A RAND NOTE

Health Economics and Econometrics

Joseph P. Newhouse

August 1987
This publication was supported by The RAND Corporation as part of its program of public service.

The RAND Publication 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.
A RAND NOTE

Health Economics and Econometrics

Joseph P. Newhouse

August 1987
PREFACE

This was an invited paper presented at the 1986 American Economics Association meetings in New Orleans. An abridged version appears in the May 1987 American Economic Review. I am grateful to Victor Fuchs, Emmett Keeler, and Willard Manning for comments on an earlier draft, although they do not necessarily agree with these views. Support for this work was provided by The RAND Corporation and is gratefully acknowledged.
SUMMARY

This Note describes links between health economics and econometrics. It covers flows in both directions, including econometric models, robustness and specification, replication retransformation, and the allocation of subjects to experimental treatments. No new results are presented; the Note is a partial survey.
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREFACE</td>
<td>iii</td>
</tr>
<tr>
<td>SUMMARY</td>
<td>v</td>
</tr>
<tr>
<td><strong>Section</strong></td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Imports</td>
<td>2</td>
</tr>
<tr>
<td>Transshipment</td>
<td>4</td>
</tr>
<tr>
<td>Exports</td>
<td>8</td>
</tr>
<tr>
<td>Concluding Remark</td>
<td>11</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>13</td>
</tr>
</tbody>
</table>
INTRODUCTION

It seems worthwhile to pause periodically and take stock of the balance of trade between economics and health economics. Martin Feldstein (1974) did this fourteen years ago at these meetings, and sufficient time seems to have lapsed to think about the subject again.\(^1\) My purpose here, however, is narrower; the space is not available to update Feldstein's monograph. Instead I discuss the balance of trade between econometrics and health economics.

I focus on two features of health economics:

1. It is an applied field with much policy interest, in part because over 10 percent of the gross national product is spent in health care and in part because public programs account for about 40 percent of American expenditure on personal health care (the percentage is even higher in most other countries). The policy interest has at least two implications: (a) The audience for the health economist will be broad and will include more than simply economists;\(^2\) and (b) the reward for robust results is considerable, because quantitative results may well be used in making choices by both public and private decisionmakers.\(^3\)

2. It has as one of its central concerns health care expenditures, which are very skewed. The highest-spending 1 percent of individuals accounts for more than a quarter of the expenditure and the highest-spending 5 percent accounts for more than half.\(^4\)

\(^1\)Although an entire survey has not been attempted recently, two excellent but focused articles by Mark Pauly cover quite a lot of ground (Pauly and Langwell, 1983; Pauly, 1986).

\(^2\)Of course the audience for many applied fields includes more than simply economists. Whether health economists face a higher proportion of noneconomists is an issue I have not assessed, although I find it plausible in part because of the small number of health economists!

\(^3\)For example, reflecting an earlier literature's estimates of marginal cost as a fraction of average cost, outlier days under Medicare are paid at 60 percent of average cost, but there is considerable question as to whether this is too low a figure (Friedman and Pauly, 1981; Friedman and Pauly, 1983; Pauly and Wilson, 1986).

\(^4\)These results come from the RAND Health Insurance Experiment and exclude spending by the elderly; similar skewness, however, appears to hold for the elderly as well. Victor Fuchs points out that skewness would fall if a time period longer than a year were considered. This is an important observation, but few empirical studies consider periods longer than a year. I suspect longitudinal studies of health expenditures would yield important results, just as the Panel Survey of
The two features are to some degree incompatible. The skewness of expenditures obviously poses problems in obtaining robust results. Moreover, common tools for dealing with skewness make results less accessible to parts of the broader audience.

