Paul J. Feldstein and W. John Carr, "The Effect of Income on Medical Care Spending"

Gerald Rosenthal, "The Demand for General Hospital Facilities: A Principal Components Analysis"

Jerome Rothenberg, Northwestern University

A. Feldstein-Carr Study

I shall discuss the Feldstein-Carr paper first. This is a deep, well-thought out study. The field of medical economics could use many more like it. I have three comments to make, two on the permanent income adjustment, the other on the interpretation of results.

1. Appropriateness of Permanent Income Adjustment

A central part of the Feldstein-Carr approach is to regress demand for medical care on an income variable (among the several variables) which represents permanent income - i.e. a variable from which as much of the transitory component of income has been removed as possible. The removal of the transitory component from income has a most significant upward effect on the estimated income elasticity of demand. Yet one can ask whether this adjustment is warranted after all.

The rationale of the adjustment is the Permanent Income Hypothesis, which holds that consumer expenditures are more closely associated with permanent income than with current measured income. Current measured income is deemed to contain both a permanent component and a transitory component; and the correlation between consumption and transitory income is assumed to be zero. So the measured association between consumption and current measured income is an impure reflection of the primary relationship - that between consumption and permanent income.

But in this formulation, consumer purchases of durables - i.e., consumer investments - are not considered a part of consumer expenditures. They are considered rather a portfolio management decision, affecting the form in which consumers hold their wealth, and only the flow of services from these durables during the accounting period is treated as current consumption. The portfolio decision is deemed to be significantly affected by transitory income. Consumer purchases to develop productive skills (as, for example, through education) are generally considered to represent one form of consumer investment - an investment in human capital - and would therefore warrant treatment as portfolio decisions. But so too is at least part of the investment in health. All medical care is directed toward future well-being. Of course, most of the consumption of so-called "necessities" have a future orientation too, in the sense that grossly inadequate provision for food, clothing or shelter will prevent a consumer from having any future. But besides this sense, some medical

care is so clearly predicated toward future productivity and well-being, with long-lasting effects (as, for example, correction of defects, preventive care, direct life-saving services), as truly to constitute another form of consumer investment in human capital. This means that purchase of such care should be sensitive to transitory components of income.

If this is so, then it is indeed total measured income, with both permanent and transitory components, that is more closely determinative of the demand for medical care than is permanent income alone. Thus, for at least some portion of total expenditure on medical care, the Feldstein-Carr adjustment to permanent income is unwarranted. The appropriate income variable may well be current measured income after all; and with it, the substantially lower income elasticity of demand.

2. Approximation of Permanent Income

If, notwithstanding my first comment, the authors wish to use permanent income as their income variable, this can be accomplished somewhat more directly and with more flexible results (since, unlike their present methods, this can be applied to individual families). Instead of, or in addition to, the averaging techniques used in the paper to eliminate transitory elements in income, permanent income can be approximated by using total per family consumer expenditures (after subtracting out consumer investments and adding in the value of imputed consumer services from the stock of consumer durables). Under the Permanent Income Hypothesis, total annual consumption is asserted to be proportional to permanent income.

3. Interpretation of the Income Elasticity of Demand

In order to interpret the paper's results, the very concept of income elasticity of demand for medical care must be given deeper consideration than is accorded it in the paper. The incidence of illness - which generates the "need" for medical care - is generally highly irregular among families at every income level, showing more variation within income levels than between income levels; and introducing significant slippage between income and demand despite any tendency for the <u>perception</u> of "need" to be related to income level. Moreover, in no other consumer sector is consumer sovereignty on how to meet the "need" so substantially delegated by the consumer to someone else, nor is the resultant provision of care so apparently income-neutral. Income-tailored features of medical care seem to be relegated to

peripheral frills rather than to the essentials of treatment. Further, serious illnesses seem generally to receive treatment, regardless of the income level of the patient, through an important tradition of charity, sliding scales of payment or public provision of care.

