PARTICIPATION RATES, EFFICIENCY, AND CHARACTERISTICS OF WORKERS

R. S. Canagarajah*

* I am grateful to the Institute for Rural Health Studies for granting me permission to use their data set. I am also thankful to William Peterson, Steve Pudney, and Hussain Samiei for helpful critical comments on an earlier draft. Usual disclaimers apply.

ISSN 1936-5071

The Cornell Food and Nutrition Policy Program (CFNPP) was created in 1988 within the Division of Nutritional Sciences, College of Human Ecology, Cornell University, to undertake research, training, and technical assistance in food and nutrition policy with emphasis on developing countries.

CFNPP is served by an advisory committee of faculty from the Division of Nutritional Sciences, College of Human Ecology; the Departments of Agricultural Economics, Nutrition, City and Regional Planning, Rural Sociology; and from the Cornell Institute for International Food, Agriculture and Development. Graduate students and faculty from these units sometimes collaborate with CFNPP on specific projects. The CFNPP professional staff includes nutritionists, economists, and anthropologists.

CFNPP is funded by several donors including the Agency for International Development, the World Bank, UNICEF, the Pew Memorial Trust, the Rockefeller and Ford Foundations, The Carnegie Corporation, The Thrasher Research Fund, and individual country governments.

 ISBN 1-56401-130-5

This Working Paper series provides a vehicle for rapid and informal reporting of results from CFNPP research. Some of the findings may be preliminary and subject to further analysis.

This document was word processed by Marnie Fulton. The manuscript was edited by Elizabeth Mercado. The text was formatted by Gaudencio Dizon. The cover was produced by Jake Smith.

For information about ordering this manuscript and other working papers in the series contact:

CFNPP Publications Department
1400 16th Street NW, Suite 420
Washington, DC 20036
202-822-6500

or

308 Savage Hall
Cornell University
Ithaca, NY 14853
607-255-8093

CONTENTS

LIST OF TABLES

LIST OF APPENDIX TABLES

LIST OF FIGURES

ABBREVIATIONS

EWT	Efficiency wage theories
IRHS	Institute for Rural Health Studies
LL	Log likelihood
LM	Lagrange multiplier
LR	Likelihood ratio
ML	Maximum likelihood
OLS	Ordinary least squares
PR	Participation rates
RDA	Required dietary allowance

FOREWORD

The increased availability of large data sets for many developing countries has brought with it new statistical and econometric challenges. These challenges are not only to use existing techniques correctly, but also to develop new methods to accurately deal with the problems of, and questions posed by, researchers and policymakers. This study shows how the appropriate data set, combined with the laborious process of constructing specific models and related diagnostic checks, pays off in the creation of unbiased and robust estimators, which is necessary for meaningful policy studies.

In particular, this paper takes up the issue of "X-efficiency," which was originally raised almost four decades ago in relation to developing countries, and tries to shed light on the link between efficiency of workers and their nutritional status. The lack of variation in the wage data from most surveys has hindered the study of this relationship and hence the policy prescriptions resulting from many of these previous studies. The author shows how the slack season participation rate of workers in a demand-constrained economy can be related to their individual and household characteristics to shed light on the propositions about efficiency of workers developed in the efficiency wage literature. The unique characteristics of the data set used for this study necessitated the use of a Two Limit Tobit Model for empirical estimation and hypothesis testing. In order to test the efficacy of the model, the author constructs and implements a Lagrange Multiplier test to check for heteroscedasticity. The Likelihood Ratio test is used to test alternative specifications. The findings correspond to the literature on rural labor markets of developing countries, and especially with the propositions of efficiency wage theory. The author also points out that the undernourished poor are caught in a vicious cycle: employers are less willing to hire undernourished workers, so they remain undernourished. Some external assistance is needed to provide nourishment to enable these workers to get back into the labor market. Hence, the study shows that in order for the state to help alleviate poverty, it must first help poor workers meet their basic nutritional requirements to enable them to compete in the labor market. The study also demonstrates the efficient application of existing tools to use available data to examine the complex relationships within the labor market.

Washington, DC
June 1992

David E. Sahn
Director, CFNPP

1. INTRODUCTION

Because the wage variations in rural labor markets are too limited to offer sufficient incentive to efficient workers, employers resort to nonwage mechanisms and other related quantity rationing mechanisms to provide incentives. We can use the number of hours of employment that a worker obtains in a demand-constrained, surplus labor market as a proxy for his or her efficiency. In general, a profit maximizing employer, both under perfect competition and monopsonistic conditions, would employ only the efficient workers in the slack season, when the demand for labor is low while labor supply is in surplus (see Canagarajah 1991). Thus, those workers who are employed in a slack, demand-constrained season would reveal their relative work efficiency, and the corresponding preference shown toward them by the employers. Thus, participation rate (PR) of workers is bound to reveal the relative efficiency of workers. Studies in rural labor markets in the past provide incontrovertible evidence of the presence of the above mentioned features (see Rudra 1982; Dreze and Mukherjee 1987; Bharadwaj 1974).

The objective of the present paper is to relate participation rate directly to characteristics of workers and indirectly to some of the propositions established in efficiency wage literature.[1] Although we do not set out to formally test efficiency wage propositions, we are able to make substantial progress in making sense of most of them through analyzing participation rate of workers in relation to their characteristics.

The present paper is organized as follows. In the next section the village and the data set used in the analysis are introduced. The third section deals with the important econometric issues, such as the features of the chosen econometric model and the related diagnostic tests. The fourth section presents an outline of a simple theoretical model that serves as a motivation for the empirical analysis that follows it. In the fifth section the results of the empirical investigation are presented and their implications in understanding efficiency of workers in terms of their personal and household characteristics are discussed. The final section summarizes the main findings.

---

1 A good survey on efficiency wage models is described by Akerlof and Yellen (1986). For a discussion on efficiency wage theories, which are relevant for developing country rural labor markets, see Canagarajah (1991).

2. THE VILLAGE AND THE DATA SET

We use information from the South Indian village of Dokur, which is situated in the Mahbubnagar District in the State of Andhra Pradesh. The village has a population of around 4,000 people who are mainly involved in agricultural activities. It shares cultural and social characteristics with other Indian villages: the caste hierarchy plays an important role in deciding the economic status and activity of individuals. Hence Reddies, who are farming caste, control most of the land, while Harijans, who are the lowest caste, are mostly landless laborers.[2]

The Institute for Rural Health Studies (IRHS) conducted a survey from April 1982 to March 1983 for 52 weeks, collecting information on household composition, asset ownership, labor force participation, health records, nutritional consumption, income and expenditure flows, and anthropometric variables on 349 individuals from 40 sample households. Of these, 172 individuals were identified as actively participating in the agricultural labor market. Hence this is a rich data set for analyzing the participation behavior of workers in the context of their personal and household characteristics. For purposes of this analysis, the 52 consecutive weeks are divided into peak and slack seasons of the agricultural activities of this village. Thus, weeks 1 through 11 and 46 through 52 represent the slack season giving a total of 16 weeks of labor force participation information. This information is very useful for studying the participation behavior of workers to understand their underlying efficiency. We define the average worker participation rate by the average number of days each individual worked, of the total number of days available in the season, which is used as the dependent variable in the analysis. Thus, lack of participation due to sickness is also used to indicate less efficiency.[3]

We generate person-specific and household-specific variables for each individual and have indicators of these for peak and slack season wherever possible and meaningful. Person-specific variables include age, gender, wage income, kilocalorie consumption of food as a percentage of individual's requirement, and anthropometric indicators such as weight, height, weight-for-height, and arm circumference. Household-specific variables include per capita landholding, family size, number of children, number of working members, and the

---

2 The interested reader is referred to Canagarajah (1991) for more detail on the village characteristics and the data set used in this paper.

3 This is in line with the explanation that illness is a result of lower resistance, which is a result of low past participation and wage income (see Behrman et al. 1988).

dependency ratio. Table 1 presents summary statistics of some of the important variables used in the analysis.[4]

In our data set we note that the dependent variable, which is a measure of participation rate of individuals, assumes a value between zero and one, with zero indicating nonparticipation and one indicating full participation. We hypothesize that higher participation rates indicate higher efficiency levels, while lower participation rates imply lower efficiency levels; this is assuming that in a demand-constrained labor market the employer chooses the workers according to their relative levels of efficiency.

However, in our case since PR is measured at a particular point in time, those who fall in the extremes have a different and more complicated relationship to efficiency than those in between these values. It would be expected that the latter would in general have a more or less monotonic relationship between PR and efficiency. Zero values may arise for different reasons — some of these workers may be handicapped, while others may have been unemployed continuously beyond the slack season and would otherwise have negative PRs. On the other hand, many of those who fall at the value of one would have entered into a binding contract that left them no choice, while some may have continued to work beyond this season and therefore have a PR value greater than one. Thus, a mass of observations collapsed at these extreme points lend themselves to different possible interpretations.[5] The distribution of workers in the survey village according to slack season participation rate confirms this pattern, as can be seen in Figure 1.

Since the values of our dependent variable cannot be directly related to the relative efficiency of workers, especially in the extreme values of zero and one, we require special econometric techniques to consider the underlying problems in using this information to relate efficiency through participation rate to the person-specific and household-specific explanatory variables. Therefore, we propose an econometric technique that comes under the class of censored models to meet this need. In the following section we discuss the characteristics of this model along with related diagnostic issues.

---

[4] The definition of all the variables generated from this primary set of variables and used in the analysis is given in Appendix A, along with descriptive statistics for all the variables used in the econometric analysis in Appendix B. The large amount of missing information on anthropometric variables precludes any investigation into the probable relationship between anthropometric status and participation rates.