IMPORTS

Applied fields will by definition import tools and techniques from economic theory and econometrics. Health economics is no exception. Looking back over the past 20 years or so, it is not difficult to find examples of tools developed in other applied fields, or in econometrics per se, that health economists have imported. Indeed, as some of these tools have declined in popularity in other applied fields, they have also declined in popularity in health economics. A good example is the waxing and waning of econometric models of the health care sector (see, for example, Paul Feldstein and Sander Kelman, 1970; Martin Feldstein, 1971; Paul Feldstein, 1973; Yett et al., 1979).

I pause on econometric models because of their prominence in Feldstein's earlier survey. Indeed, Feldstein suggests that the refinement of such models can be used as a benchmark to measure progress in health economics. For example, he says:

Although a consensus has not yet emerged on the correct general specification of a model of the health care sector, a framework for research has been defined....Although most studies have dealt with only single aspects of the health care system, the "invisible hand" that guides researchers to fill existing lacunae is collecting the pieces that will comprise an econometric model of the entire sector. This process is still far from complete.

Income Dynamics and the National Longitudinal Survey yielded important results in labor economics. Such studies would appear to be a promising research target.

Of course, health economics, like all applied fields, makes use of bread and butter tools for single (and small multiple) equation models; by bread and butter tools I mean those discussed in every elementary econometrics textbook. These tools are not my concern here. Virtually all of them were imported long ago and are widespread and unremarkable.
It seems doubtful that Feldstein or many others would write in such a tone today, or, if they did, that they would give more emphasis to the last sentence. Judging from the submissions to the *Journal of Health Economics* and a nonsystematic search of citation counts, few if any economists are trying to build a complete econometric model of the health care sector. Because it meshes with some points I make below, I digress at this point to ask why the bloom is off the rose of econometric models.

Of course, faith in the large macroeconometric models of the 1960s and 1970s has waned generally, so developments in health economics may reflect imported skepticism. But one can be more specific. The stated purpose of most models was not to advance theory but rather to assist in the policy process. To be successful in that aim, however, the predictions needed some tolerable degree of accuracy. Unfortunately, the track record of many models was, in retrospect, none too high. For example, a model of dental manpower estimated in the late 1970s predicted something that was contrary to prevailing opinion at the time, namely, that we were not training too many dentists (Mocniak 1981). Yet subsequent market behavior supported prevailing opinion; the applicant/acceptance ratio at dental schools fell from 2.5 in 1976 to 1.3 in 1985, whereas the number of first-year dental students fell from 5935 in 1976 to 4843 in 1985 (American Dental Association, 1986).

I am not trying to pick on this particular model (many other examples could be found), but I do note that the model’s estimated price elasticity of demand of the mean was -4 (!), which, when combined with the projected spread of dental insurance (a largely accurate prediction) buoyed the predicted demand for dental services and thereby the demand for dentists. One might have thought a price elasticity of -4 would, on its face, have aroused some suspicions about specification or the influence of certain observations, but robustness was not so much in fashion then. If it had been, most econometric modelling efforts may well not have been undertaken.  

---

6Another reason why the bloom is off the rose is that the Lucas (1976) critique of macroeconomic models applies in health as well, insofar as the health models are predicting the effects of policy changes. For example, entry of new manpower (e.g., dentists) is
Besides the large econometric models, there are numerous examples of other imports into health economics, a few of which I note in passing: selection models from labor economics (e.g., McGuire, 1981; Adamache and Sloan, 1982), methods for panel data (Newhouse et al., 1981; Manning et al., 1981; Wilensky and Rossiter, 1981), and latent variable methods (van de Ven and van der Gaag, 1982; and Wolfe and van der Gaag, 1981). Just like econometric models, the popularity of some of these methods may well decline for lack of robustness.

TRANSSHIPMENT

It seems more interesting, though also more presumptuous, to ask not only what econometrics has done for health economics, but also what health economics can or might do for econometrics. Because health economics is an applied field, most trade flows are one way; i.e., true exports are relatively rare. More common is encouragement or amplification that health economists can give to certain trends that arise elsewhere in economics and econometrics. In this section I focus on two developments that health economics might want to transship.

