ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 55-56 – 1999

Rethinking Within and Between Regressions: The Case of Agricultural Production Functions

Yair MUNDLAK, Donald F. LARSON, Rita BUTZER *

ABSTRACT. – The canonical form of regressions estimated with panel data consists of within-unit-time, between-unit and between-time regressions. These components represent different outcomes of the underlying economic processes. This is demonstrated in the paper by presenting results for the cross-country agricultural production function where the choice of inputs is convoluted with the choice of techniques. The results shed new light on resource productivity in agriculture and its relationship with the implemented technology.

Un réexamen des régressions «inter » et «intra » : le cas des fonctions de production agricole

RÉSUMÉ. – La forme canonique des régressions estimées sur des données de panel se décompose en des régressions sur observations individuelles temporelles recentrées (par rapport à des moyennes individuelles et temporelles), des régressions interindividuelles et des régressions intertemporelles sur des individus moyens. Ces composantes représentent différents résultats du processus économique sous-jacent. Ceci est montré dans cet article à travers la présentation des résultats obtenus sur des fonctions de production agricole de différents pays pour lesquelles le choix des intrants est croisé avec le choix des techniques. Les résultats donnent un nouvel éclairage sur la productivité des ressources en agriculture et sa relation avec les technologies installées.

* Y. MUNDLAK is a professor (emeritus), the Hebrew University of Jerusalem and the University of Chicago; D.F. LARSON is an economist in the World Bank's Research Group; and R. BUTZER is a consultant at the World Bank. The authors would like to thank A. CREGO for valuable research assistance. This research was conducted at the World Bank as part of a larger research effort on the determinants of agricultural growth (RPO 679-03).

1 Introduction

Much of the econometric work on panel data is concerned with inference on a process described by:

$$Y_{it} = X_{it} + u_{it}, \quad (1)$$

where u_{it} is decomposed to i and t components: $u_{it} = a_i + b_t + e_{it}$. The inference is carried out within the framework of the assumptions on the distribution of $\{u, X\}$. A critical element in this specification is the constancy of β over i and t. An extension of this specification is to allow β to vary over i or t. In this case the strategy can be to focus the interest on the average of β over i or over t. Each of the specifications is appropriate for some specific cases. Often, statistical tests can rank the likelihood of one specification against alternatives.

In the analysis of production, since, at any point in time, output (Y) is postulated to be explained exclusively by inputs (X), it is customary to specify the production function according to (1). However, such specifications do not take into account the fact that the choice of the coefficients in (1) is an economic decision. In their decision on a production plan, producers choose not only the level of inputs and outputs but also the technique of production to employ. The choice should be guided by the same considerations that we attribute to the choice of inputs. For instance, when a farmer decides on the composition of crops to grow he takes into consideration the physical environment such as the soil type, climate, and distance from the market, and forms expectations about uncertain matters such as prices and weather. The same decision process applies to the choice of methods of production (techniques). Differences in such attributes result in differences in the crop compositions and as such in the employed production functions that constitute the implemented technology. Thus, even though the available technology is the same for all farmers, the implemented technology is not. Consequently, we should expect the production functions to vary over farmers. The different techniques may differ in their factor intensity which, in terms of (1), will imply different β's. This calls for studying the following process:

$$Y_{it} = X(s_{it})\beta(s_{it}) + u_{it}, \quad (2)$$

where s is a vector of state variables that potentially vary over time and location.

The same logic is carried over to the aggregate production function. To detect its relevance it is desirable to have a sample with a large spread in the state variables. This is where the use of country-panel data becomes useful, in spite of the fact that its use is not free of problems. Cross-country productivity analysis has been practiced in agriculture for some time and has been extended recently to the economy at large as a result of the revival of interest in economic growth. The discussion in this paper has some implications for this more general application, but we will concentrate here on agriculture.

Knowledge of the production structure, as reflected in the production elasticities, is essential in the discussion of several key topics such as:

1. The contribution of inputs to output and, in a dynamic context, to growth.
2. The existence of returns to scale.
3. The sensitivity of the cost of production to changes in factor prices.
4. The relationships between factor prices and their productivity.

The discussion in this paper and the results bear directly on the first three subjects and indirectly on the last one.

The plan of the paper is as follows: we begin with a brief literature review of cross-country studies of the agricultural production function. This is followed by an outline of the model, a discussion of the statistical aspects, data description, presentation of the results, a discussion of the results, and conclusions.

2 Cross-Country Studies[1]

Analysis of agricultural production functions began in 1944 with the work of TINTNER [1944], TINTNER and BROWNLEE [1944], followed by HEADY [1944]. These studies were based on farm data. Subsequent work was extended to cover aggregate data, and in 1955 BHATTACHARJEE presented the first analysis based on cross-country data. The underlying notion for these early studies was that all observations were generated from the same production function. In an effort to get a definitive statement on the agricultural production function, HEADY and DILLON ([1961], Chapter 17) compared the result of BHATTACHARJEE's study with numerous other studies and discovered that the notion of a homogenous technology was elusive. Thus they concluded "*Still the variations shown among the elasticites... bears witness to the dangers associated with the use of any such global production function*". (p. 633).[2]

The use of cross-country data to estimate a global production function gained impetus with the work of HAYAMI [1969, 1970], and HAYAMI and RUTTAN [1970] – studies which sought to explore the causes of cross-country differences in agricultural productivity. Again, the assumption underlying these studies is that all countries use the same production function. This assumption does not stand, and we find considerable disparities between their results and those obtained in country studies.

Since BHATTACHARJEE's study, with the passage of time, it was possible to increase the sample size and thus estimate the functions for different periods and additional countries. Also researchers have introduced new subjects,

1. For a survey of agricultural production functions see MUNDLAK [1998].
2. CLARK [1973] assembles many results of factor shares in an informal framework but with extensive international coverage. It is very clear that the estimates depend on the economic environment which is a major theme of the discussion.

TABLE 1

Comparison of Results

	BHATTACHARJEE	HAYAMI & RUTTAN	EVENSON & KISLEV	YAMADA & RUTTAN	ANTLE	HAYAMI & RUTTAN	NGUYEN	EVENSON & KISLEV	MUNDLAK & HELLINGHAUSEN
Date of study	1955	1970	1975	1980	1983	1970	1979	1975	1982
Sample:									
Number of countries	22	37	36	41	43	36	40*	36	58
Time period	1949	1960	1955,60, 65,68	1970	1965	1955, 60	1955, 60, 65, 70,75	1955, 60, 65,68	1960, 65 70,75
Estimation method	OLS	OLS	OLS	OLS	PCR	OLS	OLS	OLS	PCR
Date specification	S; N	M; N	M; N	M; N	S; N	M; PW	M; N	M; N	M; N
Fixed effects included						year	year	country	country#
Elasticities									
Structures & equipment/Machinery/Tractors		0.12	0.10	0.11		0.11	0.14	0.06	0.07
Livestock & orchards/Livestock		0.23	0.30	0.23	0.14**	0.28	0.33	0.35	0.19
Land	0.42	0.08**	0.04**	0.02**	0.16	0.07	0.02**	0.14	0.16
Labor	0.28	0.41	0.23	0.33	0.38	0.40	0.39	0.03**	0.46
Fertilizer	0.29	0.12	0.10	0.24	0.007**	0.14	0.10	0.09	0.11
Irrigation									0.01
Schooling/General education		0.32**		0.08**	0.25**	0.24	0.10**		
Technical education		0.14	0.04	0.14		0.12	0.17	0.00**	
Research and extension			0.14		0.17			0.07	
Infrastructure					0.21				
sum of input elasticities	0.99	0.96	0.77	0.93	0.75	1.00**	0.98	0.67	1.00**

*sample is not balanced, $n = 183$ for Nguyen study
** not significant at $P = .05$ for one-tailed test
*** homogeneity constraint imposed
# Country effect on slopes and intercept
OLS and PCR are ordinary least squares and principal components regressions.
S and M represent single year observations and multi-year averages. PW represents per-worder averages of national aggregated data, N represents national aggregates.

including: checking the robustness of the estimates and the returns to scale, improving the specification by adding state variables and using different methods of estimation, and doing away with the assumption of constant technology.