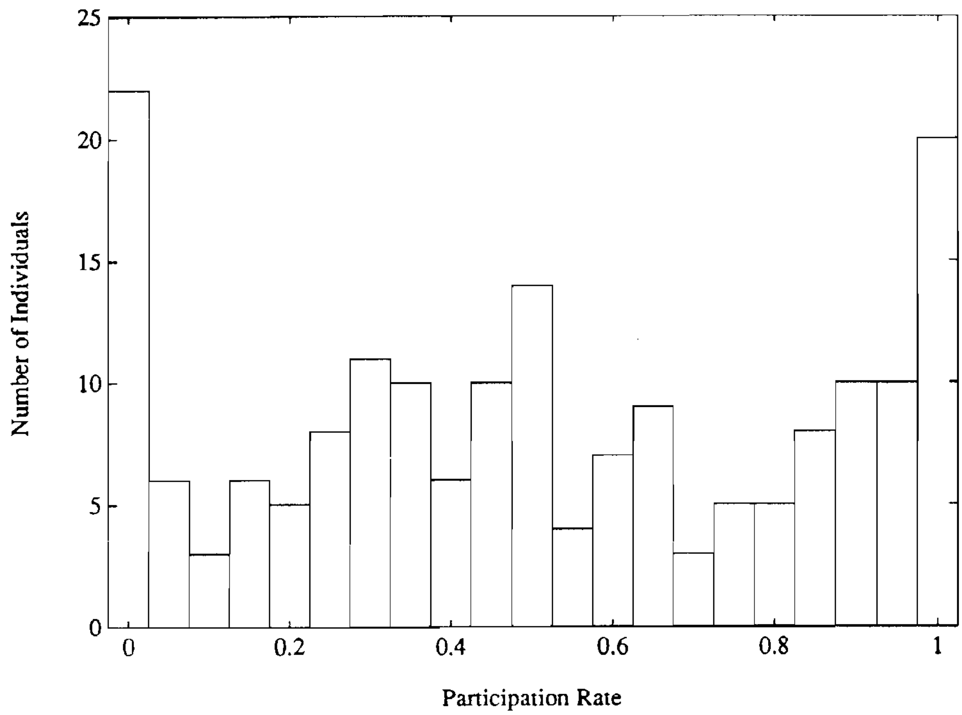
[5] For an extended discussion on this point and related economic justification, see Canagarajah (1991).

Table 1 — Dokur: Summary Statistics for Important Variables Used in the Analysis

Variables	**Mean**	**Standard Deviation**
Participation rate	0.528	0.323
Nutrition	100.660	15.410
Arm circumference	83.550	7.520
Weight-for-height	82.830	9.600
Age in years	32.800	15.700
Household size	10.050	4.230
No. of children	3.770	1.800
Dependency ratio	0.385	0.129
Nuclear family	0.314	0.465
Landholding	8.389	10.160
Per capita landholding	0.732	0.719
Wage income in cash	117.820	172.100
Per capita income	1443.300	961.590

Source: Author's calculations of IRHS data set.

Figure 1 — Dokur: Histogram of Participation Rates



Source: Author's calculations of IRHS data set.

3. ECONOMETRIC ISSUES

Modeling individual behavior from micro survey data to test theories from microeconomics has become an increasingly challenging task because of the difficulties in adapting formal techniques of statistical analysis to such issues (see Pudney 1989). Most of the consumer behavioral issues in economics involve studying the choices people make from a limited number of alternatives and attempting to relate the conditional probability of a particular choice to various explanatory factors, which include the attributes of the alternatives as well as the characteristics of the decisionmakers. In some cases the alternative values that the dependent variable can assume are limited by some exogenous considerations, based on either theoretical or intuitive arguments. For instance in *truncated* models we do not have information either on the x variables or the y variable when the value of y is above or below a certain point. On the other hand, in *censored* models we do have data on all x variables for all the observations; but, except for some values of y, in other cases we only know whether the observations are above or below a certain threshold (see Maddala 1983). Pioneering work on these models has been done by Tobin (1958), Amemiya (1973), and Heckman (1974). These models could be generally classified as regression models in which the dependent variable can only be observed in a limited range or limited way, and in which the dependent variable can assume either discrete values or continuous values. Such a dependent variable is referred to as a *latent* variable in the literature. The observable range of the dependent variable can be defined by the upper or lower (or both) limits, where these limits are defined either by the values of variables or the presence or absence of certain states of nature. The model in which the dependent variable assumes a continuous value above a certain limit has been named the Tobit model, after the pioneering work in consumer durables expenditure analysis by Tobin (1958). In contrast to the standard Tobit model, in which there is single censoring, there can be double censoring as well. Models of the latter type are called *Two Limit Tobit* models or *Double Censored* models.

The latent nature of PR, the dependent variable in our analysis, necessitates a treatment that could provide the required special consideration to the observed values, especially to those that are displayed as mass points at the extremes. One could argue that the above relationship could be treated with a logistic function. But the logistic function would truncate the zero and one values at the extremes and also could not handle the concentration of probability mass at the extremes. Hence, the logistic function not only does not use the available information efficiently, but also provides us with imprecise parameter estimates. On the other hand, ordinary least squares (OLS) estimation lacks the ability to treat the observed and latent variable structure that exists in the present problem and provides inconsistent and biased estimates. Hence we choose to estimate the preferred Two Limit Tobit model using a maximum likelihood estimation method in order to establish the relationship between the participation rate of workers and person-specific and household-specific explanatory variables.

The histogram in Figure 1 reveals that a substantial number of individuals are at the extremes with values of zero and one for PR. Thus, we censor at values zero and one to allow an analysis of the efficiency of individuals without the distortion caused by misspecification of the extreme values of the dependent variable. The statistical techniques relating to these models to which we turn our attention next are discussed by Maddala (1983) and Pudney (1989).

TWO LIMIT TOBIT MODEL

Let us consider a model in which the dependent variable is observed and relates to the phenomena of interest only if it lies between the upper and lower limits, namely, α_1 and α_2, defined by some theoretical propositions. The model then becomes

$$y_i^* = \beta' x_i + u_i \tag{1}$$

where y_i^* is the latent variable and u_i is the random error term, i.e., $u_i \sim N(0, \sigma^2)$. If y_i denotes the observed values of interest for the particular phenomena under consideration, then

$$\begin{aligned} y_i &= \alpha_1 \ if\ y_i^* \leq \alpha_1 \\ &= y_i^* \ if\ \alpha_1 < y_i^* < \alpha_2 \\ &= \alpha_2 \ if\ y_i^* \geq \alpha_2 \end{aligned} \tag{2}$$

where α_1 and α_2 are determined based on some a priori reasoning. The likelihood function for the model is given by

$$\begin{aligned} L(\beta, \sigma \mid y_i, x_i, \alpha_1, \alpha_2) = &\prod_{y_i = \alpha_1} \Phi\left[\frac{\alpha_1 - \beta' x_i}{\sigma}\right] \\ &\prod_{y_i = y_i^*} \frac{1}{\sigma} \varnothing \left[\frac{y_i - \beta' x_i}{\sigma}\right] \prod_{y_i = \alpha_2} \left[1 - \Phi\left[\frac{\alpha_2 - \beta' x_i}{\sigma}\right]\right] \end{aligned} \tag{3}$$

where Φ and $\varnothing$ signify the distribution and density functions, respectively, of the standard normal distribution.

We can show diagrammatically the various components of the likelihood function to be estimated and the logic behind the Two Limit Tobit model as depicted in Figure 2. As in the standard Tobit model, one can derive the first and second derivatives of the log likelihood (hereafter LL) function (see Maddala 1983). The concavity of the LL and therefore the convergence of the iterative maximum likelihood solution to a consistent and asymptomatically normal estimate can also be shown (see Pratt 1981).

This model must be estimated using maximum likelihood methods since OLS, using either the entire sample or the subsample of complete observations, produces biased and inconsistent parameter estimates (Maddala 1983).[6] To see this, first note that the conditional expectation expression for y_i is

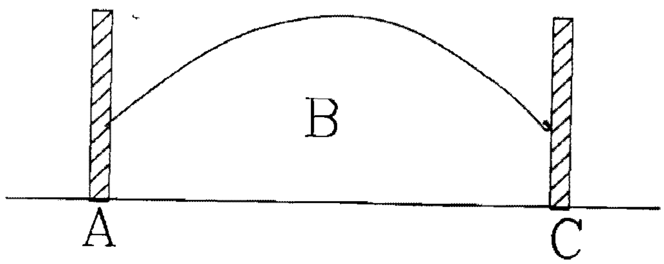
$$E\,(y_i \mid \alpha_1 < y_i^* < \alpha_2) = \beta' x_i + E\left[u_i \mid (\alpha_1 - \beta' x_i) \le u_i \le (\alpha_2 - \beta' x_i)\right]$$
$$= \beta' x_i + \sigma \frac{\varnothing_{1i} - \varnothing_{2i}}{\Phi_{2i} - \Phi_{1i}} \tag{4}$$

and that the unconditional expectation of y_i is

$$E(y_i) = P[y_i = \alpha_1]\,\alpha_1 + P\,[\alpha_1 \le y_i^* \le \alpha_2]\,E(y_i < \mid \alpha_1 \le y_i \le \alpha_2)$$
$$+ P[y_i = \alpha_2]\alpha_2$$
$$= \Phi_{1i}\alpha_1 + \beta' x_i(\Phi_{2i} - \Phi_{1i}) + \sigma(\varnothing_{1i} - \varnothing_{2i}) + (1 - \Phi_{2i})\alpha_2 \tag{5}$$

[6] Greene (1981) and Goldberger (1981) consider in more detail the properties of the OLS estimators in Tobit models in the special case when the explanatory variables are multivariate normal. However, this assumption precludes the use of dichotomous variables.

Figure 2 — Dokur: The Two Limit Tobit Model



$$A = \sum_{y_i = \alpha_1} \Phi \left(\frac{\alpha_1 - \beta' x_1}{\sigma} \right)$$

$$B = \sum_{y_i = y_i^*} \frac{1}{\sigma} \varnothing \left(\frac{y_i - \beta' x_i}{\sigma} \right)$$

$$C = \sum_{y_i = \alpha_2} \left[1 - \Phi \left(\frac{\alpha_2 - \beta' x_i}{\sigma} \right) \right]$$

where $\varnothing\left[\frac{\alpha_1 - \beta' x_i}{\sigma}\right]$ and $\varnothing\left[\frac{\alpha_2 - \beta' x_i}{\sigma}\right]$ are denoted by $\varnothing_{1i}$ and $\varnothing_{2i}$, respectively, with corresponding definitions for Φ_{1i} and Φ_{2i}.

Thus if one used either the observed y_i^* such that $\alpha_1 < y_i^* < \alpha_2$, or all the observations including the extreme values in an OLS regression on the x_i's, one would obtain biased and inconsistent estimates of the βs since the equation error terms will not have a zero mean, as seen from Equations 4 and 5. Unlike the case of standard Tobit model, a straightforward Heckman two-step estimation technique (Maddala 1983), which corrects for the nonzero expectation of the error term, is not available to us because of the doubly censored nature of the data.

