Vol XXXV
No. 2

ISSN 0019-5014

APRIL-
JUNE
1980

INDIAN JOURNAL OF AGRICULTURAL ECONOMICS

INDIAN SOCIETY OF
AGRICULTURAL ECONOMICS,
BOMBAY

Part I—Methodological Issues

METHODOLOGICAL ISSUES IN MEASURING AGRICULTURAL GROWTH: LESSONS OF RECENT INDIAN RESEARCHES

V. M. Rao*

The main source of methodological problems in the economist's measurement of agricultural growth is simply the facts of life which compel the researcher to begin his task from the wrong end. The researcher would have few methodological headaches if he were to start by working out the 'Output-input-weather' relationship at the level of a single crop grown by a homogeneous group of farmers in a homogeneous region. This is the level at which the 'text-book' statistical methodology would be most effective and, once the cropwise relationships are in hand, their aggregation over crops, groups of farmers and regions would provide the framework for analysing agricultural output changes including detection and measurement of trends over time. In practice, it is usual for the researcher to begin by looking at aggregate agricultural growth and his probings remain limited to a few broad decompositions and disaggregations by States and major crops; clearly, he operates in a zone where methodological criteria could be continual irritants except when he is reviewing other researchers' work !

This paper takes a brief look at the recent Indian researches concerned with the methodological points. By 'recent' we mean the post-Independence years and, particularly, the latter half of the period which has witnessed interesting debates on the visibility or otherwise of 'green revolution'. Brisk research activity to measure agricultural growth is largely a post-Independence phenomenon in India both because of substantial improvements in data base and the noticeable upward movement in agricultural production as compared to the situation in the pre-Independence period.[1] The methodological issues which have emerged in the course of these researches are typical of the problems met with in separating growth from fluctuations in relatively short spans of time-series data allowing little room for high-powered statistical techniques to move in. Our limited purpose here is to review the experience gained on these issues without getting into the deeper waters of problems faced in decomposition analysis and in estimating 'crop-input-weather' relationships which are being covered by other contributors.

What usually catches one's eye in an area of econometric research is the glittering display of alternative functional forms. They thrive best in situations of plentiful data for statistical play or where the researcher can employ *a priori* theoretical reasoning to pin down the forms of relationships. Analysis

* Senior Fellow, Institute for Social and Economic Change, Bangalore.

1. For a succinct account of problems of comparability, etc., arising in the course of improvement in data base, see George Blyn, "Measurement of Growth Rates in Agriculture", *Indian Journal of Agricultural Economics*, Vol. XXII, No. 1, January-March 1967.

of agricultural growth offers neither of these advantages. 25 or 30 annual observations, the current span of our production series, is hardly the kind of material to excite a statistician. Similarly, it is inconceivable that one can discern a stable 'law of growth' in changes in agricultural output affected by diverse influences—weather, technology, inputs, etc.,—having dissimilar and, in some cases, irregular time-profiles of change.

Given this context, there would be some temptation to researchers to look for only summary measures of agricultural growth without adopting formal statistical frameworks for estimation and interpretation. It is easy to see that this has indeed been the case with the researchers in India. Computations of annual linear or compound growth rates of output, area and yields now stand on par with the working out of averages and percentages and have percolated down to the level of routine reports and compilations which, in a strict sense, are not research writings. This is not to criticise the trend but only to point out that the approach to measuring agricultural growth is still largely 'arithmetic' rather than 'statistical' in nature; it is not unusual for the growth rates to be worked out on the basis of comparison of only 'end' years which should make this point all the more clear.

A risk attending this 'arithmetic' approach is a tendency on the part of researchers to be laymanish in using and interpreting growth rates. Blyn[2] points out that the linear growth rate (parameter b in the trend-equation $O_t = a + bt$) is often divided by the output level at the beginning of the period to get the compound rate which over-estimates the latter; the proper thing to do is to divide b by the average output over the period. In his Rapporteur's Report on the papers presented to the "Measurement of Agricultural Growth" session of the Twenty-sixth Annual Conference of the Indian Society of Agricultural Economics, Minhas[3] makes the point that, in strict logic, the average used for this purpose should be the harmonic mean of the series and not just any average! More interesting, Minhas notices lapses in computations and interpretations in practically every paper presented to the session. Clearly, this is a disturbing symptom since uncritical and mechanical minds would also be the ones having rather weak sensitivity to methodological issues and limited capacity to absorb good methodology.

