The Demand for Episodes of Medical Treatment

Interim Results from the Health Insurance Experiment

Emmett B. Keeler, John E. Rolph,
with Naihua Duan, Janet Hanley, and Willard G. Manning, Jr.

Rand

HEALTH INSURANCE EXPERIMENT SERIES

The research reported herein was performed pursuant to Grant No. 016B-80 from the U.S. Department of Health and Human Services, Washington, D.C.

Library of Congress Cataloging in Publication Data
Main entry under title:

The Demand for episodes of medical services.

"R-2829-HHS."

Bibliography: p.
1. Medical care, Cost of--United States. 2. Medical care--United States--Utilization. 3. Insurance, Health --United States. I. Keeler, Emmett B. II. United States. Dept. of Health and Human Services. III. Rand Corporation.
RA410.53.D45 1983 362.1 82-21554
ISBN 0-8330-0463-8

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

Published by The Rand Corporation

R-2829-HHS

The Demand for Episodes of Medical Treatment

Interim Results from the Health Insurance Experiment

Emmett B. Keeler, John E. Rolph,
with Naihua Duan, Janet Hanley, and Willard G. Manning, Jr.

December 1982

Prepared under a grant from the
U.S. Department of Health and Human Services

Rand
SANTA MONICA, CA. 90406

PREFACE

This report contains a statistical and economic analysis of data on the demand for medical care from the Rand Health Insurance Experiment (HIE). The HIE, funded by a grant from the U.S. Department of Health and Human Services, is a large-scale social experiment designed to assess how varying patients' cost of health services affects their use of services and their health status. The experimental design for estimating the effects of financing on demand for health care is described in J. P. Newhouse, "A Design for a Health Insurance Experiment," Inquiry, Vol. 11, March 1974, pp. 5-27. Other economics papers include J. P. Newhouse et al., "Some Interim Results from a Controlled Trial in Health Insurance," New England Journal of Medicine, December 17, 1981, and Naihua Duan et al., A Comparison of Alternative Models for the Demand for Medical Care, The Rand Corporation, R-2754-HHS, January 1982. In addition, many Rand Papers describe the measurement and analysis of health effects, details of the experimental design, and data reliability issues.

This report presents methods and interim results for medical care spending data organized by episodes, a powerful and fairly new approach to the study of demand. It applies the theory presented in E. B. Keeler, J. P. Newhouse, and C. E. Phelps, "Deductibles and the Demand for Medical Services: The Theory of a Consumer Facing a Variable Price Schedule Under Uncertainty," Econometrica, April 1977. The methods should interest health economists and health services researchers, and the results should be useful to persons involved in designing, choosing, or evaluating health insurance schemes.

SUMMARY

Those designing or choosing health insurance schemes need to know how cost-sharing by users affects the use of health services. Studies of the effects of copayment that are based on nonexperimental data are flawed because people in chronic or frequent poor health may choose fuller coverage than most other people, and because of inevitable gaps in such data. To overcome these data problems, The Rand Corporation has been conducting a social experiment, the Health Insurance Experiment (HIE). The HIE assigned families to 14 different insurance plans, balancing the plan groups in terms of nonprice characteristics that affect use. Much effort has gone into obtaining complete and accurate data on the participants' use of health services.

The initial analysis of the experimental data estimated the effects of plan on annual expenses. These estimates are unbiased and accurate, and can be extrapolated to plans similar to the 14 tested. However, annual estimates cannot readily be used to assess the effects of insurance with a different scope of coverage, a different basis for copayment, or a different deductible. To estimate the effects of such insurance, we must know how decisions to buy medical services during the year are actually made.

We show that such decisions can be analyzed in terms of episodes that contain all the spending associated with a given bout of illness, chronic condition, or well-care procedure. Since most participants have several episodes of various types in a year, information at this level may also increase precision, untangle competing effects, and generate new ideas on how to use copayments to meet the goals of health insurance more effectively.

To use the episodic approach, we developed procedures for grouping expenses into episodes. Grouping was based on diagnosis, treatment history (initial, repeat, or routine), referral and other linking information (for drugs, tests, and supplies), and the amount of time between possibly related claims. Episodes were categorized (hospital, acute ambulatory, chronic, well, or dental care), and spending was

summed and dated to the first time the patient could have anticipated it.

The analytic methods and results presented here are a first attempt to determine the effects of price on behavior. The results are based on data from the first three years of the experiment in Dayton, Ohio. More refined methods based on more data will be used in subsequent analyses and reports.

The major finding is that although price affects the number of episodes chosen by participants, it has little effect on the cost of each episode. Several approaches were taken to detect differences in the cost of episodes across plans, but the only differences found were somewhat larger dental and chronic episodes on the full coverage plan. Differing rates of episodes, then, explain almost all the differences in spending between plans.

Episode rates differ most between full coverage and the other plans as a group, with less difference between low and high levels of coinsurance. Acute episode rates are more affected by price than are the rates of dental, chronic, hospital, or well-care episodes. Health status, age, sex, and MD visits in the year before the experiment are the most important determinants of episode rates. Income, insurance plan, and having a regular doctor and dentist are highly significant statistically, but less powerful than the first group of variables. The occurrence of dental and well-care episodes was strongly associated with education and income, but not with health status. A negative binomial model in which variables affect the rate of episodes multiplicatively fit the data well. Family members tend to be more alike in usage than the explanatory variables they share would predict.

There were small "sales," or surges of demand, for well care and dental care on the full coverage plan at the beginning and end of the three-year experimental period. Episode rates of those on the pay plans who exceeded their maximum-dollar-expenditure limit (MDE) on out-of-pocket spending during the year showed no immediate "sale" surge, but instead rose slowly toward free plan rates for acute care spending, and more rapidly for chronic, well care, and dental care spending.

Possible anticipation of exceeding the limit for families who were close to it occurred too rarely to be important in the large deductible plans. The \$150 individual deductible effectively restrained demand for ambulatory care, as most participants never spent more than \$150 a year on ambulatory care nor appeared to anticipate that they would.

People do not appear to change the timing of medical purchases to reduce costs. They may find that it is too much trouble to keep track of sale periods (when the price of care is temporarily low), or they may not want to wait for relief.

The result that the cost of episodes is about the same on all plans is unexpected and important, and it simplifies the task of simulating and predicting medical expenses. If the episodes treated on the pay plans are more severe than average, then the equality of cost reflects some economizing beyond the decision to treat. If the treated episodes are equally severe on all plans, then the approximate equality of cost implies that insurance is not affecting the choice of doctor or the quality of care provided by that doctor. The only question about the effects of insurance on health status then becomes whether the higher full coverage rate of episodes or the lower partial coverage rate is better for health.

ACKNOWLEDGMENTS

Creating the episode data file from the claims was a large and difficult job. We would like to thank George Goldberg for clinical help with the rules, and Joan Keesey, William Fowler, and Helene Mills for the episodes program. Daniel Relles and William Rogers provided useful statistical software. Through many drafts, Martha Cooper did her usual fine job of typing. Gus Haggstrom and Mark Pauly provided challenging and helpful reviews. We also benefited from discussion with Robert Bell, Robert Brook, Carl Morris, and Charles Phelps. We would especially like to thank Joseph Newhouse for support, encouragement, and advice.

CONTENTS

Appendix

FIGURES

TABLES

Chapter 1

INTRODUCTION

The soaring costs of medical care have made health insurance more necessary, but insurance is an important cause of these rising costs. Cost-sharing in health insurance is often proposed as a way out of this bind, but until recently there has been little quantitative evidence on the effect of cost-sharing. Studies of the effects of cost-sharing based on nonexperimental data are flawed because people in chronic or frequent poor health may choose fuller coverage than most other people, and because of inevitable gaps in such data. Using surveys has the drawback that people may not know or may forget what they spent, while insurance companies have little data on uncovered expenses. To overcome these problems, the federal government sponsored a large-scale social experiment, the Rand Health Insurance Experiment (HIE). The HIE is intended to be a definitive study of the effects of alternative financing arrangements on the use of services and on health status. It assigned families to 14 different insurance plans, balancing the plans in terms of preexperimental use and other characteristics known to affect use. Much effort has gone into obtaining complete and accurate data on the use of health services.

Data from the HIE are free from the inherent flaws of nonexperimental data. The initial analysis of the HIE data yielded estimates of plan-effects based on annual spending (Newhouse et al., 1981). These estimates are unbiased, but they waste information by aggregating across many decisions that participants make over the course of the year. In this report we use episodes of spending to analyze the experimental data. This is a powerful and fairly novel approach that better reflects the actual decisions to spend money on medical care than does analyzing annual expenses.

The methods and results presented here are preliminary. They are based on three years of experience in Dayton, Ohio--only about 15 percent of the final fee-for-service experience in the HIE. More refined methods based on more data will be used in subsequent analyses.

Still, the results are already interesting, and we hope that they will stimulate work on the hypotheses they suggest, and on refinements of the methods.

When people become sick, they must decide whether to get medical care. They consider the severity of the illness, the perceived value of treatment (this perception might be based on a phone call or previous contact with a physician), and the out-of-pocket and time-costs of going to a doctor. If the costs of treatment seem higher than the benefits, sick people may decide to give time and home remedies a chance to work. A recent survey of episodes of illness among government workers found that in three-eighths of the episodes with disability days, no medical contacts were made (Riedel et al., 1982). Spending on chronic, dental, or well-care episodes of treatment may not be triggered by an acute problem, but entails a similar decision. Once a patient has decided to see a doctor, that doctor helps to decide how much to spend on care for the duration of the episode. Thus, any episode of treatment contains two decisions of interest: The first is whether to seek care at all; the second, after conferring with the doctor, is to decide on the level of treatment.

Although episodes clearly reflect behavior more closely than do annual totals, expense data grouped into episodes have only rarely been available (Kilpatrick, 1977; Stoddart and Barer, 1980; Lohr, Brook, and Kaufman, 1980; Pederson and Christiansen, 1982). Hence, most previous economic analyses of demand for health care have been performed on annual expenses, which are aggregates of many such episodic decisions. Because the HIE collected extensive information on all filed claims for health spending, we were able to organize spending into episodes of the different basic types: dental, hospital, and outpatient acute, chronic, and well care. Consequently, we can study spending on each separately.

The rest of this chapter explains the potential advantages of such episodic analysis and the rationale for economic episodes. Chapter 2 describes the data, and Chap. 3 describes the assumptions and procedures used to group claims into episodes. Chapter 4 gives our analysis of the effects of price and other covariates on the cost per episode, and the number of episodes per year. Chapter 5 shows how occurrence rates change over the year. Finally, Chap. 6 discusses the consequences of these results for economic and health services research.

POTENTIAL ADVANTAGES OF EPISODIC ANALYSIS

Because all of the HIE experimental insurance plans have an annual accounting period, it is natural to analyze the participants' expenses on an annual basis. This analysis gives unbiased and reasonably precise results about annual spending for each plan in the experiment.

We collect much more detailed information from the participants, however. Indeed, we have reports for each claim they file. This information, organized into coherent spending episodes, has three major advantages over annual totals. Episodic analysis permits

- More precise estimates of demand,
- Better understanding of the effects of deductibles, and
- Estimation of transient spending changes following a price change.

PRECISION

Disaggregation into episodes increases the number of observations to analyze, since participants average about four episodes per year.[1] The price of care and other circumstances may change as decisions are made through the year. The detailed study of individual line charges (for specific services for which charges are filed) can be used to group expenses into chronic, acute, and well categories. Since decisions within categories may be made differently, separating them for analysis may permit better modeling. Since the patient is the one who starts the episode, the number of episodes may be more related to patient characteristics than is total spending, which can be influenced by the supplying physicians (Stoddart and Barer, 1980). In sum, the greater number of episodes and the closer relationship between their generation and behavioral circumstances should allow more precise estimates of demand.

---

[1] However, because individual propensities to have episodes persist over time (are correlated), there is less than four times as much information in the episodic data as in annual totals.

THE EFFECTS OF DEDUCTIBLES

By studying episodes, we can improve our ability to model and predict the effects of deductibles. A deductible is the amount a policyholder must pay before insurance takes over; deductibles are attractive to economists because they can preserve some consumer market incentives at a low cost in risk. Arrow has shown that when the price of care does not affect spending, the optimal insurance policy for risk-averse consumers has no coinsurance after a deductible (Arrow, 1963, 1973). We have modified his argument for the case where demand is price-elastic to compute illustrative optimal deductibles that balance the value of risk-reduction with the inefficiencies resulting from lower prices and administrative costs (Keeler, Newhouse, and Phelps, 1977).

The situations arising through the year are natural experiments for different sized deductibles. To keep the number of different experimental plans small, we limited the deductible plan types to one individual deductible at $150 and three family deductibles at the minimum of $1,000 or 5, 10, and 15 percent of income.[2] During the course of the year, individuals and families will find themselves with different remaining amounts of deductible. By analyzing their behavior at those points, we may be able to better extrapolate our results to insurance plans that were not in our design. Even plans with the same nominal deductibles may not be directly predictable from the annual results, since inflation or differences in the scope of coverage may greatly change the probability that such a deductible be exceeded.

The episode is the natural sales unit for analyzing the effects of price, since estimates of the effect of price on annual expenditures are biased when a deductible is present. This bias explains why annual analysis on the HIE has been performed on plan differences rather than price differences. Economic theory usually deals with a fixed price or a one-period model, but in the experiment, the coinsurance rate falls to zero after families spend a certain amount. There is, therefore, no straightforward definition of the marginal price for annual spending.

---

[2] Our deductible or maximum dollar expenditure (MDE) is the point at which the coinsurance rate shifts from positive to zero. The traditional definition of deductible is the point where the coinsurance rate shifts from 100 percent to some smaller percentage.

Three natural proxies for the true unobserved price are the coinsurance rate at the start of the year, the coinsurance rate at the end of the year, and the average coinsurance rate. None of these is a valid measure of marginal price. The starting price ignores the effects of the zero-price period on those who exceed the deductible, and therefore is higher than the true average price. The finishing price is the marginal price under perfect foreknowledge, but illness is not completely predictable. Thus, the finishing price ignores the initial-price period and therefore is lower than the true average price.

In addition, because participants whose out-of-pocket price dropped to zero during the year obtained that status by spending on health care, they are naturally prone to spend more than the others. Although the plans are balanced for propensity to go to the doctor, use of the within-year price, instead of plan, to group observations destroys that balance. This selection problem will be discussed in more detail when its solution is presented in Chap. 5.

The nominal average price (total spending divided by total out-of-pocket spending) is biased for a more subtle reason. Families may not know for sure that they will exceed a deductible over the course of the year, but as the year goes by, they will have some information on their chances. Everything spent before the deductible is exceeded has a bonus of reducing the deductible for the rest of the year. This bonus makes the true cost of care fall away from the nominal cost as the chance to exceed the deductible increases. Thus, before the deductible is exceeded, the real price is somewhat less than the nominal price. How much less depends on the amount of time left in the year, the distribution of expected family expenses, and the amount of deductible remaining (Keeler, Newhouse, and Phelps, 1977).

In particular, families close to the deductible in the early part of the year may anticipate they will probably go over, and may then start to spend freely.[3]

---

[3] This behavior should lead to gaps in the distribution of family expenses in the region of the deductible. Another reason for such gaps is that, late in the year, families with a bent price line must have extremely inelastic demand for medical care to want to spend an amount close to the bend. (Keeler, Newhouse, and Phelps, 1977.)

Scheduled future expenses, such as a delivery fee for a mother-to-be, create a stronger form of anticipation, which we handle by dating the expenses to the first time they would be known. In the weaker form, a father of three small children may not know when or for exactly what he will be paying the pediatrician, but knows from experience that he probably will do so.

CHANGES IN SPENDING OVER TIME

Spending dynamics when insurance plans change are important in resolving issues related to the phase-in of more complete insurance, and in choosing the scope of coverage. Since all plans have a limit on out-of-pocket spending, any family that exceeds that limit receives all medical care free for the rest of the year. To the extent that people can invest in their health, such a "sale" temporarily offers more incentive to spend than if they had free care forever, and it is interesting to see how much they try to take advantage of sales.

Catch-up demand occurs when people change permanently to new and fuller coverage. This factor can be significant in planning for the period when new insurance is being phased in. For example, in the first two years of the British National Health Service (NHS), the number of dentures supplied in England and Wales rose from 200 thousand in the year before NHS to 2.3 million and then 3.6 million, and the number of spectacles supplied rose from 1.5 million in the year before NHS to 6 million and then 10 million (Abel-Smith and Titmuss, 1956).

Understanding these dynamic changes is especially important to analysis of scope-of-coverage questions, since many of the medical services whose coverage is most debatable--eye glasses, hearing aids, psychiatric services, dentures, and elective surgery--may be strongly influenced by sales. Sales and catch-up are important in another way to the experiment: Those people with the most generous family health protection plans may feel that the experiment itself is a prolonged sale. Episodic analysis of the other plans can provide additional evidence bearing on transient demand at the beginning and end of the experiment.

Chapter 2

DESCRIPTION OF THE DATA

The work presented here on the demand for medical services is based on data from the Health Insurance Experiment. This chapter describes the study, focusing on the data used in our analysis. After a brief introduction, we discuss the sample studied here, the experimental insurance plans, and the other covariates used to explain demand.

THE HEALTH INSURANCE EXPERIMENT

How the extent of health insurance coverage affects the demand for medical services has been a key issue in the American debate over national health insurance. Past studies of this issue have used nonexperimental data that suffer from several flaws: Insurance is potentially endogenous; existing policies are difficult to describe parametrically; data on use are frequently based on recall and therefore are subject to reporting biases; and coinsurance rates and deductibles often vary little for a given service, such as a hospitalization.

The HIE is an effort to overcome these problems. In late 1974, the HIE enrolled 390 families living in Dayton, Ohio, and in 1976, enrolled 2370 additional families in five other sites (Seattle, Washington; Fitchburg, Massachusetts; Franklin County, Massachusetts; Charleston, South Carolina; Georgetown County, South Carolina). The families were given health insurance plans with differing coinsurance rates and deductibles; in return they assigned to the experiment the benefits of any nonexperimental plans for which they were eligible. If the assignment could make the family worse off, the family was given a lump-sum payment equal to the maximum it could lose from participating. Seventy percent of the families were enrolled for three years, the remainder for five years.

The experimental design included a novel method to ensure balance among the various treatments (Morris, 1979), and split samples to control for limited duration and the effects of physical examinations and illness-reporting requirements. These methods are described in

Newhouse (1974) and Brook et al. (1979). Since our work is concerned with estimation of demand for health care services, we will discuss only the sample and insurance plan treatments here.

EXPERIMENTAL INSURANCE PLANS

The families were assigned plans by a new method of randomization that maximized balance across the 14 different insurance plans (Morris, 1979). No choice of plan was offered; the family could either accept the experimental plan offered or decline to participate. About one-third of the sample received all services free (their coinsurance rate was zero), and nearly one-fifth paid 25 percent coinsurance subject to an upper limit on annual out-of-pocket family expenditures of 5, 10, or 15 percent of the previous year's income, or $1,000, whichever was less. In analysis, plans that differed only by their percent-of-income limit are usually grouped. This limit or deductible is called the Maximum Dollar Expenditure (MDE). Just under one-twelfth of the sampled families had a 50 percent coinsurance rate, also subject to the MDE (in Dayton this plan accounted for about one-fifth of the sample). One-fifth had a 95 percent coinsurance rate, subject to the MDE. In effect, this last group of families had an income-related family deductible. Finally, around one-fifth of the families were enrolled in a 95 percent coinsurance plan that limited annual out-of-pocket outpatient expenditures to $150 per person ($450 per family), providing in effect an individual outpatient deductible.[1] In this plan the cost-sharing applied only to ambulatory services; inpatient services were free.

All plans covered a wide variety of services. Medical expenses included services provided by nonphysicians such as chiropractors and optometrists, and prescription drugs and supplies. The only significant exclusions were outpatient mental health services in excess of 52 visits per year, nonpreventive orthodontia, and cosmetic surgery unrelated to accidents occurring after the start of the experiment. Dental services and outpatient mental health services were, however, treated somewhat differently in the first year in Dayton.[2] The same coinsurance rate

---

[1] The coinsurance rate for the family and individual deductible plans was 100 percent in Dayton Year 1. The rate was changed to 95

applied to all medical services, except for the individual outpatient deductible plan.[3] Claims filed by participants, including those for unreimbursed expenses, provide data on the amount and type of expenses.

THE SAMPLE

The sample was a stratified random sample from each site, but the following groups were not eligible: (1) those 62 years of age and over; (2) those with incomes in excess of $25,000 (1973 dollars), which excluded approximately the upper seven percent of the income distribution; (3) those eligible for the disability Medicare program; (4) those in the military or in institutions such as jails; (5) veterans with service-connected disabilities. Over 80 percent of the population was eligible. In Dayton, the lower third of the income distribution was (intentionally) oversampled and the middle third undersampled.