DIAGNOSTIC TESTS

Tobit models are usually estimated using cross-section data, so that the specification errors most likely to occur in these types of models include omitted variable bias, heteroscedasticity and non-normality. The misspecification tests for these models are cumbersome without special subroutine programs. Here we test the alternative specifications using a Likelihood Ratio test and heteroscedasticity using a Lagrange Multiplier test as explained below.

In order to choose among alternative specifications, we use a Likelihood Ratio (LR) test which is the asymptomatic equivalent of the Lagrange Multiplier (LM) test (Godfrey 1988). The Likelihood Ratio test, which compares the maximum likelihood estimates of restricted and unrestricted versions of the model, determines the choice of the equation in relating PR to individual and household characteristics.

In the standard least squares regression models, heteroscedasticity leads to unbiased and consistent estimates but biased standard errors. In nonlinear models such as ours, if not properly controlled, heteroscedasticity also results in biased and inconsistent parameter estimates (see Godfrey 1988). The sensitivity of maximum likelihood (ML) estimators to errors in the specification of the error function in Tobit models has come under increasing scrutiny and analysis in recent years.[7]

Since heteroscedasticity is a potential problem in cross-sectional analysis (especially in censored models) and as there is no existing test for heteroscedasticity in Two Limit Tobit models, we define a Lagrange Multiplier test using the guidelines provided for defining similar tests in standard Tobit

---

[7] Also, heteroscedastic errors make the ML estimates of Tobit inconsistent, although the problem is more serious in truncated models than in censored ones (Judge et al. 1985).

models. We discuss the nature of the problem of heteroscedasticity, along with the details of the construction of the test statistic, in Appendix C. This test is applied to the models estimated in this paper. Finally, it is worth noting that rather difficult tests have been derived for non-normality by Jarque and Bera (1982), and for misspecification by Nelson (1981), in the context of standard Tobit models. If such tests were developed and conducted in the case of the Two Limit Tobit model then this would increase the robustness of our estimates.

4. A FORMAL MODEL FOR ESTIMATION

The model described below is intended to give a theoretical motivation for the empirical exercise that follows and is therefore based on the stylized facts of rural labor markets in general and of the present study village in particular (see Canagarajah 1991; see also Canagarajah and González 1991). We postulate that the quality of a worker depends on certain household-specific characteristics (X_h), such as the amount of land owned and type of family and dependency ratio, among many other variables, along with some of the observable person specific variables (X_j) like age, nutrition, gender, and reputation of the worker. We can assume the following form for the function:

$$e_i = \beta' X_{hj} + \mu \tag{6}$$

where e_i refers to the efficiency or quality of the i^{th} worker and μ is a random disturbance with zero mean and a finite variance. Unobservable worker characteristics can also be included as part of μ. If any function such as the one displayed above can be established, then it would mean that such information would be used to discriminate between workers. This is because the expected quality of workers would be a well-defined function of $\beta' X$, and workers would be ranked and hired according to that relationship. If it is a continuous function, then absolute discrimination on the basis of its value will take place.[8]

Laborers not hired in a theoretical starting period, say $t = 0$, will on average have a value of $\beta' X$ less than those who had been hired. However in the subsequent period, say $t = 1$, the employer may find that some of the hired laborers already have a lower efficiency than those outside and thus would fire them first. The process will continue until all laborers outside the pool are less productive than those in the pool. This trajectory may be because the information on β is imperfect and therefore necessitates revision every period through the information that becomes available with time. On the other hand, other unobservable effects may come to light only with time. This means that the employer would base the choice of workers on the amount of information available. This information would lead to a statistical discrimination against the lower efficiency workers. The actual mechanism of this process can take many different

---

8 However, the process of convergence of the selection process to a stationary equilibrium may not be easy to understand. Since we do not have panel data to analyze any process over time, we do not address the issue here.

paths. It is this process that we want to understand and we therefore use the PR of workers, which in reality should reveal the relationship between the observed characteristics of workers and the optimal decisions of employers. An important advantage of the method adopted here is that we are able to consider the work efficiency of the self-employed and own farmers as well, since we are using PR as opposed to wage rates in determining worker efficiency.

The ongoing research on efficiency wage theories (EWT) has yielded a plethora of results that identify variables that determine and influence the work efficiency of individuals. One can broadly classify these variables as person-specific and household-specific. By relating these variables to the slack season participation rate of workers we would be able to get a better understanding of the relationship between worker characteristics and worker efficiency. The next section presents the results of this investigation.

5. EMPIRICAL RESULTS

In the following empirical section a general-to-specific modeling technique[9] is used in constructing a very general model based on economic and statistical criteria. We begin with a general model incorporating the main person-specific and household-specific variables. Here we also test various combinations of occupational dummies in order to determine the ones that should enter the general specification along with some of the important person-specific and household-specific variables. Then we proceed to extend the preferred specification in many different ways. Initially, the preferred specification is tested to see whether males and females have different variances for the preferred specification by dividing the sample on the basis of gender. Next we try to incorporate various combinations of household composition effects, not included in the general model, which are significant and economically meaningful in understanding the work efficiency of individuals. Then we turn to the individual-specific effects, such as age and nutrition. Since nutrition is hypothesized to have a nonlinear relationship with worker efficiency (and therefore in our framework with participation rates) in the nutritional version of EWT, we test for nonlinear effects with polynomial structures and piecewise linear variations of nutrition variable. Similarly we also test for suspected nonlinear age effects using polynomials and piecewise linear variations.

In order to test the propositions on which the efficiency wage theories found in the rural labor markets literature are based, we estimate initially the *Two Limit Tobit* model with PR as the dependent variable and age, nutritional consumption, gender, and individual occupational category dummies (using agricultural laborer as the base category) as person-specific explanatory variables, and per capita landholding of the family and type of family as household-specific explanatory variables in a single equation (Equation 1, Table 2). The choice of variables was determined by what had already been identified as important in past empirical research in these fields. For example, per capita kilocalorie consumption (nutritional versions of Leibenstein 1957; Bliss and Stern 1978; Behrman et al. 1988), gender (Bardhan 1979; Bliss and Stern 1978), age (Immink and Viteri 1978), and individual work pattern (Strauss 1986; Immink and Viteri 1978) were incorporated as the main person-specific variables, while the landholding capacity of the household (Dasgupta and Ray 1986), and household composition and its effects on the welfare of the household members (Mirrlees 1975) were incorporated as household-specific variables in the general specification.

---

9 Maddala (1988) presents a brief exposition on various model selection criteria.

Table 2 —Dokur: Toward a General Specification; Dependent Variable = Participation Rate (slack season)

Variables	Equation Numbers				
	1	2	3	4	5
Constant	0.340 (2.27)	0.469 (3.18)	0.523 (3.58)	0.430 (2.96)	0.414 (3.15)
Age in years	-0.0004 (0.35)	-0.0016 (1.34)	-0.0016 (1.29)	-0.0003 (0.26)	
Gender	-0.34 (7.87)	-0.38 (8.67)	-0.402 (9.62)	-0.381 (9.71)	-0.38 (9.70)
Per capita landholding	0.055 (1.98)	0.027 (0.98)	0.35 (1.28)	0.712 (2.68)	0.071 (2.68)
Nutrition	0.0026 (2.11)	0.0022 (1.75)	0.0021 (1.61)	0.0024 (1.96)	0.0025 (2.02)
Nuclear family	0.66 (1.59)	0.583 (1.38)	0.053 (1.23)	0.058 (1.41)	0.057 (1.40)
Farmers	0.079 (1.46)				
Permanent servants	0.328 (3.71)				
Agricultural workers	0.349 (2.95)				
Shepherds and business	0.235 (2.17)				
Domestic workers	-0.142 (2.50)	-0.138 (2.39)	-0.167 (2.94)	-0.191 (3.78)	-0.189 (3.77)
All nondomestic workers		0.129 (2.39)			
All agricultural workers			0.082 (1.62)		
Permanent servants and agricultural workers				0.265 (3.78)	0.270 (3.95)
σ	0.226 (16.81)	0.233 (16.75)	0.236 (16.75)	0.230 (16.81)	0.230 (16.81)
Log likelihood	-12.216	-19.166	-20.663	-14.801	-14.834
χ^2 LM test	37.08 (10)	28.58 (7)	25.51 (7)	19.76[a] (7)	21.81 (6)

Source: Author's calculations of IRHS data set.

[a] Denotes nonrejection of null hypothesis between 1.00 and 0.05 percent.

Notes: Absolute t-ratios are given in parenthesis. χ^2 denotes LM test for heteroscedasticity, cf. Appendix C. Degrees of freedom in parenthesis.

We use LIMDEP (Greene 1989), a limited dependent variable models estimation package, to estimate our model.[10]

TOWARD A GENERAL SPECIFICATION

The large number of occupational dummies with different intensities and levels of importance compelled us to identify the important occupational dummies and to choose the appropriate specification, including person-specific and household-specific variables, which could simplify the general functional form. As a first step toward a general specification, we incorporated some important person-specific and household-specific variables with the five main occupational dummies to identify the most significant dummies. We arrived at a LL of -12.216, with coefficients significant at .05 level for almost all the variables except age and farmers dummy, as is shown in Equation 1. Next we incorporated an occupational dummy, which combined the first four dummies, i.e., all nondomestic workers dummy. All of the latter had previously shown a positive relationship with PR. Domestic workers dummy was included separately as it had assumed a negative relationship. The results are shown in Equation 2. The LL of -19.166 implies that this restricted specification can be rejected using the LR test, since -2(-19.116 + 12.216) = 13.8 is greater than the critical value of $\chi^2(3,.05) = 7.91$. It also reduced the significance of all the variables that had been previously observed to be significant. We further tested by combining only the first three dummies, which are agriculture related, and by dropping DUM4, which relates to shepherds. We arrived at a LL of -20.663, with insignificant coefficients similar to those in Equation 2. Comparing Equations 1 and 3, we find that the likelihood ratio is equal to 16.8, while the critical value is 7.91. After dropping the farmers dummy, we tried the collective dummy variable combining permanent servants and agricultural workers, with domestic workers as a separate dummy variable. This gave not only a LL of -14.801, but significant estimates for most variables, except for age as well. Using the LR test, we find that this test specification (Equation 4) is preferred to the one that includes all the dummies, since the critical value for $\chi^2(3,.05) = 7.91$, which is greater than twice the LL difference between Equation 1 and 4, namely, 2(-14.801 + 12.216) = 4.62. It is also worth noting that the LM test for heteroscedasticity does not reject the null hypothesis for Equation 4, between 0.01 and 0.005 significance levels, while the null is rejected in the other three equations. Thus we chose Equation 4 as the preferred specification relative to the one that incorporates all the dummies and the other variants tried.

