Irving Leveson and Milton Brafman\* The RAND Corporation

## Introduction

Past studies of the demand for medical care have focused on medical care as a whole or on hospital inpatient care. However, ambulatory medical care has been growing rapidly. Indications are that it will continue to grow rapidly as we pursue policies intended to promote alternatives to expensive inpatient care. There is, in addition, a widespread belief that ambulatory medical care may have a particularly high payoff in improved health. Until recently, work on the geographic distribution of medical care has centered on supply and/or on large areas, except for case studies providing results in forms which are difficult to generalize. With heightened interest on health services for the poor and concern with reaching ghetto populations, analysis of health care for small areas takes on increased importance. The demand for ambulatory care for neighborhoods has been neglected in economic analyses. Programs will increasingly be judged in the future by how successfully they reach their target populations. and in order to improve outreach, demand studies are needed. Hopefully these studies will throw light on the demand for other social services as well.

Significant developments have taken place in thinking about demand determinants in recent years. Furthermore, there are important differences in the role of variables influencing a particular type of service at the neighborhood level from their effects on broader aggregates. The choice of medical services can be treated as an investment decision in which the costs of experimentation with new sources of care weighed against the expected benefits. Both private and social benefits are relevant for government decisions. Health status is best considered explicitly as a determinant of the demand for medical care, rather than being partially represented by other variables categorized as "tastes". With the growth of Medicaid and

\* Any views expressed in this paper are those of the authors. They should not be interpreted as reflecting the views of the RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors.

The material in this paper was prepared while at the New York City Health Services Administration. It was incorporated in a report entitled <u>Medical</u> <u>Care and Income, An Economic Analysis</u>, published by the Gouverneur Economic Research Project. The effort was financed under Public Health Service Research Grant No. ROI CH00239 and New York City Health Research Council Contract U-1153, with additional support from the Carnegie Foundation. We would like to thank Dr. Paul Densen, Dr. Nicetas Kuo, Mrs. Ellen Jones and Jerry Weston for making materials available. Sid Binder contributed freely of his time in facilitating the data processing activities. other forms of "insurance", the traditional demand determinants of income and money price lose importance. Great weight falls on the role of non-money "price" variables, which may have a great impact on accessibility of care, such as opportunity costs of the consumer's time, inconvenience due to factors such as limitations to physical mobility, and consumer information. At given levels of other variables, "quality of care" becomes a "price" variable also.

The statistical analysis considers an entire population - the residents of a housing project for the elderly in which a clinic was established. Determinants of whether or not people use the clinic are analyzed in Chi-Square tests and a dummy variable multiple regression analysis is performed. Some evidence from other clinics is also examined.

The availability of services does not necessarily mean that they will be used, even if properly tailored to the needs of the community. Information about the determinants of demand for medical care in facilities will help in predicting utilization and identifying needs, and hopefully also in devising ways to encourage utilization. Here we concentrate on the primary aspect of utilization -- whether or not services are used at all. One of the methods of organizing medical care, which has great potential for reaching the community and economizing on costly inpatient and nursing home care, is the provision of services in clinics established in housing projects for the elderly. One such unit, established with these points in mind is the Queenbridge Health Maintenance Service (QHMS), which was set up in November of 1961 by the New York City Departments of Health, Hospitals, Welfare and other agencies. The unit offered a variety of services, including home care without charge to the 1,400 residents of Queensbridge Housing Project in Long Island City, New York. Extensive efforts were made to inform residents of the on site availability of free services which previously required travel of four miles or more. An evaluation of the project has been conducted by Kuo<u>et.al</u>.<sup>1</sup> Here the determinants of whether or not persons registered for the service are analyzed. Particular attention is paid to the role of health status in the demand for medical care, to separating the effects of income and education in disaggregated data, and to non-money "price" and information variables.

