Vol XXI
No. 4

ISSN 0019-5014

CONFERENCE NUMBER

OCTOBER-DECEMBER 1966

INDIAN JOURNAL OF AGRICULTURAL ECONOMICS

INDIAN SOCIETY OF AGRICULTURAL ECONOMICS, BOMBAY

RAPPORTEUR'S REPORT

ON

MEASUREMENT OF AGRICULTURAL GROWTH

RAPPORTEUR : B. S. MINHAS

Planning Unit
Indian Statistical Institute, New Delhi

The agreed scope of this subject was sharply delimited to quantitative assessment of agricultural growth. The participants were asked to present analyses of the record of growth in this sector as a whole or its various sub-sectors. Comparative quantitative studies of the growth experience of different nations, different regions within a nation or among smaller geographical units within different regions such as districts and blocks were solicited. Methodological contributions would have been just as welcome. Assessment of the relative contributions of various factors to agricultural growth and the quantification of inter-relationships among inputs and outputs were to form a part of the subject-matter of this session. Almost all the papers submitted for discussion fell within the scope of the subject-matter described above. However, no contributions were received on the important subject of the relationship between agricultural growth and the growth of the rest of the economy. Equally neglected was also the related question of the growth of demand for agricultural products and its significance for the rate and pattern of agricultural development.

Out of a total of 13 papers received on this subject, nine were recommended for full reproduction. The remaining four papers were to be printed only in summary form.

To facilitate critical scrutiny and avoid repetition in discussion, the nine fully reproduced papers can quite conveniently be classified into a few categories.

1. *Linear Trend Fitting :* Papers by Amiyamoy Chatterji, J. L. Kaul (in part), Niranjan Saha and D.A. Patel comprise of simple linear trend analysis of different sets of data on area, productivity and production either of single crops or a group of crops.

2. *Exponential Trend Fitting :* Papers by H. K. Das Gupta and (a large part) R. Dayal are concerned with exponential trend fitting.

3. *Arithmetical Accounting of Growth Components :* Parts of papers by Kaul and Dayal attempt to split up crop output growth arithmetically into relative contributions of area, cropping pattern and yield changes. Computations are based on differences between triennial end point averages of time series of crop production. Kaul uses an additive decomposition scheme on the difference between the average output of six foodgrains in 1950-53 and 1962-65 for 16 districts of the Punjab;

whereas Dayal splits up the difference in total crop output of 12 assorted countries on the basis of triennial end point averages of data pertaining roughly to the decade 1950 to 1960. He makes use of a multiplicative model.

4. *Fitting of Production Functions:* The only paper that set itself to the task of quantifying inter-relationships among inputs and total agricultural output is by R. Giri, A. V. K. Sastri and D. S. Somayajulu. They have experimented with a number of multiple regression equations and tried to estimate 4 to 5 parameters on the basis of 12 time series observations. These estimated coefficients of the time series production function are made use of in deriving the relevant growth equations.

5. *Miscellaneous :* Two papers fall in this category. They present no methodological problems and I am going to comment on them right away.

Umananda Phukan's paper is essentially a straightforward description of patterns of land ownership, size distribution of owner as well as operational holdings, land utilization and cropping patterns in the Lahoal development block of the Lakhimpur district of Assam. The only possible claim that the author can make for having measured the growth of agricultural output in the Lahoal development block will have to consist of the estimates of production of paddy and pulses for just two years, *i.e.*, 1961-62 and 1965-66. He finds a difference of 10 per cent in the output of foodgrains between the two years and notes its inadequacy in relation to population growth. These estimates, of course, are not his own : they have been prepared on the basis of average yield computed by the Statistical Department of the State of Assam for the Lakhimpur district. I very much hope that the multiplications by area have been done correctly. To justify my fear about arithmetical mistakes in Phukan's paper, I report some of them for correction. In his Tables I and IV even simple proportions are not correctly worked out. In the fifth row of Table I, only the first entry is correct. The second, third and fourth entries should have been 32.4, 61.0 and 39.0 in place of 36.0, 64.7 and 35.3 per cent respectively. In Table IV, the numbers in the last column should be 74.28, 20.00 and 5.72 in place of 70.42, 23.87 and 5.71 respectively. In the text of his paper he gives a wrong calculation about population growth in the Lahoal block. The rural sector of this block in 1961 had a population of 43,312 and this growing at a rate of 3.5 per cent in the year 1965-66 is estimated to be around 58,000. This is not correct, it should be more like 51,500.

One last paper, which clearly falls outside the scope of the present subject-matter but has been fully reproduced, is by N. S. Jodha. It is concerned with the pattern of agricultural growth in the Arid Zone of Western Rajasthan. The paper is largely descriptive but provides a good illustration of how the subsistence orientation of the farmer and the acceleration of the rate of growth of population in the past couple of decades have produced a pattern of land utilization and agricultural growth which has seriously disturbed the ecological equilibrium of nature in this region. In the author's view, left to itself the present pattern of growth of agriculture in the Arid Zone may lead to a permanent and irreparable loss of natural resource base. The situation calls for the introduction of resource conserving economic organization and technology into the Arid Zone agriculture.

I address myself now to each category of problems listed above and offer my comments on treatment given them by different authors participating in this discussion.

LINEAR TREND AND LINEAR GROWTH RATES

The two studies[1] on growth of Indian agriculture put out by the Directorate of Economics and Statistics, Ministry of Food and Agriculture can certainly claim to have stimulated a good deal of work on measurement of agricultural growth. The honest consumers of this research output of the Directorate, however, have proved to be over-eager customers. They have carried the errors and omissions from these studies over into their own work. The identification of constant absolute increase per unit time with the average linear growth rate in the papers by Chatterji, Kaul, Saha, Patel and to some extent, Giri, Sastri and Somayajulu, can be traced back to these studies. In a linear trend equation, $Y(t) = a + b \,.\, t$, b means b tons of wheat increase per year in wheat output when Y stands for tons of wheat and b bales of cotton increase per year in cotton output when Y stands for bales of cotton. Therefore, with the exception of indexed variables, b is always in terms of units of the commodity that figures on the left hand side of the equation; whereas r, per cent rate of growth per annum, if free of units in terms of which the commodity that grows is measured. Let us denote the time rate of change of Y by $\dot{Y}$; and $\dot{Y} = \frac{dy}{dt} = b.$

When we talk about per cent rate of growth, r, we mean $\frac{\dot{Y}}{Y}$.

