\rho}$ appear in Table 4.1. Standard statistics nominally having chi-square or F distributions, but computed assuming $\tau^2_s = 0$ for all s, should be divided by $\hat{D}$. Only then do they have approximately the claimed chi-square or F distribution. A t-statistic would be divided by $\hat{D}^{1/2}$.

Appendix

A DESCRIPTION OF THE HEALTH INSURANCE EXPERIMENT DESIGN

The HIE was undertaken to address questions of financing health care through alternative insurance plans, and in particular to deal with policy-related issues concerning the relationships between use of health services and both cost-sharing (coinsurance or deductible payments) and practice organization (fee-for-service or prepaid group practice). Original research questions focused on microeconomic supply and demand models of health care, emphasizing measurement of how different health care financing arrangements affect the use of services. Particular attention was paid to the poor and near-poor (the disadvantaged).

Available nonexperimental data on insurance and use of medical services were inadequate for the level of analysis necessary to address issues of financing health care, particularly those concerning the effects of generosity of insurance on the quality of care and health status. Consequently, a longitudinal experiment was planned to collect more complete and accurate information. The experimental part of the HIE provides direct estimates of the effects of differences in cost-sharing on health status, recognizing that improved health status is an important goal of efforts to expand and make more equitable the financing of medical care services. The experiment also examines the effects of differential cost sharing on the quality of care and patient satisfaction with care, because these variables are policy relevant and may influence the relationship between generosity of health insurance and health status.

SPECIFIC POLICY QUESTIONS ADDRESSED BY THE HEALTH INSURANCE EXPERIMENT

Seven major policy research objectives were initially defined for HIE investigation. All relate to the effects of altering cost to the patient of health care on the use of services, health status, quality of care, and patient satisfaction. The objectives are:

1. To estimate how alternative cost-sharing arrangements affect the demand for health care services. If several groups of similar people are covered for the same health care services and the cost of these services varies from group to group (through different levels of coinsurance, deductibles, or maximum out-of-pocket payments), how does the use of or demand for health services vary across those groups?
2. To assess the effect of varying the out-of-pocket cost of health services on individual health status. If the use of health services differs as a function of cost, what is the effect on health?
3. To determine whether and by how much cost-sharing arrangements affect low-income families more than higher-income families.
4. To learn whether the quality of the medical care process differs for individuals with various health insurance financing plans.
5. To ascertain how the ambulatory care system responds to varying levels of demand or stress. Differences in the use of services, quality of care, and patient satisfaction can be examined as a function of such factors as delays in making appointments, waiting time in physicians' offices, or referral patterns.
6. To compare utilization, health status outcomes, quality of care, and patient satisfaction in an existing prepaid group practice and in the fee-for-service system.
7. To gain familiarity with the difficulties of administering health insurance plans that relate the degree of cost-sharing to the patient's income.

METHODS

Selection of Families and Sites

A total of 2,756 families consisting of 7,706 persons have been enrolled in one of several different health-insurance plans: 70 percent of them for three years, and the rest for five years. The details of the experimental design as of late 1973 appear in Newhouse (1974). The

families come from six areas: Seattle, Washington; Dayton, Ohio; Charleston, South Carolina; Fitchburg and Leominster, Massachusetts; Franklin County, Massachusetts; and Georgetown County, South Carolina. Families who moved after enrollment were kept in the study as long as they remained within the United States.

Within these six sites, families were selected at random, with selection subject to certain restrictions. Those who were eligible for Medicare at the beginning of the study (or who would become so by virtue of age before the end of the study) were excluded. Hence, HIE results may not apply to the aged population. Families with incomes in excess of $25,000 (in 1973 dollars) were also excluded (5 percent of all families).

The sites themselves were determined as follows. The optimal number of sites for the given budget was calculated to be six. The six sites were selected to represent all four census regions, in order to account for regional variation; to obtain a spectrum of city sizes, because the complexity of the medical-care delivery system could vary with city size; to achieve variation in the degree of pressure on the ambulatory-care delivery system (e.g., in waiting times for an appointment and in the proportion of primary-care physicians accepting new patients), because the response of the type of plan could vary according to the pressure already on the system; to include both northern and southern rural areas, because these areas tend to differ in economic and racial characteristics; and to ensure that one site had a well-established prepaid group practice. The actual sites chosen satisfy these criteria. The characteristics of HIE samples, when averaged across the sites, do not differ markedly from those of the nation, save for the intentional departures such as exclusion of the aged.