The work reported here is based on the first three years' experience in Dayton. It includes enrollees who participated for a full year in any of the first three years of Dayton, plus those who died or were born during the three-year period. Excluded from this analysis are other individuals with partial years of participation: participants who were suspended (e.g., for joining the military), who voluntarily quit, and who were involuntarily terminated for noncompliance during the year.[4] But a person who left in Year 2 was included in Year 1 if he participated for all of Year 1. Moreover, the hospital expenditures on newborns were entirely allocated to the mother for the purpose of this

---

percent to increase the incentive to file in all other site-years, although there was no statistical evidence of underfiling.

[2] Dental services for adults were covered only in the plan with a zero coinsurance rate; expenditures on outpatient mental health services did not count toward satisfying the MDE. After Year 1 in Dayton and in all other sites, dental services for adults and outpatient mental health services (up to 52 visits per person annually) were covered like any other service in all plans.

[3] In other sites not analyzed here, some plans had a 50 percent coinsurance rate for dental and outpatient mental health services and a 25 percent rate for all other services.

[4] We also excluded four cases who participated in the disability Medicare program. They were enrolled because the HIE expected to receive a waiver from the Social Security Administration (SSA) permitting participation. When SSA denied the waiver, individuals eligible for such benefits became ineligible for the HIE.

analysis. The exclusions comprise about five percent of the total sample.

COVARIATES

Insurance plans are grouped into five groups and coded with indicator variables, each for: a family medical coinsurance rate of 25 percent (P25); a family medical coinsurance rate of 50 percent (P50); a family medical coinsurance rate of 95 percent (behaviorally the same as a family deductible) (PFD); the individual deductible of $150 per person or $450 per family for outpatient care (ID); and the free-care plan (FREE).

Three other coded experimental treatment variables were described above, namely: whether a household was given a preenrollment screening examination (EXAM) or not; whether the family reported disability-days weekly (WEEKLY) rather than biweekly in the first year of Dayton; and whether the family was enrolled for 3 (YR3) rather than 5 years.

The remaining variables measure other factors. Since the analysis of annual demand for medical care had already been completed, we used the set of variables found useful in predicting annual spending for our episodic analysis. These include preexperimental income, education of head of household, family size, preexperimental contact with the medical system, race, age, sex, self-reported health, pain, and worry about health. Preenrollment interviews provided data on all of these characteristics.

Chapter 3

GROUPING CLAIMS INTO EPISODES

The Episode Processing System (EPS) transforms the different claims (called MERs, short for Medical Expense Reports) into dated and typed episodes. The four types of MERs reflect the four types of services for which subjects can make a claim: (1) hospital, (2) physicians and supplies, (3) pharmacy, and (4) dental. A physician MER is shown in Fig. 3.1. As can be seen, each MER may contain charges for several different services. Each of these charges is referred to as a "line charge." Note that providers are asked to date, describe, and link each charge to a diagnosis and treatment history. The episodic programs use these line charges as the basic building blocks of the episodes. Below we discuss how episodes are linked, dated, and categorized.

LINKING

Line charges are linked according to the probable units of decision, rather than by the location or type of services provided. The first move in an episode of spending is up to the patient, although physicians can encourage that move in previous contacts, or over the phone. The total cost of the episode then is decided by the patient and physician jointly, although an unusual course of illness can also have an effect.

Our grouping is designed to illuminate these two decisions. For example, all expenses surrounding a hospital episode--office visits leading to an admission, and visits and drugs following the admission--are considered to be part of one hospital episode. Similarly, restorations following a dental exam are included with the exam in one episode of spending. Tests and drugs are linked to the visit in which they were ordered. A later discussion of results assumes in effect that at the initial visit, the participant is told of the need for additional care, and the approximate total cost of the episode is negotiated.

FAMILY HEALTH PROTECTION PLAN

PHYSICIANS, DOCTORS, SUPPLIERS AND OUTPATIENT MEDICAL EXPENSE REPORT

(Use this form for *all* outpatient charges, clinics, surgery, emergency, etc.)

MAIL TO: FAMILY HEALTH PROTECTION PLAN, P.O. BOX 2076, Oakland, CA. 94604

PART 1 PARTICIPANT TO FILL IN ITEMS 1 THROUGH 14 PLEASE PRINT OR TYPE

1. Last Name of Patient	First	M.I.	2. Sex	3. Age	4. Patient's Family No.

5. Patient's Address	City, State, Zip Code	6. Patient's Individual No.

7. What Was The Major Reason or Symptom For This Visit To The Doctor?	8. Was Illness or Injury Employment Related? YES ☐ NO ☐	9. Was Illness or Injury Accident Related? YES ☐ NO ☐	10. Date of Injury or Accident / /	11. Describe how and where accident occurred
7A. Date You First Noticed This Symptom: (For Illness or Accident)	12. Name of Doctor, Supplier or Outpatient Facility		13. Has the Patient Ever Visited This Doctor, Supplier or Outpatient Facility Before? YES ☐ NO ☐	

14. I authorize any holder of medical or other information about the patient to release to the Family Health Protection Plan or its intermediaries any information needed for this or related medical reports. I permit a copy of this authorization to be used in place of the original. In conformance with the Family Health Protection Plan Enrollment Agreement, all health care benefits covering the Patient are hereby assigned to the Family Health Protection Plan.

SIGN HERE ▶ Signature of Adult Participant or Guardian of Minor Participant	Print Adult's Name	Date Signed

PART 2 DOCTOR OR SUPPLIER TO FILL IN ITEMS 15 THROUGH 29 PLEASE PRINT OR TYPE

15. Full Name of Referring Doctor. IF NONE, WRITE NONE.	16. Full name(s) of Providers to Whom You Referred Patient for Consultation, Lab Tests, or Other Services. IF NONE, WRITE NONE.

17. Describe the Primary Problem or Diagnosis Which Brought the Patient To Your Office and Any Other Problem(s) for Which You Supplied Treatment. Please List Primary Problem or Diagnosis on Line A.	18. Type of Problem (check one)		19. Treatment History (omit if well care or pregnancy)
A.	☐ Acute ☐ Well Care (or pregnancy)	☐ Flare-up of Chronic ☐ Chronic (not flare-up)	☐ Initial Visit for this episode ☐ Repeat Visit for this episode
B.	☐ Acute ☐ Well Care (or pregnancy)	☐ Flare-up of Chronic ☐ Chronic (not flare-up)	☐ Initial Visit for this episode ☐ Repeat Visit for this episode
C.	☐ Acute ☐ Well Care (or pregnancy)	☐ Flare-up of Chronic ☐ Chronic (not flare-up)	☐ Initial Visit for this episode ☐ Repeat Visit for this episode
D.	☐ Acute ☐ Well Care (or pregnancy)	☐ Flare-up of Chronic ☐ Chronic (not flare-up)	☐ Initial Visit for this episode ☐ Repeat Visit for this episode

KEY **Place of Service Codes:** O = Doctor's Office. IL = Independent Laboratory. H = Patient's Home. IH Inpatient Hospital. NH = Nursing Home or SNF. ER = Emergency Area. OH = Outpatient Hospital, including Hospital Clinic and Outpatient Surgery. SC School Clinic. CC Company Clinic. OL = Other Location, Including Other Non-Hospital Clinic.
Type of Visit Codes: 1 Minimal Service. 2 = Brief Examination. 3 Limited Examination. 4 = Intermediate Examination. 5 Extended Examination. 6 = Comprehensive Examination. 7 Unusually Complex Examination. **SEE DETAILED INSTRUCTIONS ON REVERSE SIDE**
For **Inpatient Services**, Omit 18, 19 and 21

20	A Date Of Service	B Place of Service Use code above	C Describe Each Medical or Surgical Procedure and Other Service or Supplies Furnished For Each Date, Including Specific Lab Tests and the Specific Name of Any Drug Injected.	D Type of Office Visit Use code above	E Relate Treatment to Problem by Ref. to 17 A, B, C or D above	F Charge
1						
2						
3						
4						
5						

21. Were Any Drugs Prescribed? Were any Supplies Prescribed or Suggested? ☐ Yes ☐ No

A. If yes, specify drug(s) and/or supply(ies).	B. Relate to Problem by Reference to 17 A, B, C or D above

22. Name and Address of Doctor or Supplier	23. Social Security or Provider Tax ID Number	24. TOTAL CHARGE	
	Telephone Number	25. AMOUNT PAID IF ANY	
		26. BALANCE DUE	

27. I hereby certify that the services and/or supplies listed above have been provided on the date(s) shown.

PROVIDER'S SIGNATURE ▶ Date Signed

28. I hereby authorize payment directly to the above-named provider of the benefits otherwise payable to me but not to exceed the charges shown. I understand that I am financially responsible for any charges not covered by the Family Health Protection Plan.

ADULT PARTICIPANT'S SIGNATURE ▶ Date Signed

HBE #971 REV. 1-77

MAIL TO FHPP

Fig. 3.1 -- Physician MER

Details of the algorithms and methods used to process the claims are given in Fowler, Keeler, and Keesey (1981). The programs are long and complex because of the great variety of types of charges and claims, and because of the complex rules necessary to decide whether a new charge is or is not related to some set of previous charges.[1] The grouping is based primarily on the diagnosis, time since last charge, and Treatment History Code, which classifies diagnoses listed on the claim as acute (initial or repeat), chronic (routine or flare-up), well, or unobtainable. These codes are normally filled in by personnel of the billing physician, but are also filled in or corrected by clerks at Glen Slaughter and Associates according to guidelines supplied by Rand physicians.[2]

Analyzing spending by line charge would greatly increase the sample size; although the median episode has fewer than two charges, the mean number of charges per episode is around seven. The main problem with analyzing spending on the basis of line charges is the close connection between charges in the same episode. Treatment is often a package. Operations are often explicitly sold as a package, but even medical treatment of mild acute problems may involve a drug purchase and follow-up visit. The high correlation of charges (e.g., drug charges are much more likely in the days following an office visit than at other times) within an episode complicates analysis, and makes the apparent gains in precision from the larger number of charges an illusion. (Some hospitalizations have hundreds of charges, and these reflect behavior that is anything but a series of independent decisions.) Grouping spending into episodes avoids some arbitrariness in how follow-up charges, tests, drugs, and supplies are priced, because aggregating over the episode abstracts the problem away. Finally, aggregating to episode reduces the expense of later processing--A sevenfold increase in sample size is no boon to the computer budget!

---

[1] See Lohr, Brook, and Kaufman (1980) for a much simpler way of defining episodes.

[2] Glen Slaughter and Associates act as Rand's subcontractor in the administration of the insurance plans and in the collection of related data.

In theory, barring within-year price effects discussed later, the illnesses that might lead to acute episodes are supposed to be independent over time, and we assume that decisions for different chronic and well-care episodes will also be independently made. Thus, on the free plan, in steady state, we would expect acute and well-care episodes to be randomly distributed over time. The independence tests discussed in App. F were used as a check of claims grouping, and led to improvements in the grouping algorithms for drugs. A theoretical advantage of independence was that it simplifies statistical analysis of the results, and the simulation of the effects of different plans.

EPISODE DATES

From their responses to a questionnaire, we can assume that most participants understand their experimental insurance policies (Marquis, 1981). Most patients know their share of expenses as well as which medical services are covered. Since each plan has a limit on out-of-pocket spending, the price depends on the amount of medical services already purchased (except on the full coverage plan). The amount already purchased is provided by Glen Slaughter and Associates with each claim processed, and is thus available to participants even if they have not kept a record themselves.

The rational economic person, however, does not change his or her other decisions on medical spending on the date of receiving a bill, but on the date he or she[3] decided to incur the expense. Suppose a woman learns she is pregnant and will deliver well before the end of the accounting year. Suppose the hospital bill alone at that time will exceed her deductible. She may not wait to get her new glasses, and she may go more readily to the doctor for mild acute conditions. Since she knows the deductible eventually will be exceeded, she can act as if it already has been because she faces an effective price of zero.

Thus, for purposes of economic analysis, we try to date each charge to the earliest time that the participant would have committed himself to paying it. Continual spending for chronic conditions such as

---

[3] For expository convenience, we shall refer to the individual patient henceforth as "he," which we intend as a genderless pronoun.

diabetes or hypertension is presumably anticipated from the first visit of the accounting year; such episodes are dated to that time. Routine pregnancy costs are dated to the first prenatal visit. Drugs and follow-up visits for acute conditions are dated to the first office visit for that condition. Expenses carried over from an episode in a preceding accounting year are dated to the first of that year. Dental work is dated to the first in a series of procedures or the preceding examination. We distinguish elective surgery and well care from acute care by the criterion of deferrability, calling it acute if it loses most of its value if deferred.

We assume that the whole cost of the episode is known at the time of the first charge. This is not always true, but claims records do not reveal how much the doctor told the patient or, indeed, how much the doctor knew about the course of the episode at that time.

We could examine the sensitivity of this assumption by comparing the results with those under two extreme assumptions. At one extreme, the participants know at the beginning of each year whether or not they will exceed the deductible (i.e., we assume that all people who do exceed the deductible knew that they would from the beginning); at the other extreme, we could analyze charges by billing date, assuming that patients have no foreknowledge whatsoever.[4] We have not yet performed the latter test.

EPISODE TYPES

Deciding to get a physical examination is different from deciding to get treatment for a broken leg or for arthritis. Thus, to understand the effects of price on health care behavior, each type of decision should be analyzed separately. Without analyzing homogeneous categories of episodes separately, aggregating decisions that are affected differently by price and other covariates such as age will muddy the findings. In addition, there are reasons why different categories might be subject to different coinsurance rates in a national health insurance plan. First, there is the standard economic result that more price-

---

[4] Even this may be inadequate if participants do not know the charge until they receive a bill. In most cases, however, they should be able to make an adequate guess.

responsive categories should be subsidized at a lower rate (Baumol and Bradford, 1970). Second, the state, acting _in loco parentis_, may want to subsidize such expenditures as well care for small children. Finally, to reduce exposure to risk, it is sensible to insure rare catastrophes more fully than common smaller expenses.

Because of numerous differences from ambulatory medical episodes, we consider hospital episodes and dental episodes separately. We classify ambulatory medical episodes by how foreseen the problems were and how deferrable the treatment. We wish to group episodes that are similar with respect to the effect of price on the timing of purchases. A classification based instead on severity or necessity might show which kinds of episodes were being treated on the free plan but not on the other plans. Unfortunately, such distinctions are hard to make from the claims records. Severity is closely connected to the amount of treatment, but using this measure would confound the analysis of price effects. Necessity of treatment will be studied as part of future analysis of the quality of care on different plans. We will investigate the effects of price on the number and size of episodes for each common diagnosis. Even so, it will be difficult to separate severe and mild cases of such diseases as upper respiratory infections. Thus, we will neglect severity here and restrict our analysis to the standard medical categories: acute, chronic, and well care.[5]

Unforeseen and undeferrable treatment opportunities define acute episodes. From an economic point of view, spending on these episodes will only occur when the patient is temporarily sick.

Chronic "episodes" comprise foreseen and continuing expenses. Treatment for most chronic diseases is designed to ameliorate the consequences, rather than to cure, and people getting such conditions treated should be able to budget their routine expenditures from the start of the year. All the routine care for each chronic condition will be considered as one episode. With some conditions, there is a reasonably high chance of some complications during the year, just as

---

[5] We study episodes of treatment, realizing that not all illnesses get treated, even with free care. However, we presume that the main reason for fewer acute episodes on plans where care is expensive is that price is deterring people from treating illnesses that they might treat if care were free.

the owner of a decrepit car knows that sometime during the year his mechanic may well have to do something about it. However, we will distinguish these flare-ups as unforeseen, and treat them as we do acute spending. They are economically different from the drugs and checkup routines normally followed by those with hypertension or diabetes.

Finally, well-care episodes deal with conditions that are deferrable without great loss. Examples are gynecological examinations and immunizations. It seems inappropriate to call this category elective, because it implies a contrast with necessity, and care for minor acute conditions such as colds may be more elective than such well procedures as immunizations. The deferrability of the procedures is what distinguishes them economically, not their degree of necessity. Indeed, many of the well procedures may be medically deferrable, but not legally or socially. Examinations and immunizations for school, and pap smears for birth control pills are examples.

Chapter 4

COST AND ANNUAL FREQUENCY OF EPISODES

INTRODUCTION

In this chapter, we analyze the cost and annual frequency of episodes. Together, episode cost and frequency determine annual medical spending. That is,

$$\text{Annual Spending} = (\text{Number of Episodes}) \times (\text{Cost per Episode}). \quad (4.1)$$

We will describe how the cost and annual frequency of episodes vary by episode type, by insurance plan, and by the characteristics of individual participants. Using Eq. (4.1), we can then see how predictions of annual medical spending deduced from the episodic approach compare with the annual spending predictions obtained from data on annual medical spending by individual (Newhouse et al., 1981; Duan et al., 1982). This comparison will serve as a check on the validity of both methods.

Information about the distribution of the cost and frequency of episodes is needed for simulation models that use the episodic approach. In these models, individual streams of medical expenses can be created by first generating episode occurrences and then generating the cost of each episode. To the extent that episode size is independent of occurrence pattern and insurance plan, modeling how individuals behave is simplified. More generally, it simplifies the overall analysis when we can decouple the cost-of-episode analysis from the number-of-episodes analysis.

We deal first with the cost per episode. We tabulate how the average cost per episode varies by episode type and insurance plan. Only the average cost of dental episodes varies systematically across insurance plans. Next, we model an individual's cost per episode as a multiplicative function of his or her attributes. After adjusting for

individual characteristics, chronic episode costs also vary systematically with insurance plan.

The final section of this chapter reports our analysis of the number of episodes per year, and, in particular, how numbers vary with episode type, insurance plan, and participant characteristics. This analysis parallels the common approach to estimating annual spending but uses episodes rather than dollars. We first give descriptive statistics on how the annual number of episodes varies with plan. Episode occurrence rates differ most between full coverage and the other plans, with smaller differences between the low and high levels of coinsurance. Differences are larger for acute episodes than for the other episode types. Next, we analyze how well participant characteristics predict annual rates of episodes for each type. Most of our results are based on a negative binomial regression model of the episode counts. This model has the natural interpretation that individual propensities to have episodes vary more than participants' measured variables can predict. The major factors associated with varying episode rates are self-assessed health, age, sex, and physician visits in the year preceding the experiment. Income, plan, and having a regular physician are also important, but less so.

We defer discussion of how individuals' behavior varies within the year, to Chap. 5. There we describe the effects that within-year price-changes (caused by a participant's expenses reaching the MDE before the end of the year) seem to have on medical spending.

COST OF EPISODES

As described earlier, we have divided episodes into five types. There are at least two reasons for analyzing these types separately. First, episode categories may be affected differently by price changes. Second, to the extent that episode types have markedly different distributions of cost and frequency, analytical modeling can be simplified by treating them separately.

The first episode type is inpatient episodes or hospitalizations, which are much rarer and costlier than other episodes. Dental episodes, our second type, are distinguished from medical outpatient episodes. The three remaining types, acute, chronic, and well-care outpatient

episodes, are more alike. We analyze episode cost for each of the five types in varying levels of detail: the distribution of costs by episode type and by insurance plan, and finally, the cost per episode by type as a function of insurance plan and other participant attributes. We begin with the simplest data--episode costs in aggregate--and then adjust for other factors.

Some Descriptive Statistics on Cost

The overall distribution of episode cost varies with type of episode. No one family of distributions fits the cost of all episode types well, but the lognormal distribution is a good approximation except for dental episodes. The same conclusion holds when episodes are divided by insurance plan and the distributions are tabulated separately. For each of the five episode types (hospital, dental, acute, chronic, and well care), there is a small proportion of relatively large episodes that makes fitting distributions very assumption-dependent.

We now describe our analysis of the distribution of episode sizes for each episode type, ignoring differences across plans, site-years, and other covariates. This analysis consisted of various probability plots, computation of empirical moments and quantiles, and the fitting of several different distributional forms to the size distributions of the five episode types.[1]

---

[1] Besides fitting the lognormal distribution and the gamma distribution, we used Hinkley's quantile method (Hinkley, 1975) to estimate a power transformation from the Box-Cox family (Box and Cox, 1964), which makes the distribution of each episode type approximately symmetric. The fitted values of the power parameter λ are nonzero (not lognormal) for all episode types. Nonetheless, the lognormal distribution was chosen because it yielded a fit that was practically indistinguishable from that of the best-fitting Box-Cox. While this is the first attempt to model the cost per episode, there has been considerable work in modeling annual medical expenses. Catastrophic episodes and hospitalizations make annual expenses skewed strongly to the right. Nonzero annual expenses for Health Insurance Experiment participants follow approximately a lognormal distribution (see Manning et al., 1981; Duan et al., 1982). Other workers studying annual expenses have found that the lognormal or gamma distributions fit their data well (Friedman, 1974; Keeler, Morrow, and Newhouse, 1977). Newhouse et al. (1980) use Box-Cox power transformations to model annual expenses.