That is $r = \dot{Y}/Y = dy/dt \,.\, 1/y = b \,.\, 1/y.$

In Chatterji's Table IV, Kaul's Table I, Saha's Tables I and II, Patel's Tables I and II, b is clearly confused with r. None of these authors show awareness of the fact that b has to undergo some operations on it before it turns itself into r, the growth rate in per cent per annum. Giri, Sastri and Somayajulu seem to have derived the 2nd and 4th entries in column 3 of their Table IV by dividing the appropriate b by the average of 1st and 2nd number of column 3 of their Table I.[2] This certainly gets rid of the Y-units from b, but the estimate of the average linear (%) rate of growth in Y will be an over-estimate. This is particularly the case with fast rising series of Y values. Oftentimes people divide b by the arithmetic or the geometric mean of Y. This is like becoming victims of a routine. What average one should use depends on the particular purpose in question. In spite of the fact that I have not found much use for it elsewhere, I regard the harmonic mean of Y as the only appropriate candidate here. The validity of this assertion can be proved.

The linear trend equation, $Y(t) = a+bt$, gives us an estimate of the absolute increase per unit time. However, we are interested in finding one single per cent

1. Growth Rates in Agriculture, December, 1964; Growth Rates in Agriculture, 1949-50 to 1964-65, March, 1966.

2. If this guess of mine is not correct and if they have not divided the b in this way, then their calculation of $b(=3.9)$ in the equation $Y=a+bt$ itself is incorrect.

rate (r) of growth such that Y growing at this particular rate (r), starting from y(0) = a, reaches a value equal to a + bT in T years. One assertion is that this rate r is obtained when b is divided by the harmonic mean, H_y, of Y. In other words, $r = b \cdot \frac{1}{H_y}$. Expressed in continuous form, we want to find an r such that;

$$a\, e^{rT} = a + bT.$$

Dividing through by a and taking logs we get ;

$$r = \frac{1}{T} \cdot \log_e \left(\frac{a + bt}{a}\right).$$

In the discrete case the inverse of the harmonic mean is defined as :

$\frac{1}{H_y} = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{y_t}$; and in its continuous counterpart,

$$\frac{1}{H_y} = \frac{1}{T} \int_{0}^{T} \frac{dt}{a + bt}.$$

This integral can be easily evaluated :

$$\frac{1}{H_y} = \frac{1}{T} \int_{0}^{T} \frac{dt}{a + bt}$$

$$= \frac{1}{T} \left[\frac{\log (a + bt)}{b}\right|_{0}^{T}$$

$$= \frac{1}{bT} \left[\log (a + bt) - \log a\right]$$

$$= \frac{1}{bT} \left[\log \left(\frac{a + bt}{a}\right)\right].$$

Multiply the expression on the right hand side with b and obtain

$\frac{1}{T} \cdot \log_e \left(\frac{a + bt}{a}\right)$, which is the definition of the particular r that we were after.

Hence, the assertion that $r = b \cdot \frac{1}{H_y}$ stands proved.[3]

With these general remarks on linear trend fitting out of my way, I now turn to specific comments on the contributions of Chatterji, Kaul, Saha and Patel.

The last sentence in section 3.4 of Chatterji's paper—"Annual increase in per acre of output of all food items, excepting that of gram which had a higher rate of increment, was small in comparison to those in the base year"—is extremely

3. This proof was suggested to me by my colleague Dr. T. N. Srinivasan. For this and the discussion of many other points in the preparation of this report, I gratefully acknowledge his help.

puzzling. He claims (vide his section 3.3) to be comparing the one single increase in yields per acre between 1949-50 and 1950-51 with the trend rates of increase worked out from 13-year data. This does not seem to be legitimate. It is quite likely that he is doing something else. I suspect that his computational blunder in the case of gram has led him on to making the statement cited earlier. In his Table IV the entry in column 4 against the gram row looks an impossibility. Between 1950-51 and 1962-63, the area under gram went up from 91.6 to 123.6, the total gram production during the same period has come down from 98.5 to 83.9. How can yield per acre have a trend rate of increase of 17.65 points per year? The section on ratios of trend coefficient along with Table V is hard to make sense of and section 4.0 (d) is no better. In section 4.0 (e), the author approaches the limit of naivetè : his simple computations mislead him to make absured pronouncements on public policy. Does the author think that the larger area expansion under wheat in comparison with barley and gram is necessarily a bad thing? If so, why? What specific planning or other policy lapses by government, in his opinion, were responsible for the bad result that this trend analysis has unearthed for him? Rather than spend his time in tackling policy issues, I wish he had given the standard errors of his estimated trend coefficients.

Saha, as previously indicated, also wrongly identifies constant absolute increase per unit time with the average linear (%) growth rate. Though the hypothesis of no trend in area and production of rice as well as total foodgrains for the period 1951-52 to 1961-62 in Assam plain division or any of its seven constituent districts cannot be rejected, Saha spends a third of the text of his short paper comparing inter-district trend rates and trend in foodgrain production and population growth. This is sheer wastage of labour. There is no trend in area or production anywhere in Assam plain division. Why compare a string of zeros ?