Recent years have brought some relief from the monotony of linear and compound growth rates. The trend-equations underlying these rates are unsuitable when one has ground to suppose that growth may be proceeding at accelerated or decelerated rates rather than at constant rate. For example, it is plausible that the 'green revolution' in Indian agriculture has imparted a fresh momentum to growth which would suggest an acceleration in growth rates; equally plausible, the declining importance since the sixties of the contribution to growth by expansion in areas sown could have brought about deceleration. This is the kind of situation requiring relatively flexible func-

2. *op. cit.*
3. B. S. Minhas, "Rapporteur's Report on Measurement of Agricultural Growth", *Indian Journal of Agricultural Economics*, Vol. XXI, No. 4, October-December 1966 (see also the papers presented on "Measurement of Agricultural Growth" in the same issue).

tional forms for the trend-equation which could leave the question of acceleration or deceleration open for the data in hand to decide. Construction of such forms is a technical task which, additionally, has to conform to the requirements of estimation such as keeping the number of parameters to be estimated small and choosing forms amenable to relatively easy and standard techniques of parameter estimation. It is enough for our purpose to take a brief note of the functional forms used by the Indian researchers and their properties (see the Appendix at the end). Also, we content ourselves by noting the summary finding of the most recent study covering the period, 1949-50 to 1977-78—"while there has been a decline in the rate of growth of gross sown area, in particular under non-food crops in the decade starting from 1967-68 compared to the fifteen years ending in 1964-65, the output (and yield per unit area) of food crops and all crops grew more or less uniformly over the entire period with no evidence of either acceleration or deceleration since 1967-68."[4]

More relevant to note here are the methodological insights provided by these exercises. Proposing alternative functional forms and choosing between them in the light of data is the principal mode of application of mathematical-cum-statistical methods in empirical economics research. It brings into open the complexity of statistical reasoning involved in making the choice as also the fact that, frequently, even elaborate exercises end on a note of doubt and indecisiveness. These aspects are particularly worth noting by those who assume that there is nothing more to growth rates than a few calculations.

The findings of a recent study by Reddy[5] bear closely on the question of methodology. He fits all the four functional forms indicated in the Appendix to three time-series covering the period 1950-51 to 1973-74—(*i*) net national product, (*ii*) industrial production, and (*iii*) agricultural production. He notes at the outset that "Since we do not have any theoretical justification for the choice of a particular equation, we have to fall back on an entirely empirical approach." He discusses the limitations of criteria like minimum Residual Sum of Squares and highest $\bar{R}^2$; for example, all the four forms fit the NNP series equally well with high $\bar{R}^2$ values ranging between 0.983 and 0.991: "The implications of these models (for growth rates) are quite opposite in nature. . . . (and going by $\bar{R}^2$ alone) the researcher can get away with whatever he wanted to show." Faced with this situation, Reddy goes by the criterion that "it is ideal to choose that equation which provides the estimates of growth rates for the given period with minimum possible standard errors. . . Since the calculations are based on the sample, the interpretations have to be taken with care because the computed standard errors do not necessarily reflect the true precision of the estimates which is measured by the theoretical variance (which is) unknown." Reddy finds that in some cases this criterion also fails and, in these cases, he turns to "the criterion of accu-

4. T. N. Srinivasan, "Trends in Agriculture in India, 1949-50—1977-78", *Economic and Political Weekly*, Vol. XIV, Nos. 30, 31 and 32, Special Number, August 1979.

5. V. N. Reddy, "Growth Rates", *Economic and Political Weekly*, Vol. XIII, No. 19, May 13, 1978.

racy of prediction of growth rates." The following observation sums up the strategy he suggests in choosing among alternative functions: ". . . . a single rigid criterion may not resolve the issue of making a choice of growth curve. We have to take into account the totality of information regarding statistical measures of goodness of fit, accuracy of the growth rates, predictive capability of the estimated growth curve, etc., before choosing a particular growth curve." The point that we would like to stress is that such choices depending on multiple criteria are frequently difficult to make and involve the exercise of researcher's subjective judgment, a faculty nourished by research experience and subject-matter knowledge and not so much by grasp of formal statistical theory alone.