Our preferred specification for the relationship of PR to the person-specific and household-specific variables also gives meaning to most of the postulated relationships, especially those that are important in developing country labor markets, in the efficiency wage theory literature. First, the

---

[10] The program uses Newton's iterative technique of maximum likelihood estimation and analytical second derivatives of the LL function to estimate the variances of the parameter estimates (see Rosett and Nelson 1975; Olsen 1978).

gender of the individual assumes a significant negative coefficient, which implies that females tend to participate less in the slack season. This is not very surprising in a patriarchal society, such as the Indian rural sector, where heavy tasks and the role of family head are assumed by males, who have to search for work in the slack season to support the family, as opposed to the female members of the family, who tend to do the household chores. We take up the issue of the gender-based differences in PR in the extensions. Next the household composition variables included, namely, per capita landholding power of the family and type of family, assume significant positive coefficients. Hence, agricultural laborers from families with large landholdings tend to participate more in agriculture and related activities (see Dasgupta and Ray 1986, 1987). As the ability to self-insure is an important determinant of nuclearization, one should expect individuals from nuclear families to be more able workers (see Newbery 1989). Since there was a danger of including too many household-specific variables in the specification, we did not incorporate all the household-specific variables, but have chosen the most obvious ones through economic and statistical considerations. However, we address this issue in a subsequent section and test how various other household composition variables perform in relation to the postulated relationship.

Turning next to the person-specific variables, other than the gender variable, it can be observed that the nutritional status of the individuals also consistently assumes a positive and significant coefficient. This is a very interesting finding since the nutritional version of EWT would predict this type of relationship. We test later for the nonlinear *nutrition-efficiency* effect hypothesized by Leibenstein (1957), Bliss and Stern (1978), and Dasgupta and Ray (1986), among many others. The age variable does not assume a significant coefficient (see also Equation 5); we suspect this may be because only certain age groups are important, and also because, as the literature suggests, there may be a nonlinear relationship between age and PR. We investigate this in detail later. Checking the dummy variables, we find that permanent servants and agricultural workers' dummy constantly assumes a positive significant coefficient, which implies that those who classify themselves as primarily agricultural workers and permanent servants have higher participation rates, while the consistent negative and significant coefficient of domestic workers' dummy implies that those who classify themselves as domestic workers tend to have lower participation rates in the agricultural labor market. The simplification achieved in characterizing the occupational dummies, as in Equation 4, holds consistently when alternative specifications and extensions are investigated.

EXTENSIONS

Testing for Gender-Based Differences in PR Effects

We were interested in assessing whether different participation rates were primarily a result of gender difference of the individuals and also whether the preferred specification shows any gender-based segregation of the sample, since the gender variable was always significantly related to the observed PR. Thus,

Table 3 —Dokur: Testing for Gender-Based Differences; Dependent Variable = Participation Rates

	Equation Numbers		
Variables	6	7	8
Constant	0.175 (1.21	0.347 (1.59)	0.021 (0.11)
Age in years	0.0031 (1.83)	0.00024 (1.26)	-0.0020 (1.33)
Per capita landholding	0.102 (2.65)	0.096 (2.34)	0.052 (1.53)
Nutrition	0.363 (2.76)	0.0023 (1.26)	0.0032 (1.94)
Nuclear family	0.0092 (0.15)	0.005 (0.009)	0.102 (1.95)
Permanent servants and agricultural workers	0.32 (4.29)	0.305 (3.89)	
Domestic workers	-0.21 (2.10)	-0.232 (2.17)	-0.165 (2.96)
Age x sex	-0.0056 (2.58)		
Kilocalorie consumption x sex	-0.0017 (1.81)		
Domestic workers x sex	0.020 (0.18)		
Per capita landholding x sex	-0.053 (1.00)		
Nuclear family x sex	0.08 (1.00)		
σ	0.226 (16.81)	0.236 (11.59)	0.216 (12.19)
Log likelihood	-12.226	-10.108	-1.1824
χ^2 LM test	32.10 (11)	16.52[a] (6)	15.96[a] (5)

Source: Author's calculations of IRHS data set.

[a] Denotes nonrejection of the null hypothesis between 1.00 and 0.05 percent.

Notes: Absolute t-ratios are given in parenthesis. χ^2 denotes LM test for heteroscedasticity. Degrees of freedom in parenthesis.

in order to test for the effect of gender, we first assumed a common variance of males and females and then interacted all the variables in Equation 4 with the gender dummy variable. The resulting equation (shown as Equation 6 in Table 3) not only failed to pass the LR test as LR = 2(-14.801 + 12.226) = 5.20 as opposed to $\chi^2(5, .05) = 11.07$, but also most of the interaction variables were extremely insignificant. Along with the LR test results, the LM test (which rejects the null hypothesis of heteroscedasticity in the case of Equation 6) enables us to prefer Equation 4 with gender as a single variable to the interacted form of Equation 6. When we allowed for separate variances in the specifications for the 84 males (Equation 7) and 88 females (Equation 8) in the sample, the significant variables were different for the different genders. The collective LL was -1.182 + -10.108 = -11.290, which was not a substantial improvement over the equation with the gender variable interaction, which had a LL of -12.226, and also could not be preferred to Equation 4, with the gender variable as a simple dummy variable. Equations 7 and 8 do not reject the null hypothesis of homoscedasticity in terms of the LM test, as the χ^2 values indicate. Comparing these equations with that which includes the gender variable as an explanatory variable (Equation 4, with a LL of -14.801), we conclude that the gender-based segregation of the sample is unwarranted. Thus, both males and females have a more or less common structure in the hypothesized PR relationship.

Variations of Household Composition and Land Ownership Variables

Next we inquire into the role of household composition variables in the activity patterns and welfare of household members. The literature on agricultural household models provides evidence of increasing appreciation of the necessity of incorporating household composition variables in the analysis of household production and consumption decisions (Singh et al. 1986). Along the same lines, we also incorporate variables representing household characteristics as explanatory variables in our specification of the individual participation rate. Not only are these household-specific variables important in determining the efficiency of individuals, but they provide important information about workers in the rural labor market that employers can observe. Employers can use this information for screening workers for their ability and efficiency to help determine whom and when to hire. In the general specification we have already incorporated per capita landholding and nuclear family dummy.

We also tried some of the other possible family background variables. When the dependency ratio, defined as the number of children as a ratio to the working members of the household, was incorporated instead of per capita landholding, the variable assumed a negative significant coefficient and provided a LL of -14.910, while increasing the significance of nuclear family dummy (Equation 9, Table 4). When the total number of members in the family was incorporated in Equation 4 instead of nuclear family dummy, it gave a LL of -14.700 and a negative coefficient for household size significant at .10 level, without a substantial reduction in the significance of other variables as shown in equation 10. Therefore we could say that the larger the size of the family the smaller the potential PR of individuals belonging to it.

Table 4 —Dokur: Variations of Household Composition Effects; Dependent Variable = Participation Rates

	Equation Nos.		
Variables	**9**	**10**	**11**
Constant	0.686 (4.25)	0.544 (3.66)	0.569 (3.75)
Age in years	-0.0008 (0.64)	-0.0002 (0.18)	-0.0003 (0.26)
Gender	-0.393 (10.05)	-0.385 (9.81)	-0.386 (9.85)
Per capita landholding		0.073 (2.72)	0.062 (2.45)
Nutrition	0.0023 (1.82)	0.0021 (1.72)	0.002 (1.60)
Nuclear family	0.094 (1.99)		
Permanent servants and agricultural workers	0.216 (3.16)	0.249 (3.60)	0.239 (3.46)
Domestic workers	-0.193 (3.81)	-0.184 (3.62)	-0.219 (3.66)
Dependency ratio	-0.448 (2.62)		
Household size		-0.0066 (1.48)	
Number of children			-0.017 (1.68)
σ	0.231 (16.82)	0.229 (16.81)	0.229 (16.80)
Log likelihood	-14.910	-14.700	-14.385
χ^2 LM test	18.29[a] (7)	19.35[a] (7)	18.67[a] (7)

Source: Author's calculations of IRHS data set.

[a] Denotes nonrejection of the null hypothesis between 1.00 and 0.05 percent.

Notes: Absolute t-ratios are given in parenthesis. χ^2 denotes LM test for heteroscedasticity. Degrees of freedom in parenthesis.

As another extension, we incorporate the number of children in each household in the specification of Equation 4, instead of nuclear family dummy, giving a LL of -14.385 as shown in Equation 11. The negative coefficient indicates the possible negative relationship that the number of children variable has with PR. However, as can be seen from Equation 11, this formulation reduces the significance of the nutrition variable. All three specifications accommodating household composition effects do not reject the assumption of homoscedasticity in terms of the LM test, as can be observed from the χ^2 values reported with the respective equations.

Nutrition Effects

Next we turn to the person-specific variables incorporated in the general specification in order to see how alternative specifications, using the existing variables, perform with respect to the hypothesized relationship.

First, we test the most important of all the person-specific variables: the nutritional status of the individuals. This has been hypothesized in the EWT literature as having a nonlinear relationship with the efficiency of workers, and we assume a similar pattern in relation to participation rates, and we test for them. We propose to adopt two methods to allow for the possible nonlinearity of the relationship, the diagrammatic forms of which are shown in Figures 3A and 3B. They are:

1. *Piecewise Linear*: i.e., interacting nutrition with various dummy variable classifications of the variable according to nutritional measures (Figure 3A).

2. *Polynomials*: i.e., nutrition, "nutrition^2" and/or "nutrition^3" instead of nutrition alone (Figure 3B).