Patel's otherwise interesting study of the growth of groundnut in four regions of Gujarat is marred by some serious technical errors. Since the area and production of groundnut in Gujarat has had very sizable trends during the period of this study, the linear trend rates presented by Patel are likely to produce a mistaken impression of the rates of growth. The evaluation of these trend rates at the harmonic means of the corresponding variables would have given the correct estimates of the rates of growth.

That linear trends fit the groundnut production and area extremely well is evident from his paper. However for yield per acre he advocates fitting a second degree curve. For a small value of t, one can obtain a reasonably well fitting second degree curve for trend in yield per acre. However, looking at his per acre yield graph and using some logical reasoning, I am inclined to believe that Patel is beset by an optical illusion. Let A stand for area under groundnut, P for total groundnut output, then yield per acre, y, is equal to P/A.

$$\text{Now if } A = a + bt, \text{ and } P = c + dt \text{ ; then } Y = P/A = \frac{c + dt.}{a + bt} \cdot$$

For finite values of t, yield per acre (y) would reach an asymptote at d/b. If $d\,a > b\,c$, Y will reach this asymptote from below. If however $d\,a < b\,c$, that is, $c/a > d/b$, then the asymptotic value of Y will be reached from above. However, Patel's graph curves upwards and this possibility is ruled out. I suggest that Patel should try to estimate trends in groundnut yield per acre with the use of a power function, *i.e.*, $Y = at^c$. With c positive but less than unity, this function will graduate Patel's data just as well as the quadratic. Moreover, unlike the quadratic case, the mere passage of time will not force him to mechanically project sizable per acre yield declines by the end of the Fourth Plan. Gujarat groundnut yields will not travel on the downward sloping part of Patel's quadratic : he will get rid of his optical illusion. Let us remember that extrapolation is a tricky business. Extrapolation with time as a variable is dangerous : it is suicidal when a quadratic (polynomial) involving time is used; exponentials, when force of growth is greater than unity, are just as troublesome. I recommend that Patel may revise his calculations in the light of these comments and better work with unsmoothed data. Moving averages introduce serial correlation and trend fitting to smoothed data produces spuriously high correlation coefficients.

Kauls' share in common legacy of confusion between constant absolute linear trend and the growth rate apart, I find that his estimates of the rates of value productivity growth as they stand (knowing full well that they are over-estimates) are rather on the low side for the Punjab. Is it that the six food crops studied by him have been growing at a much slower rate than the non-food crops during the period 1950-51 to 1964-65 ? I very much hope that his computations do not suffer from some systematic error. His expression for value productivity per acre should be $\Sigma_i C_{ij} P_i/\Sigma_i A_{ij}$ rather than his $\Sigma_i C_{ij} P_i/A_j$; because the relevant net area sown in the district is the area sown to the six crops under study rather than the area under all crops.

EXPONENTIAL TREND FITTING

Das Gupta's paper is concerned with fitting of exponential trends to area under rice and yields per acre in Orissa during the six-year period 1959-60 to 1964-65. Out of total area of 10.5 million acres under rice, summer and autumn rice account for approximately 5 per cent of it. The winter crop is the main rice crop and the author finds that both area and production of this crop have expanded during the period 1959-60 to 1964-65. Yield per acre has been estimated to rise at a compound rate of 1.37 per cent per annum. The author finds a significant relation between yields per acre and total June to October rainfall in these years and reports that the simple analysis of variance does not reveal any significant difference in the yield rates between the first two and the last four years of the period under study.

A large part of Dayal's paper is also concerned with exponential trend fitting to total agricultural production of 57 countries during the period varying roughly from 1952-53 to 1962-63. Remember that total agriculture comprises temporary crops, permanent crops, livestock (including poultry) and dairy products. The author is using the F.A.O. data. He mentions two types of data on total agricultural production available with the F.A.O.—the index numbers of total agricultural production and the value aggregates in terms of wheat equivalent tons. Nevertheless, he does not make clear which of these form the basis of his growth rate estimates. He finds that in more than 25 per cent of the countries in his sample, the growth rates are not significant. Working with such crude data (whose merits he does not care to discuss) which is bound to have very large margins of error, one can never over-emphasize the virtues of caution. The author is making all kinds of generalizations without as much as a murmur about the inadequacy of his data base.

ARITHMETICAL ACCOUNTING FOR GROWTH COMPONENTS

Kauls' results in his Table II are based on the additive decomposition scheme used by Minhas and Vaidyanathan.[4] However the use of tonnage in the last column of this table is open to very serious objection. He should have instead used constant price value sums. The split up of tonnage into components in proportion to percentage increases shown in parentheses can be done only if the share of different grains in the total grain output of every district is constant. This assumption is highly unplausible. The top line of numbers against each district should better be dropped.

It must be recognized that the usefulness of this additive model is questionable if the interaction or residual element is rather large. In order to bring the maximum extent of the residual effect out in the open, soon after sending our just cited paper for publication I did an expanded set of calculations in the framework of a seven factor additive model for every State. The net import of these calculations was to reduce the pure effects of yield and cropping pattern change. Two new interactions, one between area and yield and the other between area and crop pattern, were identified and the four factor model's single interaction element was split into two, *viz.*, yield-crop pattern interaction with area fixed at base year level and a three factor interaction between changes in area, yield and crop pattern. Notational representation of this model is as follows :

$$P_t \equiv A_t \sum_{i=1}^{n} w_i \; C_{it} \; Y_{it}$$

$$P_o \equiv A_o \sum_{i=1}^{n} w_i \; C_{io} \; Y_{io}$$

$$\begin{aligned} P_t - P_o \equiv\ & (A_t - A_o) \; \Sigma \; w_i \; Y_{io} \; C_{io} + A_o \; \Sigma \; w_i \; (Y_{it} - Y_{io}) \; C_{io} \\ & + A_o \; \Sigma \; w_i \; (C_{it} - C_{io}) \; Y_{io} + (A_t - A_o) \; \Sigma \; w_i \; (Y_{it} - Y_{io}) \; C_{io} \\ & + (A_t - A_o) \; \Sigma \; w_i \; (C_{it} - C_{io}) \; Y_{io} \\ & + A_o \; \Sigma \; w_i \; (C_{it} - C_{io}) \; (Y_{it} - Y_{io}) \\ & + (A_t - A_o) \; \Sigma \; w_i \; (Y_{it} - Y_{io}) \; (C_{it} - C_{io}) \end{aligned}$$

