ECONOMIC GROWTH CENTER

YALE UNIVERSITY

Box 1987, Yale Station
New Haven, Connecticut

CENTER DISCUSSION PAPER NO. 411

THE HETEROGENEITY OF FAMILY AND HIRED LABOR IN AGRICULTURAL PRODUCTION:

A Test Using District-Level Data from India

Anil B. Deolalikar

and

Wim P.M. Vijverberg

June 1982

Notes: Comments received from participants at the Northeastern Universities Development Consortium Conference, Boston University, May 1982, are gratefully acknowledged. In addition, Robert Evenson and Brian Wright provided useful comments.

1. INTRODUCTION

There is no dearth of production function studies on the agricultural sectors of less-developed countries (Heady and Dillon, 1961; Rao, 1965; Yotopoulos, Lau, and Somel, 1970; Bardhan, 1973; Barnum and Squire, 1979). However, with the exception of Bardhan (1973), most of these studies have failed to distinguish between family labor and hired labor inputs, thus implicitly maintaining the questionable assumption of homogeneity of the two types of labor in agricultural production. Since family labor, unlike hired labor, is often entrusted with managerial tasks on the family farm, it is quite likely that the two kinds of labor are heterogeneous and may have different effects on agricultural output. The central point of this paper is that it is incorrect to simply assume away the heterogeneity of family and hired labor by treating them as identical and perfectly substitutable inputs in the production function, as previous studies have done.

District-level data from India are used in this paper to test the hypothesis of homogeneity of labor in agricultural production. The production function we employ is general enough to permit family and hired labor to have different effects on output as well as any constant elasticity of substitution between each other. Nested within the general model are several other more restrictive models, including the Cobb-Douglas production function having total labor as one input and the Cobb-Douglas production function having family and hired labor as two separate inputs. This makes it possible for us to test the general model against the conventional production functions that have been commonly estimated in the literature.

In addition, although this is not a central concern of this paper, we also test the heterogeneity of irrigated and unirrigated land using the same models. Most empirical studies normally treat the two types of land as separate inputs in production functions. We test whether such a specification is correct.

To anticipate our findings, we generally reject the conventional Cobb-Douglas production function which does not distinguish between family and hired labor. However, we fail to reject the Cobb-Douglas production function with family and hired labor as two separate inputs. Further, family labor is consistently observed to have a larger impact on output than hired labor. This suggests that family and hired labor are heterogeneous both in the sense of being imperfect substitutes for each other and in the sense of having different effects on agricultural output. This finding has important implications for the interaction of labor demand and labor supply in the agricultural sector of LDCs as well as for fertility among farm households, as is pointed out later in the paper.

The results relating to the heterogeneity of land are almost the opposite. Here, the hypothesis of perfect substitutability between irrigated and unirrigated land cannot be rejected, although the hypothesis of both inputs having identical effects on output can be. The Cobb-Douglas production function with irrigated and unirrigated land as two separate inputs is thus rejected in favor of the Cobb-

Douglas production function, in which a weighted sum of irrigated and unirrigated land is entered as a single input.

The plan of the paper is as follows. Section 2 discusses previous tests on the heterogeneity issue. Section 3 elaborates on some implications of heterogeneous labor. In section 4 we specify the functional forms of the production function, with which we test for heterogeneity. Section 5 reports the results, while section 6 concludes.

2. HOMOGENEITY VERSUS HETEROGENEITY OF LABOR IN PREVIOUS RESEARCH

Previous researchers have typically estimated Cobb-Douglas production functions having total (hired plus family) labor as a single input. The assumptions underlying such a production function are (i) family and hired labor are symmetric in terms of their effect on output, and (ii) family and hired labor are perfect substitutes (in the sense of having an infinite elasticity of substitution between them) in agricultural production. To see this, let the agricultural production function be:

$$(1) \qquad Y = C L^{\beta} \prod_{i=1}^{n} X_i^{\beta_i},$$

where i indexes non-labor inputs, Y = output, and L = labor services. C, β, and β_i are parameters to be estimated. Labor services are assumed to be "produced" according to a linear production function:

$$(2) \qquad L = FL + HL,$$

where FL and HL are quantities of family and hired labor used. As is obvious from equation (2), the coefficients on hired and family labor are identical (and equal to one), implying equal effects of the two inputs on output. Further, since the elasticity of substitution is always infinite in a linear production function, the relationship between family and hired labor in equation (2) is that of perfect substitutability.

To our knowledge, Bardhan (1973) is among the few researchers to have included family and hired labor as separate inputs in an agricultural production function. However, while Bardhan permitted family and hired labor to have different coefficients in the production function, he did not systematically test for the elasticity of substitution between the two types of labor. If family and hired labor were simply introduced as two separate inputs in a Cobb-Douglas production function, the elasticity of substitution between them would have been unity. However, Bardhan's specification was:

$$(3) \qquad Y = C\, L^{\alpha}\, L_h^{*\gamma} \prod_{i=1}^{n} X_i^{\beta_i},$$

where L = total labor used on the farm, and L_h^* = proportion of total labor that is hired, i.e., $L_h^* = HL/HL+FL = HL/L$.