First, a dummy structure was tried, with nutritional status being divided into the three groups according to the concept of Required Dietary Allowance (RDA), which defines the relative levels of nutritional welfare of individuals:[11]

a. Below 80 percent standard: "nutrition < 80" = 1 if nutritional standard is less than 80, and 0 otherwise.

b. Between 80 percent and 110 percent: "80 < nutrition < 110" = 1 if nutritional standard lies between 80 and 110, and 0 otherwise.

c. Above 110 percent standard: "nutrition > 110" = 1 if nutritional standard is greater than 110, and 0 otherwise.

---

[11] see "Incidence of Undernutrition" in Canagarajah (1991), Appendix A of Chapter 1.

Figure 3 — Hypothesized Nonlinear Relationship of Age or Nutrition with PR

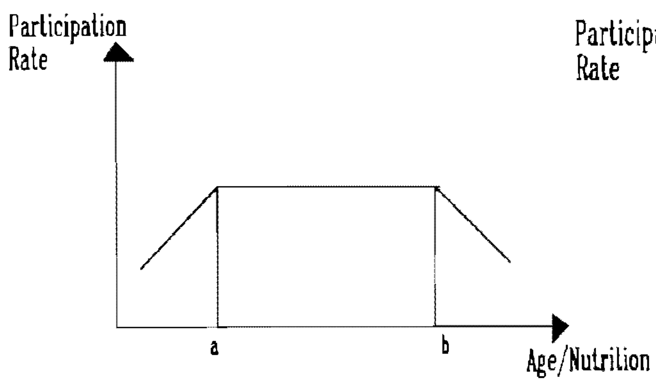


Figure 3A

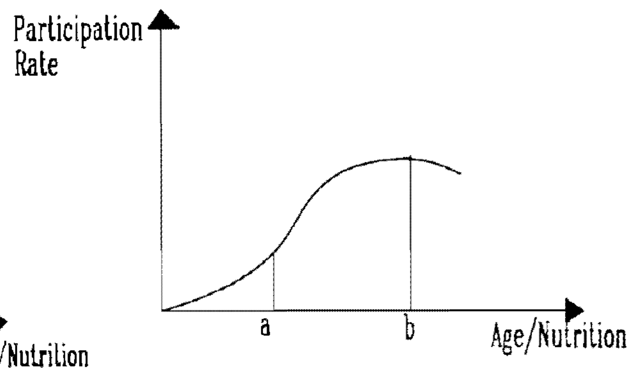


Figure 3B

Each of these dummies was interacted with the respective nutrition level of individuals to form the nutrition variables, namely, "nutrition < 80," "80 < nutrition < 110," and "nutrition > 110." This was done to allow for the probable nonlinear effect of nutrition with respect to PR within these three groups. These nutrition variables were incorporated instead of nutritional standard, and the specification was tested using LR. First, we introduced all three variables, namely, "nutrition < 80," "80 < nutrition < 110," and "nutrition > 110," instead of nutrition, and we found that none of the nutrition variables did assume significant coefficients individually. When we tested for the collective significance through the LR test we found that the LR = -2(-14.801 + 13.722) = 2.158 is less than $\chi^2(2, .05) = 5.99$. Thus, Equation 12 is not a significant improvement over Equation 4. However, when "80 < nutrition < 110" was incorporated in Equation 4 instead of nutrition variable (with age variable), the LL became -15.047, as can be seen from Equation 14 with a negative coefficient for "80 < nutrition < 110" significant at 0.10 level. This implies that the individuals with a lower nutritional status on average have a lower participation rate. On the other hand, as can be observed from Equation 15, when "nutrition > 110" was included instead of nutrition, the LL increased to -14.087, and the nutrition effect term had a positive and significant coefficient, indicating the preferential position they enjoy in terms of participation rates. But when "nutrition < 80" was incorporated, the LL became -16.669, which is a substantial reduction in the LL and the nutrition variable became completely insignificant (see Equation 13 in Table 5). Thus, "nutrition > 110" performs better than any other variation. This indicates that those who are well-nourished have the most significant relationship with PR, while those in the lower spectrum do not display such a pattern.

Next the polynomial variants of nutritional consumption — nutrition, "nutrition^2," and "nutrition^3" — were incorporated with the objective of accounting for the potential nonlinear effect nutrition is hypothesized to have on efficiency and the resulting PR of workers, following the proposals of Leibenstein (1957), Bliss and Stern (1978), and Strauss (1986), among many others. First, we incorporated all three nutrition polynomial variables, with age accounting for age effects. As shown in Equation 16, when nutrition and "nutrition^2" were incorporated, the LL became -11.046, which using the LR test is a significant improvement over Equation 4. Both nutrition and "nutrition^2" are extremely significant, strongly supporting the propositions of the nutritional version of EWT. Similarly, when "nutrition^2" and "nutrition^3" were incorporated, the LL became -10.927, also a substantial improvement over Equation 4. Again, both the nutrition variables assume significant coefficients.

Since the polynomial nutrition variables assume a combination of negative and positive coefficients, we checked whether the marginal effect of nutrition

Table 5 —Dokur: Variations of Nutritional Consumption Effect on Participation Rates; Dependent Variable = Participation Rates

	Equation Numbers						
Variables	**12**	**13**	**14**	**15**	**16**	**17**	**18**
Constant	0.46 (1.79)	0.69 (11.51)	0.748 (11.12)	0.671 (11.29)	1.84 (3.49)	1.08 (4.61)	0.69 (11.51)
Age in years	-0.0006 (0.54)	-0.0006 (0.49)	-0.0011 (0.87)	-0.0008 (0.67)	-0.00071 (0.59)	-0.0008 (0.68)	-0.0007 (0.55)
Gender	-0.375 (9.57)	-0.378 (9.47)	-0.371 (9.41)	-0.375 (9.62)	-0.376 (9.85)	-0.375 (9.79)	-0.377 (9.49)
Per capita landholding	0.066 (2.45)	0.073 (2.72)	0.065 (2.41)	0.064 (2.39)	0.0614 (2.35)	0.059 (2.52)	0.073 (2.71)
Nuclear family	0.056 (1.38)	0.05 (1.20)	0.048 (1.17)	0.052 (1.28)	0.042 (1.05)	0.041 (1.01)	0.049 (1.19)
Permanent servants and agricultural workers	0.265 (3.82)	0.267 (3.76)	0.267 (3.81)	0.265 (1.28)	0.261 (1.05)	0.263 (1.01)	0.267 (1.19)
Domestic workers	-0.195 (3.82)	-0.219 (4.43)	-0.223 (4.61)	-0.265 (4.27)	-0.261 (3.98)	-0.263 (4.02)	-0.267 (4.51)
Nutrition < 80	0.0029 (0.87)	-0.0002 (0.26)					
80 < nutrition < 110	0.0020 (0.82)		-0.007 (1.83)				
Nutrition > 110	0.0025 (1.20)			0.0008 (2.31)			
Nutrition					-0.027 (2.53)	-0.0002 (2.35)	
$(\text{Nutrition})^2$					0.00015 (2.77)		
$(\text{Nutrition})^3$						0.00001 (2.58)	
σ	0.228 (16.77)	0.223 (16.81)	0.230 (16.80)	0.228 (16.81)	0.224 (16.80)	0.224 (16.80)	0.223 (16.81)
Log likelihood	13.722	-16.669	-15.047	-14.087	-11.046	-10.927	-16.703
χ^2 LM test	28.95 (9)	25.59 (7)	27.22 (7)	26.52 (7)	25.12 (8)	26.31 (8)	23.32 (6)

Source: Author's calculations of IRHS data set.

Notes: Absolute t-ratios are given in parenthesis. χ^2 denotes LM test for heteroscedasticity. Degrees of freedom in parenthesis.

on PR is positive or not.[12] The effective influence, or the marginal effect, of nutrition on PR was calculated from estimates obtained from Equation 16; the effective influence at the mean value of nutrition (i.e., 100.66) was 0.003. Individuals who fell below approximately 90.75 of their nutrition standard had a negative relationship with PR. This indicates that nutrition generally positively influences PR, as the figure for mean nutrition shows. A similar figure for the nutrition effect was obtained for Equation 17, although at 0.03 this was somewhat higher. These results are in line with the general nutritional version of efficiency wage theory.

The LR test suggests that these specifications are superior to one in which nutrition alone is incorporated in the specification. We also found that dropping nutrition from Equation 4 indicates that the specification is worse by an LR test, since without nutrition Equation 4 has a LL of -16.703 (see Equation 18), as opposed to the LL of -14.801. The nonlinearity from the nutritional version of the EWT (see Leibenstein 1957; Bliss and Stern 1978) can be substantiated by the significant coefficients (at .05 level) of "$nutrition^2$" and "$nutrition^3$." This type of nonlinearity performs better than the piecewise linear specification using "nutrition > 110" and "80 < nutrition < 110." However, it is important to note that most of these equations fail marginally the LM test for homoscedasticity, as the χ^2 statistic reveals.

Age Effects

Next, alternative representations of the age variable were incorporated into the general specification to see whether they increased the explanatory power of Equation 4 and whether they have a specific functional relationship with the PR of individuals. The negative coefficient indicates that the higher the age the lower the participation rate of individuals, which is probably because people become physically weak with higher age; although for the most part the variable is insignificant. Thus, when we tested for the importance of the age variable by dropping it from the specification altogether, it not only affected the significance of the other variables incorporated in the equation, but the null of the LM test of heteroscedasticity was rejected for the equation, as can be seen from Equation 5. This indicates the necessity of incorporating age, as well as varying its form to capture its significance in the specifications. However, as we can observe from Equation 6, both age and (age x gender) assume significant

---

12 This effective influence of nutritional standard is measured as follows:

$$\frac{\partial PR}{\partial\, nutrition} = coeff.\ of\ nutrition + 2\left(coeff.\ of\ [nutrition^2]\right)\ nutrition$$

$$\frac{\partial PR}{\partial\, nutrition} = coeff.\ of\ nutrition + 2\left(coeff.\ of\ [nutrition^2]\right)$$
$$nutrition + 3\left(coeff.\ of\ [nutrition^3]\right)(nutrition)^2$$

coefficients, and thus we observe that females display a marked negative effect with PR, while for males there is not such a strong negative relationship. Employers may prefer men to women because, as Bardhan (1979) argues, women's labor supply is irregular due to village social customs and conventions, which make them less flexible in meeting weather-dependent and intermittent labor market demands. Therefore, in line with EWT, the profit-maximizing employers are justified in preferring younger males to work on their farms and in other agriculture-related activities in a rural economy.