Table 1 summarizes results obtained in a series of cross-country studies of agricultural production function where the quantities (outputs and inputs) are expressed either as country totals or in per worker terms.[3] Studies using data for a single period provide estimates for the between-country regression for that particular year. This is also the case for studies which use panel data and introduce time dummies (within-time regressions) but not country dummies.

Studies with panel data which contain country dummies (with or without time dummies) provide estimates for the within-country (or within-country and time) regression. We elaborate on this point below.

We will try to characterize some of the results. The study by BHATTACHARJEE does not include any measure of capital items nor shifters, technology or others, and therefore the results cannot be compared meaningfully with the other studies. The remaining studies have some shifters such as schooling, research and extension, infrastructure, and the like. In a way of generalization, the table shows that the between-country estimates of land elasticity are low in absolute terms and relative to estimates obtained from the within regression (compare the two versions of the EVENSON and KISLEV study). Two measures of capital have been used in most studies, machinery and livestock. The elasticity of machinery varies around 0.1 (a little higher for the between-country regression) and that of livestock concentrates in the range of 0.2-0.3. The estimates for the labor elasticity are less stable. Of course, this is a very general evaluation but sufficient to provide a background for the discussion.

3 The Model

The underlying premise is that producers at any time face more than one technique of production, and their economic problem is to choose the techniques to be employed together with the choice of inputs and outputs. The outline of the approach follows MUNDLAK [1988, 1993]. Let x be the vector of inputs and $F_j(x)$ be the production function associated with the jth technique, where F_j is concave and twice differentiable, and define the available technology, T, as the collection of all possible techniques, $T = \{F_j(x); j = 1,\dots,J\}$. Firms choose the implemented techniques subject to their constraints and the environment. We distinguish between constrained (k) and unconstrained (v) inputs, $x = (v,k)$, and assume for simplicity, without a loss of generality, that the constrained inputs have no alternative cost. The optimi-

3. We do not include studies where the quantities are expressed as averages per farm because this would raise issues unessential to our discussion.

zation problem calls for a choice of the level of inputs to be assigned to technique *j* so as to maximize profits. To simplify the presentation, we deal with a comparative statics framework and therefore omit a time index for the variables. The extension to the intertemporal version is conceptually straightforward. The Lagrangian equation for this problem is:

$$L = \sum_j p_j F_j(v_j, k_j) - \sum_j w v_j - \lambda\Big(\sum_j k_j - k_0\Big) \tag{3}$$

subject to $F_j(.) \in T$; $v_j \geqslant 0; k_j \geqslant 0$,
where p_j is the price of the product of technique *j*, *w* is the price vector of the unconstrained inputs, and *k* is the available stock of the constrained inputs. The solution is characterized by the Kuhn-Tucker necessary conditions. Let $s = (k, p, w, T)$ be the vector of state variables of this problem and write the solution as: $v_j^*(s)$, $k_j^*(s), \lambda^*(s)$, to emphasize the dependence of the solution on the state variables. The optimal allocation of inputs v_j^*, k_j^* determines the intensity of implementing the *j*th technique. To the extent that the implementation of a technique requires positive levels of some inputs, when the optimal levels of these inputs are zero, the technique is not implemented. The optimal output of technique *j* is $y_j^* = F_j(v_j^*, k_j^*)$, and the implemented technology (IT) is defined by

$$IT(s) = \{F_j(v_j, k_j);\ F_j(v_j^*, k_j^*) \neq 0, \quad F_j \in T\}.$$

The empirical analysis can provide estimates of the production function that corresponds to the implemented technology. The aggregate production function expresses the aggregate of outputs, produced by a set of micro production functions, as a function of aggregate inputs. This function is not uniquely defined because the set of micro functions actually implemented, and over which the aggregation is performed, depends on the state variables and as such is endogenous. Let the aggregate production function be written as:

$$\sum_j p_j y_j^*(s) \equiv F(x^*, s) \equiv \varphi(s). \tag{4}$$

This production function is defined conditional on *s*, but changes in *s* imply changes in x^* as well as in $F(x^*, s)$. It is therefore meaningless in this framework to think of changes in *x*, except by 'error', which are not instigated by changes in *s*. This means that it is impossible to reveal a stable production function from a sample of observations taken over points with changing available technology. Consequently, in general, the aggregate production function is not identifiable.

The empirical aggregate production function can be thought of as an approximation in a specific way. For (4) to be a production function in the usual sense, *x* should be disjoint from *s*. Such a separation requires a discrepancy between *x* and x^* that will allow us to write for the observed output:

$$\sum_j p_j y_j \cong F(x, s). \tag{5}$$

Strictly speaking, $F(x,s)$ is not necessarily a function since x can be allocated to the various techniques in an arbitrary way. It is only when we have an allocation rule leading to x^* that uniqueness can be achieved. With this caveat, we view $F(x,s)$ as a function of s, since s determines the techniques to which the inputs are allocated. A discrepancy, $x - x^*$, produces information on a given implemented technology, and such a discrepancy is also the source of information for identifying a given production function.

It has been shown (MUNDLAK, [1988]) that $F(x,s)$ can be approximated by a function which looks like a Cobb-Douglas function, but where the elasticities are functions of the state variables and possibly of the inputs:

(6) $$\ln y = \Gamma(s) + B(s,x)\ln x + u$$

where y is the value added per worker, $B(s,x)$ and $\Gamma(s)$ are the slope and intercept of the function respectively, and u is a stochastic term.[4] Under this framework, there is no single production function that can be estimated from a cross section of firms. This was experienced by MAIRESSE and GRILICHES [1990].

Variations in the state variables affect $\Gamma(s)$ and $B(s,x)$ directly as well as indirectly through their effect on inputs. The state variables may not be independent; a change in one state variable may be associated with a change in the others. This is illustrated by evaluating the elasticity of average labor productivity with respect to a given state variable (says s_i):

(7)
$$\partial \ln y/\partial s_i = \sum_h \{\partial\Gamma/\partial s_h + \ln x[\partial B(s)/\partial s_h] + B(s)(\partial \ln x/\partial s_h)\}\partial s_h/\partial s_i.$$

The first two terms in the brackets show the response of the implemented technology to a change in the state variables, whereas the last term in the brackets shows the output response to a change in inputs under constant technology. The elasticities in (7) have a time index, which is suppressed here, indicating that they vary over the sample points. The innovation in this formulation lies in the response of the implemented technology to the state variables. To isolate this effect, we rewrite (7), assuming independence between the state variables, namely $\partial s_h/\partial s_i = 0$ for $h \neq i$, and holding x constant to yield the elasticities

(8) $$E_i = \partial\Gamma(s)/\partial s_i + \ln x[\partial B(s)/\partial s_i].$$

The effect captured by (8) is part of the unexplained productivity residual in the standard productivity analysis under the assumption of constant technology.

When the available technology consists of more than one technique, a change in the state variables may cause a change in the composition of techniques in addition to a change of input used on a given technique. In this case, the empirical function is a mixture of functions and as such may violate the concavity property of a production function.

4. This expression can be given a more descriptive structure which leads to an approach in its estimation. However, we do not have the data on factor shares needed for this approach, and therefore we do not go into it. For the utilization of factor shares, see MUNDLAK, CAVALLO, and DOMENECH [1989], COEYMANS and MUNDLAK [1993], and LACHAAL and WOMACK [1998].