Table 4.1 gives descriptive statistics on the distribution of cost per episode for the five episode types. While the table shows the second year of Dayton, summary statistics for the other years of Dayton, Seattle, and Fitchburg are similar. The five episode types differ considerably both in frequency and cost per episode. Hospitalizations are the most costly, averaging about $1,800 with the median episode costing about $1,300. Other episode types are more than an order of magnitude smaller, with acute episodes being the least expensive on average (mean = $33, median = $16), followed closely by well-care episodes (mean = $36, median = $21). A comparison of means and medians, as well as an examination of the quantiles, reveals that all episode types are skewed to the right. The bottom four rows of the table give the fitted lognormal distribution parameters for the five episode types. The much smaller values of skewness and kurtosis indicate that the distribution of episode costs is roughly lognormal.

Plan Differences

We turn now to analyzing how episode cost varies with insurance plan--the primary goal of this section. For each of the five episode types, we looked at whether the cost per episode is the same across the five insurance plans. The results were consistent: There appear to be small plan differences for dental episodes but not for any other episode types for the three years of Dayton data. Tables 4.2 and 4.3 give comparisons by plan of the cost per episode and the logarithm of cost per episode, respectively, for all five types of episodes. The average dollar cost per episode data in Table 4.2 are more easily comprehended, but the average of the logarithm of episode cost in Table 4.3 is given because the skewed nature of the cost distributions makes comparisons of logarithms statistically more stable. We discuss both tables below.

There is not very much variation in the cost of hospitalizations across plan. Differences by plan in both average hospitalization cost and average log hospitalization costs do not seem to fall in a meaningful pattern. Because of the presence of a few very large hospitalizations, statistical tests were carried out using both the logarithm of hospitalization costs given in Table 4.3 and the relative

Table 4.1

SUMMARY OF COST PER EPISODE DISTRIBUTION BY TYPE,
DAYTON YEAR 2, ALL PLANS

(In $ rounded)

Item	Hospital	Outpatient			
		Dental	Acute	Chronic	Well Care
Number of Episodes	118	1,275	2,355	990	801
Quantiles					
100 percent (maximum)	12,000	2,900	758	1,300	750
90	3,400	200	66	160	82
75	2,200	54	32	62	44
50 (median)	1,300	23	16	25	21
25	700	15	9	10	15
10	360	10	4	5	7
0 (minimum)	15	2	1	1	1
Moments					
Mean	1,800	89	33	65	36
Standard deviation	1,900	227	65	123	49
Skewness[a]	3.0	6.2	6.3	5	7
Kurtosis[a]	11.5	51	50	42	88
Lognormal parameters					
Mean	7.1	3.5	2.8	3.3	3.2
Standard deviation	1.0	1.2	1.1	1.3	0.9
Skewness[a]	-0.9	1.1	0.4	0.3	0.1
Kurtosis[a]	3.1	1.2	0.5	-0.3	0.5

[a] The skewness and kurtosis are based on the third and fourth moments of the empirical distribution. They can be regarded as measures of how nonnormal the distribution is. Values of zero correspond to normality. See Kendall and Stuart (1961), Vol. 1, Chap. 3, for details.

Table 4.2

COST PER EPISODE BY TYPE AND BY PLAN (ROUNDED)
(DAYTON: YEARS 1, 2, 3)

Episode Type	Plan[a]				
	Free	ID	25	50	95
Hospital					
Mean	1505	2161	2023	2366	1530
(Standard deviation)	(1445)	(3592)	(1899)	(6615)	(1477)
[Number]	[108]	[31]	[78]	[40]	[83]
Dental[b]					
Mean	86	50	90	68	58
(Standard deviation)	(291)	(82)	(195)	(182)	(156)
[Number]	[837]	[178]	[576]	[364]	[616]
Acute[c]					
Mean	30	29	32	27	33
(Standard deviation)	(32)	(52)	(35)	(30)	(68)
[Number]	[885]	[164]	[510]	[310]	[486]
Chronic					
Mean	62	60	73	45	55
(Standard deviation)	(108)	(102)	(173)	(76)	(109)
[Number]	[1005]	[251]	[707]	[355]	[617]
Well care					
Mean	36	39	33	34	35
(Standard deviation)	(38)	(63)	(32)	(46)	(39)
[Number]	[818]	[189]	[525]	[425]	[492]

NOTE: The top entry in each category is the mean episode cost in dollars. The second entry is the standard deviation while the bottom entry is the number of episodes in this cell.

[a] The 11 plans studied here are grouped into the five plan categories. Free and individual deductible (ID) plans are their own category. The three coinsurance categories (25, 50, 95) each consist of plans with that coinsurance rate and one of the three MDE's (5, 10, 15 percent of income).

[b] Dental episode data are for years 2 and 3 only. The year 1 statistics are difficult to interpret since only the Free Plan covered adults, and there appears to be underreporting by adults on the pay plans.

[c] Acute cost per episode statistics are for Dayton year 2 since this sample size is sufficient to get reliable estimates. The other years are similar.

Table 4.3

LOG COST PER EPISODE BY TYPE AND BY PLAN (ROUNDED)
(DAYTON: YEARS 1, 2, 3)

	Plan[a]				
Episode Type	Free	ID	25	50	95
Hospital					
Mean	6.87	7.07	7.15	6.99	6.90
(Standard deviation)	(0.91)	(0.85)	(0.83)	(0.89)	(0.81)
[Number]	[108]	[31]	[78]	[40]	[83]
Dental[b]					
Mean	3.44	3.26	3.55	3.31	3.19
(Standard deviation)	(1.14)	(1.00)	(1.19)	(1.07)	(1.09)
[Number]	[837]	[178]	[576]	[364]	[616]
Acute[c]					
Mean	2.77	2.70	2.76	2.63	2.76
(Standard deviation)	(1.03)	(1.11)	(1.07)	(1.08)	(1.13)
[Number]	[885]	[164]	[510]	[310]	[486]
Chronic					
Mean	3.24	3.23	3.35	3.04	3.09
(Standard deviation)	(1.31)	(0.28)	(1.29)	(1.17)	(1.33)
[Number]	[1005]	[251]	[707]	[355]	[617]
Well care					
Mean	3.13	3.20	3.13	3.10	3.13
(Standard deviation)	(0.97)	(0.89)	(0.90)	(0.91)	(0.90)
[Number]	[818]	[189]	[525]	[425]	[492]

NOTE: The top entry in each category is the mean log episode cost. The second entry is the standard deviation while the bottom entry is the number of episodes in this cell.

[a] The 11 plans studied here are grouped into the five plan categories. Free and individual deductible (ID) plans are their own category. The three coinsurance categories (25, 50, 95) each consist of plans with that coinsurance rate and one of the three MDE's (5, 10, 15 percent of income).

[b] Dental episode data are for years 2 and 3 only. The year 1 statistics are difficult to interpret since only the Free Plan covered adults, and there appears to be underreporting by adults on the pay plans.

[c] Acute episode size statistics are for Dayton year 2 since this sample size is sufficient to get reliable estimates. The other years are similar.

ranking of average per-plan costs of hospitalization. We compared all five plans separately and also grouped the three "pay plans" together and the two "free plans" together and compared. (Recall that the ID plan provides for free hospitalizations.) None of the statistical tests for differences in cost per hospital episode across plans yielded values that were significant at the conventional 5 percent level but sample sizes are relatively small.[2] The similarity in costs implies that we can model size of hospital episodes as being independent of plan.

From Table 4.1 it was clear that there are substantial differences among the episode cost distributions of the four nonhospital episode types. Tables 4.2 and 4.3 show how mean episode cost and log cost vary by plan and type of episode. Unlike hospitalizations, many people have more than one episode of any given outpatient type per year. But like hospital episodes, the distribution of outpatient episode costs has some outliers. On an absolute scale, the costs of these large outpatient episodes are considerably smaller than those for hospitalizations, but their costs relative to the average cost for that episode type are still considerable. When the percentage of outpatient episodes above $500 are cross-tabulated by plan and episode type, the episodes costing more than $500 are scattered across the different insurance plans.

For dental episodes, Table 4.2 and Table 4.3 show the cost per episode falling with coinsurance. The free plan has the costliest episodes, with the 25, 50, and 95 plan costs declining in order. Because the ID plan has cost-sharing for outpatient episodes, costs on this plan should be similar to those on the 95 plan--if there is indeed a price effect. For dental episodes, the mean ID plan log cost of 3.26 falls between 50 and 95 plan means.

---

[2] We computed both the parametric F-test and the nonparametric Kruskal-Wallace Rank Test to test for differences across the five plans; the hospital episode data are consistent with no plan differences at the 5 percent significance level. When we contrasted the pay plans (25, 50, 95) against the free plans (free, ID), the difference again was not statistically significant at the 5 percent level. These tests do not account for correlations across episodes for the same individuals. This is taken into account in the next section, where other covariates are adjusted for.

For the other outpatient episodes, the patterns are less clear. However, statistical tests for plan differences for each of the episode types reveal no statistically significant difference except for dental episodes.[3] Thus we conclude that, at least when looking at plan averages, we cannot detect any differences in episode costs across insurance plan for any episode types except for dental. We next explore how this conclusion holds up when we explicitly adjust for other factors.

Plan Effects Controlling for Other Factors

Since the groups of health insurance experiment participants assigned to each insurance plan are approximately balanced with respect to their relevant measured characteristics (Newhouse, 1974; Morris, 1979), the plan differences in episode costs reported earlier should in principle be approximately correct, whether or not one controls for other factors. Still, there are at least three reasons for carrying out a more detailed analysis.

First, one use of our models will be to simulate streams of medical episodes. If the cost per episode varies systematically with participant characteristics, it is important to include this variation in any simulation model. If there is appreciable variation in the cost per episode across participants, modeling this variation will help us to estimate distributional effects of national health insurance plans more accurately. Second, to the extent that we can model how costs depend on individual characteristics, that model can be used to estimate the particular group's portion of the total costs of a National Health Insurance (NHI) Plan. This information might well be an important contribution to the political debate on NHI. Similarly, in the private sector, such a model might help in devising more competitive health insurance premiums. A third reason for including covariates is accuracy of estimation. That is, even though plan groups may be balanced, controlling on an individual basis for the effect of covariates will

---

[3] We computed the same parametric and nonparametric tests for outpatient episode types as for hospitalizations. See previous footnote (2) for details and caveats. The dental episode costs were statistically significant for all tests at the 0.01 level.

lead to more accurate estimates of the plan effects, if any, on the cost per episode. (See Duan et al. (1982) for a discussion of this issue in the context of modeling annual medical expenses.)

Our primary tool for modeling the effects of participant attributes will be regression models. Considerable work has been done in modeling annual expenses as a function of various predictors (see Manning et al., 1981). Guided by this work, we fit similar regression models to the log of episode cost. The conclusions of this analysis are with one exception consistent with the previous section. The new finding is the existence of statistically significant plan differences for chronic episode costs. Except for dental and chronic episodes, there are no detectable plan differences. There are some small statistically significant coefficient estimates associated with individual participant attributes.

Table 4.4 gives the means and standard deviations of the potential predictors. For ease of interpretation our exploratory analysis elected to use the logarithm of episode cost as the dependent variable. Table 4.5 gives the results of our regression analysis. In this table we see the logarithm of episode cost being fit by a number of predictors for each of the five episode types. We will not describe in detail the various steps that were carried out in arriving at these regression equations. A discussion and interpretation of the fitted models is given below.

In addition to computing various diagnostics and residual plots,[4] we also explored the possibility that the residuals of the log episode cost of a given individual might be correlated across episodes. The question was, even after controlling for the measured attributes of individuals (including insurance plan), does the size of one episode help in predicting the size of that individual's subsequent episodes? For the Dayton data, the answer is that there is a positive correlation among residuals from the same individual for some episode types. The

[4] A goodness of link test (Pregibon, 1980) was computed to check the appropriateness of the logarithm specification of the dependent variable. Because of the extreme values noted earlier, various influence functions were computed to see whether some observations were very influential (Cook, 1977; Belsley, Kuh, and Welsch, 1980). Neither check revealed any problems.

Table 4.4

COVARIATES USED IN REGRESSION

Name of Variable	Mean[a]	Standard Deviation	Comments
"FREE"	0.28	(0.4)	The Omitted Insurance Plan variable; full coverage
ID	0.09	(0.3)	$150 Individual Deductible Plan
P25	0.22	(0.4)	25 percent Coinsurance Plan
P50	0.16	(0.4)	50 percent Coinsurance Plan
P95	0.25	(0.4)	95 percent Coinsurance Plan (Family Deductible Plan)
NEWBORN	0.02	(0.1)	1 for those born during current accounting year, 0 otherwise
SQRAGE	4.67	(1.9)	The square root of age at start of accounting year
HPW SUM	4.71	(1.6)	The sum of 4 point scales for Health, Pain, and Worry, added so that Excellent Health, No Pain and No Worry = 3, and Poor Health, Great Pain and Great Worry = 12
LMDVIS	1.1	(1.0)	The logarithm of MD visits in year preceding experiment (set at 0 if there were no visits)
LINC	9.3	(0.6)	Log Income in year preceding experiment
LFAM	1.3	(0.5)	Log Family Size
EXAM	0.51	(0.5)	Took exam at start of experiment
MAX-ED	13	(3)	Maximum of years of education of female and male heads of family
WOMAN	0.33	(0.5)	1 if female over 17 years old at start of accounting year
NOMDVIS	0.19	(0.4)	1 if no MDVIS in year preceding experiment
NO DDS	0.16	(0.4)	1 if no regular dentist at time of enrollment
NOPROV	0.19	(0.5)	The sum of "no MD" and "no DDS" at enrollment
BLACK	0.09	(0.3)	1 if head of family is black, 0 otherwise
YR 3	0.48	(0.5)	1 if in experiment for 3 years, 0 if in for 5

[a]Values in year 1 of 1074 individuals who were in Dayton sample all three years.

Table 4.5

REGRESSION EQUATIONS FOR PREDICTING LOG EPISODE COST (ROUNDED)

Episode Type Variable	Hospital Coeff.(t)	Acute Coeff.(t)	Chronic Coeff.(t)	Well Care Coeff.(t)	Dental coeff.(t)
CONSTANT	7.0 (8.5)	2.1 (5.0)	1.3 (2.6)	2.3 (6.2)	2.3 (4.0)
ID	0.006 (0.0)	-0.050 (-0.5)	-0.085 (-0.9)	0.031 (0.4)	-0.15 (-1.4)
P25	0.15 (1.2)	0.017 (0.3)	0.034 (0.5)	-0.008 (-0.1)	0.10 (1.4)
P50	0.16 (1.0)	-0.039 (-0.5)	-0.18 (-2.2)	0.027 (0.5)	-0.085 (-1.0)
P95	-0.087 (-0.7)	-0.01 (-0.1)	-0.21 (-3.1)	-0.020 (-0.4)	-0.20 (-2.8)
NEWBORN	-0.50 (-1.3)	-0.17 (-0.7)	0.073 (0.2)	0.20 (1.7)	Not Applicable
AGE	0.004 (0.3)	-0.012 (-1.8)	-0.082 (-1.0)	-0.010 (-2.0)	-0.039 (-3.8)
SQRAGE	0.077 (0.6)	0.19 (3.2)	0.21 (2.5)	0.27 (5.5)	0.52 (5.0)
HPWSUM	-0.029 (-2.0)	0.005 (0.4)	-0.012 (-0.9)	-0.027 (-2.2)	-0.053 (-2.9)
LMDVIS	0.094 (1.7)	-0.033 (-1.2)	0.067 (2.4)	-0.012 (-0.5)	-0.044 (-1.3)
LINC	-0.044 (-0.5)	-0.007 (-0.2)	0.096 (1.9)	0.018 (0.5)	-0.020 (-0.4)
LFAM	0.050 (0.5)	0.11 (2.2)	0.045 (0.8)	-0.041 (-0.9)	-0.13 (-2.1)
EXAM	-0.14 (-1.4)	-0.055 (-1.2)	-0.098 (-2.0)	-0.008 (-0.2)	0.11 (2.1)
MAXED	-0.029 (-1.6)	-0.013 (-1.5)	0.012 (1.4)	-0.006 (-0.8)	-0.028 (-2.7)
WOMAN	0.087 (0.8)	-0.049 (-0.9)	-0.21 (-3.9)	-0.20 (-3.9)	-0.002 (-0.0)
NOMDVIS	-0.028 (-0.2)	0.17 (2.1)	0.094 (1.0)	0.051 (0.7)	0.020 (0.2)
NODDS	0.010 (0.1)	0.087 (1.3)	-0.13 (-1.7)	0.017 (0.3)	0.031 (0.3)
BLACK	0.18 (1.3)	0.18 (2.1)	0.087 (1.0)	0.023 (0.3)	0.57 (5.0)
YR 3	-0.029 (-0.3)	0.038 (0.8)	0.037 (0.8)	0.034 (0.9)	0.005 (0.1)
SAMPLE SIZE	340	2355	2935	2449	2571
ESTIMATED STANDARD DEVIATION	0.81	1.06	1.27	0.96	1.33
R^2	0.17	0.03	0.04	0.11	.07
OVERALL F-TEST (D.F.)	3.8 (18,321)	4.1 (18,2336)	7.2 (18,2916)	16.3 (18,2430)	10.9 (17,2553)
PARTIAL F-TEST FOR PLAN (D.F.)	1.1 (4,321)	0.2 (4,2336)	4.2 (4,2916)	0.19 (4,2430)	4.4 (4,2553)
INTRAPERSON CORRELATION	N.A.	0	0	0.08	0.18

size of this correlation varies with episode type, but is large enough for some types to merit using a variance components or random effects model with an individual-specific error component.[5]

The first column of Table 4.5 gives the fitted regression equation to the logarithm of the cost of hospital episodes. The squared multiple correlation (R^2) is 0.17, there are no statistically significant plan effects, and a few of the individual attributes are only marginally statistically significant: whether the individual is a newborn baby, HPWSUM (a health status measure), the logarithm of MD visits, whether the person had an enrollment exam, the person's education and race. We conclude that the data are consistent with little or no consistent plan effects on the cost per hospitalization.

Plan and individual characteristics have little effect on the cost of acute episodes. While the F-test is statistically significant ($p < 0.001$), the proportion of variance explained is only 3 percent. Plan differences and between-episode correlations are not statistically significant; a few individual attributes have small but statistically significant coefficient estimates.

Predicting the cost of chronic episodes yields a similar conclusion; the proportion of variance explained is small (0.04) and the between-episode correlation appears to be zero. However, the partial F-test for plan effects is statistically significant and the estimated plan coefficients, while small, are in plausible order. The cost of a chronic episode for a participant on the stingiest plan (95) is estimated to be 81 percent of that for a similar participant on the free plan.[6] The differences that exist appear to be between the free plan and the 50 and 95 plans. Several nonplan covariates make statistically significant contributions, and in particular, women have smaller chronic episodes than men.[7] In sum, insurance plan exhibits a small but

---

[5] See Manning et al. (1981), Chap. IIIc, for a discussion of both theoretical and computational aspects of this type of model for intrafamily correlations in annual expenses from the Health Insurance Experiment. The same algorithm is used here for estimation.

[6] Most chronic episodes are dated January 1st and consist of the entire year's treatment of a chronic condition. Thus, the cost per episode is similar to annual spending on the condition and so should exhibit a plan-effect analogous to that found by Newhouse et al., 1981, on annual spending.

[7] One can speculate about reasons for this effect if it is real:

statistically significant association with the cost of chronic episodes.

The regression equation does a better job of predicting the cost of <u>well-care</u> episodes than it does for the other outpatient episodes, in the sense of explaining more of the variance of the log cost (R^2 = 0.11). Interestingly, there are no apparent plan effects on episode cost. The coefficients for age and the square root of age show that cost is an increasing but concave function of age. Women have smaller well-care episodes than men, all else equal.

Finally, while the fitted regression equation for <u>dental</u> episode costs is not a strong predictor (R^2 = 0.06), the estimated plan effects are moderate, with a moderate-size intraperson correlation (ρ = 0.18). Stingy (95) plan episode costs are estimated to be 82 percent of free-plan dental episode costs. Like well-care medical episodes, dental episodes have a statistically significant increasing age-effect function. Also of note is that blacks and participants with less well educated household heads have larger dental episodes. Surprisingly, dental episodes are an estimated 77 percent more costly for blacks than for others.[8] In our future work, we plan to look at the prevalence of different dental procedures, by plan, to explain this difference.

These fitted regression equations for the cost of episodes yield some interesting conclusions. First, only dental and chronic episodes exhibit statistically significant plan effects on the cost per episode, and even here the estimated effects are not large. For each episode type, there are statistically significant estimated coefficients associated with some participant characteristics. These appear to be strongest for hospitalizations, dental episodes, and well-care episodes. Most important, for practical purposes the cost per episode can be treated as independent of, or only weakly related to, insurance plan. It is interesting that the apparent plan effects for the cost per chronic episode only emerged after individual attributes were controlled for.