The insignificant negative linear relationship between age and PR (as in Equation 4) is too naive to accept as an explanation for the whole population, and in general one could expect a nonlinear relationship between age and PR. The relationship could take the shape depicted in either Figure 3A or 3B.[13] This could be captured either by using polynomials or by generating interaction variables with different dummies for different age group classifications, as was done for the nutrition variable. Since the LL became -14.834 as opposed to -14.801 when age was dropped (see Equation 5), it seemed sensible to test for the importance of the variable in the hypothesized relationship, and to drop it if we do not find other transformations of the same variable improving the specifications.

First, we tried a polynomial structure with age and "age^2." These did not improve the criterion (as can be seen in Equation 24 the LL was -14.296) and both age and age^2 assumed very insignificant coefficients. Similarly, when "age^2" and "age^3" were incorporated to represent the age effect (Equation 25), the LL was -14.449 and the age variables again were insignificant. Hence, the polynomial structure does not explain the hypothesized nonlinear relationship between PR and age.

The ages of individuals were therefore divided into three categories, and three dummy variables were created. They were

1. Less than 16 years: "age < 16" = 1 for those below 16 years, and 0 for others.

2. Between 16 and 50 years: "16 < age < 50" = 1 for those between 16 and 50 years, and 0 for others.

3. Greater than 50 years: "age > 50" = 1 for those above 50 years, and 0 for others.

Each dummy variable was interacted with the age of the respective individuals to form "age < 16," "16 < age < 50," and "age > 50," thus allowing them to take into account the nonlinearity. These variables were independently

---

[13] These figures are similar to Figure 1 of Standing and Sheehan (1978) for participation rates of rural males and females in Sri Lanka. The book indicates evidence for similar pattern for the PR of workers in most rural areas in Asia.

and collectively tried in the general model derived in the last section (Equation 4). First, we introduced "age < 16," "16 < age < 50," and "age > 50" instead of age in Equation 4. As can be seen from the results for Equation 19, none of these age variables were independently significant. When we tested for collective significance, we found that the LR was 3.31, as opposed to the $\chi^2(2, .05) = 5.99$. This means they are insignificant even collectively. "Age < 16" and "age > 50" were incorporated together in the general functional form, giving a LL of -13.928 (Equation 23), which is not significantly higher than the LL of "age > 50" alone which is -13.977 (Equation 22). The relative magnitude, significance, and sign of the estimates of the coefficients did not change substantially; both, therefore, were preferred and yielded interpretations that were in line with the theory behind efficiency wages. As can be observed from Equation 21, "age < 16" variable assumed an insignificant coefficient with a LL of -14.826, whereas in Equation 22 "age > 50" assumed a negative coefficient with a LL of -13.977. When "16 < age < 50" was incorporated to represent age, the LL became -13.433 as shown in Equation 20, and the age coefficient assumed a positive coefficient at 0.10 level of significance. This indicates that middle-aged individuals have a positive, well-established relationship with PR, and those above 50 years a negative relationship. Those below 16 years show no evident relationship. Using the LM test of heteroscedasticity, the null hypothesis is not rejected for most of the above variants of the age-experience relationship.

The significant, positive coefficient of "16 < age < 50" as opposed to the negative coefficient of "age > 50" indicates the employers' preference of younger individuals to older ones in the agrarian labor market. This is in line with the expectation that adults with greater physical ability, experience, and skill are preferred in the agrarian labor market. Therefore, it is optimal and rational for the employer to hire younger people in preference to older ones for the primarily manual work in the agrarian sector.

When we allowed for different combinations of polynomial and piecewise transformations for age and nutrition variables in the same equation, we got neither a substantial improvement over our previous specifications nor any further information on the relationships established through earlier equations. Thus we chose not to report those results here.

As can be seen from the results in Table 5, the "nutrition > 110" variable assumed a higher significance than nutrition, indicating that nutrition matters more for those who are relatively healthy by nutritional consumption standards than for those who are deficient in terms of the required dietary allowance (RDA) for their age and physical stature. It may also indicate that the nutrition variable, which accounts for the nutritional welfare aspect of EWT, may have less "noise" in the higher range of the variable, thus enabling us to identify the expected significant effect of the nutrition variable. The same is true for the age variable (see results in Table 6). The age variable, which did not assume a significant coefficient by itself in Equation 4, has assumed significant coefficients with the expected signs with the piecewise linear variants defined

Table 6 —Dokur: Age Effects on Participation Rate of Individuals; Dependent Variable = Participation Rates

	Equation Numbers						
Variables	**19**	**20**	**21**	**22**	**23**	**24**	**25**
Constant	0.345 (3.70)	0.372 (2.81)	0.414 (3.15)	0.446 (3.35)	0.446 (3.36)	0.335 (1.95)	0.398 (2.75)
Gender	-0.375 (9.63)	-0.374 (9.61)	-0.38 (9.65)	-0.38 (9.72)	-0.38 (9.66)	-0.38 (9.65)	-0.38 (9.65)
Per capita landholding	0.071 (2.69)	0.068 (2.61)	0.071 (2.65)	0.07 (2.65)	0.069 (2.59)	0.07 (2.65)	0.071 (2.67)
Nutrition	0.0025 (2.02)	0.003 (2.05)	0.003 (2.03)	0.002 (1.88)	0.002 (1.90)	0.003 (2.01)	0.003 (2.00)
Nuclear family	0.040 (0.96)	0.045 (1.10)	0.057 (1.40)	0.052 (1.27)	0.052 (1.28)	0.054 (1.31)	0.054 (1.31)
Permanent servants and agricultural workers	0.278 (3.94)	0.285 (4.18)	0.271 (3.87)	0.26 (3.81)	0.265 (3.80)	0.283 (3.93)	0.275 (3.89)
Domestic workers	-0.191 (3.81)	-0.193 (3.88)	-0.189 (3.78)	-0.196 (3.91)	-0.197 (3.92)	-0.194 (3.85)	-0.193 (3.82)
Age						0.0054 (0.93)	
Age^2						-0.00008 (1.01)	0.00006 (0.67)
Age^3							-0.000001 (0.75)
Age < 16	0.0051 (0.70)		-0.0006 (0.13)		-0.0016 (0.31)		
16 < age < 50	0.0028 (1.25)	0.002 (1.68)					
Age > 50	0.0003 (0.23)			-0.0011 (1.31)	-0.0011 (1.34)		
σ	0.227 (16.79)	0.227 (16.79)	0.230 (16.79)	0.229 (16.81)	0.228 (16.79)	0.229 (16.78)	0.229 (16.80)
Log likelihood	-13.146	-13.433	-14.286	-13.977	-13.928	-14.296	-14.449
χ^2 LM test	24.66 (9)	23.39 (7)	23.77 (7)	19.27[a] (7)	19.78[a] (8)	22.38[a] (8)	19.93[a] (8)

Source: Author's calculations of IRHS data set.

[a] Denotes nonrejection of the null hypothesis between 1.00 and 0.05 percent.

Notes: Absolute t-ratios are given in parenthesis. χ^2 denotes LM test for heteroscedasticity. Degrees of freedom in parenthesis.

for the higher spectrum of age category, namely "16 < age < 50" and "age > 50" as opposed to "age < 16."

6. CONCLUSION

The objective in this paper was to estimate a relationship on the demand curve between the efficiency of workers, as evidenced by their participation rates, and the characteristics of workers. Thus, given that employers in a surplus labor market will hire workers with the highest efficiency levels, we can identify the efficient workers through their participation rates (or conversely employer hiring rates) in the labor market.

We provide consistent evidence in our initial estimates for the *consumption-nutrition-efficiency* nexus by relating the kilocalorie consumption of individuals (nutrition variable and variations of it) to their participation rates. This enables us to make much stronger statements about the expected EWT propositions that one would expect to find in the LDC rural sector and that are hypothesized in the theoretical literature by Mirrlees (1975), Bliss and Stern (1978), and Dasgupta and Ray (1986) among many others.

The econometric model used here was designed to relate the efficiency of workers in a meaningful way to their respective participation rates without the probable distortionary effects of those with extreme values of participation rates about whom we lack adequate information to infer their efficiency. In terms of the diagnostics, the heteroscedasticity test developed specifically for this model, despite its poor small sample properties, did not always reject the null hypothesis. By demonstrating how the efficiency of workers can be meaningfully related to their respective characteristics and endowments through their slack season participation rates in a demand-constrained, surplus labor market of a village economy, this paper shows a new method for understanding the efficiency of workers.

APPENDIX A

DEFINITION OF VARIABLES

PR	Slack season participation rate defined as $0 \leq PR \leq 1$.
Gender	Sex of the individual (0 for males; 1 for females).
Nutrition	Kcal nutrition standard of the individual over the peak season expressed as a percentage of each individual's requirement.
Age	Age in years.
Household size	Number of members in the family.
Number of children	Number of children in the family.
Number of adults	Number of working members in the family.
Dependency ratio	Dependency (or adult-child) ratio.
Total household income	Total (wage + nonwage) income of each household (in rupees per annum).
Total wage income	Total wage income of each household (in rupees per annum).
Per capita income	Per capita income for each household (in rupees).
Nuclear family	Type of family (dummy variable: 1 = nuclear; 0 = others).
Landholding	Operational landholding last kharif (peak) period.
Per capita landholding	Per capita operational landholding last kharif period.
Individual total income	Individual total wage income (in rupees per annum) = cash + kind.
Individual cash income	Individual wage income in cash (in rupees per annum).

Individual kind income	Individual wage income in kind (in rupees per annum).
Weight-for-height	Weight-for-height expressed as a proportion of individual's requirement.
Height	Height expressed as a proportion of each individual's requirement (Indian standards).
Weight	Weight expressed as a proportion of each individual's requirement (Indian standards).
Arm circumference	Arm circumference expressed as a proportion of each individual's requirement.
Protein	Protein consumption expressed as a proportion of each individual's requirement.

OCCUPATIONAL DUMMIES

Base Category	Agricultural laborers (as primary occupation).[14]
DUM1	Farmers (as primary occupation).
DUM2	Permanent servants.
DUM3	Combination of 0, 1, and 2.
DUM4	Shepherds and business.
DUM5	Household or domestic workers.
DUM6	Construction workers.