Equation (3) can be rewritten as:

$$(4) \qquad \ln Y = \ln C + (\alpha-\gamma) \ln L + \gamma \ln HL + \sum_{i=1}^{n} \beta_i \ln X_i.$$

Thus, Bardhan's specification is equivalent to including both total labor and hired labor as separate inputs in the production function. The elasticity of substitution between family and hired labor is, therefore, neither one nor infinity. Bardhan does not calculate this elasticity of substitution, let alone test whether it is significantly different from unity or infinity. In fact, since a lower bound of zero is not set for the elasticity of substitution, we are not assured of the concavity of the production function estimated by Bardhan.[1]

For most of his samples, Bardhan obtained estimates of γ (in equation 3) that were not significantly different from zero. However, in two cases, he obtained significantly positive estimates of , implying heterogeneity of labor such that hired labor is more efficient than family labor (Bardhan, 1973, p. 1381). These results come as a surprise, since one expects family labor to be more efficient, if anything, than hired labor. However, Bardhan does not discuss this finding in greater detail.

Since Bardhan does not systematically test the hypothesis of perfect substitutability between family and hired labor, his test of homogeneity of labor is not complete. As such, his rejection of heterogeneity of labor cannot be accepted as conclusive.

In contrast to the treatment of family and hired labor, most previous production function studies have treated irrigated and unirrigated land as separate inputs. However, since this has almost always been in the context of Cobb-Douglas production functions, a unitary elasticity of substitution between irrigated and unirrigated land is generally imposed a priori. It is quite possible that irrigated and unirrigated land, although asymmetric in terms of their effect on output, may be perfectly substitutable for each other. To our knowledge, no study has attempted to test alternative functional forms for the relationship between the two types of land.

3. IMPLICATIONS OF LABOR HETEROGENEITY

The most immediate implication of heterogeneous family and hired labor is for the growing literature on empirical applications of the 'theory' of the farm household (Lau, Lin, and Yotopoulos, 1978; Barnum and Squire, 1979). These models have typically involved separate estimation of consumption and production models of a farm household and subsequent 'integration' of the estimated models to calculate the net (final) impact of prices, wage rates, and policy variables on a representative farm household. Separate estimation of consumption and production decisions has generally been justified on the grounds that there is a perfectly competitive market for labor in LDCs and that family and hired labor are homogeneous. Farm households are thus assumed to make their family labor supply decisions independently of the demand for on-farm labor, since the competitive market and homogeneous labor assumptions imply that excess family labor can always be sold in the casual labor market, or excess demand can be met by hiring in casual labor from the market, at a fixed wage rate.

If family and hired labor are heterogeneous, the labor demand and labor supply decisions of farm households cannot be so easily separated. To take an extreme example, if the elasticity of substitution between family and hired labor is zero, the supply of labor by family members cannot be determined independently of the on-farm demand for managerial and supervisory tasks, since the latter can never be performed by hired labor. Even for more plausible elasticities of substitution (i.e., greater than zero but less than infinity), the conventional models of the farm household, which assume separability of the household's production and consumption decisions, will have to be substantially revised.

A second implication of the imperfect substitutability between family and hired labor is at the labor market level and relates to rural-urban migration. If family and hired labor are perfect substitutes and if the former migrates to the city, the demand for hired labor will go up by the amount of family labor migrated. As a result, one would expect wages paid to hired labor to rise. Thus migration would benefit the population that stays behind in the agricultural sector. However, this conclusion is not so clear when family and hired labor are imperfect substitutes. Taking the extreme case in which substitutability is zero, the demand for hired labor will decrease when family labor migrates. Therefore, the landless agricultural population may actually be impoverished due to the migration of landed household members. Thus, the degree of substitutability between family and hired labor has implications for rural-urban migration and its effect on rural poverty and income distribution.

A third implication, especially over a longer run, of the imperfect substitution between family and hired labor is that variables such as farm size, irrigation, or technical change, which increase the demand for family labor on the farm, may be expected to affect fertility rates among farm households. If family and hired labor are identical and perfect substitutes for each other, fertility among farm families should not be related to these factors,[2] since the greater demand for family labor on large or irrigated farms can always be met by hiring in casual agricultural labor at a fixed wage rate.

Our central argument is that the conventional concept of 'labor demand' is invalid if family and hired labor are heterogeneous in the sense of having different efficiencies and being imperfect substitutes for each other. Instead, we need to talk about a demand function for family labor and a demand function for hired labor arising out of constrained profit maximization by farm households. In general, the wage rate paid to hired labor will not be the correct price of family labor. The latter will be the wage rate received by family workers while working away from the family farm. Hardly any study has bothered to check whether the wage rate paid by cultivators to hired workers is different from that received by them when working on other people's farms, although such information is generally available from most household surveys. Instead, most studies have simply assumed that the two wages are equal.

4. THE MODEL

We assume that the agricultural production function facing farms is of the Cobb-Douglas type:[3]

$$Y = C L^{\beta_1} \prod_{i=4}^{n} X_i^{\beta_i}, \tag{5}$$

where Y = output, L = labor services, and X_i = quantity of the ith non-labor input used. We assume that labor services L are produced using family labor FL and hired labor HL:

$$L = L(FL, HL). \tag{6}$$

As discussed in the previous sections, the most common functional specification of L is additive: L = FL+HL. In this section, we examine two specifications which nest the additive form as a special case and which have a variety of interesting implications.