Do women go to the doctor more often? Are their illnesses less severe?

[8] We will see later in this chapter that blacks also have <u>fewer</u> dental episodes. These countervailing differences partially offset one another in terms of total dental spending.

Refinements and Conclusions About Episode Costs

In the fitted regression models reported on above, some dependencies in the data were ignored. We have already discussed intrafamily and interepisode dependencies. It may also be worth noting two issues that we do not address in this study and defer to future reports. First, there is some evidence that episode cost and episode frequency may be related, perhaps because of MDE effects. We do not use episode frequency to predict episode cost here. Second, costs per episode may be affected by whether people have exceeded their MDE. Since costs per episode are similar on all plans, it is unlikely that this MDE effect would be large, and the statistical problems involved are great. We therefore defer any analysis of MDE effect on episode size to future work. Chapter 5 describes a model of MDE effect on episode frequency.

To summarize, we estimate plan effects controlling for other factors, by fitting regression equations to the logarithm of episode cost using "dummy" variables for insurance plan, and using other independent variables to capture participant characteristics. From these regressions, we conclude that there is no evidence of large plan effects for any episode types. For dental and chronic episodes there are small plan differences in episode cost. All other things equal, we estimate that participants on stingy (95) plans have dental episodes and chronic episodes that are about 20 percent smaller on average than they would be on the free plan.

ANNUAL EPISODE FREQUENCIES

In the previous section, we showed that except for dental and chronic episodes, the plan effects on the size of episodes are negligible. Plan-related differences in demand for medical care, then, are almost fully captured by differences in the number of episodes. We investigate how price and other independent variables affect those numbers. This section considers the problem of explaining annual rates of episode occurrence, ignoring what happens as the price changes during the year. Our analysis here parallels the common approach to estimating annual spending, but uses episode counts rather than dollars. Chapter 5

deals with the more difficult problem of estimating the effects of intrayear price changes.

We first give descriptive statistics on annual episode counts by type. We then cross-classify the counts by type and plan. Episode rates differ most between full coverage and the other plans, with less difference between low and high levels of coinsurance. Differences are larger for acute episodes then the other types. Next, we model how annual rates of episodes of various types are related to family characteristics. We first discuss the choice of regression models. Our results are based on a negative binomial regression model although we give a simple (ordinary least-squares) regression of the square root of episodes in App. A.

Episode Frequency by Type

There is great variation in the number of episodes per person. Table 4.6 shows some descriptive statistics for the 1074 participants in Dayton year 2. Although the average number was only 5, one person had 38 episodes that year. Acute outpatient episodes are the most common type. The different chronic "episodes" are the number of distinct chronic conditions under treatment. There are sizable portions of participants with no episode of any given type, and 11.8 percent of participants incurred no health care spending at all.

Episode Frequency by Type and Plan

We describe here how the average annual number of episodes varies with the insurance plan. Since, as we have seen, the cost per episode varies little with plan, the major differences in spending between plans result from the differences in the number of episodes.

Participants receiving free care have more episodes of every type than those on the other plans. Table 4.7 summarizes three years of experience in Dayton. Episode frequencies for the pay plans line up roughly in order of copayment,[9] but the biggest difference is between free care and all other plans. The standard errors of the difference

---

[9] Out-of-pocket payment for outpatient services on the individual deductible plan lie between those of the 25 and 50 percent coinsurance plans.

Table 4.6

SUMMARY OF ANNUAL EPISODE FREQUENCIES BY TYPE
(DAYTON, YEAR 2: 1074 PARTICIPANTS)

		Outpatient				
Item	Hospital	Dental	Acute	Chronic	Well Care	Total
Distribution percent						
0	91.7	41.1	32.5	58.0	52.2	11.8
1	6.5	22.3	18.8	19.9	32.2	10.0
2	1.6	22.5	13.9	11.1	11.9	10.9
3	0.2	8.6	11.6	4.6	2.0	10.8
4	0.0	3.3	7.9	3.0	1.1	10.1
5	0.0	1.3	6.4	1.0	0.3	7.8
6-10	0.0	0.9	4.9	2.2	0.2	26.9
Over 10	0.0	0.0	0.9	0.1	0.0	11.7
Mean number	0.10	1.2	2.1	0.9	0.7	5.0
Standard deviation	0.36	1.3	2.5	1.5	0.9	4.2
Maximum	3	8	29	12	8	38

between free care and any other particular plan are about 0.4 for mean total nondental outpatient and acute frequencies and from 0.1 to 0.2 for mean dental, chronic, and well-care frequencies. (As explained below, family and intertemporal correlations make these errors about twice what they would be if all observations were totally independent.) It turns out that differences across plans in the number of total nondental and acute episodes are highly statistically significant, and differences for other episode types are marginally significant.[10]

[10] Gus Haggstrom has pointed out to us that the ratios of means of different plans provide estimates of the relative average price effects when the treatment groups are assumed balanced. See App. G of Haggstrom et al. (1981) for standard errors of such estimators.

Table 4.7

ANNUAL EPISODE FREQUENCIES BY TYPE AND BY PLAN
(DAYTON: YEARS 1, 2, 3)

	Plan				
Episode Type	Free	ID	25	50	95
Hospital					
Mean	0.12	0.10	0.10	0.07	0.10
(Standard deviation)	(0.37)	(0.40)	(0.37)	(0.31)	(0.37)
Dental[a]					
Mean	1.39	0.94	1.19	1.06	1.16
(Standard deviation)	(1.28)	(1.16)	(1.29)	(1.27)	(1.30)
Total NonDental Outpatient					
Mean	4.98	3.48	3.82	3.37	3.20
(Standard deviation)	(4.26)	(4.01)	(3.78)	(3.29)	(3.40)
Acute					
Mean	3.01	1.99	2.19	1.92	1.84
(Standard deviation)	(3.07)	(2.53)	(2.47)	(2.24)	(2.23)
Chronic					
Mean	1.07	0.84	0.93	0.65	0.75
(Standard deviation)	(1.55)	(1.54)	(1.60)	(1.05)	(1.41)
Well Care					
Mean	0.89	0.65	0.70	0.79	0.61
(Standard deviation)	(1.06)	(0.95)	(0.94)	(1.07)	(0.88)
Number of participants years	930	297	763	544	821

[a] Dental episode data are for years 2 and 3 only. The first-year statistics are difficult to interpret since only the free plan covered adults, and there appears to be under-reporting by adults on the pay plans.

Copayment causes the greatest proportional differences in acute episode counts. This is not too surprising since acute episodes are nondeferrable but frequently not serious. Real emergencies are rare; only 10 percent of participants are hospitalized on average in a year, and these are classified as hospital episodes. Acute conditions such as colds are much more common, and care for them may be quite optional. By contrast, care for chronic conditions such as high blood pressure and diabetes is often very beneficial to long-run health and may respond less to plan. Even well care may not be elective, in that immunizations or physical examinations may be required by schools and employers. Thus, it is quite plausible that price would have a larger effect on the number of acute episodes than on the other categories.[11]

Changes in the episode occurrence rates from year to year were small, except for a downward shift in acute episodes on all plans in the second year and a slight decline in the other nonacute categories. Table 4.8 shows trends over the three-year period for the five plans. Perhaps the second year was a nonepidemic or mild year for flu. The pay plans did not cover adult dental care in the first year of Dayton, so their values (about 0.4 episodes on average) are low.

Predicting Annual Episode Frequencies

Good predictions of annual episode frequencies are needed in order to apply the methods in Chap. 5 for estimating within-year price effects. Moreover, episodes are a rational behavioral unit and have desirable statistical properties that make them useful in studying the effects of various personal characteristics on the use of health services.

The statistical problems associated with analyzing annual expenses are greatly reduced by using the number of episodes instead of dollars. There are no problems with outliers, the distribution of episode

---

[11] Other researchers in the Health Insurance Experiment are analyzing these data to determine how the quality of care varies by plan. They will look at exactly which diagnoses are contributing most to the difference between episode frequencies on the various plans. Even using these finer diagnostic categories, it may be hard to separate what care is essential from what is not--since it is hard to determine severity from the claims record.

Table 4.8

TRENDS IN THE AVERAGE NUMBER OF EPISODES
PER PERSON (ROUNDED)
(DAYTON: YEARS 1, 2, 3)

	Plan				
Episode Type	Free	ID	25	50	95
Hospital					
1	0.10	0.15	0.11	0.08	0.07
2	0.15	0.05	0.07	0.08	0.13
3	0.10	0.11	0.12	0.06	0.11
Dental					
1	1.5	0.2	0.4	0.5	0.4
2	1.4	0.7	1.1	1.0	1.2
3	1.3	1.1	1.3	1.0	1.1
Total nondental outpatient					
1	5.0	3.8	4.1	3.7	3.3
2	4.9	3.1	3.6	3.1	3.2
3	5.1	3.4	3.7	3.3	3.1
Acute					
1	3.0	2.1	2.4	2.2	2.0
2	2.9	1.7	2.0	1.7	1.8
3	3.2	2.2	2.2	1.9	1.8
Chronic					
1	1.1	1.1	0.9	0.8	0.7
2	1.1	0.7	0.9	0.6	0.8
3	1.0	0.7	0.9	0.5	0.8
Well care					
1	1.0	0.7	0.8	0.8	0.7
2	0.8	0.7	0.7	0.8	0.6
3	0.9	0.5	0.6	0.8	0.5

occurrence is considerably less skewed (because much of the variance of dollar expenses comes from the size of episodes instead of their number), the people with no episodes are not very distinct from the people with some (e.g., zeros need not be modeled separately as they are for annual expenses (Duan et al., 1982), and perhaps most important, experience over several years is easy to handle.

In analyzing the counts, various refinements are preferable to OLS (ordinary least-squares) regression, since the variance increases with the number of episodes. Since episodes occur roughly randomly over time (evidence for this is discussed in App. F), the number of episodes for an individual should be Poisson-distributed conditional on his expected number. In addition, individuals have different expected numbers of episodes. In Table 4.7, the mean of the number of outpatient episodes on the free plan is 4.98, with a standard deviation of 4.26. A Poisson distribution with this mean would have a standard deviation of about $4.98^{1/2} = 2.23$ rather than 4.26. Thus, a model that allows for different individual rates seems warranted.

These statistics are consistent with other analyses of episodic data that have found that a negative binomial distribution fits well (Kilpatrick, 1977). The negative binomial can be generated by a gamma distribution of underlying expected number of episodes, with the number of episodes being Poisson conditional on that value (Johnson and Kotz, 1969). This "mixing" interpretation of the negative binomial distribution has been exploited by workers in a number of fields for situations in which the group-counts result from combining counts for individuals, each generating counts according to his own Poisson process. Applications include accidents (Greenwood and Yule, 1920), crime commissions (Rolph, Chaiken, and Houchens, 1981), and insurance claims (Ferreira, 1974).

In the applications of the negative binomial model referred to above, no attempt was made to explicitly take into account how an individual's Poisson "rate" might vary with his measured attributes. In this section, we model how a participant's expected number of episodes of a particular type varies with his characteristics. To fix ideas, think of one episode type and let λ_i be the expected number of episodes

for individual i in a year. We will call λ_i his episode propensity and distinguish it from his observed rate n_i in a given year. Then, conditional on λ_i, n_i has a Poisson distribution with mean λ_i. Now since λ_i is unobservable, we model λ_i as $\lambda_i = \delta_i u_i$, where δ_i is predicted from individual i's attributes (including plan) and u_i is the random component. We assume that $\log \delta_i = X_i \beta_j$, where X_i is the vector of individual i's attributes and β is a vector of regression coefficients to be estimated. We also assume that u_i is gamma distributed with mean 1 so that the marginal distribution of n_i, given δ_i, is negative binomial with mean δ_i.

In more familiar terms we can write an individual's expected number of episodes λ_i as a multiplicative regression equation. That is,

$$\log \lambda_i = \log \delta_i + \log \mu_i = X_i \beta + \text{individual error.}$$

The error term reflects the negative binomial distribution discussed above. The vector of regression coefficients is estimated by the method of maximum likelihood.

This multiplicative form is especially convenient for comparison with results on annual spending. Taking the log of (4.1), we obtain[12]

$$\begin{aligned}\log(\text{spending}) &= \log(\text{cost per episode}) + \log(\text{number of episodes})\\ &= X\beta_1 + e_1 + X\beta_2 + e_2\\ &= X(\beta_1 + \beta_2) + (e_1 + e_2).\end{aligned}$$

Since both components are modeled multiplicatively, the proportional change in spending caused by, say, income can be obtained by adding the regression coefficients of income in the two regressions. This sum can then be compared with coefficients in a regression of log(spending).

---

[12] Strictly speaking, we usually take log(spending + c) to take care of those who spend zero.

We do not lose much information when we add together three years of episode counts and regress these on predictor variables that come from data at the start of the experiment. This procedure eliminates having to worry about year-to-year correlations as we would have to if we used each of the three years as a separate data point. Three facts justify this procedure: the stability of regression coefficients over time, the stability of independent variables over time, and the independence of episode counts over time. (See App. F for the analysis leading to these conclusions.)

The same independent variables are available for predictors of episode counts as for the cost per episode regressions. Table 4.4 gives the means, standard deviations, and descriptive comments for the relevant predictors.

We first present the results of our negative binomial regression model and then discuss extensions and caveats.[13] Table 4.9 gives the results of regressing the annual frequencies of episodes, by type, averaged over the three years of Dayton for 1074 participants on a group of individual predictors. The fitted equations vary considerably across the episode types. The coefficients should be interpreted as indicating how much the logarithm of the episode rate can be expected to change when the variable is increased one unit. Hence, exponentiating the coefficient gives the proportional change with a one-unit increase in the predictor. For example, the estimated coefficient of -0.33 on the dummy variable for black (black = 1, other = 0) in the dental episode equation means that blacks on average have 72 percent (= exp(-0.33)) as many dental episodes as nonblacks, all other predictors being equal.[14]

The estimated regression coefficients show that the raw plan differences reported in Table 4.7 hold up after individual attributes have been accounted for. As before, the free-plan participants generate more episodes of every type, and the biggest difference is between free-

---

[13] This same negative binomial regression was developed independently by Hausman, Hall, and Griliches to analyze research and development expenditures (Hausman, Hall, and Griliches, 1981).

[14] Recall from Table 4.8 that blacks have on average 77 percent higher costs per dental episode than comparable nonblacks. It would appear that this is somewhat offset by the lower frequency.

plan participants and those on any other plan. The coefficients across the pay plans differ mostly in the expected order. Plan has the largest effect on acute episodes, and the smallest on hospital.

The influence of the other predictors on episode frequency varies with the type of episode being considered. The more important determinants of number of episodes appear to be the physician's visits in the preceding year, a scale of self-reported health (HPWSUM), age (both age and the square root of age, and whether newborn, are important), and whether the participant is a woman (over 17). Generally, log income, being black, no regular provider, log family size, and maximum education are also moderately important.

Other variables were tried in developing the model, but proved to be unimportant. They include measures of overweight, smoking status, marital status, life change in the preceding year, size of MDE as percentage of income, whether the participant was in the experiment for three or five years, a depression scale, a positive well-being scale, a satisfaction with medical care index, and AFDC status. Self-reported health status was measured prior to the beginning of the experiment.

There were no significant interactions of insurance plan with income, age, or number of previous visits. The interaction with income is politically important since it measures how much more (or less) poor people are affected by coinsurance than rich people. Because the individual deductible plan is not income-related, it theoretically should affect rich people less than poor people. The other plans are income-related, although the MDE limits mean that middle- and upper-income families have the same $1,000 limit. The signs of the interactions with income are plausible: positive for the nonincome-related ID plan and negative for the income-related coinsurance plans, but the coefficients were small and not statistically significant.

The coefficients in Table 4.9 show that the five episode types fall into two groups. Hospital, acute, and chronic episodes are heavily affected by previous health and previous visits to a doctor, while well-care and dental episodes are heavily affected by education and income. Well-care episodes are concentrated in women and children, and chronic episodes in older participants.

Table 4.9

REGRESSION EQUATIONS FOR PREDICTING NUMBER OF EPISODES (ROUNDED)
(DAYTON: YEARS 1, 2, 3)

	Hospital		Acute		Chronic		Well		Dental[a]	
	Coeff.	(t)	Coeff.	(t)	Coeff.	(t)	Coeff.	(t)	Coeff.	(t)
CONSTANT	-3.7	(-2.6)	0.19	(0.3)	2.95	(-2.8)	-4.35	(-7.7)	-3.09	(-4.3)
IND	-0.33	(-1.2)	-0.61	(-5.6)	-0.52	(-3.2)	-0.44	(-3.5)	-0.45	(-3.3)
P25	-0.12	(-0.6)	-0.42	(-5.2)	-0.25	(-1.7)	-0.29	(-3.3)	-0.22	(-2.1)
P50	-0.31	(-1.3)	-0.59	(-4.7)	-0.50	(-4.1)	-0.24	(-2.6)	-0.37	(-3.5)
P95	-0.11	(-0.6)	-0.54	(-6.2)	-0.47	(-3.3)	-0.40	(-4.1)	-0.22	(-2.2)
EXAM	0.22	(1.5)	0.00	(0.0)	0.21	(2.1)	0.06	(1.0)	-0.05	(-0.7)
HDWSUM	0.17	(3.9)	0.09	(4.6)	0.10	(4.2)	-0.01	(-0.7)	-0.02	(-1.1)
LMDVIS	0.23	(3.0)	0.27	(8.2)	0.40	(8.2)	0.13	(4.1)	0.06	(1.8)
NEWMEN	0.83	(1.3)	-0.28	(-1.2)	0.27	(0.7)	0.51	(2.6)	-2.60	(-3.6)
WOMAN	0.39	(2.5)	0.36	(6.3)	0.27	(3.3)	0.77	(12.6)	0.13	(2.3)
AGE	0.00	(0.1)	0.03	(3.7)	0.04	(3.0)	0.03	(4.0)	-0.02	(-2.0)
SQRAGE	0.01	(0.1)	-0.37	(-5.3)	-0.17	(-1.4)	-0.41	(-5.8)	0.17	(2.1)
MAXED	-0.04	(-1.2)	0.01	(1.0)	0.04	(2.1)	0.06	(4.7)	0.06	(4.6)
BLACK	0.17	(0.6)	-0.41	(-3.2)	-0.37	(-2.3)	-0.50	(-3.1)	-0.33	(-2.0)
LINC	0.18	(1.1)	0.18	(2.2)	0.26	(2.2)	0.31	(4.9)	0.20	(2.7)
LFAM	0.00	(0.0)	-0.17	(-2.1)	-0.18	(-1.8)	-0.22	(-3.2)	-0.01	(-0.1)
NOPROV	-0.04	(-0.2)	-0.21	(-2.4)	-0.22	(-1.9)	-0.21	(-2.1)	-0.22	(-1.7)
(α)	1.03	(0.10)	2.00	(8.6)	1.07	(0.7)	14.70	(122)	2.94	(6.9)
SAMPLE SIZE	1074		1074		1074		1074		1074	
INTRAFAMILIES CORRELATION	0.08		0.32		0.15		0.27		0.43	

[a] Dental data from years 2 and 3 only.

Because families share propensities to go to the doctor, observations from members of the same family contain less new information than would the same number of independent observations. Correcting for this intrafamily correlation has little effect on the regression coefficients, but the t-statistics measuring significance are shrunk considerably for variables such as plan and family income, which are the same for all members of the family. The greatest effects are

seen in dental, acute, and well episodes, for which intrafamily correlation is largest, and the effective sample size is about halved.[15]

The estimated values for α in Table 4.9 show how much individuals differ in unmeasured ways. The small values for hospital and chronic show that people have strong propensities that are not captured by the independent variables, and the large value for well episodes implies that most individual tendencies are captured by the independent variables.[16]

There was substantial correlation among counts of episode types over individuals. Table 4.10 shows that acute and chronic episode-counts were particularly correlated. This may reflect the assignment of chronic flareups to the acute category, where from an economic point of view they belong. Correlation does not affect predictions of average spending on the experimental plans, which is still the sum of the averages of the individual types. The correlation is most important for simulation of different deductible plans. Since the full distribution of expenditures determines who will exceed the deductible, correlation must be considered. In simulation, we will combine episode types that are highly correlated, or use a more complex random simulation procedure that accounts for the correlation.

---

[15] We would like to thank our colleague William H. Rogers for generously sharing his software with us to do these computations.

[16] Alpha is the shape parameter of the gamma mixing distribution of unmeasured propensities for episodes. This distribution has a coefficient of variation of

$$(\alpha\beta^2)^{1/2}/\alpha\beta = \alpha^{-1/2}.$$

The "t" values test differences of α from 1, compared with a sample with only one from each family.