The sum of the last four terms on the right hand side of this equation was taken as the total interaction or residual component and naturally it came out larger than the single interaction element of the four factor scheme. Nevertheless,

4. "The Growth of Agricultural Output in India, 1951-54—1958-61: An Analysis by Component Elements," *Journal of the Indian Society of Agricultural Statistics*, December, 1965.

the residual component of the expanded model was not larger than 9 per cent in any of the 14 States studied by us in the above cited paper. In most cases it was quite small. A comparative picture of the results obtained from the two models for the Punjab for the period 1951-54—1958-61 may be of interest to Kaul and Table I is intended for this purpose.

TABLE I—PERCENTAGE CONTRIBUTION OF DIFFERENT ELEMENTS

		Seven Factor Model	Four Factor Model
Pure effects	1. Area	69.93 (3.59)	69.93 (3.59)
	2. Yield	6.17 (0.32)	7.98 (0.41)
	3. Crop pattern	17.30 (0.88)	22.38 (1.15)
Joint effects	4. Area and yield	1.81 (0.09)	—
	5. Area and crop pattern	5.07 (0.27)	—
	6. Crop pattern and yield	—0.22 (—0.01)	—0.29 (—0.01)
	7. Area, crop pattern and yield	—0.06 (—0.00)	—
	8. 4+5+6+7	6.60 (0.35)	—0.29 —(0.01)

The residual element in the four factor model was negligible and in the expanded model this element amounts to 6.6 per cent of the total. In other words, 0.35 percentage points of a growth rate of 5.14 per cent are not accounted for by pure effects of changes in area, yield and crop pattern. The empirical results obtained with the use of the expanded additive model at the State level are not very different from the corresponding results produced by the computationally simpler four factor model. It is for this reason that I have not so far published my work on this expanded model. I have also done similar computations for about 250 districts and found that the residual element at this level varies far more widely. This is also confirmed by Kaul's Table II—the residual elements for Patiala, Ambala and Kapurthala are quite large. In such situations, it is better to work with the seven factor model.

All these decomposition schemes, nevertheless, serve only to help organize data in a manner which is more suitable for the analysis of causal factors in the growth process and it is this latter aim towards which studies at the small area level should be directed. In the closing paragraph of his paper Kaul makes some observations on the effect of assured irrigation on yields in the districts of Amritsar, Ferozepur, Ludhiana, Jullundur, Sangrur and Bhatinda. This sort of casual empiricism would not do : such relations need to be studied in depth.

In the later half of his paper, Dayal uses a multiplicative split up scheme to compute the effects of area, cropping pattern and yield changes in the total crop production in 12 countries. His formula is

$$\frac{\Sigma\, Y_{ti}\, A_{ti}\, P_{oi}}{\Sigma\, Y_{oi}\, A_{oi}\, P_{oi}} = \frac{\Sigma\, A_{ti}}{\Sigma\, A_{oi}} \cdot \frac{\Sigma\, Y_{ti}\, R_{ti}\, P_{oi}}{\Sigma\, Y_{ti}\, R_{oi}\, P_{oi}} \cdot \frac{\Sigma\, Y_{ti}\, R_{oi}\, P_{oi}}{\Sigma\, Y_{oi}\, R_{oi}\, P_{oi}}$$

where Y, A, P, R stand for yield, area, prices and croping pattern respectively. Subscripts o and t stand for different points of time. This formula is the same that the Technical Committee on Index Numbers, early this year, had recommended for use in the work of the Directorate of Economics and Statistics, Ministry of Food and Agriculture. This formula is a variant of a formula that I had worked out in 1964 and had later presented it along with some computations at a seminar of the Institute of Agricultural Research Statistics in February, 1965. In the Technical Committee on Index Numbers I had agreed to the adoption of this shortened version because of its comparative simplicity. Nevertheless, it is not conceptually clean; it evaluates the crop pattern effect at final year yields. One is interested in knowing the impact of a given crop pattern change if the base year yields and prices prevailed. In my formula this is achieved but like our additive model this multiplicative model also has a residual element in it. I had split $\frac{P_t}{P_o}$, the ratio of final and base year production, in the following manner :

$$P_t/P_o \equiv \frac{A_t\, \Sigma_i\, w_i\, Y_{it}\, C_{it}}{A_o\, \Sigma_i\, w_i\, Y_{io}\, C_{io}}$$

$$\equiv (A_t/A_o)\left[\frac{\Sigma\, w_i\, Y_{it}\, C_{io}}{\Sigma\, w_i\, Y_{io}\, C_{io}}\right]\left[\frac{\Sigma\, w_i\, Y_{io}\, C_{it}}{\Sigma\, w_i\, Y_{io}\, C_{io}}\right] \times$$

$$\left[\frac{\Sigma\, w_i\, Y_{it}\, C_{it}}{\Sigma\, w_i\, Y_{it}\, C_{io}} \cdot \frac{\Sigma\, w_i\, Y_{io}\, C_{io}}{\Sigma\, w_i\, Y_{io}\, C_{it}}\right].$$

Notice that the area and yield effect terms are the same in both models; the third term on the right hand side in my formula is the effect of cropping pattern at base year yields. In my view this is the right way of looking at the problem. This identity, surely, can be split up in a number of ways but we are not interested in arithmetic as such nor in index number construction alone. One should choose a split up which is capable of giving better results analytically. Notice also the fact that my interaction term is a product of two elements, *i.e.*, the crop pattern at final year yields (the element which is taken to be the crop pattern by Dayal) and the inverse of the crop pattern at base year yields.