The situation considered above is one of several trend-equations fitting equally well the data in hand. In the presence of such a pronounced and obvious trend component, pursuing methodological issues must appear like quibbling on a relatively minor problem. We now turn to a situation where the statistical model of trend-fitting nearly reaches its wit's end in measuring growth. With all its armoury of formal techniques, the statistical model rests essentially on a simple-minded approach to measuring growth. It assumes that the profile over time of the variable under study can be decomposed into two neat components—(*a*) a systematic component representable by a smooth function of time, and (*b*) a *small* error component behaving like a random variable having a known and well-behaved probability distribution. Mukherjee and Vaidyanathan's paper[6] is of help in seeing how quickly the statistical model reaches the limits of its usefulness in analysing temporal variations in agriculture. They look at the State level yield per hectare of foodgrains for ten States with the time-series running from 1961-62 to 1975-76 for seven of these States and with the time-series including a stretch of earlier years in the case of the remaining three States. Compared to Reddy, Mukherjee and Vaidyanathan move from output to yield per hectare and step down from the aggregate all-India level to the level of States. As can be seen from their paper, both of these moves turn out to be sharply 'trend-weakening' in the sense that the $\bar{R}^2$ values of their trend-equations are in most cases too small to speak of a good fit.[7] Mukherjee and Vaidyanathan's response to this situation is to include indices of input and rainfall as arguments in their trend-equation; in effect, their analytical framework is not that of 'growth curve' but of 'yield-input-weather' relationship. This places their study outside the scope of this paper. We refer here only to what seem to us to be the two principal insights provided by their study. First, even the expanded functions which they try out do not, as a general case, achieve high $\bar{R}^2$ values—not a surprising outcome since inputs and weather would need more elaborate representation than is permitted by their data and stretch of time-series. Second, they pursue the residuals of their 'best-equations' beyond the standard sign and run tests of randomness to

6. Chandan Mukherjee and A. Vaidyanathan, "Growth and Fluctuations in Foodgrain Yields Per Hectare—A Statewise Analysis", published in this issue (see Part II).
7. Mukherjee and Vaidyanathan use as trend-equation functional form 3 shown in the Appendix.

find that, even when these standard tests are satisfied, the graphs of the residuals in several cases seem to bear traces of non-randomness. It seems to us that the Mukherjee and Vaidyanathan study brings out very well the need for careful scrutiny of the assumptions of the statistical model before accepting its findings.

Could the time-series data on agricultural production be put through a process of prior screening to bring them closer to the assumptions of the statistical model? Obviously, if feasible, such screening and the clues and criteria which help in screening would merit a place in a review of the methodological issues. Mentioned below are some of the commonly used procedures for this purpose. For periods marked by sudden and concentrated technological change, it is common for researchers to check the data for break in trends and to fit trend-equations by sub-periods. Since standardised statistical tests and procedures are available for doing this, we content ourselves by referring only to a recent study by Rudra,[8] which offers a clear illustration of the break in trend in the production of wheat in the late sixties in Punjab and Haryana and in Uttar Pradesh. It is also common for researchers to exclude weatherwise extreme years which could distort the estimated trend; for illustrations, the reader may see Srinivasan[9] and Rath,[10] particularly the latter. Exclusions are usually done on the basis of the researcher's judgment; as Rath explains: "We have used our *a priori* judgment to omit particular observations from the time-series. . . . Most of the observations excluded are (the) ones which appeared to be lower due to adverse weather, though there are a few cases of very high index numbers which have been omitted because they were very unusual or suspect."[11] Rao, Nadkarni and Deshpande[12] suggest that, in addition to such extreme years, a year with noticeably higher or lower production than its *adjoining* years may also be considered for exclusion; such 'local' peaks and troughs which may not be 'outliers' in statistical sense could still be distorting the trend. A major constraint on the screening of data for breaks in trend and for deviant weather is simply the shortness of time-series limiting both the number of sub-periods considered in estimation and the number of excluded years. Some marginal and modest help in better estimation of growth rate may be all that the researcher is able to squeeze out of such screening of data.

It is noticed that along each of the dimensions considered above—choosing among the alternative forms for the trend-equation, checking the residuals for randomness and screening the data to bring them closer to the assumptions of statistical model—there are unsettled issues and questions. It would be

8. Ashok Rudra, "Organisation of Agriculture for Rural Development: The Indian Case", *Cambridge Journal of Economics*, Vol. 2, No. 4, December 1978.
9. *op. cit.*
10. Nilakantha Rath, "A Note on Agricultural Production in India during 1955-78", published in this issue (see Part II).
11. An extreme variant tried out by Rath is to fit the trend only to the peak years.
12. V. M. Rao, M. V. Nadkarni and R. S. Deshpande, "Measurement of Growth and Fluctuations in Crop Output—An Approach Based on the Concept of Non-systematic Component," published in this issue.