The first specification is the generalized CES production function, which contains the parameters α_1 and ρ_1. By appropriately restricting these parameters, we get the following five models:

$$L = (\alpha_1 FL^{-\rho_1} + (1-\alpha_1) HL^{-\rho_1})^{-1/\rho_1}, \quad \rho_1 \geqq -1, \qquad \text{(Model A.1)} \tag{7}$$

$$L = (0.5 FL^{-\rho_1} + 0.5 HL^{-\rho_1})^{-1/\rho_1}, \quad \rho_1 \geqq -1, \qquad \text{(Model A.2)} \tag{8}$$

$$L = \alpha_1 FL + (1-\alpha_1) HL, \qquad \text{(Model A.3)} \tag{9}$$

$$(10) \qquad L = 0.5\,FL + 0.5\,HL = 0.5\,(FL + HL), \qquad \text{(Model A.4)}$$

$$(11) \qquad L = FL^{\alpha_1}\, HL^{1-\alpha_1}. \qquad \text{(Model A.5)}$$

Model (A.1) represents the most general form. In model (A.2), α_1 is restricted to be equal to 0.5, while in model (A.3) ρ_1 is restricted to be -1. Model (A.4) is the commonly-estimated additive form in which α_1=0.5 and ρ_1=-1 are imposed. Finally, in model (A.5), ρ_1 is constrained to be zero (Arrow, Chenery, Minhas, and Solow, 1961); this implies that the production function of Y is Cobb-Douglas in all inputs, including family and hired labor. Clearly, this is an interesting special case, since the Cobb-Douglas production function is easier to estimate than models (A.1)-(A.3).

We call α_1 the symmetry parameter: it determines whether the function L is symmetric in family and hired labor. ρ_1 is the curvature parameter, since the curvature of the isoquants of the labor services production function becomes sharper with increasing ρ_1. In fact, the elasticity of substitution of the function L (σ^L_{23}) is related to ρ_1 by

$$(12) \qquad \sigma^L_{23} = \frac{1}{1 + \rho_1}\,.$$

There is, however, another measure of the elasticity of substitution between family and hired labor. This is the Allen-Uzawa partial elasticity of substitution σ^Y_{23}, which measures the substitutability between FL and HL in the context of the production of Y (not L) (Allen, 1938, p. 504).[4] The relation between σ_{23}^{L} and σ_{23}^{Y} is not obvious. However, when σ_{23}^{L} approaches infinity (as in models (A.3) and (A.4)), σ_{23}^{Y} goes to infinity as well.

The second general specification of the labor services production function is the generalized linear production function (Diewert, 1971, p. 503), which has two parameters α_{21} and α_{22}. Appropriately restricting these parameters yields four models:

$$(13) \qquad L = \alpha_{21}\, FL + 2\,\alpha_{22}\, FL^{\frac{1}{2}}\, HL^{\frac{1}{2}} + (1-\alpha_{21})\, HL, \qquad \text{(Model B.1)}$$

$$(14) \qquad L = 0.5\, FL + 2\,\alpha_{22}\, FL^{\frac{1}{2}}\, HL^{\frac{1}{2}} + 0.5\, HL, \qquad \text{(Model B.2)}$$

$$(15) \qquad L = \alpha_{21}\, FL + (1-\alpha_{21})\, HL, \qquad \text{(Model B.3)}$$

$$(16) \qquad L = 0.5\, FL + 0.5\, HL = 0.5\,(FL + HL). \qquad \text{(Model B.4)}$$

Equation (B.1) represents the most general specification here. Models (B.2) and (B.4) restrict α_{21} to 0.5, while in models (B.3) and (B.4) α_{22} is restricted to be zero. Note that models (B.3) and (B.4) are identical to models (A.3) and (A.4), respectively.

In models (B.1)-(B.4), α_{21} is the symmetry parameter, and α_{22} the curvature parameter. When α_{22} is positive (negative), the isoquants of the labor services production function are convex (concave) to the origin. In this sense, the parameters α_{21} and α_{22} are analogous to the parameters α_1 and ρ_1 of models (A.1)-(A.5), respectively. The elasticity of substitution of the function L can be expressed as:

$$(17) \qquad \sigma_{23}^{L} = \frac{2\ FL^{\frac{1}{2}}\ HL^{\frac{1}{2}}}{L\ \alpha_{22}} (\alpha_{21} + \alpha_{22}\ FL^{\frac{1}{2}}\ HL^{\frac{1}{2}})\ (1 - \alpha_{21} + \alpha_{22}\ FL^{\frac{1}{2}}\ HL^{\frac{1}{2}}).$$

We thus have seven distinct models that can be tested against each other in order to obtain evidence on the labor homogeneity issue. Those models that are nested can be tested with the standard F-test. Models that are not nested can be compared using a recently developed test by Davidson and McKinnon (1981).