## The Data and Methods

The sample consists of the 1,219 of the approximately 1,400 residents of the Queensbridge Housing Project for whom information has been obtained between January 1963 and the Fall of 1964. Only persons age 60 and over are included since use of the clinic was limited to that age group. Of the total, 638 had registered (participants) with the QHMS and 581 had not. As a means of separating the effects of health status from other variables, a subsample of 343 persons who reported that they had heart conditions is examined. The group of heart patients is of particular interest in view of the emphasis of the Regional Medical Programs<sup>2</sup> and is large enough to make statistical analysis meaningful. Among the heart group, 201 persons participated and 142 did not. Information on non-registrants was collected in the Fall of 1963 in interviews conducted by the National Opinion Research Center, while registrant socioeconomic data was collected at various dates at the QHMS. The initial groups were interviewed at the time of registration. It is therefore possible that the information obtained is differently influenced by degree of satisfaction with the clinic or knowledge gained about one's health.

Variation in the percentage of persons participating in the clinic across socioeconomic groups is examined. In one-way classifications, for both the total and heart groups, we try to refute the null hypotheses that the percentage participating does not vary among socioeconomic groups by Chisquare "goodness of fit" tests. Specific hypotheses concerning the variables examined are introduced at that time and the effect of the control for health status is considered for each variable.

A multivariate analysis of the heart group is then performed using the dummy variable multiple regression technique. All variables take on values of zero or one. For example, the dependent variable is coded 1 for participants and zero for nonparticipants. For variables which cannot appropriately be dichotomized, a series of dummy variables replaces a single variable. Each variable in the set represents a class interval of the continuous form, and observations are coded 1 or 0 depending on whether or not the value of the variable falls in a given interval.<sup>3</sup> One variable is omitted from each set of dummies and becomes the base. The regression coefficient for any dummy in the set then indicates by how much each class interval of the independent variable differs from the base group in its value of the dependent variable.

The chief disadvantage of the dummy variable technique is that it does not conform with the assumption of homogeneity of variance of the regression model. This still enables estimation of unbiased regression coefficients when other regression problems are absent, but casts doubt on the validity of the usual tests of statistical significance. For this reason, the analysis utilizes a Chi-square test of the significance of sets of dummy variables which depends only on the regression coefficients and not on the distribution of the regression residuals.<sup>5</sup> As in the oneway Chi-square tests, the "expected" values are the numbers of participants and non-participants that would be observed in each class interval of the independent variable if the average percentage participated in each class interval. Instead of comparing these with true actual values, however, pseudo-actual values are derived from the set of regression coefficients for a variable. These differ from the actual values in that other variables are held constant. It should be noted that in the cases in which the Chi-square test was compared with the F-test, nearly the same levels of significance were indicated.

The dummy variable technique has certain advantages over use of continuous forms. While it is especially suited to analysis of survey data, its use for this purpose has been extremely limited. Some of these advantages are great enough so it would pay at times to create sets of dummy variablesfrom continuous forms. For these reasons it is worth stating the advantage clearly. There is no need to make an assumption of linearity, and it is easy to fit functions which are nonmonotonic. By cross-classifying variables, tests for interaction can easily be made without the usual problems of high multicolinearity. Since all observations are used, it is possible to avoid many of the limitations of cross-classification. In effect, small sample rows and columns are pooled on the assumption that the effects of variables are additive except where specific interaction terms are introduced. Another advantage is the ability to retain information for persons with unknown values of certain variables but known values of others, by including an "unknown" category in the set of dummies. Since survey data often come in class intervals, measurement error associated with a choice of midpoints is avoided. Also, since errors in the data may be concentrated at certain values of a variable, this will show up as a high standard error for that coefficient, so that it may be possible to reject one or more coefficients without losing information on the remaining set of dummies or obtaining a biased coefficient for the entire variable.

#### The Results

Table 1 shows the percentage participating by level of each variable, for all residents and for those reporting heart trouble. Table 2 indicates the levels of statistical significance in the test of whether the percentages differ from equality among levels of any variable. The percentage participating is slightly higher in the heart population than for all residents, 58.6 percent compared to 52.3. This is expected since a group known to have any given health problem will tend to have poorer health than the overall population, some of whom have no health problem, and because of the existence of a health problem contributes to the decision to seek medical care. Similarly, among the total population, older persons are more likely to be participants because they will tend to have more health problems. However, among the heart group, age is not significant. This may be because aging does not produce a worsening of heart problems which requires more medical care, the care is not sought, or when conditions become more serious it is sought in places other than the clinic. Another way of looking at the effect of presence of a heart condition is to note that

younger persons with heart conditions are more likely to use the clinic than others, but at older ages where a greater percentage of persons will have health problems requiring medical attention, the two groups participate about equally. A further indication of the role of health status is given by self-rating of health, which similarly is to age, is highly significant for the total population but not significant for heart population. Sex differences are small and not statistically significant.

