



*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

[aesearch@umn.edu](mailto:aesearch@umn.edu)

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

*No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.*

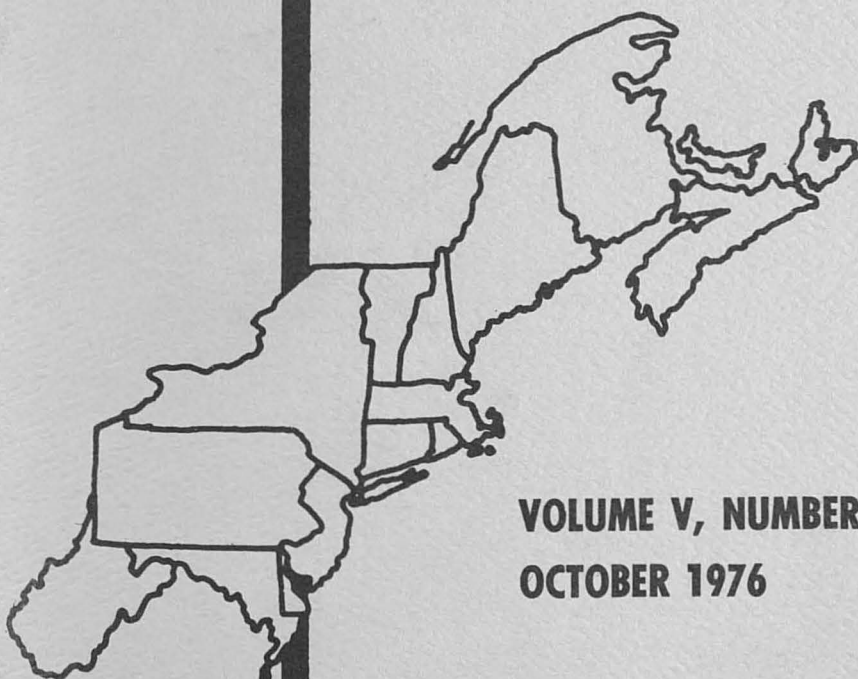
PER. SHELF

# **JOURNAL OF THE**

## **Northeastern Agricultural Economics Council**

GRANTING INSTITUTION FOR  
AGRICULTURAL ECONOMICS  
LIBRARY

1976



**VOLUME V, NUMBER 2  
OCTOBER 1976**

METHODOLOGICAL ISSUES IN ANALYSES OF USE AND ADEQUACY  
OF COMMUNITY SERVICES IN RURAL AREAS: WITH AN  
APPLICATION TO LEGAL SERVICES IN THE NORTHEAST\*

Cleve E. Willis  
Associate Professor  
Food and Resource Economics  
University of Massachusetts

Charles O. Crawford  
Associate Professor  
Agricultural Economics and Rural Sociology  
Pennsylvania State University

Introduction

Three major problems appear common to recent research efforts aimed at measurement and assessment of the adequacy of community services. These include:

- (i) difficulties and inconsistencies which relate to the definition and interpretation of utilization behavior,
- (ii) econometric misspecification—e.g., omission of critical variables affecting needs or use—, and
- (iii) choice of inappropriate estimation technique when the dependent variable (e.g., use or non-use of the service) is dichotomous.

In one of the more cogent statements relative to the first issue above, Gessaman and Rose [4] argue that adequacy measures generally appear to take the form of (minimum or maximum) standards,<sup>1/</sup> that these standards have proven useful to the communities of our nation, but that they are not sufficient for the measurement and assessment of the adequacy of our community services. As reasons for this insufficiency, they cite factors such as "... the tendency for standards to lag behind the needs of the present" and "... the tendency for standards to be set by ... officials in ways favorable to pressure groups". They recognize, however, that the

---

\* This paper emanates from the work of Regional Project NE-77 and is in support of Title V objectives. Massachusetts Experiment Station Paper 2075.

<sup>1/</sup> Set by professionals in their technical field of competence, by elected or appointed officials, by the courts, or by consumers through their actions in the political arena and in the market place.

real issue is that for researchers to measure and assess adequacy it must be made clear that a community service is in reality composed of a bundle of dimensions which include: quality aspects, quantity, cost, and mix.<sup>2/</sup>

It seems clear, given this "multi-attribute" conceptualization of service adequacy, that community service researchers ought to be drawn from the traditional theory of consumer behavior and toward the alternate approach suggested by Lancaster [11]. Applied to services, his approach views services as being demanded because they yield want satisfying characteristics to consumers. In general, a service has multiple characteristics any one of which may be shared by more than one service (the traditional consumer model is a special case which views the number of characteristics as equal to the number services). Further, services in combination may possess characteristics (attributes) different from those pertaining to the services separately. Given the multi-dimensional characterization of services, it seems clear that the point of departure of responsible research in services ought to be Lancasterian in nature.