A final note concerns the occurrence of zero hired labor inputs observed in samples of farm level data. The additive labor models (A.3) and (A.4) (and equivalently (B.3) and (B.4)) are consistent with the observation that some farms merely use family labor and do not hire outside labor. On the other hand, the CES specification does not permit zero values of inputs for positive values of ρ_1, nor does the Cobb-Douglas specification in model (A.5). For values of ρ_1 in the open interval (-1, 0), the isoquants of the production function of labor services are tangent to the FL- and HL-axes. This implies that zero inputs are consistent but will be chosen only if the price of such inputs approaches infinity. For any finite wage rate of hired labor, each profit-maximizing farm will always hire some outside labor. The generalized linear production function employed in models (B.1) and (B.2) suffers from the same problem.

This problem is not serious for the estimations reported in this paper, since we use an aggregated community-level data set which does not contain any zero values for any of the inputs. Even at the farm level, zero inputs for family or hired labor are rarely observed, at least in Indian agriculture. For instance, Rosenzweig (1978, pp. 847-848) reports, on the basis of a 1970-71 all-India survey of over 5,000 rural households, that 88 per cent of small farm households in India hire in outside labor and 85 per cent of large farm households use family labor on the family farm.

Models (A.1)-(A.5) and (B.1)-(B.4) above treat irrigated and unirrigated land as two separate inputs in the production process. It is possible that the two types of land, although very different in their effects on agricultural output, may be perfectly substitutable for each other. In that case, it may be better to aggregate the two types of land after weighing them differently. To explore the appropriate relationship between irrigated and unirrigated land, we assume that the overall agricultural production function facing farms is:

$$(18) \qquad Y = C\, A^{\beta_7}\, L(FL,\, HL)^{\beta_1} \prod_{i=6}^{n} X_i^{\beta_i},$$

where A = services from land, X_i = quantity of the ith non-labor, non-land input, and L(FL, HL) is the best labor services production function chosen from among models (A.1)-(A.5) and (B.1)-(B.4). Land services A are produced using irrigated and unirrigated land according to one of the following types of technologies:

$$(19) \qquad A = (\alpha_3 X_4^{-\rho_2} + (1-\alpha_3) X_5^{-\rho_2})^{-1/\rho_2}, \quad \rho_2 \geqq -1, \quad \text{(Model C.1)}$$

(20) $$A = (0.5\, X_4^{-\rho_2} + 0.5\, X_5^{-\rho_2})^{-1/\rho_2}, \quad \rho_2 \geq -1, \qquad \text{(Model C.2)}$$

(21) $$A = \alpha_3\, X_4 + (1-\alpha_3)\, X_5, \qquad \text{(Models C.3, D.3)}$$

(22) $$A = 0.5\, X_4 + 0.5\, X_5, \qquad \text{(Models C.4, D.4)}$$

(23) $$A = X_4^{\alpha_3}\, X_5^{1-\alpha_3}, \qquad \text{(Model C.5)}$$

(24) $$A = \alpha_{41}\, X_4 + 2\, \alpha_{44}\, X_4^{1/2}\, X_5^{1/2} + (1-\alpha_{41})\, X_5, \qquad \text{(Model D.1)}$$

(25) $$A = 0.5\, X_4 + 2\, \alpha_{44}\, X_4^{1/2}\, X_5^{1/2} + 0.5\, X_5, \qquad \text{(Model D.2)}$$

where X_4 = irrigated acreage and X_5 = unirrigated acreage. Note that models (C.1)-(C.5) and (D.1)-(D.4) are analogous to models (A.1)-(A.5) and (B.1)-(B.4), respectively. In models (C.1)-(C.5), α_3 is the symmetry (between irrigated and unirrigated land) parameter and ρ_2 is the curvature parameter for the land services production function, while α_{41} and α_{44} are the symmetry and curvature parameters, respectively, in models (D.1) and (D.2). By choosing the 'best' model among models (C.1)-(C.5) and (D.1)-(D.2), and comparing it to the best model among models (A.1)-(A.5) and (B.1)-(B.2), we can determine the most appropriate functional form for a production function with disaggregated land and labor inputs.[5]

5. ESTIMATION AND RESULTS

We have estimated the various models presented in the previous section with district-level data on 268 districts from all over India.[6] The data are for the agricultural year 1970-71, except in the case of gross value of agricultural output, which is averaged over three years (1969-70, 1970-71, and 1971-72) to eliminate short-term fluctuations arising because of abnormal weather. In calculating district output, constant all-India prices have been used to value each crop. Since the district-level variables are totals over varying numbers of farms in each district, all variables have been divided by the number of holdings (or farms) in a district before estimation.[7] Each observation is thus assumed to represent an "average" farm in a district.

The assumptions maintained implicitly in estimating an aggregate agricultural production function,[8] and the problems inherent therein, have been described by Timmer (1970), who has estimated production functions for U.S. agriculture using state-level data. Although such estimates are beset with serious theoretical complications, they serve a useful policy purpose in that they describe the aggregate response of output to changes in input levels.

The definitions of the variables used are given in Table 1. To estimate equations (A.1)-(A.5) and (B.1)-(B.4), we have added an i.i.d. disturbance term multiplicatively to each equation. Equations (A.4) (=B.4) and (A.5) have been estimated by ordinary least squares, while equations (A.1)-(A.3) and (B.1)-(B.2) have been estimated by non-linear least squares. In estimating equations (A.1) and (A.2), a lower bound of $\rho_1=-1$ (corresponding to an infinite elasticity of substitution) has been set to assure concavity (positive elasticity of substitution) of the production function. For models (B.1) and (B.2), the analogous bound was $\alpha_{22}=0$.