Table 4.10

CORRELATION OF RESIDUALS OF DIFFERENT EPISODE COUNTS

Type	Acute	Well	Chronic	Hospital
Dental	0.24	0.20	0.12	-0.01[a]
Acute		0.25	0.47	0.26
Well care			0.22	0.28
Chronic				0.26

NOTE: Residuals from the predictions discussed in App. A, shown in Tables A.1 and A.2, were correlated. Based on three years of data for 1074 stayers in Dayton. All correlations significant at 0.01 except hospital-dental.

Chapter 5

CHANGES DURING THE YEAR IN EPISODE FREQUENCIES

In Chap. 4, we studied the determinants of annual episode counts. While annual episode counts are at least theoretically a better measure of behavior than annual expenses or visits, such an aggregate analysis misses the major advantages of episodic data. If we knew how people behave as price changes through the year, we would have a much better understanding of the demand for medical care, and the effects of deductibles on demand. The results could be used to predict, with some confidence, the spending on plans with a range of deductibles. Studying spending changes during the year may also reveal the size of the transient effects in the experiment. These effects are important in judging the generalizability of results, and in planning the phase-in of insurance changes.

There are two methodological problems in analyzing changes within the year. The economic problem lies in finding the response to the marginal price of services for a family with a deductible. Below, we will show how we adapted our solution in Keeler, Morrow, and Newhouse (1977) to these data. The ability to anticipate exceeding the deductible turns out to be very rare, and by ignoring it we obtain a much more tractable estimation problem. Anticipation might be important with smaller deductibles.

The statistical problem arises because "sickly" people who tend to seek more care (and exceed the MDE) will face lower prices on average later in the year than will others. We separate price effects from sickliness effects, by assuming that unobservable individual propensities to spend are constant over the year (an assumption we test), and compare behavior before and after the MDE is exceeded with behavior on the free plan (where presumably there are no within-year price effects).

Except for small surges of demand for well care and dental care in the first three months and the last month of the experiment, episode rates for those on the free plan are constant during the year.

Empirically, the occurrence of individual episodes are random over time.[1] Episode occurrence rates of those on pay plans after they exceed their MDE during a year do not indicate much of a "sale" effect. That is, pay-plan participants faced with free care between the time they exceed their MDE and the end of the year have rates approaching free-plan rates, but do not appear to splurge on deferrable conditions.

We begin by discussing the economic and statistical problems. Our formal model of within-year behavior of pay-plan participants follows, together with the results of fitting this model and our interpretation. In App. F we discuss the assumptions underlying this model and the checks that were made on them. These include: constant episode rates during the year on the free plan, independence of an individual's episode occurrences over time, and the infrequency of situations where families might anticipate exceeding the deductible without having specific purchases of medical care in mind.

DEMAND FOR SERVICES WHEN PRICE CAN VARY: THE ECONOMIC PROBLEM

When the rate of coinsurance for medical services can change, the current out-of-pocket price is not a complete guide to action. Instead, the economically rational family looks at the effects of current spending on future prices. Before the deductible is exceeded, current spending has the bonus of reducing the remaining deductible. Thus, the true price is somewhat less than the nominal out-of-pocket price. How much less depends on the probability that the family will subsequently exceed the deductible. (See Keeler, Newhouse, and Phelps, 1977.) This is shown graphically in the left side of Fig. 5.1, in which the real price is sketched for families that are more or less likely to exceed an annual deductible.

Anticipation

If families could reasonably anticipate that they would exceed the MDE, they might start spending more from the point they first anticipated doing so, instead of from the date they actually exceeded the MDE. Thus, insurance plans with small deductibles that most

[1] That is, given a person's propensities to have episodes, the actual events occur independently over the year.

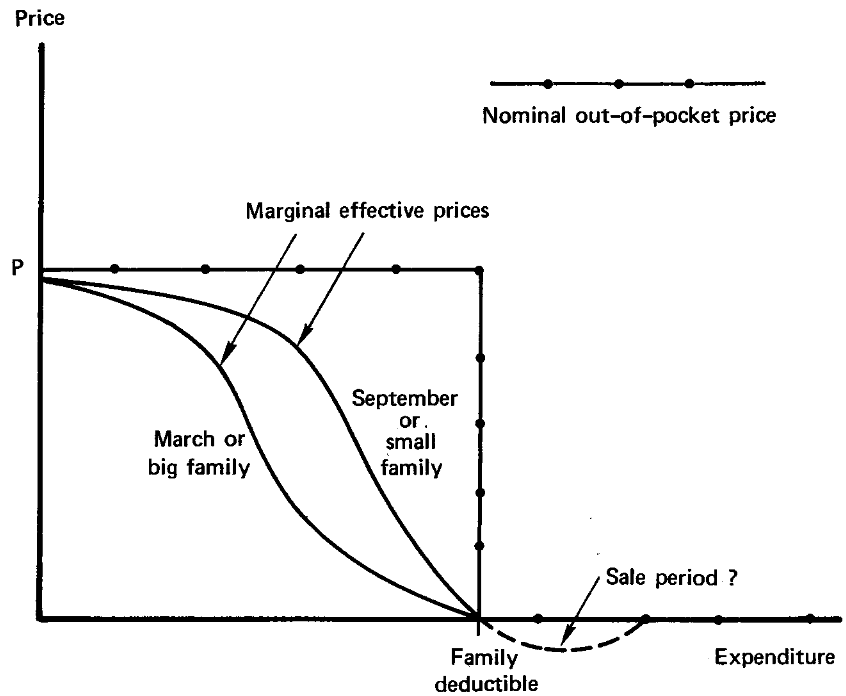

Fig. 5.1 -- Marginal effective prices when a family deductible is present

families would exceed would do little to restrain demand. In addition, such anticipation would make the effective price less than the nominal price, and estimates of price effects using average normal price would be biased.

Families might anticipate future spending in two ways. First, they may schedule spending on such things as dental problems found at a previous examination or continuing care for diabetes. Second, a large family with only a few dollars left on its MDE early in the year can be fairly sure that it will exceed the MDE with time to spare, even if it does not know precisely how.

Neither of these forms of anticipation changes our results very much. Spending on particular known problems is already taken into account by the program that links up charges into episodes. The program assumes in its linking and dating procedures that such spending is fully anticipated. The second form of anticipation cannot be important

because it occurs so rarely. We show in App. G that for both family and individual deductible plans, times at which participants might reasonably expect to exceed the remaining MDE were only one-fourth as frequent as times when families had already exceeded it. The situation for the individual deductible plan was similar to that for the coinsurance plans.

Thus, except for episodes that had already begun, participants were rarely in a position to anticipate exceeding their MDE or deductible. For this reason, we will simply compute the average rate of spending before the deductible is exceeded, in effect grouping all pre-MDE situations in the analysis. The rarity of anticipatory positions means that the bonus effect cannot be important for plans with large ceilings or caps, especially if these are combined with low coinsurance rates. The rarity also means that we will have to wait for more of the data to be processed before it is possible to attempt to distinguish between the rate of episodes in nonanticipation pre-MDE and anticipation periods.

Sales

Families with changing coinsurance rates have another interesting option. They can schedule deferrable treatment episodes to a time when out-of-pocket prices are low. Thus, families with care that is temporarily free have medical care "on sale," and have more incentive to spend than families with permanently free care. We can check this phenomenon by looking at the experience on the free plan at the start and finish of the experiment, and by studying what happens to families in the months just following the time they satisfy their deductible.

The most deferrable types of treatment are dental and well care. Three years of free-plan, quarterly data on dental and well-care episodes are shown in Fig. 5.2. We can see an initial "catch-up" surge in the first quarter of year 1 and a somewhat smaller "store-up" surge at the end of the three years. A week-by-week examination showed that the initial surge fell to the normal rate after 12 weeks. Apparently this was enough to obtain all desired deferred services. To mirror this experience on the free plan, we will select the three months following fulfillment of the MDE as the period to test for sales.

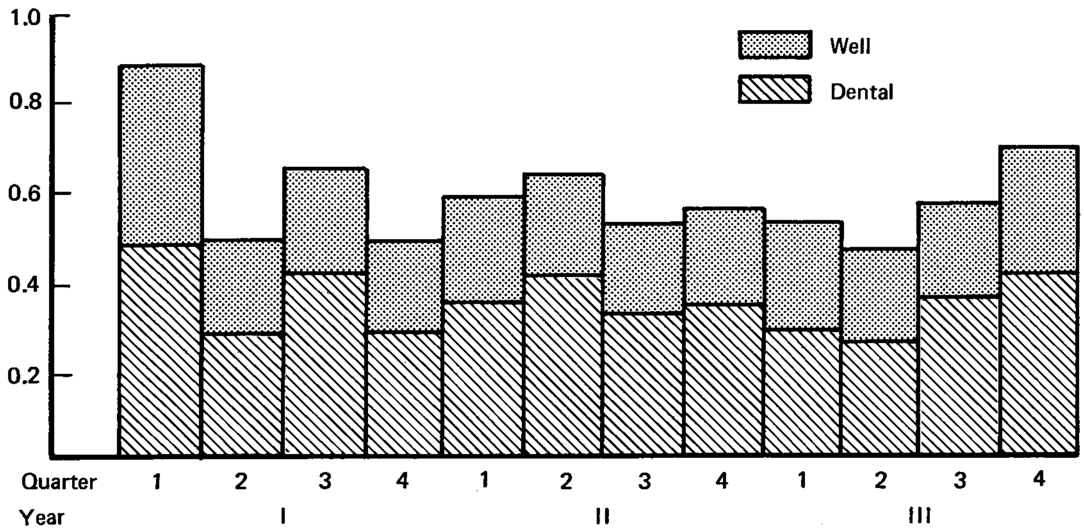

Fig. 5.2 -- Individual quarterly rate of well-care and dental episodes, free plan

The final surge is somewhat smaller because only half the participants exited after three years, and half have another two years to go. The transient effects of the experiment were very minor, representing no more than a doubling for the first quarter. This probably reflects much less pent-up demand than in the surge at the start-up of British National Health Service in 1948 noted earlier. Britain then was much poorer, and more poorly insured than we are now, and nonessential health care may have been deferred by the war.

In sum, we will check spending rates for three economic periods in each accounting year: the period before the MDE is exceeded, the next three months, and the rest of the year. We will call these the pay period, the sales period, and the final period, for short. Since only 20 percent of families exceed their MDE in any one year, these later periods often do not exist for many families.

SEPARATING SICKLINESS FROM PRICE EFFECTS: THE STATISTICAL PROBLEM

For health insurance plans with deductibles, prices are lower on the average for those who tend to have more episodes. (We will call such people "sickly.") Sicklier families tend to consume relatively more care early in the year and hence are more likely to exceed their deductibles (or MDEs). Annual analysis by plan was made easier because plans are balanced by propensities to spend. Unfortunately, plan subgroups based on price determined by spending within the year become unbalanced as the year goes on.

Some data on first-year Dayton episodes by month bear out this hypothesis. It contrasts behavior of those on the free plan with those who have a family or individual deductible plan (100 percent coinsurance in first-year Dayton). In the top display of Table 5.1, we see an apparent sales effect as the pay-plan participants who exceed their deductible (the row with 0 deductible left) then spend an average of $37 per month for the rest of the year. This is $10 more than average monthly spending of those participants on the free plan. However, let people in a given month be categorized by their previous spending. (In this case, called "high" if the previous monthly average exceeds $15 per month and otherwise called "low.") Then the lower two displays in Table 5.1 show that the big differences are between those with high and low previous spending, with only small differences depending on whether the MDE is exceeded. The apparent sale effects have almost disappeared. This is fairly convincing evidence of fixed individual propensities to generate episodes.

Spending during the period before the MDE is not a perfect measure of the natural propensity to spend, since it includes regression to the mean. In other words, people have high expenses early in the year both because they were sickly and because they were temporarily unlucky. The sickliness should persist, but the bad luck may not. Appendix C gives some data that show how large the regression to the mean is by looking at free-plan behavior. We show there that the presumed sickliness effect as given in Table 5.1 is reduced by regression to the mean.

Table 5.1

AVERAGE MONTHLY EXPENDITURES BY DEDUCTIBLE REMAINING

Plan	Deductible Left	Number of Person Months	Monthly Expenses: Average Dollars per Month	Percent Positive	Average Log if Positive
Free plan	--	3396	$27	33%	3.4
Pay plans	0	661	$37	35%	3.5
	1-100	283	$16	23%	3.2
	101+	3472	$13	20%	3.2
Low previous expense[a]					
Free plan	--	1773	$11	24%	3.2
Pay plans	0	155	$14	24%	3.4
	Positive	2762	$ 7	16%	3.0
High previous expense					
Free plan	--	1623	$44	43%	3.5
Pay plans	0	506	$44	38%	3.5
	Positive	993	$32	34%	3.5

NOTE: Dayton, first year, excluding adult dental.

[a]Low is defined as $15 per month or less.

A MODEL OF WITHIN-YEAR PRICE EFFECTS

Our model of within-year price effects resolves this statistical problem. It allows us to compute the ratio of the episode rates in different periods defined by MDE status to what those rates would have been on the free plan. The model relies on four critical assumptions, discussed briefly in the next four paragraphs. Evidence on the assumptions is presented and discussed more fully in Apps. B and F.

Nonprice-related propensities to have episodes are fixed over the year. These propensities are determined by measured and unmeasured characteristics. We will estimate the expected number of episodes on

the free plan for each individual, using nonplan individual covariates. We can add these up to get a total for family k that we will call d_k. Since our information about families is limited, they will also have propensities that will not be captured in the statistical predictions. (Over time, we would expect individual families to consistently be over- or underpredicted by the d_k.) To account for this unmeasured variation, we will assume that each family has an unmeasured propensity u_k that is constant from year to year. The families' actual propensity is the product $d_k u_k$. The u_k will be assumed to be drawn from a gamma distribution. The gamma is a fairly rich distribution (it can take a variety of shapes), with computational advantages.

The year is split into three periods by MDE status, and within these periods the effects of price are constant. Above, in the economic section, we justified splitting the year into the period before the deductible (or MDE) was exceeded (the "pay" period), the three months just after (the "sale" period), and the rest of the year (the "final" period). In effect, we will estimate three ratios (π_{ij}; i = 1, 2, 3) for each coinsurance plan j. These price ratios are defined as the ratio of the expected number of episodes in period i for a family on plan j to the expected number of episodes in period i for that family if care were always free. That is, it is the ratio of pay-plan propensity to the free-plan propensity to generate episodes.

Episode rates on the free plan can be considered constant over time. We show in App. B that episode rates for all types of episodes are uniform except for sales effects, the first day of the year, and a tendency for chronic episodes to occur early in the year. Except for sales effects, these patterns persist in all three years. We compare pay-plan episode rates with the nonsale average, assuming that this represented the steady-state rate. The surge of carry-over episodes on the first day of the year, and the tendency for chronic episodes to be dated earlier in the year, follow from our grouping conventions. Thus, these trends should be the same on all plans. In the analysis, we transform time so that rates are uniform.[2]

---

[2] For example, the first day has as many acute episodes as 22 ordinary days. We say that someone who exceeds the MDE on the third day has had (22 + 2)/(22 + 364) of the year's experience in their pre-MDE stage.

Occurrences of episodes are independent over time. Individuals have different propensities, but conditional on these propensities, episode occurrences do not appear to be bunched or spread more than random. Thus, they can be modeled as being Poisson distributed.[3] This is convenient for purposes of comparison because the Poisson distribution is infinitely divisible, i.e., the episodes in any fraction of the year also follow a Poisson distribution.

Thus, summarizing the four assumptions, a family k on plan j in MDE status i for an effective period of time t_{ik} will have n_{ik} episodes, where n_{ik} follows a Poisson distribution with mean $\pi_{ij}\delta_k u_k t_{ik}$. Table 5.2 summarizes the notation and assumptions made thus far in specifying the model; others will be made in fitting the model. For those who are still confused, we offer Fig. 5.3. This depicts episode occurrence rates for two hypothetical families on one of the pay plans. These are not the actual rates, but the assumed true long-run average rates for those families in that position.

Family 1 begins the year with expected occurrence rate $\pi_1 f_1$, where π_1 is the price ratio for period 1 (the "pay" period) and f_1 is the Poisson occurrence rate for Family 1 if it had been on the free plan. (Mathematically, one can think of f_1 as the Poisson occurrence rate for Family 1's "treatable illnesses on the free plan" with π_1 being a 0-1 Bernoulli selection function determining whether spending occurs for that illness.) When Family 1 exceeds its MDE six months into the year, its expected occurrence rate jumps to $\pi_2 f_1$ for three months (the "sale" period). If π_2 exceeds 1 as in Fig. 5.3 ($\pi_2 f_1$ is larger than f_1), there is a sales effect and people are spending more than selection and free care would explain. After three months, the sale is over and Family 1's

---

[3] Families occasionally have several episodes on the day the MDE is exceeded. We assume that all such episodes belong to the pre-MDE period. Alternative assumptions that put some of the multiple episodes into period 2 have no discernible effect on pre-MDE price ratios, and raise period 2 price ratios by 0 to 5 percent, depending on plan and type. The problem of multiple episodes exceeding the MDE is similar to the case of someone with only a little MDE remaining who has a very large episode. Multiple episodes are studied in more detail in App. D. In App. E, we separately discuss multiple hospitalizations.

Table 5.2

NOTATION AND ASSUMPTIONS

Notation:

π_{ij} = ratio of occurrence rates for period i on plan j to no pay rate; i = 1,2,3 (the price index).

δ_k = predictable component of expected annual episode occurrence rate under free plan for family k (a function of the covariates of individuals in family k).

u_k = multiplicative component of expected annual occurrence rate for family k if care were free that is due to unmeasured family propensities to have episodes.

$\delta_k u_k$ = expected annual occurrence for family k if care were free (f_k).

t_{ik} = time spent by family k in period i. The period until family k exceeds the MDE is period 1, the next 3 months is period 2 while $t_{3k} = 1 - t_{1k} - t_{2k}$ is the remainder of the year, if any.

n_{ik} = number of episodes that family k has during period i.

Assumptions:

1. Episodes occur during period i to family k on plan j, according to a Poisson process with intensity function $\pi_{ij}\delta_k u_k$.
2. Thus $n_{ik} | t_{ik}, \pi_{ij}, \delta_k u_k \sim$ Poisson $(\pi_{ij}\delta_k u_k t_{ik})$.
3. $u_k \overset{ind}{\sim}$ Gamma (α,β). That is, the unmeasured component is modeled as being drawn independently from a gamma distribution independently for each family.

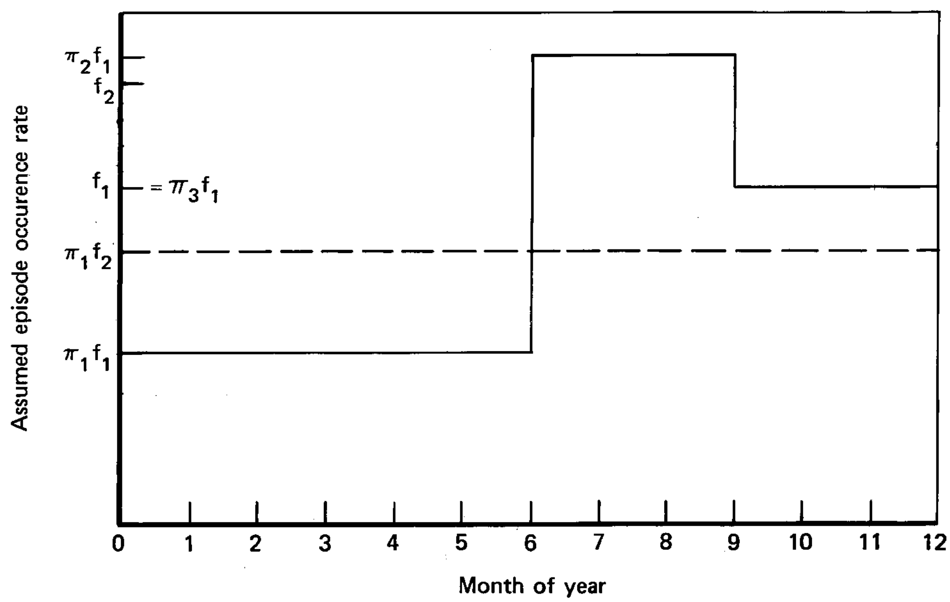

Fig. 5.3 -- Example of how expected rates change through the year

occurrence rate drops to f_1, its "free" rate for the "final" period. (In Fig. 5.3, π_3 is assumed to be 1.)

Similarly, Family 2 begins the year with occurrence rate $\pi_1 f_2$, where f_2 is Family 2's "free" rate. Since Family 2 never exceeds its deductible, its rate remains at $\pi_1 f_2$ throughout the year.