A recent article by Hershey, et. al. [7] provides solid documentation for the first and second problem areas. The paper suggests that while services<sup>3/</sup> researchers have outlined and tested numerous empirical models in recent years, little consideration has been given to exactly what behavior is being measured by the data. Their first basic hypothesis is that when utilization measures are expanded (as a health care example) to include preventive health behavior, a distinction between patient- and physician-initiated visits, and measures of access to services, the importance of various explanatory variables will shift. That is, the importance of various explanatory variables depends upon what type of utilization behavior is being measured. Secondly, the use of a limited set of independent variables (e.g., income and demographics) can lead to an incorrect interpretation of the data as compared with using an expanded set (e.g., with measures of physical and financial accessibility and social structural and social psychological variables). In brief, it is well-known that omission of important explanatory variables in least squares estimation imparts a bias in the estimated effects. Further, omission of other relations (e.g., supply) which form a system with the relation under investigation leads to simultaneous equations bias. Of the types of "problems" discussed here, this is the simplest to correct. It simply results from sloppiness in terms of model specification and use of econometric procedures—a sloppiness which can lead to rather troublesome outcomes should the results be used in any policy context.

---

<sup>2/</sup> The mix dimension reflects on adequacy, for example, in as much as the particular mix makes for weaker or stronger substitutes for the particular service.

<sup>3/</sup> Their application was to health services utilization studies, but the concepts are equally valid in other community services.



The third problem area is the most specific and is the primary focus of this paper. Survey instruments in community services research are often set up in such a way that the answers to the questions which will ultimately become the dependent variable are dichotomous (e.g., have you used a particular service in the past year—yes or no?). In this research, the rather standard approach is to let the dependent variable take a value of one if the answer is yes and zero if no and use ordinary least squares (OLS) to estimate the relation. That this approach is generally not appropriate in this situation is seldom recognized. Therefore, we organize the remainder of this paper as follows: the subsequent section provides the reasons OLS may be inappropriate and develops an alternative procedure (Probit); following this, an empirical situation regarding rural legal services in the Northeast is developed and estimated by both techniques; the final part provides a comparison of these techniques for this particular situation.

### Linear Probability Function

Consider the general linear model in the context of a dichotomous dependent variable:

$$(1) \quad Y_i = X_i \beta + u_i, \quad i = 1, \dots, n$$

where:

$u_i$  is the stochastic disturbance,

$$E(u_i) = 0$$

$$E(u_i u_j) = 0 \text{ for } i \neq j$$

$\beta$  is K dimensional column vector of unknown parameters,

$X_i$  is a K dimensional fixed row vector for observation i, and

$$Y_i = \begin{cases} 1, & \text{if the event occurs,} \\ 0, & \text{otherwise.} \end{cases}$$

Since  $Y_i$  can be only 0 or 1, and assuming that the magnitude of  $X_i \beta$  influences the probability that  $Y_i = 1$ , what we are attempting to explore is the conditional probability of an event occurring given the  $i^{\text{th}}$  values of the K explanatory variables  $X_i$ , and the way in which this conditional probability varies when the explanatory variables take on different values.

The common estimation approach is to apply OLS procedures as if the usual spherical assumptions were satisfied. Unfortunately, they are not. From (1), since  $u_i = Y_i - X_i \beta$  and  $Y_i$  is 0 or 1, there are only two values  $u_i$  can take on, viz.:  $(-X_i \beta)$  or  $(1 - X_i \beta)$ . Consequently, the disturbance is not normally distributed, but rather has a discrete distribution where:

$u_i = -X_i \beta$  with probability (p) of  $(1-X_i \beta)^{4/}$  and  $u_i = 1-X_i \beta$  with probability  $(1-p)$  of  $(-X_i \beta)$ . Thus, the variance of the stochastic disturbance is given by:

$$(2) \quad E(u_i^2) = (-X_i \beta)^2 (1-X_i \beta) + (1-X_i \beta)^2 (X_i \beta) \\ = X_i \beta (1-X_i \beta).$$

Since  $E(Y_i) = X_i \beta$ , (2) can be expressed as:

$$(2') \quad E(u_i^2) = E(Y_i) (1-E(Y_i)).$$

From (2') it is clear that the disturbance of (1) is heteroscedastic, i.e., the variance of  $u_i$  varies systematically with  $E(Y_i)$  and hence with  $X_i$ . Hence, OLS estimates, although unbiased, are not efficient (see Johnston [10, p. 215]) and the non-normally distributed  $u_i$  lead to OLS estimators of  $\beta$  which are not normally distributed so that the usual significance tests based on the students t distribution are invalid.