The data for the total population showed greater participation for whites even though non-whites might be less able to afford alternative sources of care, and a substantial proportion of the clinic personnel were non-white. It was hypothesized that this too could be attributed to health status, since non-whites surviving to age sixty or more would have to be hardier relative to their initial cohort than whites who are likely to have had more medical care when needed and a better environment. Surviving non-whites would tend to have fewer health problems and, if this were the case, the color difference would not appear among the heart group. In fact, the color difference is very much smaller for the heart group and is not statistically significant.

Not all effects of health status are to increase the demand for medical care. Mobility limitations clearly reduced participation even though home visits could have been arranged. Again the effect of health status can be associated with the presence of a disabling condition since within the heart group mobility was not significant. Mobility status can be considered a kind of a price variable, indicative of the amount of inconvenience or discomfort of getting about, or of the costs that would have to be incurred to avoid that discomfort. Another "price" variable which was included in the analysis is travel time to the person's usual source of care prior to clinic registration. Participation was greater among persons having to travel an hour or more in both groups. However, differences were not statistically significant, perhaps because the variable is poorly measured. There is some indication in the results that health insurance coverage leads to greater use of alternative sources of care, as its effect on relative prices would suggest. To some extent, both health status and price effects probably account for the large differences in participation of persons classified according to the usual source of medical care they had prior to the availability of the clinic. It is disturbing that over one-third of the total population and over one-fourth of the heart group both did not participate and did not report any prior usual source of care.

It was hypothesized that the higher the income of a person, the better he could afford to purchase medical care from alternative sources if he believed it necessary, more appropriate, of higher quality or otherwise desirable. Since much of the income for the aged represents pension payments, we probably have a measure of lifetime as well as current income. Income turns out to be a very powerful variable for both the total and heart groups.

It was possible to test the effects of several variables reflecting various aspects of the availability of information to the project residents. Both marital status and number of people in the household were intended to reflect the number of instances a person heard about the clinic, while marital status might also have operated through concern for the health of one's mate. However, neither was significant. Employment status, on the other hand, intended to reflect the awareness of alternatives, and willingness and ability to operate in the outside community, did prove important. However, this may also reflect good health and less need for medical care. Both length of residence in the project and whether foreign born, intended to reflect knowledge of alternative sources of care, were statistically significant.

It was expected that persons with more formal education would be more likely to participate because of greater concern for the future, receptiveness to information, and other factors. However, the education variable was not found significant in the univariate tests. The dummy variable multiple regression analysis of the heart group generally produces results which closely mirror those already presented. One important change was expected, however, which did materialize. Education and income tend to be highly positively correlated with each other, so that education would tend to reflect the effects of income. A zero effect of education could be produced as the resultant of a negative effect of income on participation and a positive effect of education. Table 3 presents the regression results, and the levels of significance are given in the last column of Table 2. The adjusted regression coefficients, calcu-lated by a method given by Melichar, 6 indicate the deviation of percent participating in each group from the overall mean. When income and other variables are held constant, education is seen to have a large positive effect on participation which is statistically significant.

In summary there is evidence that health status is a major determinant of whether people seek medical care. Differences in care can be discerned even when the patient himself provides broad diagnostic information. Furthermore, the tests suggest that important variation in medical care sought can be isolated by simply classifying persons as to the presence or absence of a condition, without information on its severity. The success of crude self-classification may be that such a concept influences the patient's attitudes which are relevant for behavior, though one's own evaluation may not be as well related to functional status as a physician's classification. Income appears to be a powerful determinant of patterns of care. The effects of price variables such as mobility status also appear to be important. Information variables differ greatly in their impact but as a whole seem to contribute substantially to the observed variation in use and sources of medical care. Most notably, the more

educated and recent residents of the area appear more likely to use clinic services. When other variables were held constant, color was not important. Finally, Jewish residents were more likely to use the clinic.