4 Statistical Aspects - A Review

The dependence of the implemented technology on the state variables causes it to vary across countries and over time in any given country. Countries differ in their natural endowments such as weather and soil quality – differences which remain roughly constant over time. They also vary in their constraints and the economic environment. We should therefore expect different coefficients from cross-country and time-series analyses. An examination of the empirical distribution of the various variables reveals that between-country variability is considerably larger than the variability over time. To relate the analysis to the standard properties of the linear statistical model, we summarize here some known results.[5] We label W, $B(i)$, $W(i)$, $B(t)$, $W(t)$, and $W(it)$ projection (symmetric and idempotent) matrices that generate residuals. They can be defined in terms of their operation on an arbitrary vector $x = \{x_{it}\}$ of order NT. Let $x_{.t}$ and $x_{i.}$ denote the averages of x_{it}, over t and i respectively and the terms in parentheses contain the typical elements of the vectors in question:

$$Wx = (x_{it} - x_{..}), \quad W(i)x = (x_{it} - x_{i.},), \quad W(t)x = (x_{it} - x_{.t},),$$

$$W(it)x = (x_{it} - x_{i.} - x_{.t} + x_{..}),$$

$$B(i)x = (x_{i.} - x_{..}) \text{ , and } B(t)x = (x_{.t} - x_{..}), \quad i = 1,...,N; \quad t = 1,...,T.$$

The following identities can then be derived:

(9) $$W = W(i) + B(i)$$

(10) $$W = W(t) + B(t)$$

(11) $$W = W(i) + W(t) - W(it)$$

(12) $$W = B(i) + B(t) + W(it)$$

Let y be the vector of observations of the dependent variable and X be the matrix of observations of the explanatory variables. The regression coefficients of interest can be written in a generic form for a projection matrix P as:

$$b = (X'PX)^{-1}X'Py$$

where P can be any one of the projection matrices of interest listed above with rank not smaller than rank X. To introduce the notation for the regression coefficients, we present in the brackets the projection matrix and the corresponding vector of coefficients: Pooled: [W, b]; within-time and country: [$W(it)$, $w(it)$]; within-country: [$W(i)$, $w(i)$]; within-time: [$W(t)$, $w(t)$]; between time: [$B(t)$,$b(t)$]; between-country: [$B(i)$,$b(i)$].

It is well-known (MADDALA, [1971]) that the pooled regression can be presented as a matrix weighted average of within and between regressions. For instance, use (9) and simplify:

(13) $$b = (X'WX)^{-1}X'Wy = Gw(i) + (I - G)b(i),$$

where $G = (X'WX)^{-1}X'W(i)X$.

5. For the underlying theory, see SCHEFFÉ [1959].

When, G, the share of the within component of variance X, is relatively small, the coefficients of the pooled regression will reflect largely the between-regression coefficients.

The common practice in empirical analysis is to use dummy variables for time or country and to obtain $b, w(i), w(t)$, and $w(it)$ from pooled data without dummies, with country dummies, with time dummies, and with country and time dummies respectively. Referring to the identities presented above, it appears that data generated by $W(it)$ (using country and time dummies) is cleaned from the between-time and between-country variations in the data and as such should best represent a measure of the more stable technology, referred to here as the core implemented technology. Data generated by $W(i)$ (using country dummies) are cleaned of the between-country variations but contain the variability over time. Similarly, data generated by $W(t)$ are cleaned of the time variability but contain the cross-country variability. This can be seen by rearranging the identities in (9)-(12):

$$W(i) = W(it) + B(t) \tag{14}$$

$$W(t) = W(it) + B(i) \tag{15}$$

To determine the empirical importance of this distinction, we estimate and present the three canonical regressions of this problem: $b(i)$, $b(t)$, and $w(it)$. There is however a practical problem with the estimation of $b(t)$, as we have only 21 years, and as such the sample size is small for estimating the full model. However, we can obtain an approximation of the time-effect on the regression coefficients by comparing $w(it)$ and $w(i)$.

5 Data[6]

5.1. Output and Input

We estimate a cross-country agricultural production function where agricultural output depends on inputs, agricultural technology, and the state of the economy. Agricultural *output* is measured as agricultural GDP in 1990 US dollars. Inputs to agricultural production include land, capital, labor, and fertilizers and pesticides.

Hectares of arable and permanent cropland, along with permanent pastures, are used for the measure of *land*. Agricultural *labor* is defined as the economically active population in agriculture. *Fertilizer* consumption is often viewed as a proxy for the whole range of chemical inputs and more.

Data on agricultural *capital* is taken from CREGO, LARSON, BUTZER, and MUNDLAK [1998]. The series contains two components: fixed capital, consis-

6. For a description of the data, including sources and construction, see MUNDLAK, LARSON and BUTZER [1997].

ting primarily of *structures and equipment* reported in 1990 US dollars and capital of agricultural origin, consisting of livestock and in trees. The two components differ in the method of construction, and also in terms of markets and pricing. The cost of production capital of agricultural origin is largely independent of the markets for nonagricultural inputs, which are often imperfect. As such, farmers may face noncompetitive prices for their fixed investment inputs. Fixed investments depend more on outside finance and are frequently hindered by credit constraints.

5.2. **Environment**

Agricultural production depends on the physical environment or natural conditions. We follow MUNDLAK and HELLINGHAUSEN [1982] to represent the environment by using two variables extracted from BURINGH, VAN HEEMST, and STARING [1979]. The variables are *potential dry matter* (*PDM*) and a *factor of water deficit (FWD)*.[7] The first variable is intended to measure the theoretical potential production of dry matter. The production of dry matter requires moisture. Arid areas may have a large value for *PDM*, but actual production is small due to water deficit. The relative water availability is measured by the ratio of actual transpiration to potential transpiration. The difference between potential and actual evapotranspiration is moisture deficit during the growing season. The variable, referred to as *FWD*, measures relative availability of moisture rather than deficit; the lower is the value of the *FWD* the larger is the deficit. Thus the variable is expected to be positively correlated with productivity.

5.3. **Technology**

As indicated above, the tacit assumption that all the observations are generated by the same technology is a very strong one and should be tested empirically. This can be done by introducing variables that will account for important differences in technology in the sample. To do this, we experimented with several variables.

The most common variable used in empirical studies as a carrier or representative of technology is some measure of human capital. The application of the concept of human capital in empirical analysis is quite problematic for a variety of reasons. In practice, it is mostly represented by a measure of *schooling*. The basic idea is that higher levels of education are conducive for technological progress. However, the causality could go in either direction in that economic progress generates a demand for schooling. Therefore, the interpretation of a schooling variable in empirical analysis is somewhat ambiguous. We include the mean school years of education of the total labor force, taken from NEHRU, SWANSON and DUBEY [1993] as a proxy for the human capital stock for the overall economy.

7. BINSWANGER, MUNDLAK, YANG and BOWERS [1987], used the same source to construct somewhat different variables.

In addition to schooling, we include several variables whose meanings are somewhat more specific and perhaps easier to interpret. In the case of agriculture, there is a natural variable to measure the level of technology for a given crop; this is the yield or output per unit of land. Extending this concept to aggregate output, we construct an aggregate *peak yield*. For each country and each commodity, the maximum of the past yields is computed. Country-specific Paasche indices (1990 = 1) are constructed of these peak commodity yields, weighted by land area. A Paasche index is used since changing the composition of output changes the relevance of existing technologies. This index is intended to capture country-specific variations in technology over time but, as measured, is inadequate to represent variability across countries in technology in a given year. However, we can capture the cross-country variability in growth rates. To do this, we calculate the average rate of growth in the index over the period by regressing the log of the peak yield on time. This variable is used in the between-country analysis.

Countries do not always perform on the technology frontier. Deviation from the peak yield may be due to economic considerations or to natural disturbances. Be the reason what it may, the deviation may affect the productivity of the various inputs. To allow for such an effect, we measure the difference between the actual yield and the peak yield for each crop, obtain a land-weighted sum of these differences, and divide by the peak yield. This ratio is calculated for each country in each year and referred to as the *yield gap*.

Agricultural productivity is likely to be affected by the overall technological level of the country. As economies develop generally, the physical, legal, and regulatory infrastructure and institutions which support agriculture develop as well. We measure this influence with two variables. The state of *development* is measured by the per capita output in the country relative to that in the United States. The main variability of this measure is across countries, but it also varies over time. To supplement the variability of the overall technology over time in each country, we use the maximum past average labor productivity in nonagriculture in the country and refer to this variable as *na-peak*. It turns out that this variable was only of marginal importance and will not be discussed here.