14 Some small landowners have agricultural labor as primary occupation.

APPENDIX B

FREQUENCIES AND SUMMARY STATISTICS

Table B.1 — IRHS Data Set: Frequency of Occupational Categories

Code	Description	Number	Valid Percentage
0	Agricultural laborer	34	19.77
1	On own farm	72	41.86
2	Permanent servant	12	6.98
3	0, 1, and 2	10	5.81
4	Shepherd	6	3.49
5	Other business	6	3.49
6	Household work or none	32	18.60
	Total	**172**	**100.00**

Source: Author's calculations.

Table B.2 — IRHS Data Set: Median Values of Main Variables

Variable	**Median**
Participation rate	0.496
Landholding	6.000
Household size	10.000
Number of children	3.000
Number of adults	1.000
Total household income	10,170.000
Household wage income	654.500
Total household expenditure	10,766.200
Household food expenditure	757.650
Nutrition	103.480
Protein	112.050
Arm circumference	83.840
Weight-for-height	80.850
Weight	78.410
Height	95.100
Individual wage income	10.000
Individual kind income	0.000
Individual total income	27.000

Source: Author's calculations.

Table B.3 —IRHS Data Set: Summary Statistics of Person-Specific Variables

Variable	Mean	Standard Deviation	Minimum	Maximum	Valid N
Participation rate	0.528	0.323	0.00	1.00	172
Nutrition (calorie)	100.660	15.411	43.58	127.20	172
Protein	112.627	21.636	41.71	158.51	172
Age in years	32.810	15.701	6.00	75.00	172
Arm circumference	83.552	7.522	66.06	104.35	172
Weight-for-height	82.831	9.604	57.16	109.12	172
Gender	0.512	0.501	0.00	1.00	172
Individual cash income	117.820	172.100	0.00	922.50	172
Individual kind income	42.880	91.760	0.00	517.00	172
Individual total wage Income	160.710	233.570	0.00	925.00	172
Permanent servants and agricultural workers	0.105	0.307	0.00	1.00	172
Domestic worker	0.186	0.390	0.00	1.00	172

Source: Author's calculations.

Table B.4 —IRHS Data Set: Summary Statistics of Household-Specific Variables

Variable	Mean	Standard Deviation	Minimum	Maximum	Valid N
Participation rate	0.528	0.323	0.000	1.000	172
Landholding	8.389	10.160	0.000	40.000	172
Per capita landholding	0.732	0.719	0.000	3.280	172
Household size	10.058	4.235	3.000	23.000	172
No. of children	3.773	1.807	1.000	9.000	172
No. of adults	6.289	2.943	2.000	14.000	172
Nuclear family	0.314	0.465	0.000	1.000	172
Dependency ratio	0.385	0.130	0.125	0.714	172
Per capita income	1,443.300	961.590	339.200	4,989.000	172
Total household income	16,049.000	15,883.000	1,540.000	59,870.000	172
Household wage income	1,153.500	1,558.700	0.000	6,023.000	172
Total household expenditure	17,363.000	17,869.00	921.000	68,811.500	172
Household food expenditure	890.020	431.800	256.000	1,964.250	172

Source: Author's calculations.

APPENDIX C

AN LM TEST FOR HETEROSCEDASTICITY

Since it is natural to expect the problem of heteroscedasticity in cross-section analysis of this kind, we have developed a Lagrange Multiplier (LM) Test for the Two Limit Tobit Model based on the heteroscedasticity tests that exist for the Standard Tobit models (see Godfrey 1988). We follow the standard procedure and set up an alternative hypothesis of the type employed by Breusch and Pagan (1979). We write the alternative hypothesis as

$$\sigma_i^2 = h\left(\sigma^2 + \underline{z}'_i\,\gamma\right) \qquad i = 1,\ldots,n \tag{7}$$

where $\underline{z}_i$ and γ are k - 1 dimensional vectors with the variables of $\underline{z}_i$, where $\underline{x}_j' = (1, \underline{z}_j')$ are exogenous and satisfying the necessary regularity conditions. The function $h(.)$ is assumed twice differentiable and is defined such that $h'(\sigma^2) = 1$. The null hypothesis to be tested is

$$H_0 : \gamma_2 = \ldots = \gamma_k = 0 \tag{8}$$

Thus under the null σ_i^2 is a constant and equal to $h(\sigma^2)$. Let $\theta' = (\beta', \theta^2, \gamma')$ denote the parameter vector for the alternative and $\theta' = (\beta', \theta^2, 0')$ denote the constrained MLE obtained by imposing H_0.

The Log Likelihood for the heteroscedasticity censored Tobit model defined by (1), (3), and (7) is

$$\begin{aligned} L^* &= \sum_{i=1}^{n} l_i(\theta) \\ &= \sum_{i=1}^{n} \left\{ d_{1i} \times ln\,\Phi\left[\frac{\alpha_1 - \beta' x_i}{\sigma_i}\right]\right\} - \frac{1}{2} d_{2i}\left[ln\,(\sigma_i^2) + \left[\frac{y_i - \beta' x_i}{\sigma_i}\right]^2\right] \\ &\quad + d_{3i} \ln\left[1 - \Phi\left[\frac{\alpha_2 - \beta' x_i}{\sigma_i}\right]\right] \end{aligned} \tag{9}$$

where $d_{1i} = 1 \;\; \textit{if}\; y_i = \alpha_1, \;\; d_{1i} = 0 \;\textit{otherwise}$

$d_{2i} = 1 \;\; \textit{if}\; y_i = y_i^*, \;\; d_{2i} = 0 \;\textit{otherwise}$

$d_{3i} = 1 \;\; \textit{if}\; y_i = \alpha_2, \;\; d_{3i} = 0 \;\textit{otherwise.}$

Whatever the choice of the variables for z_i, it remains to construct the asymptotic χ^2 criterion, which serves to check the significance of $\partial l(.) / \partial\gamma$. The form of the *efficient score statistic* criterion would be

$$d(\hat{\theta})' . [I(\hat{\theta})]^{-1} . d(\hat{\theta}) \qquad \textit{where} \qquad d(\hat{\theta})' = \left[0, 0, \frac{\partial l_i(\hat{\theta})'}{\partial\gamma} \right] \tag{10}$$

$I(\hat{\theta})$ is an estimate of the information matrix based upon θ. The LM statistic would be the same for all alternative hypotheses that are locally equivalent with respect to the null model, provided that a common consistent estimate of the information matrix I is used. Efron and Hinkley (1978) and Brendt et al. (1974) propose alternative consistent estimates of I.

Since difficulties in calculating the second order partial derivatives are unavoidable, tests based upon either the Hessian $D(\theta)$ or its expected value are unattractive. Thus, we consider an estimate of I, which is derived from the fundamental information matrix equality

$$- \mathrm{E}[D(\theta)] = \sum_i \mathrm{E} \left[\frac{\partial l_i(\theta)}{\partial\theta} \right] \cdot \left[\frac{\partial l_i(\theta)}{\partial\theta} \right]' \tag{11}$$

which requires only first order partial differentiation. This is equal to

$$\sum_{i=1}^{n} \left[\frac{\partial l_i(\hat{\theta})}{\partial\theta} \right] \left[\frac{\partial l_i(\hat{\theta})}{\partial\theta} \right]' . \tag{12}$$

It can be shown that the LM statistic would reduce to

$$LM = \sum_i \left[\frac{\partial l_i(\hat{\theta})}{\partial \theta}\right]' \left\{ \sum_i \left[\frac{\partial l_i(\hat{\theta})}{\partial \theta}\right] \left[\frac{\partial l_i(\hat{\theta})}{\partial \theta}\right]' \right\}^{-1} \sum_i \left[\frac{\partial l_i(\hat{\theta})}{\partial \theta}\right]$$

$$= \left[\sum_i \hat{\underset{\sim}{g}}_i\right]' \left[\sum_i \hat{\underset{\sim}{g}}_i \hat{\underset{\sim}{g}}_i\right]^{-1} \left[\sum_i \hat{\underset{\sim}{g}}_i\right] \sim \chi^2(k-1)$$

$$= i' \, G'(\hat{G}'\hat{G})^{-1}\hat{G}'i \tag{13}$$

where i is the n dimensional vector with every element equal to unity, $(k-1)$ is the number of restrictions, and $\hat{G}$ is a matrix whose n rows consist of the vectors $(\partial l\ (\theta)\ /\ \partial\ (\theta))'$ estimated under H_0.[15] This variant of the LM procedure is referred to as the *Outer Product of the Gradient* (OPG) form of the LM test.

The n vectors $\underset{\sim}{g}_i$ are the derivatives of the *heteroscedastic censored Tobit* model and can be calculated as follows:

$$\hat{\underset{\sim}{g}}_i = \begin{bmatrix} \dfrac{\partial l_i}{\partial \underset{\sim}{\beta}} \\[2ex] \dfrac{\partial l_i}{\partial \sigma^2} \\[2ex] \dfrac{\partial l_i}{\partial \underset{\sim}{\gamma}} \end{bmatrix}_{\substack{\underset{\sim}{\beta} = \hat{\underset{\sim}{\beta}} \\ \sigma^2 = \hat{\sigma}^2 \\ \underset{\sim}{\gamma} = \underset{\sim}{0}}} \tag{14}$$

---

15 Godfrey (1988) demonstrates how the statistic would reduce to the above form.

where the individual derivatives are given below, estimated under the null.