Robustness and Specification

Insofar as the results of studies in health economics are intended to influence private and public actors' decisions, the current emphasis on robustness and specification is of obvious importance. Footnote 7 Under the heading of robustness and specification I group several issues.

Specification Tests. One commonly used specification test is RESET, although it seems underused. Footnote 8 A second classical specification
test is split-sample analysis. Surprisingly, not much use is made of this tool. The Health Insurance Experiment analysis employed it (Duan et al., 1983), and from that experience two findings of more general interest emerged. First, choice of transformation on the left hand side is not innocuous. One can calculate mean expenditure by experimental insurance plan using either raw expenditure or the logarithm of expenditure. In part because of the skewness of expenditure, these two specifications of the dependent variable yield substantially different estimates of responsiveness to plan, although the same data and the same right hand side variables (specified in the same way) are used in both equations (Table 1). Moreover, the standard errors around the estimates using the logarithmic transformation were relatively much smaller than those using the raw dollars, making those estimates superficially more attractive to report. Ultimately the estimates using the raw dollars proved to be more accurate (using the split-sample technique).

Second, analysis of variance (i.e., mean expenditure by insurance plan) had lower mean square error in the forecast sample than analysis of covariance with simple demographic covariates such as age, sex, race, physician visits in the prior year, self-perceived health status, and education of head of household. Put another way, including these common covariates made one worse off in predicting expenditure. This result, which might be incomprehensible to a person who had simply taken an introductory econometrics course, obviously did not happen because these variables are in fact not related to the dependent variable. Rather, the problem stemmed from the skewed distribution of expenditure and overfitting, even in a sample of several hundred observations. That is, the covariates tended to fit extreme observations in the estimation sample and hence did not fit the forecast sample very well.

Neither of these findings would have been uncovered using the standard significance tests of the coefficients to test specification. Indeed, in both cases the standard tests pointed the other way. The t-statistics are more impressive for the logarithmic specification and, of course, the standard demographic variables, when included, are quite

---

9A constant must be added to avoid taking the logarithm of zero. The constant $5, approximately 1 percent of the mean, minimizes the skewness and has been used to derive the values in Table 1.
Table 1
INDEXES OF RESPONSIVENESS OF MEDICAL CARE USE TO INSURANCE PLAN, TWO ALTERNATIVE DEPENDENT VARIABLES

(Free Care = 100; t-statistics on contrast with free plan in parentheses)

<table>
<thead>
<tr>
<th>Plan</th>
<th>First Site Year</th>
<th>First Nine Site Years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw $</td>
<td>Log ($ + 5)</td>
</tr>
<tr>
<td>Free Care</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>25% Coinsurance</td>
<td>109</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>(1.36)</td>
<td>(2.37)</td>
</tr>
<tr>
<td>95% Coinsurance</td>
<td>57</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>(2.99)</td>
<td>(4.30)</td>
</tr>
<tr>
<td>N</td>
<td>841</td>
<td>841</td>
</tr>
</tbody>
</table>

SOURCE: Derived from Duan et al. (1983), Appendix E.
NOTE: Values are unweighted averages of nine site years, using ANOCOVA and 1 part model. I have not calculated standard errors for the fourth column because of the difficulty of correcting for intertemporal correlation, but the same pattern is found in each site year, so t-statistics would be quite significant.

"significant." Hence, both findings underscore the importance of more omnibus specification tests than the usual t-statistics.

Robust Estimators. Another development that health economists should find congenial is use of robust estimators. Robust estimators are similar to transformations as a technique for analyzing skewed data because both effectively downweight extreme observations. Of course, extreme observations can be very informative, but with the skewness of health expenditure data, there is serious danger of overfitting using standard least squares methods. One recent health application of robust estimators is Krasker (1986). Some use of robust estimators was also

Krasker's inference from the use of the estimator is, however,
made in the analysis of the Health Insurance Experiment data
(specifically, they were used for predicting the logarithm of
expenditure for those with inpatient admissions, the most skewed
distribution, which was thicker tailed than lognormal (Duan et al.,
1983).