5.4. **Incentives**

We introduce two measures of incentives to allow for the direct effect of incentives on productivity over and above their indirect effect that comes through resource allocation and accumulation. The measures are the terms of trade between the agricultural and manufacturing sectors obtained as the ratio of the sectoral prices, or the *relative price*, (agricultural GDP deflator to manufacturing GDP deflator, lagged one period) and its fluctuations, calculated as a moving standard deviation from the three previous periods. For the between-country analysis, we modify the measures, and use the average rate of growth in the relative price over the period, calculated by regressing the log of the price ratio on time. Also the standard deviation of the relative price over the entire period is used in place of the moving standard deviation. The *variability in agricultural prices* reflects the market risk faced by agricultural producers. In addition to the sector-specific risk, there is an economy-wide market risk, that of price volatility for the economy as a whole as measured by the rate of *inflation*. This is calculated as the rate of change in the total GDP deflator.

Expected improvement of future profitability encourages investment and thereby augments the capital stock which appears as a variable in the analysis. The regression coefficients of the incentive variables represent only the direct effect of prices that is not captured in input changes. To obtain the full impact of the incentives on productivity, it is necessary to add their indirect effect through investment, but this is not done here.

5.5. **Samples Description**

The sample was determined by the data availability and the preference for a balanced data panel in order to simplify the analysis. It consists of annual data from 37 countries for a 21-year period (1970-90). The information conveyed by the sample is summarized in Table 2. The first column presents the average annual growth rate of the variables over the sample period. Output grew at a rate of 3.82 percent. Capital has the highest growth rates among the inputs,

TABLE 2
Growth Rates and the Decomposition of the Sum of Squares

	Average Annual Growth Rate	**Decomposition of the Sum of Squares (expressed as a percentage of total)**		
Variable	**(%)**	**SSB(t)**	**SSB(i)**	**SSW(it)**
Output:				
GDP	3.82	2.49	96.17	1.34
Inputs:				
Capital	4.25	2.67	95.77	1.56
Structures & equipement	5.42	3.00	95.24	1.76
Livestock & orchards	2.17	0.96	97.83	1.21
Land	0.12	0.00	99.95	0.05
Labor	– 0.04	0.01	99.35	0.64
Fertilizer	3.04	1.14	96.76	2.09
Technology:				
Schooling	1.80	4.14	93.48	2.38
Peak yield	1.90	58.10	24.64	17.27
Development	– 0.29	1.41	94.18	4.41
Yield gap		8.97	30.91	60.12
Prices:				
Relative prices	– 0.30	3.01	41.98	55.00
Price variability		2.48	15.78	81.75
Inflation		2.38	10.65	86.97
Per Labor Output and Inputs:				
GDP		2.67	95.00	2.33
Capital		2.14	95.46	2.40
Structures & equipment		2.17	95.86	1.98
Livestock & orchards		1.14	96.78	2.07
Land		0.01	99.41	0.58
Fertilizer		0.98	97.78	1.24

4.25 percent. This is a weighted average of its two components, where the growth rate of structures and equipment is 5.42 percent. On the other hand, the growth rate of schooling is 1.8 percent and that of the peak yield is 1.9 percent. The gap in the level of development compared to the US has widened over the period and the terms of trade of agriculture deteriorated at an average rate of 0.3 percent. There has been little change in land and less so in the labor force. It should be noted that these rates are for the sample as a whole, and there are differences among countries as we can learn from the decomposition of the sum of squares.

The remaining columns of the table present a decomposition of the total sum of squares to its components that corresponds to equation (12) above. To standardize the results, we divide the components by the total sum of squares so that the numbers give the percentage of each component in the total sum of squares. The terms are

$$SS_{\text{total}} = SS(x_{it} - x_{..}),$$

$$SSW(it) = SS(x_{it} - x_{i.} - x_{.t} + x_{..}),$$

$$SSB(i) = SS(x_{i.} - x_{..}),$$

$$SSB(t) = SS(x_{.t} - x_{..}),$$

where, for any variable z, we use the notation: $SS(z) = \sum_i \sum_t Z_{it}^2$.

The between-country differences account for most of the variability of the output and the inputs. Thus, just by allowing for a country effect, and without introducing any input to the regression, the R^2 is 0.962 so that the unexplained residual from country averages accounts for only 3.83 percent of the total sum of squares of output. Similarly, the between-country variability accounts for 95 to 98 percent of the variability of capital and practically all of the variability in land and labor. The situation is similar when the output and inputs are measured per worker.

The relative importance of the country and time components is different for the state variables; the between-country component is important in schooling and development and less important in the other variables. In part, this difference is due to the way the variables are measured. Schooling and development are measured in units that allow cross-country comparisons, and interestingly, the relative importance of the between-country component in the total sum of squares is similar to that of inputs. On the other hand, peak yield, prices, and price fluctuations were measured differently for the variability over time and for that over countries as explained above. In short, peak yield varies largely over time. This variable is measured as an index for each country, so that the between-country variability reflects differences in the rate of growth of the peak yield rather than differences in the level. The yield gap is a derived measure from the peak yield, but here most of the variability is in the transitory component of within-country-time. This suggests that the deviations from the peak yields are affected considerably by local conditions, some of which are weather triggered, while the others can be attributed to the economic environment. Be the case what it may, these variables show variability over time and also have a strong transitory component.

To sum up, the relative importance of the between-country component is dominant. This can lead to the erroneous conclusion that the within analysis has little to contribute. As a matter of principle, this conclusion is not well founded because the precision of the estimated coefficients depends not only on the spread in the regressors but also on the variance of the equation shock which usually contains a component that is time invariant. Consequently, the variance of the within component is considerably smaller than the total variance. Indeed, as we see below, the within estimates are meaningful empirically and informative substantively.

6 Empirical Results

The estimates for a base model are given in Table 3 which, in line with the specification in equation (12), consists of three blocks. The first block presents the within-country-time estimates, $w(it)$, which summarize the changes that took place over time and over countries after allowing for country and time effects. As such, these estimates are based on observations taken from the more stable, or core, technology. The second block presents the between-time estimates, $b(t)$, obtained from a short time-series of the sample means for each year. This represents the time-series component, common to all countries, and as such it captures the impact of changes in the available technology. The last block presents the between-country estimates,

TABLE 3
Base Model

	Within time and country		Between time		Between country	
Variable	**Estimate**	**t-score**	**Estimate**	**t-score**	**Estimate**	**t-score**
Inputs:						
Capital	0.37	6.90	1.03	6.01	0.34	13.13
Land	0.47	3.78			– 0.03	– 2.82
Labor	0.08		– 0.16	– 0.16	0.26	13.67
Fertilizer	0.08	1.53	0.14	0.33	0.43	21.91
Technology:						
Schooling	0.09	0.55	– 0.28	– 0.06	0.02	0.52
Peak yield	0.83	3.80	– 0.32	– 0.07	0.06	4.19
Development	0.52	3.36	– 0.21	– 0.33	0.31	2.97
Prices:						
Relative prices	0.04	1.78	0.02	0.09	0.01	1.95
Price variability	– 0.03	– 0.97	– 0.07	– 0.26	– 0.08	– 2.82
Inflation	– 0.00	– 0.75	0.04	0.71	0.07	4.25
Environmental:						
Potential dry matter					0.16	2.68
Water availability					0.44	7.96

Note : R-square for 777 obs. = .9696

$b(i)$, based on the between-country variations which constitute the major component of the total sum of squares. It summarizes the locus of points that go across the different techniques implemented by the countries, all operating under the same available technology.

Variables that remain constant over time, such as the environment, are not included in the first two blocks. By construction, variables in different blocks are orthogonal to each other, and therefore the estimates in one block are unaffected by the omission or addition of variables in the other blocks.