$$\frac{dL^*(\theta)}{d\gamma} = \sum_{i=1}^{n} \frac{dl_i(\theta)}{d\gamma}$$

$$= \sum_i \left[\frac{\partial l_i}{\partial h_i}\right] \left[\frac{\partial h_i}{\partial \gamma}\right]$$

$$= \frac{h'}{2} \sum_i \left\{ -d_{1i} \frac{\varnothing_{1i}}{\Phi_{1i}} \left[\frac{\gamma_1 - \beta' x_i}{\sigma_i^2}\right] \underset{\sim}{z_i} + d_{2i} \left[\frac{u_i^2 - \sigma_i^2}{\sigma_i^3}\right] \underset{\sim}{z_i} \right\}$$

$$+ \left\{ d_{3i} \frac{\varnothing_{2i}}{1 - \Phi_{2i}} \left[\frac{\alpha_2 - \beta' x_i}{\sigma_i^2}\right] \underset{\sim}{z} \right\} \tag{15}$$

where $\quad u_i = y_i^* - \beta' x_i \;\; for \; \alpha_1 < y_i^* < \alpha_2$

and $\quad h_i = \sigma_i^2 = h(\alpha^2 + \gamma' \underset{\sim}{z_i})$

and $\varnothing(.)$ and $\Phi(.)$ denote, respectively, the density function and the distribution function of the standard normal distribution. Similarly,

$$\sum_{i=1}^{n} \frac{dl_i(\theta)}{d\beta} = \sum_i \left\{ d_{1i} \frac{\varnothing_{1i}}{\Phi_{1i}} \left[-\frac{x_i}{\sigma_i}\right] + d_{2i} x_i \left[\frac{y_i - \beta' x_i}{\sigma_i^2}\right] + d_{3i} \frac{\varnothing_{2i}}{1 - \Phi_{2i}} \left[\frac{x_i}{\sigma_i}\right] \right\} \tag{16}$$

and

$$\sum_i \frac{dl_i(\theta)}{d\sigma^2} = \sum_i \frac{\partial l(\hat{\theta})}{\partial h_i} \frac{\partial h_i}{\partial \sigma^2} = \sum_i \frac{\partial l(\hat{\theta})}{\partial \sigma_i} \frac{h'(\sigma^2) 2\sigma}{2\sigma_i}$$

$$= \sum_i \left\{ - d_{1i} \left[\frac{\varnothing_{1i}}{\Phi_{1i}} \right] \left[\frac{\alpha_1 - \beta' x_i}{\sigma_i^3} \right] \sigma h' \ (\sigma^2) \right\} -$$

$$\frac{\sigma}{\sigma_i^2} d_{2i} h' \ (\sigma^2) + d_{2i} \frac{\sigma}{\sigma_i^4} (y_i - \beta' x_i)^2 h' (\sigma^2)$$

$$\left\{ + d_{3i} \frac{\varnothing_{2i}}{1 - \Phi_{2i}} \left[\frac{\alpha_2 - \beta' x_i}{\sigma_i^3} \right] \sigma h' (\sigma^2) \right\}. \tag{17}$$

Following Jarque and Bera (1982) we assume that $h'(\sigma^2) = 1$ under the null, which would enable us to get the partial derivatives into a tractable and estimable form with no loss of generality. The statistic is equal to the uncentered R^2 x n, where R^2 is calculated by regressing i_i ($\equiv 1$) on $\underset{\sim}{g}_i$. However, this is not usually available in standard econometric software used for estimation, as it is conventional to estimate only the centered R^2 in most regression analysis. We therefore estimate this LM test using a program specially written for the present exercise.

The small sample properties of the LM test in general, and the OPG variant in particular, are very weak; and the test is very sensitive to problems such as non-normality and omitted variable bias, which are distinct from the heteroscedasticity problem the test is expected to identify. Even if we can identify that a particular model fails the diagnostic test, it is difficult to correct it without adequate information on the exact nature of the problem. Davidson and Mackinnon (1983) argue that the inefficiency in the estimation of the information matrix will tend to result in the OPG variant of the LM tests having relatively poor small sample properties. They therefore suggest an alternative form of the LM test that they refer to as *Double Length Regression* (DLR) test. However, the OPG version can generally be more easily implemented than the DLR version. Since we do not have the possibility of using DLR in the case of limited dependent variable models we conduct only the OPG version of the LM test in our analysis.[16]

---

[16] Godfrey (1988) provides a discussion of the relative merits and demerits of the OPG and DLR methods.

REFERENCES

Akerlof, G. A., and J. L. Yellen. 1986. "Introduction." In *Efficiency Wage Models of the Labor Market*. Cambridge, UK: Cambridge University Press.

Amemiya, T. 1973. "Regression Analysis When the Dependent Variable is Truncated Normal." *Econometrica*. 41(6): 997-1016.

Bardhan, P. K. 1979. "Wages and Unemployment in a Poor Agrarian Economy: A Theoretical and Empirical Analysis." *Journal of Political Economy*. 87(3): 479-500.

Behrman, J. R., A. B. Deolalikar, and B. L. Wolfe. 1988. "Nutrients: Impacts and Determinants." *The World Bank Economic Review*. 2(3): 299-320.

Berndt, E. K., B. H. Hall, R. E. Hall, and J. A. Hausman. 1974. "Estimation and Inference in Non-linear Structural Models." *Annals of Economic and Social Measurement*. 3(4): 653-665.

Bharadwaj, K. 1974. *Production Conditions in Indian Agriculture*. Cambridge, UK: Cambridge University Press.

Bliss, C. J., and N. H. Stern. 1978. "Productivity, Wages and Nutrition: The Theory" and "Some Observation." *Journal of Development Economics*. 5(4): 331-362; 363-398.

Breusch, T. S., and A. R. Pagan. 1979. "A Simple Test for Heteroscedasticity and Random Coefficient Variation." *Econometrica*. 47(5): 1287-1294.

Canagarajah, R. S. 1991. *Employment and Consumption Behaviour in a Village Economy: Issues in Imperfect Information and Uncertainty.* Ph.D. dissertation. Cambridge, UK: University of Cambridge.

Canagarajah, R. S., and P. Gonzâlez. 1991. Non-Price Rationing and Efficiency in Casual Labor Markets. Cambridge, UK: University of Cambridge. Photocopy.

Dasgupta, P., and D. Ray. 1986. "Inequality as a Determinant of Malnutrition and Unemployment: Theory." *Economic Journal*. 96 (December): 1011-1034.

Dasgupta, P., and D. Ray. 1987. "Inequality as a Determinant of Malnutrition and Unemployment: Policy." *Economic Journal*. 97 (March): 177-188.

Davidson, R., and J. G. MacKinnon. 1983. "Small Sample Properties of Alternative forms of the Lagrange Multiplier Test." *Economic Letters*. 12(3-4): 269-275.

Dreze, J., and A. Mukherjee. 1987. "Labor Contracts in Rural India: Theories and Evidence." Discussion Paper 7. London: School of Economics, Development Research Programme.

Efron, B., and D. V. Hinkley. 1978. "Assessing the Accuracy of the Maximum Likelihood Estimator: Observed versus Expected Fisher Information." *Biometrika*. 65(3): 457-482.

Godfrey, L. G. 1988. *Misspecification Tests in Econometrics*. Cambridge, UK: Cambridge University Press.

Goldberger, A. 1981. "Linear Regression After Selection." *Journal of Econometrics*. 15(3): 357-366.

Greene, W. 1981. "On the Asymptotic Bias of the Ordinary Least Squares Estimator of the Tobit Model." *Econometrica*. 49(2): 505-513.

Greene, W. H. 1989. *LIMDEP*. Version 5.1. New York: Econometric Software, Inc.

Heckman, J. 1974. "Shadow Prices, Market Wages and Labor Supply." *Econometrica*. 42(4): 679-694.

Immink, M.D.C., and F. E. Viteri. 1981. "Energy Intake and Productivity of Guatemalan Sugar Can Cutters: An Empirical Test of Efficiency Wage Hypothesis." Part 1 and Part 2. *Journal of Development Economics*. 9(2): 251-287.

Institute for Rural Health Studies. 1984. *Dokur 1982-1983 Household Survey*. Hyderabad, India: IRHS.

Jarque, C. M., and A. K. Bera. 1982. "Efficient Specification Tests for Limited Dependent Variable Models." *Economic Letters*. 9(2): 153-160.

Judge, G. G., W. E. Griffiths, R. Carter Hill, H. Lutkepohl, H. and T. Lee. 1985. *The Theory and Practice of Econometrics*, 2nd ed. New York: John Wiley & Sons.

Leibenstein, H. 1957. "The Theory of Underdevelopment in Backward Economies." *Journal of Political Economy*. 65(2): 91-103.

Maddala, G. S. 1977. *Econometrics*. New York: McGraw-Hill.

Maddala, G.S. 1983. *Limited-dependent and Qualitative Variables in Econometrics*, Cambridge, UK: Cambridge University Press.

Maddala, G. S. 1988. *Introduction to Econometrics*. New York: Macmillan Publishing Company.

Mirrlees, J. A. 1975. "A Pure Theory of Underdeveloped Economies." In *Agriculture in Development Theory*. G. Reynolds, ed. New Haven: Yale University Press.

Nelson, F. D. 1981. "A Test for Misspecification in the Censored Normal Model." *Econometrica*. 49(5): 1317-1329.

Newbery, D.M.G. 1989. "Agricultural Institutions for Insurance and Stabilisation." In *The Economic Theory of Agrarian Institutions*. P. K. Bardhan, ed. Oxford: Clarendon Press, 267-296.

Olsen, R. 1978. "A Note on the Uniqueness of the Maximum Likelihood Estimator for the Tobit Model." *Econometrica*. 46(5): 1211-1215.

Pratt, J. 1981. "Concavity of the Log Likelihood." *Journal of the American Statistical Association*. 76(373): 103-106.

Pudney, S. E. 1989. *Modelling Individual Choice: The Econometrics of Corners, Kinks and Holes*. Oxford: Basil Blackwell.

Rosett, R. N., and F. D. Nelson. 1975. "Estimation of the Two-Limit Probit Regression Model." *Econometrica*. 43(1): 141-146.

Rudra, A. 1982. *Indian Agricultural Economics: Myths and Realities*. New Delhi, India: Allied Publishers Pvt. Ltd.

Singh, L., L. Squire, and J. Strauss. 1986. "Introduction." In *Agricultural Household Models: Extensions, Applications, and Policy*. L. Singh, L. Squire, and J. Strauss, eds. Washington, DC: World Bank, 3-14.

Standing, G., and G. Sheehan. 1978. "Labor Force Participation in Sri Lanka." Case Study No. 10. In *Labor Force Participation in Low-Income Countries*. G. Standing and G. Sheehan, eds. Geneva: International Labour Organisation.

Strauss, J. 1986. "Does Better Nutrition Raise Farm Productivity?" *Journal of Political Economy*. 94(2): 297-320.

Tobin, J. 1958. "Estimation of Relationships for Limited Dependent Variables." *Econometrica*. 26(1): 24-36.