Replication. Replication is a time-honored method for enhancing
confidence in both qualitative and quantitative findings. One might
suppose that contact with those in medicine and their traditions would
make health economists particularly sensitive to those encouraging
replication (DeWald, Thursby, and Anderson, 1986), but I doubt that the
past record of health economists in this regard is any better than other
applied fields of economics. It is interesting to speculate why there
may be more replication in medical research than in health economics,
assuming for the sake of argument that there is. Possibly in medical
research it is easier for researcher 2 to generate researcher 1's data
without researcher 1's cooperation (by rerunning the experiment) than in
health economics. Of course, this merely raises the question of why
data sharing of the kind Dewald, Thursby, and Anderson (1986) discuss is
not more widespread in economics.

Experimentation

A second development in economics that health economists can
encourage is experimentation. Contact with those in medicine also ought
to make health economists sensitive to experimentation; the randomized
controlled trial is the gold standard of clinical research. And, of
course, health economics can point to examples of such trials, the

peculiar. Kraker is attempting to determine if for-profit hospitals
are more efficient and regresses a measure of cost on several
explanatory variables including a for-profit dummy. When using a robust
estimator, the coefficient on the for-profit dummy is more negative than
when using ordinary least squares. In both cases, however, the standard
error of the for-profit dummy is larger than the coefficient. Kraker
interprets the more negative coefficient in the robust regression as
evidence supportive of a relationship between the two key variables, but
the result is clearly consistent with no relationship.

11For a useful attempt by health economists to replicate and
extend, see Cromwell and Mitchell (1986).
largest being the RAND Health Insurance Experiment (Newhouse, 1974, Manning, Newhouse, Duan, et al., in press). But before dismissing that undertaking as an isolated example, one should consider that there are not only smaller scale experiments (Greenberg and Robins 1986; Burtless and Orr, 1986), but also an increasing number of laboratory experiments (Plott, 1982). For example, Plott and Wilde's (1982) experiment on demand with asymmetric information seems like a promising avenue with which to investigate the issue of supplier-induced demand.

Controlled experiments, of both the field and laboratory variety, have several well-known advantages over observational studies. In particular, by keeping the treatment exogenous and (in general) uncorrelated with other explanatory variables, they yield more precise results than observational studies for a given (finite) sample size. The usually cited disadvantage is cost, but this applies only to field experiments, and even in that case the additional cost of an experiment relative to a prospective observational study is likely to be relatively small.

EXPORTS

I close with two examples of tools developed to solve problems in health economics that appear to have wider applicability. Both came from work done for the RAND Health Insurance Experiment.

Retransformation

Long before anyone heard of health economics, econometricians were using logarithmic and other transformations. Generally the issue of retransforming to the original scale did not arise. For example, if a Cobb-Douglas production function was estimated with the log of quantity on the left hand side, econometricians rarely predicted quantity in raw units. Usually the emphasis was on the elasticities, and one did not need to retransform to estimate the elasticities.

\footnote{In light of the prior social experiments in income maintenance, labor economists might say health economics imported field experiments.}
If, however, one transforms to minimize the effect of skewness (e.g., takes the logarithm of dollars expended), but still wants to know the response surface in dollars, a retransformation is necessary. In this example, if the error term is lognormally distributed, one can exponentiate the predicted log and then multiply by half the variance exponentiated.\footnote{One, for example, might specify a wage equation with the logarithm of wages on the left hand side but want an estimate of the difference in dollars between the wages of two groups.} Estimates using this formula, however, are quite sensitive to seemingly minor departures from lognormality (because those departures are exponentiated).