idle to hope that improvements in data and statistical techniques, by themselves, would in due course resolve all of them. The indications, on the contrary, are that the statistical model will work with relative ease and validity at only two levels: (*i*) the aggregate all-India level when there is a dominating trend component in agricultural output, and (*ii*) the farm-cross-section level and experimental-data level analysis of 'output-input-weather' relationships. The intermediate levels of disaggregation by crops and States may always remain more difficult and inconvenient to handle with purely formal statistical approaches to measurement of growth. It is tempting to point out that even careful researchers capable of sophisticated econometric analysis often opt for a more informal approach when measuring growth at these intermediate levels; for example, the studies by Rath,[13] and Venkataramanan and Prahladachar[14] leave completely out the questions of goodness-of-fit and significance levels. Apparently—and, in our view, very rightly—, these researchers believe that it is important to judge one's findings in terms of their reasonableness and plausibility rather than by looking at the statistical caste-marks put on them by the computer.

A way to strengthen such informal statistical approaches is to give larger role and weight to items of information which could be used to supplement the 'hard' statistical data on area, production, rainfall, etc. Economic journalism now yields a plentiful harvest of narrative accounts of agricultural seasons and crops, weather conditions, market arrivals and prices, off-take by farmers of inputs like fertilizers and improved seeds distributed by public agencies, etc. A hint of drought or scarcity usually brings forth a spate of field reports on how the Government and the rural communities meet such critical times. Production drives, campaigns and crash programmes are the order of the day and so are the reports and post-mortems on them. Finally, even an esoteric subject like agricultural research continually endeavours to reach out to a wider readership through popular articles and write-ups. In fact, if only he would try with some patience and persistence, an agricultural economist should normally be in a position to put together a good descriptive account of how the production of a crop in a given State or region changed over recent years and what have been the main factors shaping its time-profile. This is particularly true of the extreme years which, if looked into in detail as individual instances, should provide helpful qualitative insights and findings on interactions among weather, controlled variables and policies. Researchers measuring agricultural growth are seen rarely to move away from the 'hard' data on area, production, etc., and, when they do so, it is usually only to eliminate an extreme year or to decide on sub-periods. A more positive approach would be to treat measurement of agricultural growth as a part of a larger task of building up a narrative account of behaviour of agricultural production over time. The task cannot be quite assigned to a computer but this could be its strong point in the sense of making

13. *op. cit.*
14. L. S. Venkataramanan and M. Prahladachar, "Growth Rates and Cropping Pattern Changes in Agriculture in Six States: 1950 to 1975", published in this issue (see Part II).

the researches on agricultural growth more substantive, meaningful and interesting to pursue. Interestingly, Reddy[15] and Krishnaji[16]—both statisticians themselves—seem to reach a similar position. Reddy suggests two non-parametric tests to bring out the behaviour of agricultural growth rate over time, *i.e.*, tendency on its part to accelerate or decelerate. Taking a hard look at the assumptions of the statistical model and their validity in agriculture, Krishnaji recommends that the best one can do in measuring agricultural growth is to describe the year-to-year changes. We can only exclaim: How true !

APPENDIX

Alternative Functional Forms Used for the Trend-Equation

1. *Linear Trend-Equation:*

Form: $O_t = a + bt$

Remarks

It is usual to estimate a and b by the Ordinary Least Squares (OLS) method. When the Durbin-Watson statistic indicates auto-correlation in the residuals or falls in the inconclusive region, the OLS method does not apply. In such cases, Reddy[17] re-estimates the equation by *assuming* a first-order auto-correlation in the residuals and using the *search technique.* For details, see Reddy's study. This remark also applies to other forms described below. Interestingly, Reddy does not scrutinise further the assumption of first-order auto-correlation in residuals indicating that it is rarely feasible in statistical practice to hope to check every assumption that one makes. In this form, b is in the nature of a constant absolute increment in output per unit of time and, in the normal case of positive b, implies decelerating compound growth rate of output. However, the prevailing research practice is to convert b into an *average* compound growth rate over the study period by dividing it by the output level. As has been indicated earlier in the text, the proper output level to be used for this purpose is the harmonic mean of the output series. According to Minhas,[18] the compound growth rate obtained thus is usually quite close to the growth rate given by the exponential trend-equation.