The first question is whether we can do away with any of the three aforementioned regressions. The answer to this question is given in Table 4 which presents results for several F-tests. The null-hypotheses that blocks can be omitted are rejected, and therefore we should interpret the information embedded in all of the blocks. The second question is whether the coefficients of the variables common to the various equations are the same. Casual inspection indicates that they are quite different, confirming the basic initial hypothesis that the regressions summarize the combined effect of changes in inputs and technology, and therefore the within and between regressions summarize different processes. We now turn to interpret the results.

TABLE 4
Statistical Tests

Tests of significance of blocks of variables						
	***R*-square**	***k***	***h***	***F*-statistic**	**5% critical**	**outcome**
Full regression	0.9696	31				
Null hypothesis is that the following block is not significant						
within time & country	0.9643	22	9	14.45	1.89	reject
between time	0.9448	22	9	67.62	1.89	reject
between country	0.0302	19	12	1921.03	1.77	reject
technology block	0.9676	22	9	5.45	1.89	reject
price block	0.9685	22	9	3.00	1.89	reject
environmental block	0.9670	29	2	31.90	3.01	reject

Number of observations, n, is 777.
k represents the number of parameters estimated.
h represents the number of constraints imposed by omission of a block of variables.

Tests of constant returns to scale for within-time and country analysis	
Input	**Estimate**
Capital	0.36
Land	0.42
Labor	0.08
Fertilizer	0.08
Sum	0.94
F-Statistic	0.04
5% critical value	3.84

The null hypothesis of constant returns to scale is not rejected.

6.1. **Inputs**

The results are striking in several respects. Perhaps the most interesting result is the magnitude of the elasticity of capital, 0.37 in the within regression, 0.34 in the between-country regression, and 1.03 in the between-time regression. Thus, it is high and significant in all three blocks, which is not the case with the other inputs. The between-time estimate is particularly high, and it indica-tes that the implementation of changes in the available technology, which can not be observed directly, was strongly affected by investment in agriculture.

The effect of capital becomes even stronger when capital is disaggregated to the two components as seen in Table 5. The within coefficient is 0.29 for structures and equipment and 0.13 for livestock and orchards, a sum of 0.42. The relative importance of the two components is almost the same in the between-time regression, and their sum is similar to that of the elasticity of total capital. The relative importance takes on a different form in the between-country regression, where livestock and orchards is the dominant component.

The land elasticity in the within regression is 0.47 for the aggregate capital (Table 3) and 0.44 for the disaggregated capital (Table 5). Thus we find no empirical support for the idea (see for instance KAWAGOE and HAYAMI [1985], p. 91) that land has lost importance in modern agriculture. Indeed, the land coefficient is negative, though small, in the between-country regression. This indicates that the techniques used by the more productive countries were land saving.[8] However, with a given technology, the marginal productivity of land is positive.

Finally, land is not included as a variable in the between-time regression in view of the low variability of land over time as indicated in Table 2.

The coefficient of fertilizers is particularly high in the between-country regression and much lower in the other two. A value of 0.08 obtained in the within regression for the elasticity of fertilizers may seem to be low, but this is not the case. A point estimate of 0.08 means that about 8 percent of the within changes in agricultural output are to be attributed to fertilizers. It is to be noted that this result is obtained for the aggregate agricultural output, whereas fertilizers are used only on plant products. It is likely that a production function for plant products alone would show a larger elasticity for fertilizers. Thus, a value of 0.08 for aggregate output may even be a bit high. One possibility is that fertilizers capture the impact of other chemicals and

8. To illustrate possible consequences of imperfect knowledge of the production process we bring an example from a study of cross-country differences in agricultural productivity. "*Depending on the weight to be applied to land, the total productivity of new continental countries becomes higher or lower than other DCs*." (KAWAGOE and HAYAMI [1985], p. 91). They use three alternative weight systems in their productivity exercise, where land weights (elasticities) are 0.1 (obtained from cross-country production function), 0.3 (less developed countries "*in which factor shares are high for land and low for modern inputs such as fertilizers and machinery*"), and 0.05 the "*situation in the highly advanced stage of economic development*" relying on SCHULTZ [1953] that at this stage the importance of land declined). (KAWAGOE and HAYAMI [1985], p. 88). There is no theoretical-basis for this classification, but we bring this reference to indicate that the high value is close to that obtained in the within analysis and the low value is not far from that obtained from the between-country analysis. There is similar ambiguity with respect to the other weights. Therefore, without gaining a clearer view with respect to the appropriate weights, the exercise does not convey a clear message.

TABLE 5
Alternative Model, Disaggregated Capital

Variable	Within time and country Estimate	t-score	Between time Estimate	t-score	Between country Estimate	t-score
Inputs:						
Structures & equipment	0.29	6.51	0.58	2.95	0.11	5.62
Livestock & orchards	0.13	2.09	0.44	1.81	0.29	– 9.04
Land	0.44	3.62			– 0.06	– 4.35
Labor	0.10		– 0.58	– 0.55	0.22	11.05
Fertilizer	0.04	0.79	0.09	0.22	0.42	20.41
Technology:						
Schooling	– 0.07	– 0.43	1.08	0,26	0.05	1.07
Peak yield	0.78	3.60	– 1.64	– 0.33	0.01	0.69
Development	0.46	3.01	– 0.02	– 0.02	0.28	2.73
Prices:						
Relative prices	0.04	1.78	0.01	0.03	0.05	5.91
Price variability	– 0.02	– 0.76	– 0.12	– 0.42	– 0.15	– 4.87
Inflation	– 0.00	– 0.84	0.01	0.22	0.11	6.15
Environmental:						
Potential dry matter					– 0.03	– 0.49
Water availability					0.55	9.41

Note : R-square for 777 obs. = . 9703

more generally, the modern inputs, as indicated above. As such, it may be that the variable also captures some inter-technology effects. Another possibility is that the elasticity reflects a high shadow price of fertilizers, which is a signal for constraints that prevent optimal use.

The striking result is the high value of the fertilizer elasticity obtained from the between-country regressions. This means that the locus of country means represents a changing technology package where the improvement in the implemented technology is fertilizer using. At the same time it is also capital using but land saving.

Referring to Table 3, the elasticity of labor, obtained in the within regression by imposing the assumption of constant returns to scale, is relatively low, 0.08, and it is 0.26 for the between country regression. The corresponding results for Table 5 are 0.10 and 0.22 respectively. Thus, the variations of output in each country are largely accounted for by variations in capital and land and less by labor.

The within regressions were obtained under the constraint of constant returns to scale. The constraint was tested empirically, and it is not rejected, as can be seen from the results in Table 4.[9] The between regressions are unconstrained, and it is interesting that the sum of the elasticities of the inputs of

9. We tested for evidence of constant returns to scale on the inputs only, not across the state variables.

the between-country regressions is practically one in Tables 3 and 5. This is in contrast to the results of those cross-county studies which show increasing returns to scale (for instance, KAWAGOE, HAYAMI and RUTTAN [1985], p. 120). This indicates that our specification succeeds in capturing the impact of cross-country differences in technology and thus eliminates the spurious result of increasing returns to scale.[10]

6.2. **Technology**

The technology variables play a dual role in the analysis. First they serve as technology shifters and as such reduce, or eliminate completely, the bias of the estimated input coefficients that is caused by the correlation of inputs and technology. Second, we can examine empirically how well they describe the data and thereby guide us in the search for appropriate technology indicators. The test of the null hypothesis that the technology block can be omitted is rejected, as reported in Table 4.

Turning to the individual components, the estimated elasticities in Tables 3 and 5 give the same message. The peak yield serves well as a shifter of the agricultural productivity, as measured by the core technology, with an elasticity of about 0.83. The level of development of the country relative to the US is also an important explanatory variable of agricultural productivity. Note that the contribution of this variable is over and above that of the peak yield which shows that the yield level is not the only indicator; first, the yield variable does not represent the productivity in livestock production which accounts for about one third of output, and second, there is a scope for improving efficiency under a given technology by coming closer to the frontier, as represented by the performance of the US.