# Comparisons With Other Evidence

Some additional data have come to my attention which permits further examination of the role of information. Efforts were made to encourage utilization of the newly established Bedford Health Center in the Bedford section of Brooklyn. New York. Some health aides were sent around to residences to inform persons of the availability of free services at the general medical clinic, at various dates during the period. Eligibility for the services was restricted to those residing in four health areas, and within those areas, aides were instructed to "knock on every door." The aides conducted a lengthy interview to determine needs for a broad range of social services and make necessary referrals, so that the information about the clinic would not have been foremost in the minds of many residents. This would presumably lead to lower estimates of the effects of health information than if counselling were more limited. Complete information on the impact of the counselling was not available, and only two of the four areas had enough data for adequate testing./

Table 4 shows the mean number of admissions per week for the three health areas in which there was counselling, before, during and after the period in which counselling took place. In only one case was data for the period after counselling available and even then it only covered two weeks. To allow for a lag between counselling and admission the mean is also computed for the period during which counselling took place, excluding the first two weeks. In health areas 20 and 21 combined, admissions per week were 1/3 higher in the period after the first two weeks of interviewing than before interviewing began. This counselling would, therefore, increase the total patient load at the clinic by about 2 1/2 percent, assuming that the admission rate fell back to earlier levels after interviewing terminated. The counselling covered an area with about 40 percent of the households in the four health areas.

I know of few studies for which suitable comparisons with the present research can be made. Two recent studies of younger populations in different institutional settings are of interest. One is the portion of the Yale Ambulatory Care Studies which pertains to emergency room use which was recently published by Weinerman, <u>et. al</u>. The other is a just published study of utilization of prenatal clinics and the Judson Health Center in New York City conducted by Morton Silver as part of the Gouverneur Economic Research Project.<sup>O</sup> Both studies were analytical rather than descriptive, formulating hypotheses in advance, and subjecting them to multivariate analysis.

The Weinerman group examined variation in the percentage of emergency room visits in the Yale-New Haven Medical Center which were classified as non-urgent by physicians. This can be taken as a

measure of the demand for clinic services. Over 2,000 visits were analyzed. Because users of a specific institution rather than complete population were studied there is uncertainty as to what extent non-users receive better care or no care at all. Age, a measure of health status, was found to be quite important, as were measures of health status in the present study. In the Yale study it reflected the greater need of the youngest and oldest groups for true emergency care. Negroes were found to have a greater percent of visits nonurgent, perhaps because of poorer health. While lower income groups tended to have a higher percent non-urgent, the difference was only weakly statistically significant. However, no result for income is given with education held constant. It may be for this reason that educational differences are not important. A number of information variables were found to be significant. Persons with a short length of residence at their current address, self-referred persons, and persons without a regular physician were more likely to use the emergency rooms for non-urgent problems. There were no significant effects of sex or religion. Zbrowski found that "the Jew tends to manifest a future oriented anxiety as to the symptomatic and general meaning of pain experience."  $^{9}$  Therefore it is not surprising that Jewish persons were more likely to use a health maintenance clinic than emergency room, relative to persons of other faiths.

Morton Silver investigated the determinants of number of prenatal care visits and month of first visit for a sample of 142 women using seven prenatal clinics in low income areas of New York City. Regressions with number of first visits as dependent variables were run with and without holding constant month of first visit. Two variables which can be run interpreted as measures of (actual or expected) health status -- gestation period and unsuccessful outcome of prior pregnancies, were very important. An income measure was not available and education, which was associated with fewer visits and a later month of first visit may reflect income. Education's association with fewer visits probably reflects use of other sources. When the dependent variable was month of first visit anywhere, education was positive but not significantly related to month of first visit. Whether or not a woman had children, a "price" variable reflecting the difficulty of making a visit, adversely affected care. Travel time had the expected sign but was only significant in separate regressions for women without children. Whether or not a woman was working, in part interpreted as reflecting information and awareness about the importance of care, was associated with greater demand for prenatal care.