Methodology for Fitting the Price Effects Model

There are several pieces to put together in order to estimate the parameters of the price effects model. Define $f_k = u_k \delta_k$, where δ_k is a predictable family component and u_k is the remaining unpredictable component. First, the predictable family component δ_k must be estimated from individual covariates. Next, the maximum likelihood procedure for estimating the price indices π_i and the parameters of the distribution of u_k will be laid out. Finally, we describe the way pay-plan episode rates are scaled to make them directly comparable with free plan rates. We describe our methods before presenting the numerical estimates.

Recall that δ_k is the predicted annual number of episodes for family k if care were free. This is estimated by d_k, which is obtained by summing individual regression predictions of annual episode rates (from App. A) over the individuals in the family.[4] We assume there are no interactions between within-year price effects and other covariates, so that all the effects of covariates other than price on numbers of episodes are summarized by δ_k.

Analogously to our negative binomial regression model in Chap. 4, we assume family k has an unmeasured multiplicative factor u_k, with u_k having a gamma distribution with parameters α and β. Given the value u_k, the annual rate for family k is assumed to be $u_k \delta_k$ if care were free. This can be regarded as the Poisson regression analog to random effects models in the analysis of variance.

The parameters α and β of the gamma mixing distribution measure by how much the underlying episode rates of families differ from the rates predicted from their demographic characteristics. The larger the unobservable or unmeasured differences between families, the higher the variance $\alpha\beta^2 = \beta$ of the mixing distribution, and the lower the value of α.[5]

Appendix H gives the likelihood function and technical derivations of the maximum likelihood estimates presented in the next section.

RESULTS

We first give results on acute outpatient episodes. These are the most frequent type and hence generate estimates that are the most precise statistically. As noted earlier, the year is split into three periods and the sale period is assumed to last three months after the

---

[4] For cost reasons, we chose the more standard OLS regressions predicting the square root of individual episode frequencies given in App. A in preference to the negative binomial regressions given in Chap. 4. It turns out that both models give substantially the same predictions.

[5] Since the regression estimates d_k of rates based on demographic characteristics are unbiased (the sum of the d_k equals the sum of the n_k), the estimated mean of the Gamma distribution should to be close to one. The mean of a gamma is given by $\alpha\beta$ ($\simeq 1$) and the variance by $\alpha\beta^2$, where α is the shape parameter and β the scale parameter.

MDE was exceeded. Table 5.3 shows that the pay-period rates for each plan for the first three years of Dayton are all substantially less than the free plan rate. The price indices for 95 and 50 percent coinsurance plans are similar, being between 50 and 60 percent, while the price index for the 25 percent coinsurance plan is somewhat larger (about 70 percent). Depending on the plan and the site and the year, only 15 to 30 percent of participants exceeded their MDE on these plans in the first nine site-years (including sites other than Dayton) of the experiment. Thus, most of the data are in the pre-MDE range and the results are quite stable over time. In Table 5.3, the values for the free-plan index are set at one and the estimated occurrence ratios p_i on the other plans are all relative to the free plan.

In the sale period after the MDE is exceeded, the estimated price indices (p_2) on the other plans are closer to the free-plan rates than was p_1. To compute more precise estimates of these indices, we based

Table 5.3

ESTIMATED PERCENTAGE PRICE INDEX FOR ACUTE EPISODES BEFORE THE MDE IS EXCEEDED

Year	Free	25	50	95	b
1	100	78	67	60	0.45
2	100	66	47	56	0.44
3	100	68	50	50	0.35
Standard error (S.E.)		(10)	(10)	(7)	(0.06)
All three years	100	71	54	55	0.40
S.E. for 3 years		(5)	(4)	(4)	(0.03)

NOTES: Here and in other tables, rates are given relative to the free plan so that the free-plan rate is 100. The values tabled are 100 p_1 for each plan, where p_1 is the maximum likelihood estimate of π.

The entries in parentheses are the standard errors of 100 p_1 based on the asymptotic variance of the maximum likelihood estimate. Including the effects of random variation on the reference free plan increases these by about 1. Details of the error calculation are in App. F.

Table 5.4 on the combined data for all three years. We also show estimates for the individual deductible plan here. They are very close to those for the family deductible (95 coinsurance) plan. No sales effects for acute conditions are observed ($p_2 < 1$), as expected, since there are few investment possibilities in acute care. The indices do not jump to 1.0 in either of the last two periods, perhaps because of habit, perhaps because participants may not know that they exceeded the MDE. Unless they keep personal records, they will not be informed of their MDE status until their insurance receipt arrives, that is, on average, about a month after the billing date. By contrast, people on the free plan, who did show pronounced sale behavior at the beginning of the experiment, were aware from their enrollment interviews that their insurance was changed by the experiment. Finally, we have assumed that people on the pay plans have incorporated the full cost of an episode into their planning as soon as it starts, and they probably did not do so.

Table 5.4

ESTIMATED PERCENTAGE PRICE RATIOS FOR ACUTE EPISODES, THREE YEARS COMBINED

	Plan			
Period	25	50	95	Individual Deductible
Pay (before MDE)	71 (5)	54 (4)	55 (4)	57 (4)[a]
Sale (next three months)	79 (11)	57 (10)	75 (9)	73 (10)
Final (rest of year)	97 (10)	71 (11)	85 (9)	91 (10)
Estimated b = 0.4 (0.03)				

[a] The entries in parentheses are the standard errors of 100 p as given by the asymptotic variance of the maximum likelihood estimate. Including the effects of random variation on the reference free plan increases these by about 1. The details of the error calculation are in App. F.

Episode linking conventions may be responsible for some of the difference between price ratios in the second and third periods. To deal with visits for the same illness that might have been diagnosed differently, we put acute spending within a week of an earlier acute episode into that episode. Thus, there would be little chance of two acute episodes occurring a week apart.[6] One week missing from three months is about 0.08 of the total of randomly occurring illness.

We tested our methods on the actual free-plan experience split into three pseudo-periods by an imaginary $500 deductible. The price ratios for the three periods were 1.04, 0.88, 1.00. Since the three periods were artificially imposed on the data, the differences from 1.0 were either random or artifacts of the methods. The drop in the second period was not quite statistically significant, but it may reflect the acute episode one-week gap.

For acute episodes, the estimate variance b of unmeasured differences between families was 0.4. This small variance means that our measures capture most of the difference between families, and that residually higher spending in the sale periods is mainly due to price effects. If b were larger (if families differed more in unmeasured characteristics), those going over the MDE would be quite different from the others, and the increased spending would reflect selection effects more than price effects. The price ratios in the sale periods would then have been estimated smaller than in Table 5.4. In fact, the uncertainty in the estimate of b is quite small (0.03), and the values in Table 5.4 would not be greatly affected by much larger errors.[7]

---

[6] Our analysis of the independence of episodes was based on monthly aggregates, so this very short-term correlation did not appear strongly there. We plan to investigate this issue further when we look at interarrival times of episodes.

[7] Setting b at ± 11 standard variations from its estimated value of 0.4 leads to changes of 10 to 15 percent in sale period ratios and 2 percent changes in pay period ratios.

Nonacute Episodes

After participants exceed the MDE, the rate of episodes of other types goes up, but just as with acute episodes, there is not a large "sale" splurge. After splitting the year into three parts for these episodes, the estimate p_2 is higher than p_3 five times out of twelve (3 years times 4 episode types: chronic, well care, dental, hospital). That is, the price index for the sale period exceeded the price index for the final period in about half the cases. This contrasts with acute episode behavior and might be partially due to sales effect on stored-up problems (which are more likely for well-care or dental episodes than for acute). Sales for some participants might be balanced by other participants being unaware that they had exceeded the MDE. Another possible contributory factor is the larger error in the estimates. These other episode types are less frequent than acute episodes. In all cases p_2 and p_3 were close. Because of the lack of pattern and the imprecision in the three part estimates, we combine all the experience after the MDE is exceeded in Table 5.5. Even with this aggregation, the hospital episode estimates are very imprecise.

Except for the 25 percent coinsurance plans, there are substantial increases in spending after the MDE is exceeded. We would expect effects to be smallest on the 25 percent plan, since the financial change from coinsurance to free care is smallest there. For other pay plans, episodes after the MDE is exceeded occur mostly at rates close to the free-plan rate. Even for those episode types where the timing is flexible, coinsurance appears to be restraining overall demand. Sales effects, where they occur, seem much too small to overcome the initial low-demand period. With more data we will get better precision, but for the time being, the data are roughly consistent with demand after the deductible is exceeded occurring at the free-plan rate.

Table 5.5

PERCENTAGE PRICE INDICES FOR EPISODES
IN NONACUTE CATEGORIES

	Individual Deductible		95 Percent Coinsurance		50 Percent Coinsurance		25 Percent Coinsurance		Dispersion b	
Type of Episode	Pay Period	Rest of Year	Pay Period	Rest of Year	Pay Period	Rest of Year	Pay Period	Rest of Year	Pay Period	Rest of Year
Well	72 (6)[b]	87 (13)	66 (5)	101 (10)	84 (5)	110 (20)	86 (6)	56 (14)	None[a]	
Chronic	67 (8)	84 (12)	62 (5)	100 (11)	55 (6)	78 (17)	82 (6)	81 (12)	0.49 (0.05)	
Dental[c]	70 (10)	92 (20)	79 (9)	112 (18)	74 (9)	88 (20)	90 (9)	104 (26)	0.50 (0.06)	
Hospital	64 (19)	212 (60)	88 (15)	102 (20)	67 (14)	83 (29)	87 (15)	152 (49)	0.55 (0.22)	

NOTE: Based on three years of data.

[a] There are no detectable unmeasured differences between families in propensity to use well care. To get convergence of the estimates, α was fixed at 9.0.

[b] The entries in parentheses are the standard errors of p_1 as given by the asymptotic variance of the maximum likelihood estimate. Including the effects of random variation on the reference free plan increases these by about 1. Details of the error calculation are in App. F.

[c] Based on second and third years.

Chapter 6

CONCLUSIONS

The episodic analysis was designed to take a closer look at behavior that might not be apparent in the aggregated annual analysis. Much effort and resources went into creating the programs that split expenses into episodes of different types and placed them within the year. Most of the data have not yet been analyzed, but these interim results have important implications for insurance design, for other HIE analysis, and for further work.

Plan differences in expenses appear mainly due to different numbers of episodes--plan effects on episode size are small. The two-part model of annual expenditures (Manning et al., 1981) showed that people on the free plan were more likely than pay-plan participants to have some positive spending, and to spend more when they did spend. Our results show that both of these effects are due mainly to the higher expected number of episodes on the free plan. This higher expected number of episodes results in a lower probability of no episodes, and a higher expected number given that there are some episodes.

We have not studied whether illnesses that would be treated on the free plans, but are not treated on the stingy insurance plans, represent mostly illnesses that are milder than average. It seems plausible that they could be. If so, the equality of episode size on different plans is the result of two effects canceling each other out. The more severe illness seen in stingy plan episodes may be treated less intensively because of the cost to the patient. Differences in severity will be tested when we study plan effects on the occurrence of different diagnoses, although we will still be unable to infer severity within the diagnosis categories.

If the severity mix proves to be similar on all plans, the effects of price on medical care costs must be limited to the patients' initial decision to seek care. We would then be unlikely to find plan differences in type of provider chosen, tendency to use generic drugs, outpatient surgery, or other possible actions that economize on the

treatment of a given illness. This would imply either that patients do not ask doctors to keep uninsured expenses down, or that doctors do not adjust treatments toward that goal. Also, if price affects only the decision to seek care, it will be important to supplement quality-of-care analyses with studies of the appropriateness of behavior by the patient in initiating episodes. Studying whether each hypertensive patient episode has enough checkups will miss hypertensive patients who never get into the system. Fortunately, we have criteria for the presence of tracer diseases in the HIE that are independent of care received in the study and are based instead on the entrance screening examination.

The various episode types clearly represent different decisions. Acute and chronic episodes are related to health variables, whereas dental and well care are more related to income and education. Overall price has slightly more effect on the occurrence of dental and acute episodes than on the other kinds. Within the year, the effects of exceeding the MDE are smallest for the acute episodes, as expected. Further analysis of particular subtypes such as pregnancy, or vision and hearing, may reveal more coherent categories, and shed light on scope-of-coverage issues.

The within-year price analysis shows that deductibles may be more effective than expected in restraining demand. For the medium-sized HIE deductibles, unspecific anticipation (where one expects to exceed the small amount of deductible remaining, but has no particular illness in mind) is too rare to be important. Surprisingly, observed sales effects within the year were small. Apparently people do not care enough about reducing the cost of medical purchases to keep up-to-date records and defer deferrable care until they exceed their MDE.

These results support the economic appeal of such income-conditioned, medium-sized deductibles as were offered in the experiment. Two theoretical problems (sales and anticipation) with such deductibles apparently are not important. Sales would show that treatment demands were being delayed from their optimal (from the point of view of health) time of purchase. Both sales and anticipation would reduce the effects of the deductible in keeping demand close to its uninsured level.

A major goal of this analysis was to acquire information that would permit us to simulate the effects of plans other than the 14 experimental plans. In particular, we would like to develop tables or formulae for the expected effects of various-sized deductibles on demand. The results of the analysis are encouraging, because they point to a fairly simple simulation procedure. First, since occurrences of episodes are independent over time, we can use the nice convolution properties of the Poisson distribution, and simply use the characteristics of individuals in a family, and the family coinsurance rate, to compute a family base rate. Because the numbers of different outpatient episode types are correlated over individuals, it is probably better to have only two distinct medical episode generators, one for hospital episodes and the other for all outpatient episodes. For medium-sized deductibles or MDEs, we need not worry about weak anticipation, so we can simply generate episode dates for the family using the initial (pre-MDE) rate, and generate costs for each episode,[1] adding them up to see when the MDE is exceeded. At that point, we can switch to the free-plan rate for the rest of the year.

We cannot yet model small deductibles very well, since we do not yet have enough data to enable us to estimate the effects of weak anticipation.

Only about 20 percent of participants exceeded their deductible or MDE, and often did so late in the year. Hence estimates of how they then behaved are based on much less abundant data than the before-exceeding estimates. Next year, with more years of data, we will be able to estimate the anticipation and sales effects more accurately. For now, we have a solution to the vexing theoretical problem of untangling sickliness and price effects in those who exceed, and a suggestion that sales effects may not negate the advantages of deductibles.

---

[1] We are still uncertain about the relation between numbers and costs of episodes, and whether there is any sale effect on the costs of episodes. However, based on the lack of plan effects on costs, we may assume for now that episode costs are independent of number or time in the year.

Besides analyzing more data, we hope to improve these estimates by developing more powerful techniques for within-year analysis. These will be adapted from multivariate survival methods, in which we analyze the time between episodes rather than episode rates. More data will also permit us to create finer categories of episode types and utilize a finer division of time within the year. For now, we are encouraged that these interim results are both plausible and interesting.

Appendix A

REGRESSION MODEL FOR SQUARE ROOT OF EPISODE COUNTS

In Chap. 5 we use a square root regression model to predict individual annual counts on the free plan. These predictions are then summed over each family to get a family prediction. Here we give the development and results of this regression model. Since ordinary least squares (OLS) is inexpensive and has more supporting software than the negative binomial, it was used in preliminary analyses leading up to the specification of the negative binomial model. We begin with a brief description of some theory about the choice of a transformation for episode counts and then show why we chose the square root. Finally, we give the results of fitting this model.

Because of the possible interest in modeling episode counts in a more conventional way, we report our analysis leading to using the square-root transformation in some detail here.

CHOICES OF TRANSFORMATION

Common assumptions for a linear model are that $y_i = \mu_i + e_i$, where $\mu_i = X_i\beta$ is the expected number of episodes for person i, and that the errors e_i are independent with a common variance. Since even after conditioning on personal characteristics, the numbers of episodes do not have equal variances, we will try to generalize the model in three ways. Our discussion follows the theory given in Pregibon (1979). First, we could try a transformation $y = g(n)$, where n is the number of episodes, with the hope that y would then satisfy the linear model assumptions. Second, we can assume that the error distribution is the one-parameter exponential family with density of the form

$$f(y;\theta) = \exp\{y\theta - a(\theta) + b(y)\}$$

which includes the normal density as a special case, but also includes the Poisson and inverse Gaussian. Finally, we can let $E(y_i) = \mu_i =$

$g^{-1}(X_i\beta)$, so that mean of y is a monotone transformation, not necessarily the identity, of the weighted sum of individual characteristics.

Ideally, one would like a transformation $y = g(n)$ to satisfy a number of conditions. In simple linear regression, the mean of the transformed variable is assumed to be a weighted sum of the explanatory factors (i.e., the transformation $g(\mu) = x'\beta$ is the identity $\mu = x'\beta$). The error should be independent and identically distributed for all people. Precision falls and confidence intervals are harder to compute when OLS is used if, for example, those people with higher means also have higher variance. Moreover, the usual computed significance probabilities in least-squares assume normal errors; skewed error distributions lead to outliers that may have undue influence on the results and bias the t-tests of significance.[1] Finally, since these results will be used to predict numbers of episodes, we would like retransformation to be simple.

To make retransformation simple, we confined our choice of transformation to four candidates:

$$y, \quad (y + 0.375)^{1/2}, \ (y + 0.5)^{1/3}, \ \log(y + 0.5).$$

The square-root and log transformations are the commonest for studying the Poisson, and as described in the text, it appears that occurrences of episodes for individuals are, in fact, independent over time.

After rating these transformations along all the criteria, it appears that $(y + 0.375)^{1/2}$, Anscombe's variance-stabilizing square-root transform, is best. Table A.1 summarizes the theoretical and

---

[1] It can be shown that

$$E(t) \simeq -\gamma/2/n \text{ and } Var(t) \simeq 1 + 2(1 + \gamma^2)/n,$$

where t is Student's t and γ is skewness. These formulas show that moderate skewness has only small effects on t.

Table A.1

EFFECTS OF VARIOUS TRANSFORMATIONS OF THE NUMBER OF OUTPATIENT EPISODES

	Transformations			
Criterion	Y	$(Y + 0.375)^{1/2}$	$(Y + 0.5)^{1/3}$	$\log(Y + 0.5)$
Linearity of link	No	Significantly nonlinear ($t = 1.7$), but after removal of one outlier, t drops to 1.1	Yes	Significantly nonlinear at 1 percent. Test indicates larger Box-Cox λ
Approximate theoretical variance[a]	m	$1/4 + 1/(64m^2)$	$Cm^{-1/3}$	$1/m$
Actual variance	*(b)*	About 0.5 for all m	Slightly smaller for large predictions; 0.3 on 95 plan vs. 0.26 on free	Marked heteroscedasticity
Actual skewness, kurtosis	*(b)*	$S = 0.27$ $K = 0.1$	$S = 0.06$ $K = -0.2$	Negatively skewed
Retransformation	Not necessary	Easy; $\mu^2 + \sigma^2 - 0.375$	$\mu^2 + 3\sigma^2\mu + S\sigma^3 - 0.5$	If lognormal $\exp(\mu + \sigma^2/2)$

[a] Calculations as in Anscombe (1948), assuming $Y_i|X_i$ is Poisson with mean m.
[b] Not computed.

empirical properties of each transform. Not much detail is given for the identity transformation, since OLS is clearly inappropriate here.

The link transformation was tested using the goodness-of-link test (Pregibon, 1980). This test regresses residuals from a nonlinear function of $(X'\beta)$ on residuals from the OLS regression of Y. If the coefficient is significant, then the identity link is flawed. (Note that this assumes the $X'\beta$ form of covariates.) The test also indicates whether the transformed variable should be more compressed or spread out to improve the fit. The test showed that the cube root of n might be preferable to the square root. The log fit was even worse and the test pointed to a less severe transform. The cube-root identity link looked good, as expected. However, further analysis of the square-root link showed that most of the significantly poor fit came from the one significant outlier, a woman with 103 outpatient episodes. When this case was removed, the misfit of the square root was no longer significant ($t = 1.1$).

The variance of various transforms of the Poisson can be approximated using calculations such as those done by Anscombe (1948). These calculations show that heteroscedasticity should not be present for the square root (the constant 0.375 is chosen to take out the 1/m term in the expansion). The variance declines slightly with larger m for the cube-root transform and is inversely proportional to m for the log. These theoretical computations are mirrored by the data.

The variance of residuals from regression of the square root transform is the same for all plans, for other major dependent variables, and for big or small predicted values. The variance of residuals ranged from 0.50 to 0.56 on the five plans, with the free plan in the middle. This variance is the sum of 0.25 (the Poisson variance) plus differences between individuals that are not explained by the regression. (These differences are modeled by the u_k in Chap. 5.) Such differences are roughly balanced across plans. All the other plans had sample variances within one standard error of the free-plan estimate.

The residuals were checked for outliers, skewness, and kurtosis. For both the square-root and cube-root transform, there were no significant outliers. The distribution of square-root residuals is

positively skewed ($S = 0.27$), which is significant statistically ($p = 0.04$), but is not big enough to have much effect on the significance tests. The cube-root residuals are fairly symmetric and the log residuals have marked negative skewness. Finally, retransformation is easy for the square root, and requires more assumptions and work for the cube-root and log transformations.