The between-time regression shows that, for the sample as a whole, none of the technology variables were important in accounting for the changes in agricultural productivity over time. The work is done by physical capital. The implication is that even though schooling and peak yields increased with time (Table 2), we get no evidence that they contributed to the benefits harvested from improvements in the available technology. It is the changes in the available technology that caused the increase in these variables, at least in peak yield and perhaps in schooling. But it was capital availability that was crucial for the countries to take full advantage of the available technology. This sheds light on the importance of physical capital in accounting for the changes in agricultural productivity in the study period.

The results are different for individual countries, as seen from the between-country regression, where the level of development is important in accounting for the productivity variations. This is a statement of the importance of the various attributes of the overall level of development of a country in determining the level of agricultural productivity. This may also be the reason that

10. The finding of increasing returns to scale in cross-country analysis found justification of similar finding by GRILICHES [1963b] in cross-regions analysis for the US. For more evidence, see KISLEV and PETERSON [1996].

TABLE 6
Alternative Model, Includes Yield Gap

Variable	Within time and country Estimate	t-score	Between time Estimate	t-score	Between country Estimate	t-score
Inputs:						
Capital	0.37	7.13	1.03	6.02	0.39	14.04
Land	0.44	3.55			– 0.05	– 3.85
Labor	0.12		– 0.18	– 0.18	0.24	12.52
Fertilizer	0.07	1.44	0.17	0.38	0.44	22.66
Technology:						
Schooling	0.11	0.65	0.06	0.01	0.02	0.57
Peak yield	0.93	4.24	– 0.82	– 0.15	0.07	4.91
Development	0.54	3.55	– 0.29	– 0.39	0.20	1.95
Yield Gap	– 0.42	– 2.90	0.17	0.20	1.24	4.39
Prices:						
Relative prices	0.04	1.68	0.02	0.12	0.01	2.24
Price variability	– 0.02	– 0.74	– 0.05	– 0.20	– 0.05	– 1.92
Inflation	– 0.00	– 0.63	0.05	0.73	0.06	4.04
Environmental:						
Potential dry matter					0.18	3.09
Water availability					0.46	8.45

Note : R-square for 777 obs. = .9707

schooling appears to be irrelevant. To the extent that schooling matters, it may have an indirect effect through the development variable. However, to what extent schooling matters and how it can be measured using aggregate data is still an open question and was recently highlighted by PRITCHETT [1996].

The peak yield is a measure of the frontier of the implemented technology, but countries do not always operate at the frontier. We thus introduced the yield gap variable described above, whose average for the sample as a whole is 9 percent of the peak yield. As seen in Table 2, this variable varies considerably within country and time. We add this variable to the regression and report the results in Table 6. The variable has a negative effect on output in the within regression. This is of no surprise, operating off the frontier has a negative effect on productivity. The reason for introducing the variable is to allow for a shift of the function to account for the unfavorable environment and thereby relieve the inputs from accounting for this outcome. Interestingly, the effect on the input elasticities is minor, except for labor, as can be seen by comparing Tables 6 and 3. It thus leads to the conclusion that the within-country-time yield gap was caused by transitory effects that could not be anticipated or were anticipated to be transitory, and therefore there was no reason to make major changes in the inputs. The main change in the within analysis is in the peak yield coefficient, which increases from 0.83 (Table 3) to 0.93 (Table 6). The latter number is obtained by freeing the variable from the noise of the transitory conditions and thus is thought to be more reliable.

The results are different for the between-country regression where the coefficient of the yield gap is positive, significant, and sizable, 1.24, higher than the coefficient of peak yield in the within regression. To understand this result, we have to go back to the definition of the variables. In the between-country regression, we use the average growth rate of the peak yield in place of the peak yield index, and therefore we do not control for the level of technology as we have done in the within analysis. Thus, the gap captures some of this effect. An increase in the gap occurs either by an increase in the peak yield, which is of a permanent nature, or by a transitory decline in the performance in a given year. In the between-country regression the relative gap is measured as the country average gap over the 21 years of the study period, and as such it summarizes permanent effects, rather than the transitory ones which dominated the within variations. Therefore, a larger permanent increase in the gap seems to reflect larger peak yields, and this explains the positive regression coefficient of this variable in the between-country regression. What countries find it difficult to stay closer to their own frontier? The answer can be found by examining the correlation coefficients of the gap and the other variables.[11] The correlation coefficients are close to zero for the within deviations, which is consistent with the interpretation that these deviations in the yield gap were transitory. On the other hand, they are significantly negative for the country averages, particularly with capital and fertilizers, and to a lesser degree with schooling and development. It is the poorer countries that had on average a larger yield gap.

6.3. **Prices**

The test of the null hypothesis that the price block can be omitted from the analysis is rejected. It appears however that the allocation of the effect to the individual components is problematic. On the whole, the signs of the coefficients are in line with expectations, but the precision is low. The coefficient of relative prices is positive and that of its variability is negative. The magnitude of the price elasticity is small, 0.04 in the within regression. This indicates a small quantitative effect on agricultural productivity, but note that this effect is obtained conditional on given inputs and on technology. Thus, there is little scope for additional price effects. The fact that this effect is at all detected is of prime importance. The channels for the price effect are the level of inputs and the choice of technology, and these are represented by explanatory variables. Therefore, when evaluating policies that alter the relative returns to economic activity, it is important to realize that the legacy of past policies become enshrined in current stocks.

Two measures of market-risk, inflation and relative price volatility, dampen agricultural production (as seen by the within estimates); however both effects are quantitatively small. The coefficient of the measure of price volatility is negative, but it is significantly different from zero only in the between-country regressions. The effect of inflation is ambiguous in that it is negative and insignificant in the within regression, and it is positive and significant in the between-country analysis.

11. The results are reported in MUNDLAK, LARSON and BUTZER [1997].

6.4. **Environment**

The two environment variables introduced to the between-country regression have a positive and significant effect in the analysis with aggregate capital, but the effect of potential dry matter becomes unimportant when capital is disaggregated. This may suggest that this measure is not sufficiently robust.

6.5. **Comparing to Conventional Analysis**

Often, regressions are presented with either time or country dummies, which amounts to running a within-time or within-country regression. The relationships between such regressions and ours is established by combining equations (14) and (15) with one similar to (13). This will show that $w(i)$ is a matrix-weighted combination of $w(it)$ and $b(t)$, and similarly, $w(t)$ is a matrix-weighted combination of $w(it)$ and $b(i)$. The implication is that allowing for only one effect results in regression coefficients which reflect some between effects. The importance of the between component is determined by its relative weight in the total sum of squares (see Table 2). Thus, the regression coefficients obtained with time dummies reflect largely the between-country regression because the relative weight of the between-country sum of squares is dominant. Similarly, the regression coefficients obtained with country dummies reflect the between-time regression, but not to the extent of the previous case because the variations in the variables over time are much smaller than those across countries as can be seen from Table 2. In any case, the mere fact that a regression is done with some dummies does not guarantee that it identifies a stable function.

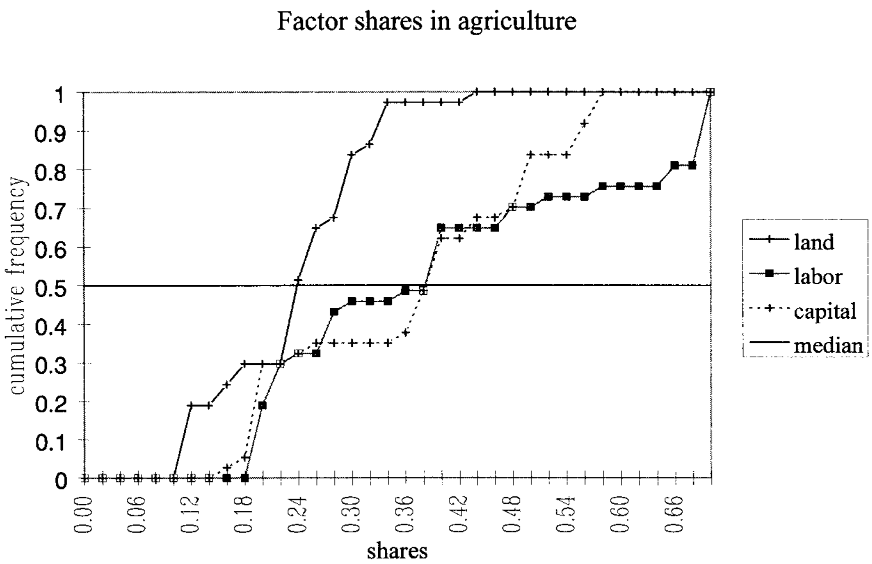
Table 7 presents within regressions for the specifications in Tables 3 and 5 respectively. Comparing the within-country regression in the second column in Table 7 to that of the within-country-time regression in Table 3, we see the impact of the between-time regression. In a way of generalization, the between-time sum of squares of output and input is of the order of magnitude of twice as much as that of the within-time-country. The coefficient of capital increases from 0.37 to 0.53, and many of the coefficients of the within-country regression are insignificantly different from zero reflecting the influence of the between-time coefficients. This comparison provides a framework for interpreting empirical results obtained under different specifications of effects.