The problems due to this heteroscedasticity condition can be circumvented by the use of Aitken's Generalized Least Squares procedures (see Goldberger [5, p. 250]). A difficulty remains, however. The linear probability function formulation permits  $E(Y_i)$  to fall outside the interval 0 to 1 which is inconsistent with both the definition of  $Y_i$  and the interpretation of  $E(Y_i)$  as a probability. That is, since the linear function is unbounded, there exist values of  $X_i$  such that  $X_i \beta = E(Y_i)$  exceeds one and others such that  $X_i \beta = E(Y_i)$  is negative.

Probit analysis is one alternative approach to the dichotomous regressand problem in which  $E(Y_i)$  is kept within the range zero to one from the beginning. While the technique is not new to agricultural and applied economists (e.g., see Dagenais [2], Hill [8], Hill and Kau [9], and Tomek [13]), its advantages have been all but overlooked by researchers in community services analyses.

Very briefly, the approach involves the following. Let  $I$  be an index which is a linear function of the regressors ( $I_i = X_i \beta$ ) and let  $I^*$  be a standard normal variable. Then let the value of  $Y_i$  be determined as follows:

$$(3) \quad Y_i = \begin{cases} 1, & \text{if } I_i \geq I_i^* \\ 0, & \text{if } I_i < I_i^* \end{cases}$$

---

<sup>4/</sup> Since  $E(u_i) = 0$ , by assumption,  $(-X_i \beta)p + (1-X_i \beta)(1-p) = 0$ , and  $p = 1-X_i \beta$ .

Thus,  $Y_i$  depends on the  $X_i$  via  $I_i$  and on  $I_i^*$ . The  $I_i^*$  may be regarded as critical (threshold) values of the index (e.g., if  $Y$  = use of legal services and  $X$  = age, and individual ( $i$ ) with a high  $I_i^*$  would seek legal services only if his age were so high that  $I_i \geq I_i^*$ ). Finally, denoting as  $F(Z)$  the value of  $Z$  of the standard normal cumulative distribution, we can express

$$(4) \text{ Prob } \{ Y = 1/I \} = \text{Prob } \{ I^* \leq 1/I \} = F(I)$$

and

$$\text{Prob } \{ Y = 0/I \} = \text{Prob } \{ I^* > 1/I \} = F(I)$$

since  $I^* \sim N(0, 1)$ . Since  $I$ , and hence the probabilities, is a function of  $\beta$ , Probit analysis is simply a direct application of maximum likelihood estimation using this specification. The following section illustrates the use of Probit analysis for estimating a linear probability model in rural legal services and contrasts this with the results of OLS.

#### Empirical Application

Of the various human services for rural areas, perhaps the least attention has been paid to the adequacy of legal services in rural areas. Derr [1973] asserts that the legal needs of the poor have been one of the most neglected of all human service areas and that legal problems of the rural poor have received even less study and attention than those of the urban poor. Some indications of inadequacies for rural areas, she suggests, include: lack of traditional Legal Aid Societies, inadequate number of rural private attorneys, inadequate transportation network for utilizing the services of urban attorneys, and documented reluctance of local rural attorneys to handle intraclass legal problems for the rural poor. Haveman [6] adds that rural attorneys tend to be older, less well trained, and less likely to be attuned to current legal and social developments than their urban counterparts. Some of the selected results of a large survey by American Bar Association [1] tend to support some of these notions. The data suggests rather strongly, for example, that socio-economic characteristics are important determinants of the need for various kinds of legal assistance and the ability to those people to satisfy those needs.