Now let us look at Dayal's numerical work with the multiplicative model. This is given in his second table. (Typescript of Dayals' paper was a cut and paste job. It was very casually done and he did not number his tables; whereas the pages carried two sets of numbers). The first column of numbers in this table gives the rate of growth in crop output. And since the numbers on the right hand side are obtained by splitting an identity into parts, the sum of numbers in columns 2, 5 and 6 of any row should not differ from the number in the first column, *i.e.*, the overall growth rate, by more than an infinitesimal amount. As to why this should be so, I shall explain below. However, as it is clear from Table II, for a number of countries these differences are rather large.

TABLE II

Country 0	Growth rate % Dayal's col. 1 1	Sum of Dayal's cols. 2, 5 and 6 2	1 — 2 %
India	3.9	3.6	0.3 or .003
Thailand	5.0	4.5	0.5
Spain	2.6	2.5	0.1
China—Taiwan	4.7	4.2	0.5
Greece	6.4	5.6	0.8
Turkey	4.9	4.7	0.2
Sudan	12.1	10.7	1.4
Egypt	2.4	1.9	0.5
Mexico	7.9	6.9	1.0
Jordan	3.0	3.5	—0.5
Pakistan	2.4	2.3	0.1
Chile	4.2	3.9	0.3

In the multiplicative model, the index of crop output is equated to a product of index numbers of area, yield, cropping pattern and interaction. (In Dayal's equation the interaction does not figure in the right hand side). In symbols:

$$P_t \equiv A_t \cdot Y_t \cdot C_t ;$$

where P, A, Y and C stand for index numbers of production, area, yield and crop pattern. If each of these index numbers obeys the usual law of growth, then we can write the above expression as follows :

$$P_o e^{rT} = A_o e^{r_1 T} \cdot Y_o e^{r_2 T} \cdot C_o e^{r_3 T}$$

The relation between the overall growth rate, r and growth rates, r_i's of the components is obvious. One can fit regressions of the type $X = ab^t$, *i.e.*, $\log X = \log a + t \log b$ to each of the index numbers and estimate their growth rates (r's). Remember that antilog of the regression coefficient will be equal to $(1 + r)$. The sum of the regression coefficients of the logarithmic equations on the right hand side of the identity will equal to the single regression coefficient on the left hand side. Or, alternatively, the following relation will hold :

$$(1 + r) = (1 + r_1)\ (1 + r_2)\ (1 + r_3) \quad \text{or}$$

$$r = r_1 + r_2 + r_3 + r_1r_2 + r_1r_3 + r_2r_3 + r_1r_2r_3.$$

A growth rate of, say, 3.9 per cent as given by Dayal in case of India, will be written as 0.039 ; $r_1 = 0.019$, $r_2 = 0.016$ and $r_3 = .01$. All these numbers are very small and terms beyond r_3 are second and third order of smalls and may contribute a significant number only in the fourth place after the decimal point. In the illustration chosen here, a difference of .0006 between r and sum of r_i's is admissible. However, Dayal's differences for all countries are too large ; in case of India, for example, three-tenths of a percentage point in a growth rate of 3.9 per cent. I am aware, however, that Dayal is not fitting regression relations for finding out the r's. He is not using all the information and is simply working with the difference between the triennial averages 8 to 10 years apart. For short series this is not bad. One can read the r values off from compound interest tables or consult log tables. In a yet unpublished paper of mine, I have used my version of the multiplicative scheme given above for the study of agricultural growth in India as a whole, 14 States of the Indian Union and the two wings of Pakistan. I have worked with all observation through the regression method as well as with the end point averages alone. In order to illustrate the point about the arithmetic equivalences that I have been talking about, I reproduce a table of my results at the all-India level for the period 1951-1961 (Table III).

TABLE III—REGRESSION RESULTS: $X = ab^t$

	Area	Yield	Cropping pattern	Inter-action	Total production
B or (log b)	.007231** (.001061)	.006514* (.002222)	.000972 (.000871)	.000111 (.000319)	.014827** (.002299)
A or (log a)	2.0068	2.0128	1.9867	2.0004	2.0066
Correlation coefficient ..	.9236**	.7197*	.3669	.1224	.9158*
a	101.568	102.998	96.9748	100.083	101.531
b=(1+r)	1.0168	1.0151	1.0022	1.00026	1.0347
	Two point estimates between averages of 1951-54 and 1958-61				
(1+r)	1.01711	1.01541	1.00285	1.0003	1.0357

* Significant at 5 per cent ; **Significant at 1 per cent.

Notice that the independently derived regression coefficient (B) for total production which is equal to .014827 is equal to the sum of the four other independently estimated regression coefficients in the top row. Total crop production has been estimated to grow at an annual rate of 3.47 per cent; whereas area, yield and crop pattern effect have been growing at the respective annual rates of 1.68 per cent, 1.51 per cent, 0.22 per cent. Interaction has been insignificant. The sum of r_i's is 0.03436 against an r of 0.0347. The results obtained on the basis of end point averages are fairly close to the regression results. The growth rate in this case comes out to be 3.578 per cent; area growth at 1.711 per cent; yield at 1.541 per cent; crop pattern effect is found to be 0.285 per cent and the interaction is again insignificant. The sum total of r_i's, which is equal to 0.0354 per cent is different from the overall by a very small amount (0.0357—0.0354 = .0003).

Large differences in Dayal's growth rates may be the result of using internally inconsistent data for different index numbers. This is very likely. Or he may not be aware of the fact that in splitting up the components of an identity and estimating each component separately, one is walking a very tight rope. Anyway in the discussion of his results, he is not at all bothered by the fact that components do not add up: this fact does not draw even a single mention from the author.