With this interpretation we can now compare our results to those presented in Table 1. In this we ignore the first study by BHATTACHARJEE because it has no measure of capital nor of state variables, and therefore it is not comparable to the other studies. Most of the studies are strictly cross-country and as such are comparable to the between-country results. The similarity is in the low land elasticity, and also the sum of the elasticities of machinery and livestock is close in most cases to the value of 0.4 we obtained for the sum of structures and equipment and livestock and orchards in the between-country regression. This similarity is consistent with our interpretation that these studies describe only the between-country changes; hence they provide a limited and incomplete picture of the production process. In any case, they do not provide coefficients of a stable production function and as such, do not

TABLE 7
Within-Time and Within-Country Analyses

	Aggregate Capital				Disaggregated Capital			
	Within time		Within country		Within time		Within country	
Variable	**Estimate**	**t-score**	**Estimate**	**t-score**	**Estimate**	**t-score**	**Estimate**	**t-score**
Inputs:								
Capital	0.44	19.04	0.53	11.51				
Structures & equipment					0.18	10.82	0.38	9.21
Livestock & orchards					0.23	10.48	0.21	3.50
Land	– 0.04	– 4.08	0.51	1.81	– 0.06	– 4.81	0.57	2.06
Labor	0.31	19.37	0.13	1.16	0.35	21.11	0.08	0.76
Fertilizer	0.29	16.70	0.03	0.68	0.28	14.65	0.00	0.07
Technology :								
Schooling	0.15	3.83	0.07	0.42	0.15	3.68	– 0.10	– 0.58
Peak yield	0.37	2.42	0.82	4.24	0.32	1.94	0.84	4.41
Development	0.34	4.45	0.21	1.60	0.66	8.80	0.22	1.69
Prices:								
Relative prices	0.06	2.71	0.03	1.05	0.04	1.80	0.02	1.01
Price variability	– 0.05	– 1.71	– 0.01	– 0.50	– 0.05	– 1.47	– 0.01	– 0.36
Inflation	– 0.01	– 1.30	– 0.00	– 1.07	– 0.00	– 0.96	– 0.00	– 0.98

FIGURE 1



provide the appropriate weights for growth accounting, as they were intended to do.

It is always useful to check the results against all available information. The Global Trade Analysis Project (GTAP) reported factor shares of land and labor in agriculture for 1992 for 24 regions (HERTEL [1997]). The data needed to compute factor shares are not available for all countries. The more available data are on labor costs, and these were used as a pivot to generate the other shares relying on "*other sources*" where available (p. 113). Applying the appropriate regional data to the 37 countries in our study, we summarize this information in Figure 1 in terms of the empirical distributions. The median values are 0.24 for land, 0.39 for labor, and 0.39 for capital.[12] Another source of information is the OECD which reports "*compensation of employees*" by sectors. Computing labor shares from these series for 19 countries for the period 1970-90 (for 7 countries the period is somewhat shorter) yields a median value of 0.19. The labor share in these statistics is higher than the estimated elasticity from the within regression, but nevertheless, these values are conveniently close to the within estimates and are conspicuously far away from the between-country estimates. This seems to provide independent support for our interpretation.

7 Summary and Conclusions

The analysis is based on the premise that the production technology is heterogeneous and the implemented technology is endogenous and determined jointly with the level of the unconstrained inputs. Therefore, a change in the state variables affects both the technology and the inputs, and consequently the production function is not identified. To overcome this problem, we decompose the changes in productivity to three orthogonal components caused by the fundamentally different processes underlying panel data and thereby gain meaningful economic insight. The between-time process captures changes that are induced by changes in the available technology (technical change). The between-country process captures the changes that take place when the available technology is held constant but other state variables differ across countries and account for their differences in the implemented technology. Finally, the within-country-time process represents the changes in outputs, inputs, and state variables when the available technology is held constant as well as the fundamental changes across countries and as such comes closest to a production function. This framework also allows us to reinterpret results from earlier studies of cross-country productivity in a new way.

The most striking result is the relative importance of capital. This result is quite robust to various modifications of the model and to the disaggregation

12. The median values obtained using the unweighted regional data were similar (0.25 for land, 0.31 for labor, and 0.42 for capital).

of capital to its two components. Capital seems to account for about 40 percent of total output in the core technology. This indicates that agricultural technology is capital cost-intensive compared to nonagriculture.[13] This result is further reinforced by the magnitude of the land elasticity in the core technology and is at variance with the view that land is not an important factor of production in modern agriculture. This view is based on an incorrect reading of the data where no distinction is made between changes in technology and a movement along a given production function. The sum of capital and land elasticities is around 0.8 in various formulations, making it clear that agriculture should be more sensitive than nonagriculture to changes in the cost of capital, and less to that of labor. The value we obtained for the sum is a bit high compared to the literature. Consistent with our view of heterogeneous technology, it is possible that a different choice of countries and time periods would lead to somewhat different results. However, a sum of 0.8 for land and capital elasticities leaves a lot of room for the conclusion on the importance of capital to remain intact.[14]

The capital elasticity in the between-time regression is much higher than in the within regression, and this is consistent with the view that physical capital serves as a constraint to agricultural growth. At the same time, the between-time regression shows that the shift to more productive techniques is associated with a decline in labor, which is an indication of the labor-saving technical change in agriculture. This is consistent with the observed slight decline of labor over time (Table 2) while output grew at a brisk pace. This is not news, but it is emphasized here because it comes out of an integral view of the process which separates between the core technology and the changes that took place over time and between countries. These results highlight the importance of capital in agricultural production, an attribute critical in the understanding of agricultural development and its dependence on the economic environment.

The introduction of the appropriate state variables to account for technology, prices, and physical environment produced a production function that displays constant returns to scale and thus avoided the pitfalls of previous studies and the misguided conclusions that followed.

The statistical framework provides an explanation for the unstable results observed in production functions derived from panel data. Technically, the results depend on how the data are projected. Underlying this is the fact that comparisons between units, over time, or of deviations from unit-means or time-means all describe different processes. This is based on theory, but has an intuitive appeal as well. In the case at hand, the spread in productivity among countries is a different economic phenomenon or process from the spread in productivity for a country through time. In turn, the factors explaining the spread will differ. Panel data measures a combination of these economic phenomena–a fact which should be recognized explicitly in the modeling approach.

---

13. We say that a technology is capital cost-intensive with respect to a reference technology if the factor share of capital is larger than that of the reference technology.

14. We have tried some modifications in the specification (introduction of cross products) and in the choice of variables (principal components). The basic result turned out to be quite robust.

While we are specifically dealing with these issues in the context of agricultural production, the arguments probably extend to other applications as well. The decomposition of the sum of squares to their canonical components most likely will tell the importance of the various processes. It is however important to emphasize that even if one of the components carries only a small weight in the sum of squares, it can still convey a great deal of meaningful information, as we have demonstrated in the case of the within-country-time regression.