Results of the least squares estimation of models (A.1)-(A.5) and (B.1)-(B.4) are reported in Table 2. Since the boundary limit on ρ_1 (of -1) was binding in the case of model (A.2), the estimated model (A.2) was identical to model (A.4). Similarly, the boundary limit on α_{22} (of 0) was binding for model (B.2), thus making it equivalent to model (B.4). Since models (A.4) and (B.4) are equivalent by construction, this means that the estimated models (A.2), (A.4), (B.2), and (B.4) are all equivalent.

Of interest in Table 2 are the parameters α_1 and α_{21}, which indicate the relative weight to be attached to family labor vis-a-vis hired labor. Both α_1 and α_{21} are consistently greater than one-half (except in models where they are constrained to be equal to one-half), implying that, in adding up family and hired labor, the former should be weighed anywhere from three to nine times as much as the latter.

In contrast to the consistency of the symmetry parameters, the curvature parameters differ dramatically across models. For instance, the elasticity of substitution implied by the estimated curvature parameter in model (A.1) is 0.6, while in model (B.1) the implied elasticity of substitution between family and hired labor is 2402. To decide which one, if any, of these estimates to accept, we need to test the two models against each other.

There are a total of five models (A.1, A.2 = A.4 = B.2 = B.4, A.3 = B.3, A.5) that can be tested against each other. Our strategy is to first test the A models against each other, next test the B models against each other, and finally test the best A model against the best B model. We test nested models with a standard F-test and non-nested models with the Davidson-McKinnon (1981) test. In Table 3 are shown the results of the model specification tests that were run. Within the A models, (A.2) and (A.4) are clearly rejected in favor of (A.1) at the 0.04 level of significance. Models (A.2) and (A.4) are also rejected against model (A.3) at the 0.04 level of significance. Model (A.3) in turn is rejected in favor of (A.1), but only at the 0.13 level of significance. However, model (A.5) cannot be rejected against (A.1). The Davidson-McKinnon specification test for models (A.3) and (A.5) indicates that the latter cannot be rejected in favor of the former, but that the former can be rejected in favor of the latter at the 0.13 level of significance. Similarly, the Davidson-McKinnon test for models (A.4) (equivalent to A.2) and (A.5) indicates that (A.4) and (A.2) can be rejected in favor of (A.5) at the 0.01 level of significance, but that the latter cannot be rejected in favor of the former. Thus, by any yardstick, model (A.5) (a Cobb-Douglas relationship between family and hired labor) emerges as the best model within the A models.

Similar tests were made for the best model within the B models. The results shown in Table 3 suggest that model (B.3) (an additive relationship between family and hired labor, but with different weights attached to each) is the best model within the B models. However, since model (B.3) is equivalent to model (A.3), and model (A.5) has already been chosen over model (A.3), model (A.5) emerges as the best model among all the A and B models considered. The vindication of model (A.5) is fortunate in a sense, since it is perhaps the easiest model to estimate. All that is required for estimating (A.5) is that data be available separately for family and hired labor and that these inputs be entered as two separate inputs in a conventional Cobb-Douglas production function. The model most frequently estimated in the literature, viz., the Cobb-Douglas production function with additive labor (A.4), is clearly rejected in favor of the Cobb-Douglas production function with family and hired labor a separate inputs.

Next, we took the Cobb-Douglas relationship between family and hired labor as given, and explored different functional forms for the land services production function. Due to high computational costs and the poor performance of the generalized linear functional form in estimating the relationship between family and hired labor, we tried only the generalized CES relationship between irrigated and unirrigated land (i.e., models C.1-C.5).

Results of the least squares estimation of models (C.1)-(C.5) are presented in Table 4. The boundary value (of -1) for ρ_2 was binding in the case of model (C.1), which made the latter equivalent to model (C.3). It should be noted that model (C.5) is equivalent to model

(A.5), since a Cobb-Douglas relationship between irrigated and unirrigated land was already assumed in the search for the best functional form for the labor services production function.

The estimate for the symmetry parameter α_3 in models (C.1) and (C.3) is 0.793, which suggests that in adding up irrigated and unirrigated land, the former should be weighed roughly four times as much as the latter. The elasticity of substitution implied by the parameter in model (C.2) (which is the only model where the curvature parameter is freely estimated) is 1.616. However, the results of the model specification tests (Table 5) clearly indicate that model (C.2) can be rejected in favor of (C.3). The Davidson-McKinnon test further indicates that model (C.3) cannot be rejected in favor of (C.2). Models (C.4) and (C.5) are also clearly rejected in favor of (C.3). Thus, model (C.3), in which irrigated and unirrigated land are simply added together, but after being weighed differently, emerges as the best model within the C models. In fact, since models (C.5) and (A.5) are equivalent and (C.5) can be rejected in favor of (C.3), it follows that model (C.3) is the best model among all the A, B, and C models considered in this paper. It has a residual sum of squares (RSS) that is over 16 per cent lower than the RSS for the model with the next lowest RSS (viz., model C.4).[9]

To conclude, the best functional form for an agricultural production function in which labor and land are disaggregated is the Cobb-Douglas form with family and hired labor as separate inputs and irrigated and unirrigated land added together as a single input, but only after being weighed differently. Unfortunately, because of its non-linearity, this may not be the easiest form to estimate. One alternative for researchers is to add up irrigated and unirrigated land by weighing the former four times as much as the latter prior to estimation. This practice is often followed in adding up bullock hours and tractor hours to arrive at a single measure of draught animal input (Barnum and Squire, 1979). Once an aggregate measure of land is constructed, a conventional Cobb-Douglas production function with family labor, hired labor, aggregate land, and other inputs can be estimated.