For present purposes, the relation used for comparison of OLS and Probit procedures involves estimating how several factors influence the likelihood of a family or individual expressing a need for legal services in a particular year. The analysis is regional and based on data gathered in regional project NE-77, for the Northeast.<sup>5/</sup> The explanatory

---

<sup>5/</sup> For more detail on this project see [12].



variables chosen include regional parameters—e.g., the status of the county in terms of population and income change (expanding, stable, declining)—as well as family characteristics—e.g., age and sex of the head of the household as well as marital status. The model is expressed as:

$$(5) \quad Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_6 X_{6i} + u_i, \quad i = 1, \dots, n$$

where:

$$Y_i = \begin{cases} 1, & \text{if the respondent indicated a need for legal services in the previous year,} \\ 0, & \text{otherwise,} \end{cases}$$

$$X_{2i} = \begin{cases} 1, & \text{if the } i^{\text{th}} \text{ respondent resides in a county in the lowest category in terms of population and income change relative to the Northeast region,} \\ 0, & \text{otherwise,} \end{cases}$$

$$X_{3i} = \begin{cases} 1, & \text{if the } i^{\text{th}} \text{ respondent resides in the middle category,} \\ 0, & \text{otherwise} \end{cases}$$

$$X_{4i} = \text{age of the head of the } i^{\text{th}} \text{ household in years,}$$

$$X_{5i} = \begin{cases} 1, & \text{if the head of the } i^{\text{th}} \text{ household is male,} \\ 0, & \text{otherwise,} \end{cases}$$

$$X_{6i} = \begin{cases} 1, & \text{if the head of the } i^{\text{th}} \text{ household is single or married,} \\ 0, & \text{if the head is divorced, separated, widowed, or deceased.} \end{cases}$$

The  $\beta_j$  are the unknown parameters and  $u_i$  are the stochastic disturbances presumed spherical.

The *a priori* assumptions were that  $\beta_2$  and  $\beta_3$  are negative,  $\beta_4$  is positive and  $\beta_5$  and  $\beta_6$  are negative. That is, relative to the deleted dummy variable (highest category in terms of income and population), it was presumed that legal needs in the other areas would be less frequent due to poorer services and less information in the rural areas. However, it seemed reasonable to expect legal services needs to expand with age. Finally, the presumption was that households with male heads and with a marital status of married or single would have less occasion to use a legal service than for the counterparts with female heads of, say, a separated or divorced household. There are a host of other explanatory factors which could be considered, of course, but the present focus is



upon comparison of two estimation techniques rather than developing a full representation of the relevant set of relations.<sup>6/</sup>

For the regional estimation of (5), both OLS and Probit procedures were used. The results using the sample of 2141 households interviewed at 13 sites in 9 participating Northeastern states are set out in Table 1. The estimates of  $\beta_2$ ,  $\beta_4$ , and  $\beta_6$  are more than double their standard error, and of  $\beta_5$  nearly so for both sets of estimations. Notice also that the values in parentheses are remarkably similar for the two estimation procedures although the estimated coefficients differ substantially. For example, the OLS estimate of  $\beta_2$  suggests that, other things constant, the odds of a family in a declining area needing legal services are 94.7 percent lower than for its counterpart in an expanding area (the deleted variable). For the Probit estimation, this effect is only 25.3 percent lower.

Notice also that the signs of the estimates of  $\beta_4$  and  $\beta_5$  are contrary to the hypotheses for both estimations. In retrospect, the observed sign on the estimate of  $\beta_4$  is not too surprising. The average age of the heads of sample households was just under fifty and it seems reasonable to expect greater needs before fifty than after. The sign of the estimate of  $\beta_5$  is somewhat more surprising. It estimates that other things constant, male heads of household tend to request legal assistance more often than female heads. The estimating effect was not especially convincing relative to its standard error, however.

A more instructive comparison of the results of the alternative estimation procedures involves comparing  $\hat{Y}_i$  for various situations with respect to the row vector  $X_i$ . Some sample values are provided in Table 2 below. As a rather extreme example, consider the situation in which the respondent resides in a declining area and the head of the household is male, fifty years of age, and divorced, separated, etc. (situation 1). The OLS estimate of the "conditional probability" of needing legal services in this case is a negative 0.72 as compared with the Probit estimate of 0.65. Likewise, for a similar situation in which the head is 25 years of age and married, OLS estimates produce a nonsensical estimate of minus 0.397 as compared with the Probit value of 0.532. As further illustration, for a declining area with a fifty year old female married (or single) household head, OLS yields a value of  $\hat{Y}$  of minus 0.76 as compared with the Probit estimate of conditional probability of needing some form of legal services.

---

<sup>6/</sup> This development is ongoing as a part of the regional project NE-77.

<sup>7/</sup> Recall that for OLS,  $\hat{Y}_i = X_i \hat{\beta}$  is the point estimator of  $E(Y_i|X_i)$  and is not required to fall within the unit interval consistent with the definition of  $Y$  and its expected value as a conditional probability. However, for Probit,  $\hat{Y}_i = F(\hat{I}_i) = F(X_i \hat{\beta})$ , as defined in (4) above, necessarily lies in the unit interval.