To sum up, we can now return to the key topics listed at the beginning of the paper. The contribution of inputs to growth should be judged by their contribution to output under a constant technology, attributing the rest of the growth to technical change. After all these years of intensive work on production functions, there is no hard evidence that estimated elasticities do a better job than the factor shares as were originally used by SOLOW [1957]. In terms of our work here, the within estimates come closest to the little evidence we have on factor shares. Also it is clear that the weights obtained from between-country regressions by HAYAMI and RUTTAN [1970, 1985], and earlier by GRILICHES [1963a, 1963b], give a distorted picture. We have pointed above to the sensitivity of the growth calculations to the weight of land, and this is where the between results erred the most. Not independently, our results indicate no evidence of increasing returns to scale. The increasing returns to scale inflated the contribution of the inputs and reduced the role of the residual technical change in the growth calculations *à la* GRILICHES and HAYAMI and RUTTAN.

Our results support the view that agriculture is capital cost-intensive as compared to nonagriculture. We provide estimates of factor productivity that can be compared to the factor price. Comparing our results to the factor shares, it seems that on the whole there is surprising agreement.

• References

ANTLE J.M., (1983). – “Infrastructure and Aggregate Agricultural Productivity: International Evidence”. *Economic Development and Cultural Change* 31 (3), pp. 609-19.

BHATTACHARJEE J.P., (1955). – “Resource Use and Productivity in World Agriculture”. *Journal of Farm Economics* 37 (1), pp. 57-71.

BINSWANGER H., MUNDLAK Y., YANG M.-C., BOWERS A., (1987). – “On the Determinants of Cross-CountryAggregate Agricultural Supply”. *Journal of Econometrics* 36, pp. 111-31.

BURINGH P., VAN HEEMST H.D.J., STARING G.J., (1979). – “Computation of the Absolute Maximum Food Production of the World”. Wageningen : Netherland Department of Tropical Soil Science, 1975 ; also published in *Moira Model of International Relations in Agriculture*, ed H. Linneman et al. Amsterdam : North-Holland Publishing Co.

CLARK C., (1973). – *The Condition of Economic Progress*. London : Macmillan & Co.

COEYMANS J.E., MUNDLAK Y., (1993). – “Sectoral Growth in Chile : 1962-82”. *Research Report 95*. Washington, D.C.: International Food Policy Research Institute.

GREGO A., LARSON D., BUTZER R., MUNDLAK Y., (1998). – “A New Database on Investment and Capital for Agriculture and Manufacturing”. *World Bank Working Paper Series 2013*. Washington, D.C.: The World Bank.

EVENSON R.E., KISLEV Y,. (1975). – *Agricultural Research and Productivity*. New Haven : Yale University Press.

GRILICHES Z., (1963a). – "The Sources of Measured Productivity Growth: United States Agriculture, 1940-1960". *Journal of Political Economy* 71 (4), pp. 331-46.

GRILICHES Z. (1963b). – "Estimates of the Aggregate Agricultural Production Function from Cross-Sectional Data". *Journal of Farm Economics*, 45(2), pp.419-28.

HAYAMI Y., (1969). – "Sources of Agricultural Productivity Gap Among Selected Countries". *American Journal of Agricultural Economics* 51, pp. 564-75.

HAYAMI Y., (1970). – "On the Use of the Cobb-Douglas Production Function on the Cross-Country Analysis of Agricultural Production". *American Journal of Agricultural Economics* 52, pp. 327-329.

HAYAMI Y., RUTTAN V.W., (1970). – "Agricultural Productivity Differences Among Countries". *American Economic Review* 60 (5), pp. 895-911.

HAYAMI Y., RUTTAN V.W., (1985). – *Agricultural Development, An International Perspective*. Baltimore : The Johns Hopkins University Press.

HEADY E.O., (1944). – "Production Functions from a Random Sample of Farms". *Journal of Farm Economics* 28 (4), pp. 989-1004.

HEADY E.O., DILLON J.L., (1961). – *Agricultural Production Functions*. Ames: Iowa State University Press.

HERTEL T., (1997). *Global Trade Analysis*. Cambridge: Cambdrige University Press.

KAWAGOE T., HAYAMI Y. (1985). "An Intercountry Comparison of Agricultural Production Efficiency". *American Journal of Agricultural Economics* 67, pp. 87-92.

KAWAGOE T., HAYAMI Y., RUTTAN V.W., (1985). – "The Intercountry Agricultural Production Function and Productivity Differences among Countries". *Journal of Development Economics* 19, pp. 113-32.

KISLEV Y. PETERSON W., (1996). – "Economies of Scale in Agriculture: A Reexamination of Evidence". In *The Economics of Agriculture: Papers in Honor of D. Gale Johnson*, Volume 2. J. M. Antle and D.A. Summer, eds. pp. 156-170. The University of Chicago Press.

LACHAAL L., WOMACK A.W., (1998). – "Impacts of Trade and Macroeconomic Linkages on Canadian Agriculture." *American Journal of Agricultural Economics* 80 (3), pp. 534-42.

MADDALA G.S., (1971). – "The Use of Variance Components Models in Pooling Cross Section and Times-Series Data". *Econometrica* 39 (2), pp. 341-58.

MAIRESSE J., GRILICHES Z., (1990). – "Heterogeneity in Panel Data: Are There Stable Production Functions?", In *Essays in Honor of Edmond Malinvaud, Volume 3: Empirical Economics*, P. Champsaur et al., eds. pp. 192-231. Cambridge, Mass. : MIT Press.

MUNDLAK Y., (1988). – "Endogenous Technology and the Measurement of Productivity". In *Agricultural Productivity: Measurement and Explanation*, S.M. Capablo and J.M. Antle, eds. pp. 316-31. Washington D.C. : Resources for the Future.

MUNDLAK Y., (1993). – "On the Empirical Aspects of Economic Growth Theory". *American Economic Review* 83 (2), pp. 415-20.

MUNDLAK Y. (1998). – "Production and Supply" *Working Paper* #9807, Rehovot, Israel: The Center for Agricultural Econometric Research.

MUNDLAK Y., HELLINGHAUSEN R. (1982). – "The Intercountry Agricultural Production Function: Another View". *American Journal of Agricultural Economics* 64(4), pp. 664-72.

MUNDLAK Y., CAVALLO D., DOMENECH R. (1989). – "Agriculture and Economic Growth in Argentina, 1913-84". *Research Report* 76. Washington, D.C.: International Food Policy Research Institute.

MUNDLAK Y., LARSON D.F., BUTZER R., (1997). – "The Determinants of Agricultural Production: A Cross-Country Analysis". *Policy Research Working Paper* 1827. Washington, D.C.: The World Bank.

NEHRU V., SWANSON E., DUBEY A., (1993). – "A New Database on Human Capital Stock: Sources, Methodology, and Results". *Policy Research Working Paper* 1124. Washington, D.C.: The World Bank.

NGUYEN D., (1979). – “On Agricultural Productivity Differences Among Countries” *American Journal of Agricultural Economics* 61 (3), pp. 565-70.

PRITCHETT L., (1996). – “Where Has All the Education Gone ?” *Policy Research Working Paper* 1581. Washington, D.C.: The World Bank.

SCHEFFE H., (1959). – *The Analysis of Variance*. New York: John Wiley and Sons.

SCHULTZ T.W. (1953). – *Economic Organization of Agriculture*. New York: McGraw Hill.

SOLOW R.M., (1957). – “Technical Change and the Aggregate Production Function”. *Review of Econometrics and Statistics* 39 (3), pp.312-20.

TINTNER G., (1944). – “A Note on the Derivation of Production Functions from Farm Records”. *Econometrica* 12, pp. 26-34.

TINTNER G., BROWNLEE O.H., (1944). – “Production Functions Derived from Farm Records”. *Journal of Farm Economics* 26(3), pp. 566-71 (a correction in *JFE*, February 1953, p.123).

YAMADA S., RUTTAN V.W., (1980). – “International Comparisons of Productivity in Agriculture”. In *New Developments in Productivity Measurement and Analysis*, J.W. Kendrick and B. N. Vaccara, eds. pp.509-94. Chicago: The University of Chicago Press.