Given this background, to be told that the income elasticity of demand for medical care is unity is surprising. It implies that a constant proportion of (permanent) income is spent on medical care. While this is not inconsistent with the background characteristics cited. it represents a knife-edge regularity that is not obvious, to say the least. To know the significance of these results for planning medical needs, or even just to understand consumer behavior, we must know much more about the composition of the demand for medical care by different income groups. (The income elasticity here is a crosssectional one referring to income differences among different families, not a temporal one referring to income changes for the same families. It may be quite amiss to predict temporal elasticity from the cross-section elasticities, because of some of the issues to be mentioned now.) We must know, for example, some of the following:

a. To what extent does the estimated elasticity reflect simply price discrimination, whereby prices charged differ for recipients of different income levels, so that the total cost of services varies by income level far more than does the quantity or even quality of services? To what extent are actual quantities or qualities of service, involved? To what extent do commodity packages going to different income groups differ chiefly in quasi-medical adjuncts of care - such as larger or more sumptuous hospital rooms or more privacy of care - rather than in the more directly medical attributes of the package?

b. To what extent does estimated elasticity reflect different compositions of care - as for example, a larger share of patent medicines for lower income groups, a larger share of physician and hospital care for higher income groups? The relative effectiveness of different bundles that is, the "real" amount of care received - may well be related to these different configurations.

c. To what extent does the elasticity reflect income-induced differences for some kinds of illnesses, but not for others? For example, do patients at different income levels get much the same care for "serious", non-discretionary ailments, with significant differences showing up chiefly for "non-serious", or discretionary problems (cosmetic operations, corrective as opposed to ameliorative, dental care - orthodontia, etc.)?

d. Is the estimation subject to measurement bias? For example, higher income groups typically obtain psychiatric care in the form of private financing of psychiatrists and private hospitals. These expenditures would be included in the Feldstein and Carr figures. Lower income groups typically obtain psychiatric care (admittedly not the same kinds of treatment, nor for the same composite of illnesses) in publicly supported mental institutions. The cost of these services would be excluded from figures on private medical care expenditures. This kind of systematic difference would produce an upward bias in the computed income elasticity.

B. Rosenthal Study

This is an imaginative, thought-provoking study. It attempts to reduce the dimensionality of an explanatory cross-section schema in order to perform time series analysis, where data availability is a critical barrier. An important collateral goal is to reduce the possibly substantial multicollinearity of the original schema. While the method and some of the findings are suggestive, some of the procedures used raise serious questions.

1. Dimensionality and Multicollinearity

The dimensionality of the model is reduced by finding five orthogonal principal components. Yet these components are not themselves used as explanatory variables in the reduced model: only one original characteristic with high factor loading (out of the twelve constituents) is chosen as a proxy for each component. One can raise the following questions:

a. The factor loadings for each principal component are not so <u>uneven</u> that any <u>one</u> factor clearly and exclusively represents it. Moreover, no single factor captures what appears to be the broader sense of any of the principal components.

b. While the system composed of the five principal components as explanatory arguments minimizes multicollinearity, since the components are orthogonal, the set of proxy variables selected from the principal components is not necessarily orthogonal; so the system employing these as arguments does not necessarily exclude multicollinearity. Specifically, the loadings on each principal component are generally quite low and diffuse; and a factor selected to proxy for one component will sometimes have significant loadings on other components (especially urbanization and education). Very likely the education, low income and urbanization variables are related; likewise, old age and marital status. Thus, in fact the reduced model may well have retained a significant degree of multicollinearity. This is an empirical question. It would be very helpful to have a correlation matrix for the explanatory variables in order to evaluate the seriousness of this problem.

c. There are devices open to Rosenthal that might decrease multicollinearity and give better explanatory power. Contrary to Rosenthal's statement about the principal components ("Since these influences are statistical creations, however, they can not be measured directly." Page 8.), they can indeed serve as arguments in a reduced model. Each is a specific linear combination of the twelve original factors and can be measured as a separate variable by measuring its constituents. Their use by Rosenthal is, however, precluded for time series analysis because of the unavailability of data. To bypass this barrier, however, a modification of the procedure can be employed: linear combinations of various representative <u>clusters</u> of three or four constituents from each principal component can be used, selected for data availability. This could give much better coverage to the sense of each principal component.