In his review of this report, Gus Haggstrom pointed out that, for moderate values of λ, the logarithm and square-root transformation are very similar. (The correlation between $\log(n + 1/2)$ and $(n + 3/8)^{1/2}$) is 0.968 for $\lambda = 5$.) He argues for using the logarithm on the grounds of its simpler multiplicative-mean structure. This argument has merit, but we marginally prefer the square-root because of its retransformation properties and traditional role as the variance-stabilizing transformation for the Poisson.

The reasoning and findings in the Chap. 4 model specification for the negative binomial regression model are virtually identical for the square-root model as described above, and so will not be repeated here.

FITTING THE MODEL

Table A.2 shows the regression equations used in predicting numbers of episode for the within-year analysis. Effects of plan are not needed, but otherwise the results are very similar to Table 4.9. The retransformed predictions are given by Y^2 + mean-square error -0.375.

Table A.2

REGRESSION EQUATIONS FOR PREDICTING NUMBER OF EPISODES

Variable	f(Acute)	f(Chronic)	f(Well)	f(Dental)	f(Hospital)
CONSTANT	0.58	0.07	0.35	-0.41	0.60
LINC	0.16	0.09	0.09	0.09	0.01
LFAM	-0.10	-0.06	-0.05	0.05	0.00
NOMD	-0.14	-0.08	-0.03	-0.05	-0.02
LMDVIS	0.20	0.13	0.03	0.05	0.01
HEALTH	0.02	0.03	-0.003	-0.02	0.01
NEWBORN	-0.63	-0.11	0.30	-0.07	0.43
WOMEN	0.23	0.08	0.24	-0.01	0.02
AGE	0.02	0.02	0.02	-0.01	0.00
SQRAGE	-0.26	-0.09	-0.22	0.12	0.00
BLACK	-0.22	-0.05	-0.10	-0.12	0.01
MAXED	0.004	0.007	0.021	0.018	-0.002
N	3343	3343	3343	3343	3343
R^2	0.16	0.25	0.20	0.10	0.11
MEAN-SQUARE ERROR	0.46	0.21	0.13	0.21	0.03

NOTE: For definition of variables, see Table 4.4.

Appendix B

ANALYSIS OF FREE-PLAN OCCURRENCE RATES

To show the trends in free-plan rates over time, we have graphed the number of episodes by quarter. Because the acute and chronic categories are consistent over the years, Fig. B.1 shows only the third year for each. The surge on the first day of the year reflects our accounting convention that episodes carried over from the preceding accounting year are dated to the first day of the next year. The rate of free-plan acute episodes is stable after that, but because of the convention that episodes are dated to their first visit or to the beginning of the year, chronic episodes are more likely to occur toward the beginning of the year.[1]

To study the effects of copayment on episode rates within the year, we will compare pay-plan episode rates at each time point with free-plan rates, because changes in the rate of free-plan episodes cannot reflect any within-year price effects. The rate of free-plan episodes

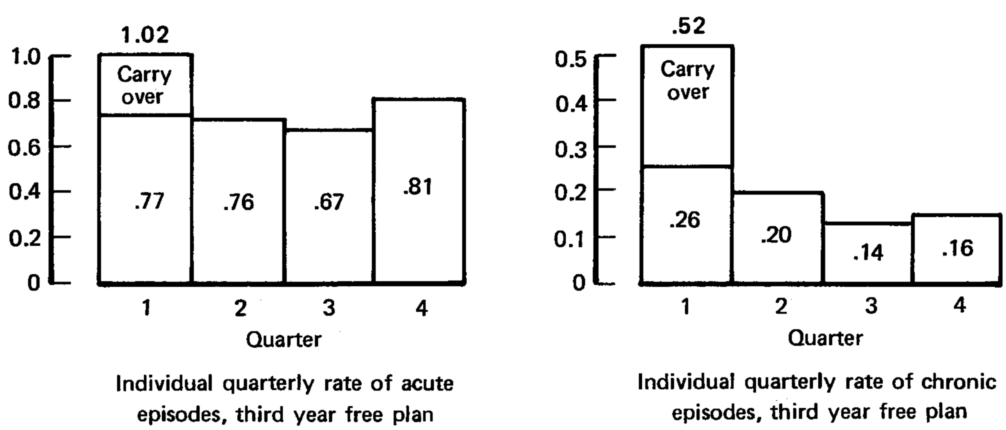

Fig. B.1 -- Quarterly rates over time

---

[1] Since each year is processed independently, unless a continuing chronic episode has a link (such as a prescription date) into the preceding year, it will be dated to its first claim in the new year. If it has such a link, it is dated to the first day in the year.

varies over time because of the way episodes are linked and dated, and possibly because of seasonal effects. By convention, all episodes carried over from the preceding year are dated to the first day, and other episodes (including chronic routine care episodes) are dated to the first claim in a year they are mentioned. A changing rate of free-plan episode occurrence affects estimation, because within-year price estimation essentially compares the observed number in a period with the number expected in that period on the free plan.

It turns out that free-plan episode occurrence rates are uniform over time with three exceptions: the first day of each accounting year, sale effects at the beginning and at the end of the experiment, and a tendency for chronic episodes to occur early in the year.[2] Since we will also show that minor variations in free-plan rates over time have very small effects on the estimates of price effects for the pay plans, we combine data from all three years, estimate the proportion of episodes on the first day, and assume constant occurrence rates for the rest of the year for all episode types except chronic. For chronic episodes, the rate is higher for the first part of the year. For our free-plan chronic episode occurrence rate, we estimate 1.3 episodes per month from days 2 through 60 and 0.7 episodes per month thereafter.[3]

Details of Free-Plan Occurrence Analysis

When the within-year patterns of free-plan episodes are compared for the first three years of Dayton, only chronic episodes have the same pattern over time. Table B.1 gives the chi-square and significance probabilities of testing whether the three years of Dayton have the same pattern. We split the year into the first day and 26 consecutive 14-day periods, and used these 27 categories for the contingency table. There are 52 degrees of freedom.

---

[2] For the free plan, the first day contains 6.9 percent of the acute episodes, 22.2 percent of the chronic, 1.2 percent of the well, and 11.1 percent of the hospital episodes in the first three years of Dayton. These percentages are approximately constant from year to year, and are computed after estimated sales effects have been subtracted out.

[3] This is the best "two-rate" fit using a minimum χ^2 technique. (χ^2 = 81 on 75 degrees of freedom.)

Table B.1

CHI-SQUARE TEST OF SIMILARITY OF WITHIN-YEAR PATTERN FOR FREE PLAN

Item	Acute	Chronic	Dental	Hospital	Well
Chi-square (52)	81	56	73	51	78
P-value	0.007	0.30	0.02	0.43	0.01
Number of episodes	2795	1004	1297	108	817

We will argue that the high chi-square values are primarily due to catch-up demand at the outset of the experiment and to within-family similarity. Since we are interested in comparisons of long-term rates of spending on the free plan, we scale spending at the nonsale rate by subtracting out the sales. Within-family similarity in episode occurrence rates increases the variance of individual counts, but it should not affect expected rates. Indeed, the cells with the high chi-square values appear to be randomly scattered throughout the year.

When the first twelve weeks of the first year and last four weeks of the last year are omitted, neither acute, dental, nor well free-plan rates are significantly different from uniform. Episode rates for these types during this omitted period are 30 to 80 percent higher than average for dental, well-care, and acute episodes, presumably reflecting catch-up demand.[4] By comparison, rates on the pay plans are only 0 to 10 percent higher in these periods than their nonsales average.

The moderately high chi-squares for dental and acute reflect within-family similarities of episode rates that could result from contagion or convenience in combining trips to the doctor or dentist. When one

[4] Compared with the nonsales average (the average excluding the first twelve weeks and last four weeks), the free plan has an estimated excess of 23 percent acute, 63 percent well-care, and 31 percent dental episodes in the first twelve weeks, and 43 percent acute, 34 percent well, and 55 percent dental in the last four weeks. This means an excess of 7 percent acute, 19.4 percent well-care, and 9.6 percent dental episodes in the first year, and 3.3 percent acute, 2.6 percent well-care, and 4.2 percent dental episodes in the last year.

Table B.2

CHI-SQUARE VALUES FOR TESTING UNIFORMITY OF EPISODE RATES UNDER THE FREE PLAN

Item	Acute	Dental	Well Care	(D.F.)
(Omitting sales period)				
All episodes[a]	87	88	71	69
One from each family[a]	30	32	25	34[b]

[a] No values significant at 0.05 level.
[b] Based on 4-week periods.

member from each family is picked at random (to eliminate within-family correlation in the data), the significance probabilities fall dramatically, as do the sample sizes, as shown in Table B.2.

Seasonal variations are a conceivable explanation for the high chi-square values, and their existence would argue for a more refined fitting method for free-plan values. However, the HIE enrolled families in each site over at least three months, so that different families have different accounting years. Thus, actual epidemics are spread out over the family accounting years, greatly reducing seasonal effects. Such effects were tested by splitting the free-plan families at random into two groups and looking at the correlation of episode rates over time. This test showed quite small correlations, which implies that the reduced seasonal (or more properly, "time in family accounting year") effects must be small.

Appendix C

REGRESSION TO THE MEAN AND SICKLINESS EFFECTS

We present our analysis here of the effect of regression to the mean on estimating the sickliness effects referred to in Chap. 5.

Spending before the MDE is exceeded may be a poor measure of the natural propensity to spend, since it includes regression to the mean. In other words, people have high expenses early in the year both because they were sickly and because they were temporarily unlucky. The sickliness may persist, but the bad luck may not. Table C.1 gives some data on regression to the mean for the free plan. Here and elsewhere, we have used the device of imagining a $500 family deductible on the free plan. Since there are no real within-year price effects on the free plan, any estimated effects will be either artifacts of the model or random fluctuations. Table C.1 shows studentized residuals from a regression of the logarithm of bimonthly spending (plus a dollar) on important predictors of spending. These residuals are averaged for people on the free plan whose families went over the imaginary $500 deductible at different times during the year. Note that the spending residuals before the family went over depend, as expected, on how quickly the family exceeded the deductible. The spending thereafter

Table C.1

SPENDING RESIDUALS BY TIME FAMILY SPENDING EXCEEDED $500

Month Family Spending Exceeded $500	Average in Months Before Exceeding $500	Average for Whole Year	Number of People
0-2	1.2	0.5	58
3-6	0.3	0.2	57
7-12	0.1	0.2	69
Never	-0.6	-0.6	91

shows a pronounced regression effect. That is, while there is still a monotone pattern from the presumed sickliness effect, it is much reduced.

Appendix D

MULTIPLE FAMILY EPISODES

Families often begin several episodes on the same day. Indeed, 25 percent of all episodes are a part of multiple episodes. Most of these (70 percent) involve more than one family member, with the remainder generally arising from a person with a chronic condition who has it looked after while having some other problem checked. Multiple episodes are 5 to 30 percent smaller (depending on type) than all episodes on average, probably because severe conditions do not allow for the slight scheduling changes necessary to combine everyone's visits. The frequency of multiple episodes shows that there must be substantial noncharged costs of going to a doctor--either in the actual time and transportation costs, costs in remembering different schedules, or perhaps baby-sitting costs for children who do not go. The patterns observed during the second year in Dayton are shown in Table D.1.

The results in Table D.1 are interesting; they illustrate the effects of the arbitrary assumptions of our expense-grouping procedures. For example, we decided to separate chronic, acute, and well-care procedures performed at the same visit into separate episodes. Consider someone who gets his high blood pressure checked during a visit for an acute condition. The person may decide to schedule care for his chronic high blood pressure at that time rather than at some other (in which case two separate episodes are appropriate). Alternatively, if the acute condition had not occurred, the person might never get the chronic care at all (in which case, one combined episode is more appropriate, although since our methods are not suited to fractional episodes, we would still have to decide what type).

Similarly, siblings having a physical or dental checkup together may reflect only one family decision, but we can imagine that a hypothetical family not concerned with the convenience of joint visits would have checkups spread throughout the period around the day the real

Table D.1

FREQUENCY OF MULTIPLE EPISODE TYPES

Type of Multiple Episode	Frequency (percent)	
Different people	70	
All dental		20
All well		14
All acute		30
Other		6
Same person	30	
Multiple chronic		8
Chronic and other		14
Other		8

SOURCE: Dayton year 2, all plans.

family went. Thus, it is probably best to assume in the overall analysis that each child's checkup counts.

Multiple chronic episodes reflect our decision to make each chronic condition a separate episode. This assumption combines with our decision to treat prescriptions from before the start of the year as starting on the first of the year, to create multiple episodes at the start of the year.

Finally, there are the cases where several family members come in together for acute conditions. In some cases this might represent conditions so mild that, if only one child were involved, the family might not go. Again, to make individual and family methods consistent, it is best to count one episode for each person. The frequency of family-combined acute conditions must reflect diseases that spread within the family.

The main problem with multiple episodes for our analysis is the multiple episodes on the day the MDE is exceeded. If family episodes

were small and independent, the problem of such borderline multiple episodes would be unimportant. In actual fact, the decision of how many episodes to put into the post-MDE period could be important for the estimated second-period price ratio. (Borderline multiple episodes are rare enough that they could have very little effect on the large pre-MDE episode category.) The estimation procedure for the first period conditions on the time to the episode exceeding the MDE, so it seems appropriate to put at least one of a set of multiple episodes in that period. However, that decision means that the pre-MDE price ratio does not exactly measure what the number of episodes would be in a pure coinsurance plan, but is a mixture of that ratio with one from a plan with an MDE. This would be true even without multiple episodes, because of the problem of assigning a large episode greatly exceeding the MDE to the pre-MDE or to the sale period. Economic theory does not deal with such lumpy purchases very well.

Fortunately, episodes are common enough that even putting all of the transitional episodes into the post-MDE period does not affect results very much. Table D.2 compares price ratios under three assumptions: all transitional episodes put into the pre-MDE period (as assumed in the body of the report), all episodes that exceed the MDE put into the second period, and by episodes ranked in descending order of expense, with the MDEs remaining subtracted off one by one, and the episodes that start with the MDE already exceeded assigned to the second period. Only the extreme marginalist assumption that all episodes that exceed the MDE be put into the sale period has much of an effect. The differences between the method chosen, and splitting of multiple episodes, are vanishingly small.

Table D.2

PRICE INDICES WITH DIFFERENT ASSUMPTIONS ON MULTIPLE EPISODES THAT EXCEED MDE

	All in 1st Period	All in 2nd Period	Split
Pay period	1.05	0.99	1.05
Next 3 Months	0.88	1.03	0.89
Rest of year	1.01	1.04	1.01

NOTE: Based on free plan in Dayton year 2.

Appendix E

MULTIPLE HOSPITALIZATIONS

Hospital episodes are much more bunched than they would be if everyone had the same propensity to have episodes and episodes were independent. In three years of Dayton, with 3355 person-years of experience, there were 235 single episodes per year, 34 doubles, 11 triples, and 1 person with 4 episodes. This distribution is clearly not Poisson. A negative binomial with n = 0.311, p = 0.755 (i.e., β = 0.325) fits it well ($\chi_2^2 = 3.0$).

The multiple episodes within a year are usually related but not completely predictable. We have examined the diagnoses at discharge for each case. Of the multiple episodes, 16 were unrelated, 8 were for the same problem but more than two months apart, 17 were for the same problem and less than two months apart, and 7 were completely foreseeable from the first hospitalization. (The numbers do not quite add, because the quadruple episode and one of the triples split into two unrelated parts.) Even in cases where the diagnoses were not related, it often seems as if people were searching for illness, instead of being victims of unrelated unlucky accidents. Thus, as with other types of episodes, it appears that people have propensities (either from condition or temperament) to have episodes and the mixture of Poissons model is appropriate.

The effects of price may be smaller on large expense items. Severe illnesses are often expensive to treat. Moreover, on the 95 plan, one hospitalization is usually enough to put a family over the deductible or MDE. Indeed, 70 percent of all families with hospitalizations in the first 9 site-years exceeded their MDE (Newhouse et al., 1981). Our data also bear out the observation that multistay hospitalizations cost more than average (Zook, 1980). In year 2, for example, 25 out of 40 (nonsingle) episodes cost over \$1,000 when the overall median was about \$800. The distribution of episode size is closer to being lognormal than is the distribution of annual hospital spending.

We could simply study annual hospital bills, ignore hospitalizations, and rely on analysis from the annual model. We will defer a decision on whether to use the simpler annual model until we have a larger number of hospital episodes to analyze.

Appendix F

THE VALIDITY OF SOME ASSUMPTIONS IN THE PRICE EFFECTS MODEL

We now justify some of the underlying assumptions made in Chap. 5 in modeling the rate of occurrence of episodes and its dependence on price. First, recall that we convert episode occurrence rates on pay plans into the "free-plan scale" by using a new time scale t^*, with $t^* = F^{-1}(t)$, where F is the cumulative free-plan episode-occurrence function during the year. Since F is estimated from free-plan data, we examine the effect of errors in estimating F on our other parameter estimates. In particular, we discuss how estimation of the decline of chronic episodes through the year on the free plan might affect other parameter estimates. Second, we test our assumption that episodes occur independently over time. Finally, we discuss the stability of regression coefficients and independent variables over time.

EFFECTS OF ERRORS IN THE FREE-PLAN SCALE ESTIMATE

What are the effects of errors in free-plan occurrence-rate estimates on annual price indices? In Chap. 5, we denoted the true "pay" price index by π, and our estimates by p. We normalized the true free-plan index to be 1. Let μ be the true underlying free-plan rate and σ^2 be the variance of observed annual number of episodes for a randomly chosen person. Let r be the observed rate in our sample. There are approximately 300 people on the free plan in Dayton, so the Central Limit Theorem allows us to assume that $X = r/\mu \sim N(1, \sigma^2/n\mu^2)$, where n = 300 and σ and μ are given by the entries in Table 4.6. As can be seen there, σ/μ is about 1 for all episode types. (Recall that these counts are assumed to be mixtures of Poisson distributions.) Now assume that p is the maximum likelihood estimate (MLE) of π conditional on the free-plan price index being 1. For a given pay plan, think of π as the proportion of the expected rate on the free plan that a person on the pay plan would be expected to have. We are interested in the variance of pX, our estimate of the pay-plan rate. Assume $p \sim N(\pi, \tau^2)$, where the estimate of τ^2 comes from the MLE procedure; and assume that p and X are independent. Then

$$E(pX) = E(p)E(X) = \pi$$

$$\begin{aligned} Var(pX) &= E(p^2)E(X^2) - \pi^2 \\ &= (\pi^2 + \tau^2)(1 + \sigma'^2) - \pi^2 \\ &= \tau^2 + \tau^2\sigma'^2 + \pi^2\sigma'^2 \end{aligned}$$

---

where $\sigma' = \sigma/\mu n^{1/2}$. The term $\tau^2\sigma'^2$ is negligible since τ varies from 0.05 to 0.1 depending on plan and episode type; σ' is about 0.06, and π about 0.8. The two remaining terms have similar magnitude so that $Var(pX) = \tau^2 + \pi^2\sigma'^2$ which is approximately bounded above by .01. The above expression is used for the error calculation in the footnotes to Tables 5.4, 5.5, and 5.8. For example, in Table 5.4, there is a 57 in the upper right-hand corner. The 4 alongside is 100τ, which was taken from the MLE printout. The total error is $\tau^2 + \pi^2\sigma'^2$, which is $0.04^2 + 0.57^2 \bullet 1/300 = (0.05)^2$. Thus, the one hundred times the total error is about 5, or one more than 4.

What are the effects of scale estimation error on estimates of within-year rates? Since the free-plan rate is assumed uniform over the year for all episode types except chronic, the error contributed by X is the same for all types except chronic. The error contributed from the MLE is higher since there is less experience on the pay plans in the subperiods than in the year as a whole. For chronic care, we have estimated the transition from high rates in the early part of the year to lower rates later in the year. Errors in this estimate can add to errors in the annual estimates. In the most extreme case, suppose that everyone goes over the MDE before the transition period. Then the errors in true free-plan rates are all in period two. If these errors are caused by a momentary rise in free-plan rates, then the effect of that rise on variance in the sale period estimates will be four times as great as its effect on annual rates, since the sale period is one-fourth of the year. In fact, since less than half of those who exceed the MDE

do it before the chronic transition point (see Table G.2), the effects of such changes are at worst twice as large on within-year chronic rates as on annual chronic rates.