FITTING OF PRODUCTION FUNCTIONS

Now I discuss the last (but very important) of the methodological problems of our subject-matter. Giri, Sastri and Somayajulu have touched upon many important technical issues. They have provided us with 12 different fits of the production functions of linear and Cobb-Douglas types with just 12 observations to serve as machine feed. The electronic computing machines are very benevolent lenders : they never bother us about the number of degrees of freedom that we owe them.

Besides time, they have used three independent variables, X_1—the gross area sown; X_2—the proportion of gross irrigated area to gross area sown (*i.e.*, X_1); and X_3—quantity of fertilizer (in terms of nitrogen) per unit of irrigated area, to explain the behaviour of total crop output, Y. The authors claim that they have used X_2 and X_3 in ratio form in order to reduce multicollinearity and specification bias. I want to examine this claim. Let us look at the latter aspects of their claim first.

I am afraid that the authors do not have a clue to the commonly understood meaning of specification bias. By specification bias we mean the followings : when the specification of the *list* of relevant factors in a relation is wrong and if, say, the n th factor is omitted—either because we cannot measure it though we consider it important or we just omit it because we do not regard it relevant though, in fact, it is—the estimate of the i th coefficient is biased by $\hat{\gamma}_{ni}\, \beta_n$ where $\hat{\gamma}_{ni}$ is the coefficient of X_i in the regression of omitted factor on the factors included in the list of independent variables . One can sketch a simple algebraic representation of the concept in the following manner :

Let the true specification be given by,

$Y = X_1 \beta_1 + X_2 \beta_2 + \mu$; whereas the maintained specification is,

$Y = X_1 \beta + v$.

$$\text{Therefore } \hat{\beta} = (X_1' X_1)^{-1} X_1' Y$$
$$= (X_1' X_1)^{-1} X_1' [X_1 \beta_1 + X_2 \beta_2 + \mu]$$
$$= \beta_1 + (X_1' X_1)^{-1} X_1' X_2 \beta_2 + (X_1' X_1)^{-1} X_1' \mu; \text{ and}$$

$$E(\hat{\beta}) = \beta_1 + \hat{\gamma}_{21} \beta_2.$$

Notice that in the regression of X_2 (omitted factor) on X_1 (the included factors *i.e.*, $X_2 = \hat{\gamma} X_1 + \xi, \hat{\gamma} = (X_1' X_1)^{-1} X_1' X_2$. In our maintained specification the expected value of $\hat{\beta}$ differs from its true value by $\hat{\gamma}_{21} \beta_2$. This value $\hat{\gamma}_{21} \beta_2$ is the magnitude of the specification bias.

If one could help it, one would never let this bias enter in one's work. When one cannot help it, one can at least recognize it. Entering the same set of variables in one way or the other, as Giri and his colleagues seem to think, does not tackle (there is nothing to "reduce") the problem of specification bias. I can mention a number of variables, labour and draft power being the two most important ones, which these authors have omitted from their equations and, in consequence, the specification bias is sitting pretty in their main entrance.

I now turn to the other half of their claim. After a good deal of work, which the authors may have never intended, I have been able to establish that this claim of the authors one can let go by as some sort of a half truth. Let us recall that Giri and his colleagues are working with Y as a function of X_1, X_2 and X_3; where $X_2 = G/X_1$ (G stands for irrigated area); and $X_3 = F/G$ (F stands for quantity of fertilizer). Their claim about the presence of multicollinearity in the non-ratio form would mean that :

(i) $G = \alpha + \beta X_1$; and
(ii) $F = \gamma + \delta G$.

Hence when Y is conceived of as a function of X_1, G and F, *i.e.*, $Y = g(X_1\ G, F)$, the variables X_1, G and F are collinear with each other. But if we consider Y as a different function of X_1, G/X_1 $(= X_2)$ and F/G $(=X_3)$, *i.e.*, $Y = f\ (X_1, G/X_1, F/G)$, then this functional representation f is devoid of multicollinearity.

Working up to first approximation, one can prove that $\mathrm{Cov}(X, G/X) = 0$.[5] One will also have to prove that :

(a) $\mathrm{Cov}(G/X, F/G) = 0$; and
(b) $\mathrm{Cov}(X, F/G) = 0$.

Proofs for (a) and (b) are very similar to the one already given. The one for (a) goes through quite easily but in case (b) (remembering that we are neglecting terms containing powers more than unity of $1/\mu$), $\mathrm{Cov}(X, F/G) = 0$, if $\alpha < \beta\mu$.

5. Let $E(X) = \mu_1$
$E(X-\mu_1)^k = \mu^k$; $k = 2, 3, \ldots\ldots\ldots\ldots$

Then $E(1/X) = 1/\mu_1 \left[1 + \frac{\mu_2}{\mu_1^2} - \frac{\mu_3}{\mu_1^3} + \ldots\ldots\ldots\right]$.

or $E(1/X) = 1/\mu_1$

Then
$$\begin{aligned} \mathrm{Cov}\,(X, G/F) &= E(G) - E\,(X)\,.\,E\,(G/X) \\ &= \alpha + \beta\mu - \mu\,(\alpha/\mu + \beta) \\ &= 0. \end{aligned}$$

In other words, our conclusion is that the claim for the absence of multicollinearity in the functional form f $(X_1, G/X_1, F/G)$ is true provided, $\alpha < \beta X_1$ for all values of X_1 or if $1/X_1 < \beta/\alpha$.

Reflecting further, I find that the fulfilment of this condition implies that changes in gross irrigated area must be proportionate to changes in gross area sown. Looking at col. 4 in Table I of Giri and his colleagues, this condition is pretty nearly fulfilled. Except for two or three observations, the index number for X_2 stays pretty nearly constant. However this accidental fulfilment of the condition for the absence of multicollinearity between X_1, X_2, and X_3 throws the entire specification of the production function out of gear.