Duan (1983) developed an estimator for retransformation (the "smearing" estimator) that is both easy to use and does not require parametric assumptions on the error term. It does, however, require that the error term be independent of the explanatory variables. Because this condition is required for consistent estimation anyway, it does not seem particularly onerous. Moreover, when the error is lognormal, the estimator loses little to the parametric estimator. The smearing estimator is now beginning to be used (Ohsfeldt and Culler, 1986).

The Allocation of Subjects to Treatments in an Experiment and the Choice of Units to Sample in an Observational Study

The classical method for allocating subjects to experimental treatments is simple randomization or randomization within strata and blocking designs. For the design of the New Jersey Negative Income Tax Experiment, two econometricians, John Conlisk and Harold Watts, developed a new method for allocating the sample to treatments (Conlisk and Watts 1969). Conlisk (1973) describes an alternative decision-theoretic approach to that described here if the designer is willing to consider a finite number of possible response surfaces.

Conlisk and Watts began with the premise, natural to economists, of optimizing subject to a budget constraint. The objective function they chose was to minimize a function of the variance of the regression coefficients they intended to estimate from the experimental data. To
some degree their model can be seen as a generalization of the optimal design for experimental data when those data will be analyzed using analysis of variance; in that case the optimal number of subjects (observations) to assign to each design point (negative income tax plan) is proportional to $\sqrt{w(i)/c(i)}$, where $w(i)$ is a weight reflecting the interest in the $i$th design point and $c(i)$ is the marginal cost of another observation at the $i$th design point.

The novelty of Conlisk and Watts's method lay in exploiting the dependence of the relative cost of an observation on demographic characteristics. In particular, families with higher incomes were relatively less costly to enroll in the more generous negative income tax plans. Conlisk and Watts specified a simple additive (main effects) model in the parameters of the negative income tax plan and income (as well as other covariates). As a result, the model yielded an unbalanced design, placing higher income families disproportionately on the more generous plans.\textsuperscript{14} Had Conlisk and Watts specified a fully interacted equation, the model would have yielded a nearly balanced design; used in this way it would have approximated a traditional stratified or blocked design.

In the interest of robustness, the Health Insurance Experiment wanted a balanced design. It could have used stratification or blocking, but these techniques had two drawbacks: (1) For computational reasons one could stratify or block on relatively few dimensions; and (2) continuous variables such as income or age were grouped into discrete intervals, hence, within group variation was lost.

The Finite Selection Model (Morris, 1979) begins with a (finite) list of persons to be allocated to plans; the intent is to allocate the list in such a way that the distribution of characteristics on each plan is similar to the distribution on every other plan. The characteristics can be treated as continuous variables (if they are continuous), and they can be weighted in importance (i.e., the distribution of some

\textsuperscript{14}A further problem arises if one oversamples on a variable that is an imperfect measure of the variable of interest (e.g., current income as a measure of permanent income). In such a case a sufficient degree of oversampling can be costly, even if one is interested only in the favored group (Morris, Newhouse, and Archibald 1979).
characteristics could be made more similar than other characteristics). For practical purposes one could take account of as many characteristics as one wanted. In the case of the Health Insurance Experiment, the design obtained using the Finite Selection Model yielded standard errors that were on average about 25 percent less than those that would have been obtained from simple random allocation.

The model can also be used in observational studies to choose which units to sample. Used in this fashion, it chooses a sample that most closely represents the distribution of the population along dimensions specified by the analyst; for example, which 50 metropolitan areas most closely represent the universe of metropolitan areas. Thus, the model affords protection against the possibility that simple random sampling yields an unrepresentative sample through bad luck.

CONCLUDING REMARK

Health economics can draw on economic theory and the tools of econometrics, as well as the empirical traditions of biomedical and clinical research. This may at times lead to a certain schizophrenia in trying to write for different audiences, but in the long run should prove to be an advantage in advancing knowledge.
REFERENCES


Pauly, Mark V., and Peter Wilson, "Hospital Output Forecasts and the Cost of Empty Hospital Beds," Health Services Research, Vol. 21, August 1986, pp. 403-428.