2. *Exponential Trend-Equation:*

Form: $O_t = Ae^{bt} = e^{(\log A + bt)}$

i.e., $\log O_t = \log A + bt$ - - (continuous t case)

$O_t = A(1 + r)^t$

i.e., $\log O_t = \log A + \log(1 + r)\, t$ - - (discrete t case)

Remarks

This form assumes the compound growth rate (b or r) to be constant over time. It is seen that researchers commonly use the form appropriate for continuous t, presumably because it gives the compound rate directly rather than in a transformed form as in the case of discrete t. In strict logic, b, by itself, cannot be treated as an *annual* rate but, according to Srinivasan,[19] the error involved would be minor when the growth process is slow. It is usual for the parameters in this form to be estimated by the OLS method applied to the logarithmic version under the assumption of an additive disturbance term. An interesting point to note is that Srinivasan mentions in passing that if exponential trend-equations are assumed for individual crops, the same form cannot be assumed, consistently, for aggregate agricultural output. This is a familiar problem in aggregation of relationships and, very likely, one would begin to hear more of it when researchers try seriously to link the behaviour of agricultural output over a large areal unit with cropwise and regionwise 'output-input-weather' relationships.

---

15. *op. cit.*
16. N. Krishnaji, "Measuring Agricultural Growth", published in this issue.
17. *op. cit.*
18. *op. cit.*
19. *op. cit.*

3. *Modified Exponential Trend-Equation:*

Form: $\log O_t = a + bt + c t^2$

Remarks

In this form, the compound growth rate changes over time being equal to b + 2 ct. Assuming positive b, the sign of c would indicate whether there is any acceleration or deceleration in the growth of output. In other words, the form leaves this question open for the estimated c (*i.e.*, for the data in hand) to decide. As regards estimation of parameters, the remarks made above apply also to this form. A problem posed by this form is that of multicollinearity because of correlation between t and t^2. Reddy,[20] and Mukherjee and Vaidyanathan[21] use the transformation $t' = t - \left(\frac{n+1}{2}\right)$ to overcome this problem (n is the number of years in the time-series data). An interesting feature pointed out by Srinivasan is that the simple average of the changing growth rates b + 2 ct equals the growth rate given by the exponential trend-equation when the period of estimation does not involve any gaps. It is worth noting that the average compound growth rates estimated for a period by all the three trend-equations considered above are constrained to be close to each other.

4. *Gompertz Trend-Equation:*

Form: $O_t = ab^{c^t}$ ......(a, b and c positive)

Remarks

Like form 3, this form also leaves open the question of acceleration or deceleration in the growth rate. It can be shown that both b and c greater than 1 implies acceleration and both b and c less than 1 implies deceleration. The literature that we consulted yields only two instances of the use of this form—Reddy[22] and Dey[23] (quoted by Rudra). The reader may see Handbook of Methods of Applied Statistics, J. Roy, R. G. Laha and I. M. Chakravarti, John Wiley & Sons, New York, 1967, for a procedure to fit the form to time-series data. Reddy, who gives this reference, points out the limitations of the suggested procedure. He works with transformations permitting the use of standard regression techniques to estimate the parameters. A point which remains obscure in the limited literature that we have consulted is whether this form has any clear advantage over the form 3 mentioned above.

Search for an appropriate functional form for a relationship needs a framework to visualise the plausible or desirable properties of the relationship as suggested by the relevant subject-matter theory and prior information. It would seem obvious that the sole idea that the agricultural growth may be proceeding at a non-constant rate could hardly be an informative enough framework to guide the search for an appropriate trend-equation for agricultural output. It would also seem from the results obtained so far that the alternative forms proposed for the trend-equation trace closely similar time-paths over the period of observation so that the data in hand afford little help in choosing between them. Apparently, the researchers have not been sufficiently ingenious and innovative in proposing the alternative forms. However, the extrapolations beyond the period of observation, differing sharply as between the alternative forms, could provide a helpful basis for choice. This underlines the importance of comparing the alternative forms in terms of their predictions, a procedure not in common use among researchers. Standard statistical tests are available for judging predictive accuracy and there would be better scope to use them with the increase in the span of the time-series.

One can still argue that the functional forms covering trend-component alone could never be fully adequate in dealing with changes in agricultural output. It is here that the researchers' understanding of and insights into the year-to-year changes in agricultural output assume importance. The strategy should be to progressively strengthen such understanding and insights and to link them up with the researches on alternative functional forms. There would then be some hope of progress towards developing a more informative framework for designing the alternative functional forms and incorporating more of agricultural reality in their properties.

20. *op. cit.*
21. *op. cit.*
22. *op. cit.*
23. A. K. Dey, "Rates of Growth of Agriculture and Industry", *Economic and Political Weekly*, Vol. X, Nos. 25 and 26, June 21 and 28, 1975.