In a non-homogeneous linear regression equation, if one of the independent variables included in it stays pretty nearly constant (like X_2 of Giri, Sastri and Somayajulu), near perfect multicollinearity enters the problem through the back door. Giri, Sastri and Somayajulu, as I have tried to explain above, were able to accidentally avoid multicollinearity between X_1, G/X_1 and F/G. It is ironic, however, that near constancy of G/X_1 which came to their rescue in one sense, later wrecks the entire game for them : in equations 1, 3, 4, 5, 6 and 7 of Table II, there does not appear a single significant coefficient. Multicollinearity entered through the back door and took its heavy toll. It is only when the variable X_2 is dropped, that a few significant coefficients appear in Table III.

The constant term in a regression equation is always of considerable interest. The authors of this paper did not present any estimates of the constant term. Their availability would have helped in the interpretation of results here. In their regression Nos. 2 and 9 the estimate of the constant term would have been of particular interest.

Grave consequences from which the work of Giri, Sastri and Somayajulu has suffered, prompts me to share a part of the received econometric wisdom with the participants of this Conference : in empirical work avoid mixing up of extensive and intensive magnitudes in one and the same relationship. In such specifications, interpretation of results is often difficult and many times they make no sense whatever. The models 4, 7, 10 and 12 of this paper belong to this latter category. Notice that I am discussing the question of specification from an altogether different angle now. I am concerned with the wider question of specifying the form in which variables should enter into models and not with specification bias. When *total crop output* (Y) is a linear function of gross area sown, X_1; the proportion of gross irrigated area to gross area sown, X_2; and amount of fertilizers per unit of irrigated area, X_3; what happens to Y really depends on, what happens to the other extensive magnitude in the relation, *i.e.*, gross area sown. Consider a year of exceptional drought. The gross sown area supposedly can get cut down by a large factor; the proportion of gross area irrigated to gross area sown can go up phenomenally. Since fertilizers are used only on irrigated lands, the fertilizer use per irrigated acre would stay more or less constant. What would happen to the total output in this year? Think through the reverse case. A year of very good rainfall; nobody irrigates any fields. The gross area sown is at its peak; the proportion of gross area irrigated to gross area sown goes down; all lands, irrigated and rain-fed, can use fertilizers; the amount of fertilizer per irrigated acre is likely to decrease. What would happen to the total agricultural output during this year? The changes in the extensive magnitude X_1, by and large, provide the answer to our question; the usual *ceteris paribus* kinds of statements do not make much sense for the intensive variables in this case. It is these kinds of commonsense considerations that speak against mixing up extensive and intensive magnitudes. In order to attach their due meaning to the partial derivatives of Y with respect to X's, the arguments of the production function, the specification of the form of X's is of great relevance. Regression cannot produce meaningful coefficients unless the variables are fed in appropriate form.

The authors of this paper have been trying different transformations of variables rather blindly. Compare, for example, their models 1 and 3. Both these models are in logarithmic form. In model 1, the dependent variable is the total crop output; whereas in model 3 they are working with output per acre. In this case estimating two separate equations is unnecessary. If model 3 were properly specified one could write it as follows :

$$\log (Y/X_1) = \log Y - \log X_1 = \log a + (b_1-1) \log X_1 + b_2 \log X_2 + b_3 \log X_3$$

Notice that all the coefficients of this equation are the same as those of model 1 except that instead of b_1, the coefficient of X_1, we have (b_1-1). Nevertheless, the authors of this paper have omitted the variable X_1 in models 3 and 6. This is as good an example of specification bias as one can get anywhere and in this case the magnitude of specification bias can in fact be measured. All this has happened right under their noses even though the authors gave the impression of being on guard right from the beginning.

Now I turn to their analysis of components of crop output growth as provided in their Table IV. I confine my attention to their growth equations 11 and 12 of this table. The authors did not provide estimates of the time rates of change of the independent variables — an estimate of the trend in Y, *i.e.*, dy/dt and of the rate of growth of Y, *i.e.* 1/Y. dy/dt was available. I present my estimates of trend rates of change and compound rates of growth of Y, X_1 and X_3 in Table IV. All these estimates are very highly significant.[6]

Notice that as against the linear and compound rates of 3.9 per cent and 3.5 per cent[7] given by Giri, Sastri and Somayajulu, my estimates come out to be 3.255 per cent and 3.317 per cent respectively (see line 1 cols. 4 and 5 in my Table IV). This substantiates numerically the already proved assertion that the estimated linear trend in a variable when multiplied by the inverse of its harmonic mean turns itself into the appropriate constant (%) rate of growth. Whether one fits exponential or linear trends, there is only one appropriate rate of growth and its estimated value under the two methods must come out very nearly the same.

6. Each of these variables is auto-correlated in time and the use of first difference transformations in properly specified models would have been the right thing to do. The authors did not pursue this line of attack because they found their coefficients of Multiple Determination getting reduced in a hurry. However, this should not have deterred them from the use of first differences. In variables subject to time trends, the non-first-difference-form analysis gives rise to spuriously high correlations and one cannot attach much significance to the magnitude of R^2. First difference equations 2 and 9 seem to be given sensible regression estimates (but for X_2 which has been dealt with extensively already); they are significant and of the right sign. However, the authors may have ignored these regressions because of other difficulties.

7. This estimate of the compound rate of growth of 3.5 per cent seems to suffer from some computational error.

Recalling that $\dot{Y}/Y = 1/Y\ dy/dt = \sum_{i=1}^{N} n_i\ g_{xi}$, where

$n_i = \dfrac{\partial \log Y}{\partial \log X_i}$ and $g_{xi} = \dot{X}_i/X_i = 1/X_i \,.\, dx_i/dt$, I find (using the estimated n_i's of model 11 of Giri and his colleagues from their Tale III) that the contributions of X_1, X_3 and t are respectively 79.7 per cent, 13.2 per cent and 6.9 per cent (or 3 per cent) in a growth rate of 3.32 per cent rather than their corresponding estimates of 75 per cent, 13 per cent and 12 per cent.