6. CONCLUDING OBSERVATIONS

In this paper, we have tested the hypotheses of homogeneity of family and hired labor and of irrigated and unirrigated land, using district-level data from India. The evidence suggests that both inputs are heterogeneous. However, the nature of heterogeneity is different in the two cases. While family and hired labor are heterogeneous both in the sense of having different effects on output and in the sense of being imperfect substitutes for each other, irrigated and unirrigated land are heterogeneous only in the former sense (asymmetry). The hypothesis of perfect substitutability between irrigated and unirrigated land cannot be rejected. Hence, while it is valid to add up irrigated and unirrigated land (after attaching different weights to each) and include the weighted sum as a single input in a Cobb-Douglas production function, it is not valid to treat family and hired labor in the same way. Our results suggest that it is better to enter family and hired labor as separate inputs in a Cobb-Douglas production function, since the hypothesis of unitary elasticity of substitution between the two types of labor cannot be rejected.

Clearly, the hypothesis of homogeneity of labor in agricultural production needs to be further tested with household-level data sets from India and other LDCs before it is completely rejected. There are several important implications of the heterogeneity of family and hired labor that make it worthwhile to explore this issue further. For instance, the entire literature on labor demand in LDC agriculture needs revision to accomodate labor heterogeneity. For instance, if family and hired labor are neither symmetric nor perfect substitutes for each other, it can no longer be assumed, as it has been by previous studies on farm households, that family labor supply decisions by cultivator households are made independently of on-farm labor use decisions. This considerably complicates the existing "theory" of the farm household, since the assumptions of homogeneity of labor and perfectly competitive labor markets have been critical in the empirical applications of this theory. Furthermore, the heterogeneity issue has implications for the effect of migration patterns and policies on the economic welfare of the rural population. Finally, heterogeneity of labor implies that factors inducing an increase in the demand for family labor (such as farm size, irrigation, and technical change) will, in the long run, increase the demand for fertility among farm households.

FOOTNOTES

1) Since Bardhan has not presented his complete estimation results or the sample means of the variables in his data set, it was not possible for us to calculate the elasticity of substitution between family and hired labor implied by his model. However, when we estimated equation (3) with our data set, we obtained a negative elasticity of substitution between the two types of labor. The isoquants between family and hired labor were thus observed to be concave, not convex (as is required by theory), to the origin. See the Appendix for more details.

2) Of course, the positive relationship between fertility and farm size or irrigation may well be due to a positive income effect of the latter variables on fertility. The positive income effect would imply that children are normal goods.

3) A Cobb-Douglas relationship between labor and other inputs and among non-labor inputs is assumed, since there is considerable empirical evidence for a unitary elasticity of substitution between land and labor and between land and capital in agriculture (Yotopoulos, Lau, and Somel, 1970). Besides, we are primarily interested in this paper in exploring the 'best' relationship between family and hired labor; as such, the relationships among other inputs are not of central concern to us.

4) If the production function of Y with inputs Z_i is written as $Y = f(Z_1, \ldots, Z_n)$, then σ^Y_{23} is defined as

$$\sigma^Y_{23} = (F_{23} \sum_i f_i Z_i)/(Z_2 Z_3 F),$$

where $f_i = \partial Y/\partial Z_i$, $f_{ij} = \partial^2 Y/\partial Z_i \partial Z_j$, F is the bordered Hessian, and F_{ij} is the cofactor of f_{ij}.

5) In what follows, the word 'best' is used to describe a model that cannot be rejected in favor of any other model on the basis of a standard F-test or the Davidson-McKinnon test.

6) The data have been compiled from a number of sources, including the various state reports of the Agricultural Census of India 1970-71, a joint Jawaharlal Nehru University-Planning Commission study entitled Foodgrains Growth: A Districtwise Study (for data on gross value of agricultural output), and Fertilizer Statistics 1972 (for data on fertilizer use).

7) The division of all variables by the number of farms in a district removes a likely source of heteroscedasticity in the residuals of the production function, thus assuring us of consistent

estimates of the latter.

8) One particularly strong assumption is that the number of days worked in agricultural production activities by a cultivator or agricultural laborer does not vary systematically across regions. This assumption is necessary because of the nature of the data available at the district level; in particular, only data on numbers of cultivators and agricultural laborers, and not on days or hours worked, are available.

9) As a contrast, consider the fact that model (C.4) has a RSS which is only 7 per cent lower than the RSS of the model with the highest RSS among all A, B, and C models (viz., model A.2 = A.4 = B.2 = B.4).