Table 1  
Probit and OLS Results\*

Estimation Procedure	Estimations of $\beta_j$					
	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$
OLS	0.74620	-0.94678 (-3.26674)	0.01083 (0.37414)	-0.00480 (-7.05947)	0.08686 (1.79986)	-0.16306 (-3.43582)
Probit	0.65699	-0.25323 (-3.30000)	0.02762 (0.36300)	-0.01275 (-6.99600)	0.24149 (1.83700)	-0.44221 (-3.43300)

\*Values in parentheses are ratios of estimated coefficients to standard errors.

Table 2

Comparison of OLS and Probit Estimates of  $\hat{Y}$  for Alternative Situations  
Regarding the Independent Variables

Situation	Value of Independent Variable						OLS	Probit	$\hat{I}_i$
	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$\hat{Y}_i$	$\hat{Y}_i$	
1	1	1	0	50	1	0	-0.721	0.653	0.394
2	1	1	0	25	1	1	-0.397	0.532	0.044
3	1	1	0	50	0	1	-0.757	0.357	-0.367
4	1	0	1	25	1	0	0.724	0.728	0.606
5	1	0	1	50	1	1	0.441	0.439	-0.154



Not all comparisons are this striking of course. For example, for a stable area and a twenty-five year old divorced male head (situation 4) or a fifty year old male married head (situation 5), the OLS and Probit results are within three or four one-thousandths of each other.

### Summary

It was suggested that three problem areas common to much of community services research involve: proper definition of utilization behavior, accurate econometric specification, and choice of estimation procedure when the dependent variable is dichotomous. The latter problem area served as the focus of this paper. In survey instruments the behavior ultimately to become the dependent variable very often takes a "yes-no" or "zero-one" form. The common procedure has come to be application of OLS and interpretation of  $\hat{Y}_i$  as estimates of  $E(Y_i)$ —the conditional probability of a "yes" response given the situation characterized by  $X_i$ . However, this linear probability function permits  $E(Y_i)$  to fall outside the range 0 to 1 which is inconsistent with both the definition of  $Y_i$  and the interpretation of  $E(Y_i)$  as a probability. The appropriate alternative, then, is Probit. Finally, a sample relation is estimated with regional data by both Probit and OLS procedures, and the results are compared.

### References

1. American Bar Association, "Alternatives: Legal Services and the Public", ABA Consortium on Legal Services and the Public, Volume 3, No. 1, January 1976, (Special Issue).
2. Dagenais, M. G., "A Threshold Regression Model", Econometrica, 37: 2:193-203, April 1969.
3. Derr, J. M., "Rural Social Problems, Human Services, and Social Policies", Working Paper 12: Legal Services, Social Welfare Research Institute, Center for Social Research and Development, Denver Research Institute, September 1973.
4. Gessaman, P. H. and G. D. Rose, "Problems of Measurement and Assessment of the Adequacy of Community Services: A Naive Viewpoint", Department of Agricultural Economics Staff Paper #21, University of Nebraska, Lincoln, 1971.
5. Goldberger, A., Econometric Theory, John Wiley and Sons, New York, 1964.
6. Haveman, Robert, *et. al.*, "Midwest Rural Poverty, Human Rights, and the Need for Legal Services", University of Kansas Law Review, Volume 15, May 1967.

7. Hershey, J. C.; H. S. Luft and J. A. Gianaris, "Making Sense Out of Utilization Data", Medical Care, 13:10:838-854, October 1975.
8. Hill, L. D., "Use of Weighted Regression in Estimating Models of Binary Choice", Canadian Journal of Agricultural Economics, 18:64-72, November 1970.
9. Hill, L. and P. Kau, "Application of Multivariate Probit to a Threshold Model of Grain Dryer Purchasing Decisions", American Journal of Agricultural Economics, 55:1:19-27, 1973.
10. Johnston, J., Econometric Methods, McGraw-Hill, New York, 1972.
11. Lancaster, K. J., "A New Approach to Consumer Theory", The Journal of Political Economy, 74:3:1966.
12. Technical Committee of the Northeast Regional Research Project, NE-77, Methodological Considerations in Researching Community Services in the Northeast, New Jersey Agricultural Experiment Station Bulletin 836, September 1975, (100 pages).
13. Tomek, W. G., "Regression Analysis With a Limited Dependent Variable", American Journal of Agricultural Economics, 50:445-447, May 1968.