INDEPENDENCE OF EPISODES OVER TIME

Episodes are said to occur independently over time if their probability of occurrence in a time interval does not vary with how far the time interval is from the preceding episode, and simultaneous episodes occur with probability zero. Independence simplifies statistical analysis. Because of these desirable consequences, the independence of episode occurrences over time was tested for each category of episode early in the analysis. Two tests were used. The episodes on the free plan were grouped by month, with continuing chronic and carry-over acute episodes assigned to month "zero." The first test added episode counts in month zero, and in each of the six bimonths (e.g., January and February comprise the first bimonth of the year). Bimonthly episode counts were transformed by taking log(x + 1) to make them less skewed. Then regressions of the transformed counts were computed with demographic variables of participants as regressors. Finally, the correlation over time of the residuals of these regressions were computed. The results are shown in Table F.1. The average of the diagonal elements is nearly constant. Such a pattern would appear if participants had unmeasured constant differences in the episode occurrence rate but episodes were independent over time, e.g., they did not occur in bunches. By contrast, bunching would make the correlations of points that were closer together in time larger than those that were more distant. This test provides additional evidence supporting another assumption--namely, the persistence of spending patterns through the year.

A second test examined runs of months in which a certain type of episode occurred. For each individual, we computed the number of months having episodes, and the number of runs of consecutive months with the same "sign" (having episodes or not). Having either too many or too few runs indicates short-term dependence.

Table F.1

BIMONTHLY RESIDUAL OUTPATIENT CORRELATIONS ON THE FREE PLAN

	0	1-2	3-4	5-6	7-8	9-10	11-12	Diagonal Average
Month 0 (carry-over)	1	0.25	0.16	0.19	0.11	0.12	0.22	
1-2		1	0.13	0.08	0.18	0.21	0.15	0.22
3-4			1	0.15	0.15	0.12	0.20	0.14
5-6				1	0.15	0.11	0.16	0.17
7-8					1	0.20	0.17	0.17
9-10						1	0.11	0.13
11-12							1	0.17

SOURCE: Dayton year 1.
NOTE: Diagonal averages are computed from upper left to lower right, e.g., 0.22 = 0.22, 0.14 = 1/2 (0.12 + 0.15), etc.

The run-test P-values from each individual are then combined into chi-square values using Fisher's method (Fisher, 1970). Because of the discrete nature of the distribution, Lancaster's correction was used in conjunction with Fisher's combination method (Lancaster, 1949). The combined result, denoted as the <u>combined runs test</u>, is then used to test for the existence of short-term dependence.

The combined runs test was done on data from Dayton year 1. We found that dental episodes on pay plans are negatively dependent, with a one-sided P-value of 0.006. There was no evidence of short-term dependence in outpatient episodes.[1] (We did not separate outpatient into well care, acute, and chronic for this analysis.)

---

[1] We also found that drug episodes had significant positive dependence, with one-sided P-value 0.004. This led to a revision of the matching rules for drugs. After the conditions for a match were relaxed, the number of distinct drug episodes fell, and they were no longer positively dependent.

It should be pointed out that the run test (and therefore the combined runs test) is based on the assumption of stationarity, namely, that the probabilities for having episodes in different months are the same over time. The initial sale effect clearly violates this assumption. However, the test is believed to be robust against the mild nonstationarity caused by the initial sales effect.

Although an individual rarely has two episodes of the same type on the same day, families often have several episodes on the same day. We have described such multiple episodes more fully in App. D, showing that they mainly reflect family convenience, but can also come from contagion. We have tested various assumptions on how to divide multiple episodes that together are enough to exceed the MDE, and have seen that the choice of assumption has little effect. Multiple episodes should also slightly raise the standard error of the estimated price ratios.

Independence and the effects of family dependence will be tested in more detail when we study the time intervals between episodes. For the present, we will abide by the results of these tests, and assume for analysis and prediction that individual episode occurrence times are independent over time.

STABILITY OF REGRESSION COEFFICIENTS OF INDEPENDENT VARIABLES OVER TIME

This stability, together with the independence of episodes, allows us to pool three years of data in the estimation. The stability of regression coefficients was tested by regressing the differences between transformed episodes in the first and second and in the first and third years on the dependent variables. If coefficients are stable, the regression coefficients for the differences should be zero. F-tests were used to check this, and, indeed, neither regression was significantly different from random at the 10 percent level.[2] Such

---

[2] The variance in transformed outpatient episodes explained by enrollment variables falls from 0.36 to 0.28 over the three years. This is somewhat due to changes in such variables as income and health, somewhat due to declining transient effects of initial exams and price, and somewhat unexplained. We plan further analysis to understand this drop.

independent variables as sex, race, age, plan, and family size are extremely stable over time. Changes in such health-related variables as usage in the preceding year and self-assessed health status are endogenous to the experiment. It would be improper to allow their variation over the course of the experiment to reduce apparent experimental price effects.

Independence of episodes within each year implies that an individual's total episodes in three years should also be Poisson. If underlying rates are stable from year to year, these three-year totals are then sufficient statistics for inference about participants' expected rate of episode occurrence. In our results, we look at the 1,074 people who were in the sample all three years and regress the total number of episodes against initial characteristics and experimental plans.

Appendix G

ANALYSIS OF PARTICIPANT MDE ANTICIPATION

Families might anticipate future spending in two ways: They might foresee spending on particular known conditions or, if a large family has only a few dollars left on its MDE early in the year, they may feel sure that they will exceed it. Neither form of anticipation changes our results very much. Spending on particular conditions is already taken into account by the program that links up charges into episodes. The program assumes in its linking and dating procedures that such spending is fully anticipated. The second form of anticipation cannot be important because it occurs so rarely. We now show that for both family and individual plans, times at which participants might expect to exceed the remaining MDE are only one fourth as frequent as times when families have already exceeded it.

We first analyzed the data to see when in fact there was a good chance of exceeding the MDE. As part of this analysis, we define the "remainder" to be the MDE remaining divided by the coinsurance rate. This term is the amount a family must spend in order to exceed its MDE. For example, a family on a 25 percent coinsurance plan with \$200 remaining on their MDE would not exceed it until they had purchased \$200/0.25 = \$800 of medical care. Obviously, the chance of exceeding the MDE increases with smaller remainders and with larger expected number of episodes before the end of the year. To study the exact relationship, and to see the possible effects of other variables, we looked at the situation of families who had not exceeded their MDE in Dayton year 2 at the beginning and at 30, 90, 180, and 270 days into the accounting year. For each of these 1141 family-date pairs, we computed the MDE remainder, their expected number of episodes for the time remaining (by multiplying their annual expected rate by the time remaining), and collected some other statistics. Only 121 (11 percent) of these family-date pairs exceeded the MDE.

To address the anticipation problem, we first looked at methods for predicting whether a family would exceed its MDE given its remainder and its expected number of episodes in the balance of the accounting year. After experimenting with a variety of independent variables and specifications of them, we found that a good predictor of the probability of exceeding the MDE is $0.65 - 0.096 X_1 + 0.015X_2$, where X_1 is the logarithm of the remainder and X_2 is the expected number of episodes remaining. This exercise in prediction led to a classification of the actual data. We used the two variables X_1 and X_2 to obtain Table G.1. The 3/4 in the upper left corner of the table means that in four cases, families had a remainder of less than \$100 and fewer than six additional expected episodes, and that in three of the four cases they subsequently exceeded their MDE. If we use the criterion that families in a certain situation might anticipate exceeding the MDE only if more than half of the families in that situation actually did, we see that families are very rarely in such situations. The upper right section of the table above the solid line shows these families. There, 25 out of 34 families later exceeded the MDE.

Table G.1

FAMILY SITUATIONS WITH THE NUMBER WHO ULTIMATELY EXCEEDED MDE, BY REMAINDER AND EXPECTED NUMBER OF EPISODES

	Expected Number of Episodes				
Remainder	0-6	6-12	12-18	18+	Total
<100	3/4	1/1	4/4	2/2	10/11
101-200	8/28	1/3	2/2	1/1	12/34
201-400	5/76	5/18	4/8	2/4	16/106
401-800	5/126	2/43	2/12	6/8	15/189
1+	7/272	14/251	20/167	27/111	68/801
Total	28/506	23/316	32/193	38/126	121/1141

Put another way, only 3 percent of the family-date pairs were in a position where they might plausibly anticipate going over the MDE in that year.[1] A breakdown of anticipation by plan shows that almost all of these situations occur on the 95 percent coinsurance plan. Only 6 out of the 34 cases occurred on the 25 or 50 percent coinsurance plans. Because the remainder is divided by the coinsurance rate, the lower the coinsurance rate, the less likely families will be in a situation to anticipate.

For individuals, just as for families, positions of anticipation occurred only 3 percent of the time. This is about one-quarter the frequency of actually being over. Participants on the individual deductible plan are somewhat more likely to exceed their MDE than on other plans, but they tend to exceed it later in the year, as Table G.2 shows.

Table G.2

TIME OF EXCEEDING SPENDING LIMIT BY PLAN

	Day of Year Exceeded							
Plan	1	2-30	31-90	91-180	181-270	271-366	Never	Total
Individual deductible	3	4	3	10	3	4	70	97
Family deductible								
95	4	4	8	6	4	2	64	92
50	0	1	3	4	1	1	61	71
25	2	1	0	2	1	3	79	88

[1] Table G.1 shows that 3 percent of family-date pairs anticipate exceeding MDE. However, proportions of families who ever anticipate exceeding MDE will be different. A family that never exceeded is counted four times. Looking at families as the unit will change the numbers in this family-date pair analysis but cannot change the qualitative conclusions. Anticipation still happens rarely.

Table G.3

INDIVIDUAL SITUATIONS WITH THE NUMBER WHO ULTIMATELY EXCEEDED MDE BY REMAINDER AND EXPECTED NUMBER OF EPISODES

	Expected Number of Episodes			
Remainder	0-5	5-10	10+	Total
< 50	6/21	5/5	1/1	12/27
51-100	4/23	5/6	0/0	9/29
101-158	24/275	26/90	4/4	54/369
Total	34/319	36/101	5/5	75/425

Situations of anticipation occur for individuals with high expected numbers of episodes. In the marked upper right corner of Table G.3, 15 of 16 individuals later exceeded the MDE. Here, the expected number of episodes is a better predictor than in the family case since the deductible is small and hospital expenses are not counted. Note that if the annual deductible were reduced below \$100, individuals expecting 5 to 10 episodes could anticipate exceeding the deductible from the beginning of the year.

Appendix H

DERIVATION OF MAXIMUM LIKELIHOOD ESTIMATES

Recall that n_{ik} and t_{ik} are the number of episodes and the time spent by the kth family in period i, respectively. (For the rest of this exposition, we drop most subscripts.) Under all these assumptions, conditional on $u_k\delta_k$, the density function of (n, t) = $(n_1,n_2,n_3,t_1,t_2,t_3)$ is given by P (time to episode n_1 is t_1, n_2 episodes in time t_2, n_3 episodes in t_3) for those who do exceed the MDE, and is given by P(n_1 episodes in t_1) for those who do not.[1] In either case,[2] the likelihood function for family k is given by

$$L_k \propto \int_0^\infty [\Pi_i \exp(-\pi\delta ut)(\pi\delta ut)^n/n!]dF(u).$$

Where the proportionality sign is for factors independent of price and

$$dF(u) = \exp(-u/\beta)u^{\alpha-1}du/(\alpha-1)!\beta^\alpha.$$

---

[1] These probabilities do not depend on how the year is partitioned into three groups. In particular, it makes no difference whether the first period ends because the MDE is exceeded in episode n_1, or (contrary to fact) the insurance plan states that price falls to zero after n_1 episodes, or even if the first period ends after the first episode with a physician whose name started and ended with a g.

[2] The probability that the time to the n_1th episode is t_1 is given by

$$\exp(-\pi\delta ut_1)(\pi\delta ut_1)^n/(n_1 - 1)!t_1$$

which is n_1/t_1 • Prob(n_1 episodes in t_1). The difference is independent of any estimated parameters, so it can be ignored.

Let

$$N = \Sigma_i n_{ik}, \text{ and } R = \Sigma_i \pi_i t_{ik} \delta_k;$$

then

$$L_k \propto \Pi(\pi\delta t)^n \bullet$$

$$\int \exp[-u(R + 1/\beta)]u^{(N+\alpha-1)}/(\alpha - 1)!\ \beta^\alpha du$$

$$\propto \Pi_i(\pi)^n \bullet [R + 1/\beta]^{-(N+\alpha)}(N + \alpha-1)!/(\alpha-1)!\ \beta^\alpha.$$

Thus up to a constant,

$$\text{Log } L_k = \Sigma_i n\log\pi - (N + \alpha)\log(R + 1/\beta) + \log(N + (\alpha-1)!) -$$

$$\log(\alpha - 1)! - \alpha\log\beta.$$

Now we put the subscripts back and differentiate. Letting $\ell_k = \text{Log } L_k$,

$$\delta\ell_k/d\pi_i = n_{ik}/\pi_i - (N_k + \alpha) \bullet t_{ik}\delta_k/(R_k + 1/\beta).$$

$$\delta\ell_k/d\alpha = -\log(R_k + 1/\beta) + \Sigma^N 1/(j + \alpha - 1) - \log\beta$$

$$\delta\ell_k/d\beta = (N_k + \alpha)/(\beta^2 R_k + \beta) - \alpha/\beta$$

These equations can be used to maximize ℓ_k with standard numerical maximum likelihood estimation programs. Since the second derivatives are not constant, we must iterate. As mentioned earlier, δ_k (which is part of R_k) is estimated from the regression in App. A and treated as fixed for the maximum likelihood procedure.

Modifications are made to reflect the time trends in episode rates through the year (caused by carryover from the preceding year, and the episodic dating and linking conventions). If F(t) represents the cumulative propensity function of episode rates within the year on the free plan, we simply change the time scale t_{ik} so that $t_{ik}^* = F^{-1}(t_{ik})$. All analysis goes through as before, provided the multiplicative model $E(n|t^*) = \delta_k u_k \pi_i t_{ik}^*$ holds. The actual transforms used are estimated from the data and are discussed in App. B.

BIBLIOGRAPHY

Abel-Smith, B., and R. M. Titmuss, The Cost of the National Health Service in England and Wales, Cambridge University Press, Cambridge, England, 1956.

Andrews, D. F., and D. Pregibon, "Finding the Outliers that Matter," Journal of the Royal Statistical Society, Series B, No. 40, 1978, pp. 85-93.

Anscombe, F. J., "Transformation of Poisson, Binomial and Negative-Binomial Data," Biometrika, Vol. 35, Parts III and IV, December 1948, pp. 246-254.

Arrow, K. J., "Uncertainty and the Welfare Economics of Medical Care," American Economic Review, Vol. 53, No. 5, December 1963, pp. 941-973.

Arrow, K. J., Optimal Insurance and Generalized Deductibles, The Rand Corporation, R-1108-OEO, February 1973.

Baumol, W. J., and D. V. Bradford, "Optimal Departures from Marginal Cost Pricing," American Economic Review, Vol. 60, June 1970, pp. 265-283.

Belsley, D. A., E. Kuh, and R. E. Welsch, Regression Diagnostics: Identifying Influential Data and Sources of Collinearity, John Wiley & Sons, New York, 1980.

Box, G.E.P., and D. R. Cox, "An Analysis of Transformations" (with discussion), Journal of the Royal Statistical Society, Series B, No. 26, 1964, pp. 211-252.

Brook, R. H., et al., "Overview of Adult Health Status Measures Fielded in Rand's Health Insurance Study," Medical Care Supplement, Vol. 17, No. 7, July 1979.

Cook, R. D., "Detection of Influential Observations in Linear Regression," Technometrics, Vol. 19, 1977, pp. 15-18.

Duan, N., W. C. Manning, Jr., C. N. Morris, and J. P. Newhouse, A Comparison of Alternative Models for the Demand for Medical Care, The Rand Corporation, R-2754-HHS, January 1982.

Ferreira, J., "The Long-Term Effects of Merit Rating Plans on Individual Motorists," Operations Research, September/October 1974.

Fisher, R. A., Statistical Methods for Research Workers, 14th ed., Hafner Press, New York, 1970.

Fowler, W., E. B. Keeler, and J. W. Keesey, The Episodes-of-Illness Processing System, The Rand Corporation, N-1745-HEW, September 1981.

Friedman, Bernard, "Risk Aversion and the Consumer Choice of Health Insurance Options," Review of Economics and Statistics, Vol. 56, 1974, pp. 209-214.

Greenwood, M., and G. U. Yule, "An Enquiry into the Nature of Frequency Distributions Representative of Multiple Happenings, with Special Reference to Multiple Attacks of Disease or Repeated Accidents," Journal of the Royal Statistical Society, Vol. 83, 1920, pp. 255-279.

Haggstrom, G. W., T. J. Blaschke, W. Chow, and W. J. Lisowski, The Multiple Option Recruiting Experiment, The Rand Corporation, R-2671-MRAL, November 1981.

Hausman, J. A., B. Hall, and Z. Griliches, Econometric Models for Count Data with an Application to the Patients' R&D Relationship, NBER Working Paper No. 17, Stanford Unversity, Stanford, Calif., November 1981.

Hinkley, D. V., "On Power Transformation to Symmetry," Biometrika, Vol. 62, 1975, pp. 101-111.

Jackson, O.A.Y, "Some Results on Tests of Separate Families of Hypotheses," Biometrika, Vol. 55, 1968, pp. 355-364.

Johnson, N. I., and S. Kotz, Discrete Distributions, Houghton Mifflin Co., Boston, 1969.

Keeler, E. B., D. Morrow, and J. P. Newhouse, "The Demand for Supplementary Health Insurance, or Do Deductibles Matter?" Journal of Political Economy, Vol. 85, 1977, pp. 789-801.

Keeler, E. B., J. P. Newhouse, and C. E. Phelps, Deductibles and the Demand for Medical Services: The Theory of the Consumer Facing a Variable Price Schedule Under Uncertainty, The Rand Corporation, R-1514-OEO/NC, December 1974; also Econometrica, April 1977.

Kendall, M. G., and A. Stuart, The Advanced Theory of Statistics: Vol. I, Distribution Theory, 2d ed., Hafner Publishing Company, New York, 1963.

Kilpatrick, S. J., Jr., "An Empirical Study of the Distribution of Episodes of Illness Recorded in the 1970-71 National Morbidity Survey," Applied Statistics, Vol. 26, No. 1, 1977, pp. 26-33.

Lancaster, H., "The Combination of Probabilities Arising from Data in Discrete Distributions," Biometrika, Vol. 36, 1949, pp. 370-382; (corrig. Vol. 37, p. 452).

Lohr, K. N., R. H. Brook, and M. A. Kaufman, "Quality of Care in the New Mexico Medicaid Program: The Effect of the New Mexico Experimental Medical Care Review Organization on the Use of Antibiotics for Common Infectious Diseases," Medical Care Supplement, Vol. 18, January 1980, pp. 1-129.

Manning, W., et al., A Two-Part Model of the Demand for Medical Care: Preliminary Results from the Health Insurance Experiment, in J. van der Gaag and M. Perlman (eds.), Health, Economics, and Health Economics, North Holland, Amsterdam, 1981.

Marquis, M. S., Consumers' Knowledge About Their Health Insurance Coverage, The Rand Corporation, R-2753, July 1981.

Morris, C. N., "A Finite Selection Model for Experimental Design of the Health Insurance Study," Journal of Econometrics, Vol. 11, No. 1, 1979, pp. 43-61.

Newhouse, J. P., "A Design for a Health Insurance Experiment," Inquiry, Vol. 11, March 1974, pp. 5-27.

Newhouse, J. P., et al., "The Effect of Deductibles on the Demand for Medical Care Services," Journal of the American Statistical Association, Vol. 75, No. 371, September 1980, pp. 525-533.

Newhouse, J. P., et al., "Some Interim Results from a Controlled Trial in Health Insurance," New England Journal of Medicine, Dec. 17, 1981, 305:1501-1507.

Pederson, K. M., and T. Christiansen, "On the Use of Treatment and Illness Episodes in Health Economics," Odense University Working Paper, Odense, Denmark, August 1982.

Pregibon, D., "Data Analytic Methods for Generalized Linear Models," Ph.D. dissertation, Department of Statistics, University of Toronto, 1979.

Pregibon, D., "Goodness of Link Tests for Generalized Linear Models," Applied Statistics, Vol. 29, 1980, pp. 15-24.

Riedel, D. C., et al. (eds.), Consumer and Provider Behavior and the Use of Health Care Resources; A Comparative Study of Two Health Plans, Health Administration Press, Ann Arbor, Michigan, 1982.

Rolph, J. E., J. M. Chaiken, and R. L. Houchens, Methods for Estimating Crime Rates of Individuals, The Rand Corporation, R-2730-NIJ, March 1981.

Stoddart, G. L., and M. L. Barer, "Analyses of Demand for Utilization Through Episodes of Medical Service," paper prepared for World Congress on Health Economics, Leyden, The Netherlands, September 1980.

Zook, C. J., and F. D. Moore, "High-Cost Users of Medical Care," New England Journal of Medicine, Vol. 302, May 1980, pp. 996-1002.

RAND/R-2829-HHS