Table 1

Variable Dictionary: Indian Districts, 1970-71

Variable		Mean	Std. Dev.
Y	(Gross Rupee value, at constant all-India prices, of output of 22 major crops per operational holding in district)	2743.30	8214.10
FL	(Number of cultivators in district per operational holding)	1.28	0.55
HL	(Number of agricultural laborers in district per operational holding)	0.71	0.45
X_4	(Hectares of irrigated cropped land in district per operational holding)	0.68	2.36
X_5	(Hectares of unirrigated cropped land in district per operational holding)	2.12	2.43
X_6	(Kilograms of fertilizer used in district per operational holding)	39.20	52.73

Notes: Number of observations is 268.

Table 2

Production Function Estimates, Models A.1-A.5 and B.1-B.4 (with Labor as a Heterogeneous Input): Indian Districts, 1970-71.

(asymptotic t-statistics in parentheses)

Parameter	A.1	A.3 B.3	A.2, A.4, B.4	A.5	B.1	B.2
C	8.215	8.237	7.969	8.225	8.238	7.969
β_1	0.450 (3.177)	0.449 (4.311)	0.380 (3.846)		0.471 (4.226)	0.380 (3.846)
α_1	0.901 (4.733)	0.758 (6.052)	0.500[a]	0.821[b]		
ρ_1	0.701 (0.413)	-1.000[a]	-1.000[c]	0.[a]		
α_{21}					0.740 (6.047)	0.500[a]
$\alpha_{22}(\times 10^3)$					0.135 (0.454)	0.[d]
β_2				0.381 (4.307)		
β_3				0.083 (2.052)		
β_4	0.220 (6.484)	0.220 (8.933)	0.236 (10.043)	0.221 (9.120)	0.219 (8.852)	0.236 (10.043)
β_5	0.252 (4.607)	0.251 (6.361)	0.273 (7.185)	0.249 (6.386)	0.248 (6.227)	0.273 (7.185)
β_6	0.133 (3.331)	0.141 (4.962)	0.122 (4.453)	0.134 (4.649)	0.142 (4.969)	0.122 (4.453)
σ^L_{23}	0.588	∞	∞	1,000	2,402.3	∞
σ^Y_{23}	0.034	∞	∞	1,000	5466.3	∞
RSS	45.430	45.833	46.579	45.509	45.760	46.579

Notes: [a]Imposed value. [b]Value calculated using formula: $\alpha = \beta_2/\beta_2 + \beta_3$.

[c]Boundary value in the case of model A.2; imposed value in the case of model A.4.

[d]Boundary value.

Table 3

Model Specification Tests, Models A.1-A.5 and B.1-B.4: Indian Districts, 1970-71[a]

Alternative Hypothesis (H_1)	Null Hypothesis (H_0)		
	A.3, B.3	A.2, A.4, B.2, B.4	A.5
A.1	F = 2.315[b] (0.129)	F = 3.301 (0.038)	F = 0.454 (0.501)
A.3, B.3		F = 4.264 (0.040)	P = -0.606 (0.544)
A.2, A.4, B.2, B.4			J = -0.733 (0.464)
A.5	P = 1.507[c] (0.132)	J = 2.524[d] (0.011)	
B.1	F = 0.416 (0.520)	F = 2.336 (0.099)	

Notes: [a]Significance levels in parentheses.

[b]Standard F-Statistic.

[c]Davidson-McKinnon's t_α of the P-test.

[d]Davidson-McKinnon's t_α of the J-test.

Table 4

Production Function Estimates, Models C.1–C.5 (with Land as a Heterogeneous Input): Indian Districts, 1970–71

(asymptotic t-statistics in parentheses)

Parameter	C.1, C.3	C.2	C.4	C.5, A.5
C	8.287	8.320	8.028	8.225
β_2	0.190 (2.323)	0.172 (2.014)	0.135 (1.537)	0.331 (4.307)
β_3	0.044 (1.196)	0.029 (0.765)	-0.035 (-0.957)	0.083 (2.052)
β_4				0.221 (9.120)
β_5				0.249 (6.386)
β_6	0.138 (5.512)	0.175 (6.929)	0.242 (11.723)	
β	0.590 (13.468)	0.590 (12.678)	0.540 (11.778)	
α_3	0.793 (11.705)	0.500[b]	0.500[b]	0.470[c]
ρ_2	-1.000[a]	-0.381 (-4.876)	-1.000[b]	0. [b]
σ^A_{45}	∞	1.616	∞	1.000
σ^Y_{45}	∞	2.008	∞	1.000
RSS	36.212	40.843	43.194	45.509

Notes: [a]Boundary value in the case of model C.1; imposed value in the case of Model C.3.

[b]Imposed value.

[c]Calculated as $\alpha_3 = \beta_4/\beta_4 + \beta_5$.

Table 5

Tests Across Models C.1-C.5:
Indian Districts, 1970-71[a]

Alternative Hypothesis (H_1)	Null Hypothesis (H_o)			
	C.2	C.3, C.1	C.4	C.5
C.2		P = -0.932 (0.351)	F = 15.085[b] (0.000)	P = 7.670 (0.000)
C.3, C.1	P = 5.625[c] (0.000)		F = 50.522 (0.000)	P = 8.433 (0.000)
C.4				J = 9.315 (0.000)
C.5	P = 2.221 (0.026)	P = 0.977 (0.328)	J = 4.087[d] (0.000)	

Notes: [a]Significance levels in parentheses.