ASSIGNMENT OF FAMILIES TO EXPERIMENTAL INSURANCE PLANS

The families selected to enroll in the experiment were assigned to an insurance plan by an unbiased method, the Finite Selection Model (Morris, 1979). The model limits the amount of randomization in order to improve on simple and stratified random assignments. It does this by making the distribution of 20 characteristics of families or individuals

as similar as possible across insurance plans (cf. Morris, 1979). The Finite Selection Model computer description is provided by Seda and Keesey (1980).

Of the contacted families, 15 percent were not considered because they refused an early screening or baseline interview; there is little way of knowing whether these families would behave differently from those who ultimately enrolled. Another 20 percent of the families refused a subsequent interview before enrollment or refused the enrollment offer itself.

DESCRIPTION OF INSURANCE PLANS

The insurance plans offered to the families varied along two dimensions: the coinsurance rate (the fraction of the bill paid by the family) and the maximum dollar expenditure (an upper limit on the family's annual out-of-pocket expenditure). The four coinsurance rates were 0 (free care) and 25, 50, and 95 percent. The maximum dollar expenditure (MDE) varied as a fraction of the family's income, either 5, 10, or 15 percent, to a maximum of $1,000 (MDE distinctions are ignored in this report). The 95 percent coinsurance plan, together with the limit on the family's expenditure, approximates an income-related catastrophe plan.

The following example illustrates how the experimental plan operated. A family assigned to a plan with a 25 percent coinsurance rate and a $1,000 maximum dollar expenditure would pay 25 percent of all medical and dental bills in each year until the total of all these bills reached $4,000. At that point, it would have spent $1,000 out of pocket; afterward, all further expenditure during that year would be fully paid (or reimbursed) by the experimental plan. At the beginning of the next year, the family would revert to paying 25 percent coinsurance until they again reached the $1,000 limit.

One plan differs somewhat from all the others. It has a 95 percent coinsurance rate, but the maximum annual out-of-pocket expenditure is limited to $150 for each person or, alternatively, to $450 for the family. This plan approximates a plan with a $150-per-person annual deductible, with a provision that no more than three people in the family satisfy the deductible. We refer to this plan as the individual deductible (ID95) plan.

Most enrolled families were permitted to seek care from any provider. Some families in Seattle, however, were assigned to a prepaid group practice, the Group Health Cooperative (GHC) of Puget Sound.

All insurance plans covered a wide variety of services, including hospitals, physician, dental, mental-health, visual, and auditory services, prescription drugs, and supplies. The only noteworthy exclusions from coverage were nonpreventive orthodontic services, cosmetic surgery for preexisting conditions, and outpatient mental-health visits exceeding 52 per year.

With two exceptions, each policy covers all services at a single coinsurance rate. First of all, three plans require 50-percent coinsurance for dental services and outpatient mental-health services, but only 25 percent for all other services. Secondly, the individual deductible plan applies cost sharing solely to outpatient services; inpatient services are free to the family. The individual deductible plan thus approximates the situation of many families who have complete or nearly complete insurance for inpatient services but poorer insurance for outpatient services. It is designed to test the hypothesis that failure to provide full coverage for outpatient services inflates total medical expenditures by inducing additional hospitalization.

Upon enrollment, families signed over (assigned) the benefits from any existing health-insurance coverage to the experiment; such policies are kept in force so that the families could return to them at the end of the experiment. Any family whose existing coverage was such that participating in the experiment could potentially make it worse off financially was paid, over the course of the experiment, an amount at least equal to its maximum possible loss--an amount paid irrespective of the family's actual use of services. For example, if the family had an experimental plan with a $450 maximum dollar expenditure and had an existing policy with a $200 deductible, it was paid at least $250 per year (250 = 450 - 200). Thus, a family always gained financially by enrolling.

ARTIFACTS OF THE EXPERIMENT

The HIE controlled for several methodologic artifacts of the experiment, including the "hold-harmless" payments just described, possible transitory effects from the limited duration of the experiment, the frequency of questionnaire administration, and the taking of an initial screening examination.