TABLE IV—TRENDS AND GROWTH RATES

Variable	Harmonic mean	Linear trend $b = dv/dt$	Exponent al trend $1/v.dv/dt = g_v$ (%)	Col. 3 ÷ Col. 2 (%)
(1)	(2)	(3)	(4)	(5)
Crop output (Y)	124.360	4.0478	3.317 (3.5%)	3.255 (3.9%)
Gross area sown X_1	112.987	1.7917	1.616	1.588
Fertilizers per unit irrigated area (X_2) ..	181.540	26.0158	13.328	14.331

I have already indicated that their estimate of the growth rate of Y in the linear trend case is wrong. Their representation (*i.e.*, equation No. 13) of the growth equation and the numerical computations of its components for the linear case (model No. 12, Table IV) is also wrong. In their model No. 12, $Y = a + b_1 X_1 + b_3 X_3 + b_4 t$.

Then $\dot{Y} = b_1 \dot{X}_1 + b_3 \dot{X}_3 + b_4$ where $\dot{Y} = dy/dt$ and $\dot{X}_i = d_{xi}/dt$.

Dividing through by Y, we obtain the growth equation:

$$\dot{Y}/Y = b_1 \dot{X}_1/Y + b_3 \dot{X}_3/Y + b_4/Y \text{ ; or}$$

$$\dot{Y}/Y = b_1 X_1/Y \,.\, \dot{X}_1/X_1 + b_3 X_3/Y \,.\, \dot{X}_3/X_3 + b_4/Y \,.$$

This is the familiar form of the growth equation : the sum of the products of the elasticity n_i ($= b_i X_i/Y$) and the exponential growth rates, g_{xi} ($= \dot{X}_i/X_i$) is equal to the growth rate of output. The expressions on the right hand side are again to be evaluated at the harmonic means of Y and X_i's. The relative contribution of X_1, X_3 and t works out respectively to be 86.2, 25.2 and (—) 12.1 per cent in an output growth rate of 3.255 per cent and *not* 85 per cent, 35 per cent and (—)20 per cent out of growth rate of 3.9 per cent.

The concluding section of Giri, Sastri and Somayajulu (but for the remark that three-fourth to four-fifth of the growth of crop output came out of gross sown area expansion which itself, by and large, is a consequence of inadequate specification of the production function) is not at all supported by their econometric findings. A good bit of growth is being attributed to factors for which they failed to get a single significant coefficient in twelve tries. The promise of God's kingdom on Indian soil—the modern technological break-through which is to be anchored on the larger use of irrigation and fertilizer may be good lay wisdom but it surely is not founded in their econometrics : a total of seven irrigation coefficients derived by them are all of negative sign. Pursuit of scientific knowledge and peddling of lay wisdom do not mix too well. I only hope that Giri and his colleagues will persist in their effort to firm up the scientific basis of lay wisdom which is the common legacy of us all.

A POSTSCRIPT

Lest they stake a claim for unfair treatment, I dedicate this short postscript to the four authors whose contributions have not received close scrutiny thus far.

Narottam Nanda's paper comprised of collecting total rice output data of 12 holdings in the village Kharadiha in the Nilgiri Block-1 of Balasore district in Orissa. The data pertain to two years, 1950-51 and 1965-66 and were obtained on the basis of a questionnaire. Out of the 12 holdings studied by the author, four augmented their size between 1950 and 1965 and the estimates of their 1950 rice output are derived by asking them what would have been their output if they had possessed the same amount of land in 1950 as they did in 1965. Such a procedure is down right absurd. One can hardly expect the peasants to correctly remember the year of birth of their third son, much less the total output of rice on their holding 15 years ago.

P. S. Sharma's "Some Aspects of the Measurement of Agricultural Growth in India" was purported to be an excercise in projecting the output of foodgrains, sugarcane, oilseeds, cotton and jute for the year 1965-66 on the basis of "best fit" trend line obtaining in 1952-53 to 1964-65 data. He sought some comfort in the fact that the figures of output for 1965-66 as assumed in the Draft of the Fourth Five-Year Plan were very close to his trend projections. He would have derived even greater comfort by computing the confidence bands around his trend. However it is just as good that he did not : the paper would have remained just as pointless.

M. L. Patel's "Choice of Criteria to Measure Agricultural Growth" was nothing more than an off-the-cuff effort in taxonomy interspersed with loose rambling discussion. The edited version of his summary quite adequately covers what was offered in the 18-page typescript. Some of the points in the summary can be discussed at the Conference.

K. Ramachandran Nair's "On India's Agricultural Growth" was a rehash of data tables from Changes in Agriculture in 26 Developing Nations, 1948-1963 of the Economic Research Service, U.S. Department of Agriculture, 1965. He culled comparative data from this study for India, Israel and Japan. In some tables the source was clearly acknowledged; whereas some others, which were put together in one piece on the basis of many different tables in the parent study, did not carry proper acknowledgement. In his discussion he went into many issues relating to the organizational base of agriculture in India, Israel and Japan but the discussion was not pointed enough to yield a central theme. All factors considered relevant to agricultural growth were touched upon at one stage or the other. His view on the organizational aspect of Indian agriculture can be best summed up by his following remark : "India cannot by-pass the labour-utilization approach to agricultural growth by jumping from institution-building approach to seed-fertilizer approach." However, this should be read along with the concluding paragraph of his paper where he stated that "What India immediately requires is a yield revolution in agriculture which is possible only with wider application of hybrid seeds, fertilizers and pesticides. Let us remember "every thing can wait but agriculture cannot." If higher yields are within our reach, let us have it." My surmise is that a string of *obiter dicta* and exhortations would not get *it* for us.