The Silver study provides additional information on amount of use which merits further attention. It is possible to crudely estimate the relative importance of medical care at early stages and frequent care, given the stage of the condition in accounting for the effects of each variable.<sup>10</sup> The main effects of education (or income) and employment status are through the frequency of visit, while the outcome of prior pregnancies and the presence of children operate relatively more through month of first visit.

Silver also examined visits of 125 children age one to five to the Judson Health Center on the lower east side of New York City. Once again an income variable was not available and the relationship of parents' education to the number of visits per child was negative. The relationships of education to the number of ill child visits and the number of illnesses were also negative, but against well child visits education was not significant. Silver notes that the classification of visits was not clear cut, depending, for example, on whether the illness was discovered during a check up. This raises the possibility that the more educated mothers more often recognized illnesses or communicated information about them. This would result in an understatement of the effect of education on ill child visits. Length of residence again was important. The foreign born tended to have significantly fewer well child visits. The new residents of the area tended to have more ill child visits.

The additional information cited, either directly or with some reinterpretation, can be considered to be very consistent with the Queensbridge data.

The foregoing analysis should be of value in determining the kinds of data which are required for planning community health services, and in making gross estimates of utilization and population coverage. For more complete planning, however, much more detailed information on the demand for different types of services is required. Furthermore, in order to actively influence the completeness with which medical care is distributed, we will have to ascertain what is behind the broad health, information, price and other variables so that specific programs can be formulated. This challenge to research is, thus far, being met with only an embryonic response.

## FOOTNOTES

1. Nicetas H. Kuo <u>et</u>. <u>al</u>., "The Queensbridge Health Maintenance Service for the Elderly, An Evaluation", unpublished manuscript, New York City Department of Health, n.d.

2. The President's Commission on Heart, Cancer and Stroke, <u>A National Program to Conquer Heart</u> Disease, Cancer and Stroke, Washington: 1964.

3. See J. Johnston, <u>Econometric Methods</u>, New York McGraw-Hill Book Company, Inc., 1963, pp. 221-228 and Arthur S. Goldberger, <u>Topics in Regression</u> <u>Analysis</u>, New York: The Macmillan Company, 1968, pp. 112-118.

4. Daniel B. Suits, "Use of Dummy Variables in Regression Equation", <u>Journal of the American</u> Statistical Association, 52 (1957), pp. 548-51.

5. The standard test would be an F-test to see if the increase in explained variance when a set of dummy variables is added is significant in comparison with the unexplained variance. See Emanuel Melichar, "Least Squares Analysis of Economic Survey Data", <u>Proceedings of the Business</u> and Economic Statistics Section of the American Statistical Association, 1965, pp. 373-85.

6. <u>Op</u>. <u>cit</u>.

7. The imposition of fees after the period analyzed would have made interpretation of subsequent data impossible.

8. E. Richard Weinerman, Robert S. Ratner, A. Anthony Robbins, and Marvin A. Lavenhar, "Determinants of Use of Hospital Emergency Services", <u>American Journal of Public Health</u>, 56, No. 7 (July, 1966), pp. 1037-56, and Morton Silver, <u>Utilization of a Neighborhood Health Center: An</u> <u>Economic Model of Demand for Services Available</u> <u>at the Judson Health Center</u>, New York: Gouveneur Economic Research Project, 1968.

9. Mark Zborowski, "Cultural Components in Responses to Pain," <u>Journal of Social Issues</u>, VIII No. 4 (1952), p. 23.

10. Let the stage in the episode of which medical care is sought be determined by a linear function such as the following:

$$s = a_{10} + b_{11} x_1 + b_{12} x_2 + b_{13} x_3$$

The number of treatments depends on a number of factors including stage of first treatment.

$$N = a_{20} + b_{21} X_1 + b_{22} X_2 + b_{24} X_4 + b_{s} S_{s}$$

Solving,

$$N = a_{20} + b_s a_{10} + (b_{21} + b_s b_{11}) X_1$$
$$+ (b_{22} + b_s b_{12}) X_2 + b_s b_{13} X_3 + b_{24} X_4$$

The proportion of the effect of a variable operating through stage of episode can be estimated by terms such as  $b_s$   $b_{11}$ 

$$b_{21} + b_{s} + b_{11}$$