[b]Standard F-statistic.

[c]Davidson-McKinnon's t_α of the P-test

[d]Davidson-McKinnon's t_α of the J-test

APPENDIX

Bardhan (1973) considers the following functional form for testing the assumption of heterogeneity of labor:

$$\text{(i)} \qquad Y = cL^{\alpha} (HL/L)^{\gamma} \prod_{i=1}^{n} X_i^{\beta_i},$$

where HL is hired labor and L is total labor (which includes hired labor). As we show in the text, this can be rewritten as:

$$\text{(ii)} \qquad \ln Y = \ln c + (\alpha-\gamma) \ln L + \gamma \ln HL + \sum_{i=1}^{n} \beta_i \ln X_i .$$

The elasticity of substitution between HL and FL (family labor) implicit in equation (i) can be derived analytically. It turns out to be:*)

$$\text{(iii)} \qquad \sigma_{23}^{L} = 1 + \frac{HL}{FL} \left[\frac{\alpha-\gamma}{\gamma} + 1\right].$$

Thus is $\gamma < 0$ (i.e., hired labor is less efficient than family labor), it is possible for σ_{23}^{L} to be negative. This will occur if $\left|\frac{\alpha-\gamma}{\gamma}\right| > 1$ an and $\left|\frac{HL}{FL} \left[\frac{\alpha-\gamma}{\gamma} + 1\right]\right| > 1$. However, as long as $\gamma > 0$ (i.e., hired labor is more efficient than family labor, σ_{23}^{L} will always be positive, assuring concavity of the production function. Since the γ's that Bardhan reports for two of his samples are positive, the implicit elasticity of substitution between family and hired labor in these samples is positive. However, Bardhan does not report the sign of γ for his other samples; he only mentions the lack of significance of the other estimates of γ. At any rate, it is important to restrict the value of σ_{23}^{L} to positive values in the estimation procedure. Due to the complicated formula for σ_{23}^{L} in equation (iii), it is not straightforward to impose such a

---

*The definitions for σ_{23}^{L} and σ_{23}^{Y} can be found in section 4 of this paper.

constraint for all observations in the sample in estimating equation (i).

The expression for σ_{23}^{Y} is much more complicated. It seems impossible to formulate a restriction that guarantees a positive sign for σ_{23}^{Y} while permitting a negative estimate for γ.

We obtained the following results when we fitted equation (ii) to our data set:

$$\ln Y = \underset{}{7.908} + \underset{(4.028)}{0.418}\ \ln L - \underset{(-1.188)}{0.064}\ \ln (HL/L) + \underset{(9.370)}{0.228}\ \ln X_4$$

$$+ \underset{(6.606)}{0.261}\ \ln X_5 + \underset{(4.597)}{0.133}\ \ln X_6$$

$R^2 = 0.613$
(t-statistics in parentheses)
(all variables as described in Table 1)

Since the ratio of $\frac{HL}{FL}$ at the sample mean was equal to 0.555 in our data set, the elasticity of substitution between family and hired labor (σ_{23}^{L}) at the sample mean is -2.625. The estimated value of σ_{23}^{Y} equals -7.933. So both measures of the elasticity of substitution are negative. However, they are based on an insignificant estimate of γ.

REFERENCES

Arrow, K.J., H.B. Chenery, B.S. Minhas, and R.M. Solow (1961). "Capital Labor Substitution and Economic Efficiency." The Review of Economics and Statistics (August), pp. 225-250.

Bardhan, P.K. (1973). "Size, Productivity, and Returns to Scale: An Analysis of Farm-Level Data in Indian Agriculture." Journal of Political Economy, Vol. 81, No. 6 (November), pp. 1370-80.

Barnum, Howard, and Lyn Squire (1979). A Model of an Agricultural Household: Theory and Evidence. Baltimore: Johns Hopkins Press.

Davidson, R., and J.G. McKinnon (1981). "Several Tests for Model Specification in the Presence of Alternative Hypotheses." Econometrica, Vol. 49, No. 3 (May), pp. 781-93.

Diewert, W.E. (1971). "An Application of the Shaphard Duality Theorem: A Generalized Leontief Production Function." Journal of Political Economy, Vol. 79, No. 3 (May), pp. 481-507.

Lau, L.J., W.L. Lin, and P.A. Yotopoulos (1978). "The Linear Logarithmic Expenditure System: An Application to Consumption-Leisure Choice." Econometrica, Vol. 46, No. 4 (July), pp. 848-63.

Rao, C.H.H. (1965). Agricultural Production Functions, Costs and Returns in India. Bombay: Asia Publishing House.

Rosenzweig, M.R. (1978). "Rural Wages, Labor Supply, and Land Reform: A Theoretical and Empirical Analysis." American Economic Review, Vol. 68, No. 5 (December), pp. 847-61.

Timmer, C.P. (1970). "On Measuring Technical Efficiency." Food Research Institute Studies, Vol. IX, No. 2.

Yotopoulos, P.A., L.J. Lau, and K. Somel (1970). "Labor Intensity and Relative Efficiency in Indian Agriculture." Food Research Institute Studies, Vol. 9, No. 1, pp. 43-55.