2. Choice of Variables

Within Rosenthal's method, his choices of proxy variables are frequently questionable. Consider some details.

a. Principal Component 1 (C_1): Low income was chosen. It had high loadings in both years but - and this is true of almost every other factor in this component - these loadings have different signs in 1950 and 1960! To interpret "low income" as representative of something that is claimed to have the sense of "economic prosperity" is most peculiar when this proxy has a high positive association with C_1 in 1950 and just as high a negative association with it in 1960.

The characterization of C₁ as an "economic prosperity" variable is not apt. Price, <u>high</u> income, young age and urbanization also have high loadings (with opposite signs between the two years). Yet, by a priori reasoning, price should have a different kind of marginal influence on hospital demand than the two income variables; and these latter two are themselves not nearly perfectly correlated: it is perfectly conceivable to have both a high percentage of the population with incomes above \$5,999 and with incomes below \$2,000. Further, the very use of these absolute figures to represent distinctive situations is itself suspect, since these absolute income levels have a very different "real income" significance for hospital demand between 1950 and 1960, because income and hospital costs changed appreciably during the period. If anything, these two income variables characterize not the degree of prosperity but the population structure within each state. This is an example of my earlier argument about the desirability of using clusters of constituents as individual variables in order to get better representation.

As a compromise here, however, one could use per family income as the proxy variable. This would incorporate information from both income variables used by Rosenthal.

b. Principal Component 4 (C_{μ}): The urbanrural variable was chosen. This is referred to by Rosenthal as reflecting the degree of crowdedness in housing. Such a characterization is misplaced. The actual association is poor: rural slums have as much crowding per dwelling unit as urban slums; perhaps more, since **rura**l family size is probably larger. Indeed, the family size variable seems the better representative of crowdedness, if crowdedness is what is wanted.

Yet crowdedness is probably not what Rosenthal wants after all. In his analysis he refers to this variable as indicating the existence of alternatives to hospital care - but not so much for convalescence, where crowdedness is relevant, but for the availability of treatment services. Degree of urbanness is relevant to the availability (or accessibility) of hospital and other medical services. Yet this fact leads to a more serious problem. The supply of hospital accommodations is related to the urban-rural ratio. Moreover, there is good reason to believe that the supply of hospital services influences hospital use. Thus, inclusion of this supply side of the market begs the whole identification question. Introducing a variable associated with supply into a demand function, without also explicitly introducing a separate function explaining supply, makes it extremely difficult to interpret the estimated function as either a demand or a supply function. It is in fact neither, but only a composite.

Thus, the urban-rural factor - which shows no significant relationship in the regressions anyway - should be excluded to preserve identification of the estimated function as <u>demand</u>; its place should be taken by either the family size or race variable, both probably a better proxy for crowdedness anyway (and crowdedness <u>is</u> a relevant argument in a demand function).

3. Interpretation of the Results

Generally, only the coefficients explaining total patient days utilization are statistically significant. With two exceptions, no coefficient referring to admissions or length of stay is significant. (The exceptions are the old age and marital status variables for admissions in 1950.) Yet non-significant coefficients are subjected to heavy interpretation as though their signs were meaningful. The urban-rural variable is so used for all three dependent variables; most of the others are used for admissions and average length of stay. Rosenthal's interesting - and by no means far-fetched - conjectures are nometheless based on non-significant findings. They are pushed much too hard.

On the other hand, no attention is paid to the extreme instability of the otherwise splendidly significant partial coefficient for the effect of education on total patient days between 1950 and 1960. It changes from 23.13 to 2.15! Some thought to the interpretation of this finding is called for.

4. Effect of Economic Constraints

Rosenthal conjectures that "economic constraints" and the **availability** of substitutes affects hospital utilization via the average length of stay rather than via number of admissions. This throws some light on my discussion of the interpretation of the results in the Feldstein-Carr paper. While the conjecture is, as I have argued, not backed up by statistically solid evidence, it is consistent with my comments on the background of the demand for medical care. It suggests that the extra medical care spending by higher income groups is not due to a greater "need" (incidence) but to greater "quality" or quasi-medical concomitants of care for each medical condition. This suggestion must of course be qualified, not only by the lack of statistical significance of the Rosenthal evidence, but also by the fact that the income variables used in the two studies are quite different. Nevertheless, a sign that these studies are useful, whatever their difficulties, is that the reader is sorely tempted to confront each with the other to generate a dialogue. So progress may be generated.